



# Evaluation of a cosmic-ray neutron sensor network for improved land surface model prediction

Roland Baatz<sup>1,2</sup>, Harrie-Jan Hendricks Franssen<sup>1,2</sup>, Xujun Han<sup>1,2</sup>, Tim Hoar<sup>3</sup>, Heye Reemt Bogena<sup>1</sup>, and Harry Vereecken<sup>1,2</sup>

<sup>1</sup>Agrosphere (IBG-3), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

<sup>2</sup>HPSC-TerrSys, 52425 Jülich, Germany

<sup>3</sup>NCAR Data Assimilation Research Section, Boulder, CO, USA

Correspondence to: Roland Baatz (r.baatz@fz-juelich.de)

Received: 22 August 2016 – Discussion started: 26 August 2016

Revised: 12 April 2017 – Accepted: 19 April 2017 – Published: 16 May 2017

**Abstract.** In situ soil moisture sensors provide highly accurate but very local soil moisture measurements, while remotely sensed soil moisture is strongly affected by vegetation and surface roughness. In contrast, cosmic-ray neutron sensors (CRNSs) allow highly accurate soil moisture estimation on the field scale which could be valuable to improve land surface model predictions. In this study, the potential of a network of CRNSs installed in the 2354 km<sup>2</sup> Rur catchment (Germany) for estimating soil hydraulic parameters and improving soil moisture states was tested. Data measured by the CRNSs were assimilated with the local ensemble transform Kalman filter in the Community Land Model version 4.5. Data of four, eight and nine CRNSs were assimilated for the years 2011 and 2012 (with and without soil hydraulic parameter estimation), followed by a verification year 2013 without data assimilation. This was done using (i) a regional high-resolution soil map, (ii) the FAO soil map and (iii) an erroneous, biased soil map as input information for the simulations. For the regional soil map, soil moisture characterization was only improved in the assimilation period but not in the verification period. For the FAO soil map and the biased soil map, soil moisture predictions improved strongly to a root mean square error of 0.03 cm<sup>3</sup> cm<sup>-3</sup> for the assimilation period and 0.05 cm<sup>3</sup> cm<sup>-3</sup> for the evaluation period. Improvements were limited by the measurement error of CRNSs (0.03 cm<sup>3</sup> cm<sup>-3</sup>). The positive results obtained with data assimilation of nine CRNSs were confirmed by the jackknife experiments with four and eight CRNSs used for assimilation. The results demonstrate that assimilated data of a CRNS network can improve the characterization of soil

moisture content on the catchment scale by updating spatially distributed soil hydraulic parameters of a land surface model.

## 1 Introduction

Soil water content (SWC) is a key variable of land surface hydrology and has a strong control on the partitioning of net radiation between latent and sensible heat flux (Brutsaert, 2005). Knowledge of SWC is relevant for the assessment of plant water stress and agricultural production, as well as runoff generation as a response to precipitation events (Vereecken et al., 2008; Robinson et al., 2008). In atmospheric circulation models, SWC is important as a lower boundary condition, while it is calculated as a state variable in land surface models. Coupling of atmospheric circulation models and land surface models allows for the quantification of the role of soil moisture on atmospheric processes such as soil moisture–precipitation feedbacks (Koster et al., 2004; Eltahir, 1998) and summer climate variability and drought (Seneviratne et al., 2006; Oglesby and Erickson, 1989). It is therefore important to improve the modelling and prediction of SWC. Data assimilation of soil moisture provides a way to improve imperfect land surface model predictions. Here, soil moisture measurements are used to update model predictions by optimally considering the uncertainty of model initial conditions, model parameters and model forcings. However, there is a lack of high-quality soil moisture data (Vereecken et al., 2016). Soil moisture measured by spaceborne remote sensing technologies provides information over

large areas but is strongly affected by vegetation and surface roughness (e.g. Temimi et al., 2014). Therefore, in this paper an alternative source for soil moisture information is explored which can measure soil moisture more accurately under dense vegetation (Bogena et al., 2013). Cosmic-ray neutron sensors (CRNSs) measure fast neutron intensity on an intermediate scale of  $\sim 15$  ha (Kohli et al., 2015; Zreda et al., 2008), which is the desired application scale of land surface models (Ajami et al., 2014; Chen et al., 2007; Shrestha et al., 2014). Fast neutrons originate from collisions of secondary cosmic particles from outer space with terrestrial atoms. Fast neutrons in turn are moderated most effectively by hydrogen because the mass of a neutron is similar to that of a nucleus of the hydrogen atom. Therefore, the corresponding fast neutron intensity measured by CRNSs strongly depends on the amount of hydrogen within the CRNS footprint, allowing for a continuous non-invasive soil moisture estimate on the field scale. The spatial extent of this measurement is desirable as it matches with the desired grid cell size of a high-resolution land surface model (Crow et al., 2012), and small scale heterogeneities are averaged over a larger area (Franz et al., 2013; Kohli et al., 2015). Vertical measurement depth ranges from a maximum of  $\sim 70$  cm under completely dry conditions and decreases to roughly  $\sim 12$  cm under wet conditions (e.g. 40 vol. % soil moisture) (Kohli et al., 2015; Franz et al., 2012). Worldwide several CRNS networks exist, such as the North American COSMOS network (Zreda et al., 2012), the German CRNS network (Baatz et al., 2014) installed in the context of the TERENO infrastructure measure (Zacharias et al., 2011), the Australian COSMoZ network (Hawdon et al., 2014) and the British COSMOS-UK (Evans et al., 2016).

In this work, fast neutron intensity data measured by CRNSs are assimilated in a land surface model, to evaluate the impact those data can have on improving soil moisture characterization and land surface model predictions. The ensemble Kalman filtering (EnKF) is one of the most commonly applied data assimilation methods (Evensen, 1994; Burgers et al., 1998). The EnKF is much less CPU-intensive compared to alternative methods such as the particle filter (e.g. Montzka et al., 2011), because for high-dimensional problems the EnKF requires a much smaller ensemble size to achieve reasonably good predictions. The ensemble Kalman filter (Reichle et al., 2002a; Dunne and Entekhabi, 2005; Crow, 2003; De Lannoy and Reichle, 2016), variants such as the extended Kalman filter (Draper et al., 2009; Reichle et al., 2002b) and the local ensemble transform Kalman filter (Han et al., 2013, 2015) were applied for updating soil moisture states in land surface models. Reichle et al. (2002a) performed a synthetic experiment using L-band microwave observations of the Southern Great Plains Hydrology Experiment (Jackson et al., 1999) to analyse the effect of ensemble size and forecast errors. Dunne and Entekhabi (2005) showed that an ensemble Kalman smoother approach, where data from multiple time steps were assimilated to update current and past states, can yield a reduced prediction error compared

to a pure filtering approach. More recently, state updates with the EnKF were tested for the Soil Moisture Ocean Salinity (SMOS, Kerr et al., 2012) mission. De Lannoy and Reichle (2016) assimilated SMOS temperature brightness and soil moisture retrievals into a land surface model with large improvements in surface soil moisture. However, localized error patterns were not captured well enough, and locally optimized EnKF error parameters would improve prediction results further.

More recent work addressed joint updating of model states and parameters in hydrologic and land surface models with data assimilation methods. Joint state–parameter estimation with EnKF is possible by an augmented state vector approach (Chen and Zhang, 2006), a dual approach (Moradkhani et al., 2005) or an approach with an additional external optimization loop (Vrugt et al., 2005). In the augmented state vector approach, parameters are included in the state vector and are updated via cross-covariances between states and parameters. The cross-covariances are estimated from the ensemble. In the dual approach, first parameters are updated by data assimilation, and the assimilation step is repeated with the updated parameters to also update the states by data assimilation. In the approach with an external optimization loop the parameters are not updated by EnKF, but in an external optimization loop. Pauwels et al. (2009) were one of the first to optimize soil hydraulic parameters of a land surface model by data assimilation, assimilating synthetic aperture radar data. Lee (2014) used synthetic aperture radar soil moisture data to estimate soil hydraulic properties at the Tibetan plateau using the EnKF and a soil–vegetation–atmosphere transfer model. Bateni and Entekhabi (2012) assimilated land surface temperature with an ensemble Kalman smoother and achieved a better estimate of the partitioning of energy between sensible and latent heat fluxes. Han et al. (2014) updated soil hydraulic parameters of the Community Land Model version 4.5 (CLM) by assimilation of synthetic brightness temperature data with the local ensemble transform Kalman filter (LETKF) (Hunt et al., 2007) and showed the potential of this approach for improving land surface states and fluxes like evapotranspiration. Shi et al. (2014) used the ensemble Kalman filter for a synthetic multivariate data assimilation problem with a land surface model and then applied it to real data (Shi et al., 2015). Both cases illustrate that parameters from different compartments can be updated successfully by multivariate data assimilation. Kurtz et al. (2016) developed a particular CPU-efficient data assimilation framework for the coupled land surface–subsurface model TerrSysMP (Shrestha et al., 2014). They successfully updated  $2 \times 10^7$  states and parameters in a synthetic experiment. Whereas these studies were made with land surface models, in soil hydrological applications recently data assimilation was also used to estimate soil hydraulic parameters. Early work was by Wu and Margulis (2011, 2013) in the context of real-time control of wastewater reuse in irrigation and also showed the potential of EnKF in soil hydro-

ogy. Montzka et al. (2011, 2013) explored the role of the particle filter for handling non-Gaussianity in soil hydrology data assimilation. They showed that the nonlinear character of the soil moisture retention characteristic is critical for joint state–parameter estimation in data assimilation systems and showed that the particle filter is an interesting alternative for soil hydraulic parameter estimation for 1-D problems. Erdal et al. (2014) investigated the role of bias in the conceptual soil model and explored bias-aware EnKF as a way to deal with it. They argued that the exact location of soil layers is often not known and that this can severely deteriorate the performance of EnKF. Song et al. (2014) worked on a modified iterative EnKF-based filter to handle the non-linearity and non-Gaussianity of data assimilation for the vadose zone. They proposed a modified procedure which avoids the high CPU need of a fully iterative method, but which still gives stable results. Erdal et al. (2015) also focussed on handling of strong non-Gaussianity of the state variable in EnKF under very dry conditions. They showed that classical EnKF fails under such conditions and proposed two alternative strategies, both involving transformation of state variables, which also performed favourably under very dry conditions with strongly skewed pressure distributions. All these studies on joint state–parameter estimation showed in general that estimation of soil hydraulic or land surface parameters improves model predictions (strongly), but can be unstable for strongly non-Gaussian distributions and nonlinear problems. For a further literature review on data assimilation in the context of hydrological and land surface models, we refer to Reichle (2008) and Montzka et al. (2012).

Shuttleworth et al. (2013) developed the Cosmic Ray Soil Moisture Interaction Code (COSMIC), which is a forward operator to be applied for assimilating neutron intensity observations from CRNS. The COSMIC code was evaluated for several sites (Baatz et al., 2014; Rosolem et al., 2014). The COSMIC operator was successfully implemented in the Data Assimilation Research Testbed (Rosolem et al., 2014) to allow for state updating by the ensemble adjustment Kalman filter (Anderson, 2001). The surface soil moisture information was propagated into greater soil depth than only the measurement depth using COSMIC in combination with data assimilation (Rosolem et al., 2014). The COSMIC operator was implemented in a python interface that couples the land surface model CLM and the LETKF for joint state–parameter updating (Han et al., 2015). Neutron counts measured by CRNSs have been used in data assimilation studies to update model states (Han et al., 2015; Rosolem et al., 2014). Soil hydraulic parameters were also updated by assimilation of neutron counts in one synthetic study (Han et al., 2016), showing its feasibility. CRNSs were also used for inverse estimation of soil hydraulic parameters of the HYDRUS-1D model (Villarreyes et al., 2014).

This work further explores the value of measured neutron intensity by CRNSs to improve modelling of terrestrial sys-

tems on the catchment scale (Simmer et al., 2015) using a land surface model. The main novelties are as follows:

- i. Data from a network of nine CRNSs were assimilated in the CLM with an evaluation of the information gain by this assimilation at the catchment scale. Until now, evaluations with CRNSs were made for a single location, but not for a complete network of CRNSs. A very important question is whether CRNSs can also improve the soil moisture characterization on the catchment scale. The high variability of soil moisture at a short distance could potentially limit the CRNS measurement value and make updating of soil moisture contents further away from the sensor meaningless. Conversely, soil moisture, soil maps and atmospheric forcings show spatial correlations over larger distances (Kirkpatrick et al., 2014; Korres et al., 2015), which suggests that CRNS measurements potentially carry important information to update soil moisture contents for larger regions (e.g. Han et al., 2012). If it is found that CRNS networks with a density such as that in this study (nine stations per 2354 km<sup>2</sup>) can improve soil moisture content characterization on the catchment scale, this is of high relevance and importance for agricultural applications, flood prediction and protection, and regional weather prediction (Whan et al., 2015; Koster et al., 2004; Seneviratne et al., 2010). The main research question addressed in this paper is therefore whether a CRNS network of the density as in this study can improve large-scale soil moisture characterization.
- ii. Soil hydraulic parameters are updated in this study together with the soil moisture states in a real-world case study. The study in this paper also allows some evaluation of the updated large-scale soil hydraulic parameters.

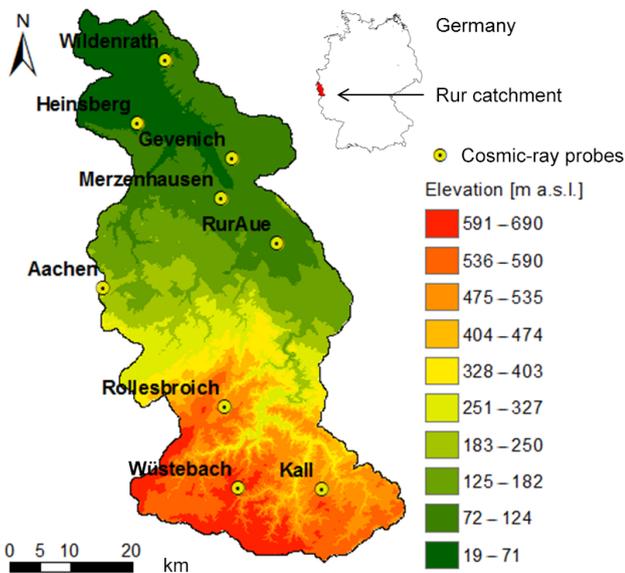
## 2 Materials and methods

### 2.1 Site description and measurements

The model domain, the Rur catchment (2354 km<sup>2</sup>), is situated in western Germany and illustrated in Fig. 1. The altitude varies between 15 m a.s.l. in the flat northern part and 690 m a.s.l. in the hilly southern part. Precipitation, evapotranspiration and land use follow the topography. The dominant land use types are agriculture (mainly in the north), grassland, and coniferous and deciduous forest. Annual precipitation ranges between less than 600 mm in the north and 1200 mm in the hilly south (Montzka et al., 2008). Annual potential evapotranspiration varies between 500 mm in the south and 700 mm in the north (Bogena et al., 2005). The Rur catchment CRNS network comprises nine CRNSs (CRS1000, HydroInnova LLC, 2009) which were installed

**Table 1.** Site information on elevation (m a.s.l.), average annual precipitation ( $\text{mm yr}^{-1}$ ), CLM plant functional type (Bonan et al., 2002), sand content (%), clay content (%) and the date of the first SWC retrieval assimilated.

Name	m a.s.l.	Precip.	Plant functional type	Sand	Clay	Date of first assimilation
Aachen	232	952	Crops	22	23	13 Jan 2012
Gevenich	108	884	Crops	22	20	7 Jul 2011
Heinsberg	57	814	Crops	18	19	9 Sep 2011
Kall	504	935	C3 grass	20	22	15 Sep 2011
Merzenhausen	94	825	Crops	21	22	19 May 2011
Rollesbroich	515	1307	C3 grass	22	23	19 May 2011
RurAue	102	743	C3 grass	19	26	8 Nov 2011
Wildenrath	76	856	Broadleaf deciduous temperate tree	65	12	7 May 2012
Wüstebach	605	1401	Needleleaf evergreen temperate tree	19	23	20 Mar 2011



**Figure 1.** Map of the Rur catchment and locations of the nine cosmic-ray neutron sensors. The hilly south of the catchment is prone to more rainfall, lower average temperatures and less potential evapotranspiration than the north of the catchment.

in 2011 and 2012 (Baatz et al., 2014). Climate and soil texture of the CRNS sites can be found in Table 1.

The CRNSs were calibrated in the field using gravimetric soil samples. At each site, 18 soil samples were taken along 3 circles with distances of 25, 75 and 175 m from the CRNS, and 6 samples were evenly distributed along each circle. Each sample was extracted with a  $50.8 \times 300$  mm round HUMAX soil corer (Martin Burch AG, Switzerland). The samples were split into 6 sub-samples with 5 cm length each and oven dried at  $105^\circ\text{C}$  for 48 h to measure dry soil bulk density and soil moisture. Lattice water, hydrogen from organic and an-organic sources, was determined for each site

using a heat conductivity detector (Ray, 1954). Soil bulk density, soil moisture, lattice water and 12 h averaged measured neutron intensity were used to determine calibration parameters specific to each CRNS and the COSMIC operator. This represents a compromise between the measurement noise (which follows a Poisson distribution) and the assumed variation of environmental variables over the averaging time window (Iwema et al., 2015).

## 2.2 Community Land Model and parameterization

The CLM was the land surface model of choice for simulating water and energy exchange between the land surface and the atmosphere (Oleson et al., 2013). Some of the key processes which are modelled by CLM are radiative transfer in the canopy space, interception of precipitation by the vegetation and evaporation from intercepted water, water uptake by vegetation and transpiration, soil evaporation, and photosynthesis, as well as water and energy flow in the subsurface. SWC in CLM is influenced by precipitation, infiltration into the soil, water uptake by vegetation, surface evaporation, and surface and subsurface runoff. To limit the scope and complexity of this study, CLM was run using satellite phenology, e.g. prescribed leaf area index data and the biogeochemical module turned off. The biogeochemical module allows CLM to model the vegetation development dynamically, but it requires a large spin-up of 1000 years, and little additional gain is expected for this study from these additionally modelled processes.

Vertical water flow in soils is modelled by the 1-D Richards equation. Soil hydraulic parameters are determined from sand and clay content using pedotransfer functions for the mineral soil fraction (Clapp and Hornberger, 1978; Cosby et al., 1984) and organic matter content for the organic soil fraction (Lawrence and Slater, 2008).

The joint state–parameter estimation used in this study updates soil texture and organic matter in CLM. Hence, param-

eter estimates directly determine soil hydraulic properties in CLM. The following equations describe how soil texture and organic matter define the soil hydraulic properties in CLM such as porosity, hydraulic conductivity, the empirical exponent  $B$  and soil matric potential. Hydraulic conductivity ( $k(z_i)$ ,  $\text{mm s}^{-1}$ ) at the depth  $z$  between two layers ( $i$  and  $i + 1$ ) is a function of soil moisture ( $\theta$ ,  $\text{m}^3 \text{m}^{-3}$  in layers  $i$  and  $i + 1$ ), saturated hydraulic conductivity ( $k_{\text{sat}}(z_i)$ ,  $\text{mm s}^{-1}$ ), saturated soil moisture ( $\theta_{\text{sat}}$ ,  $\text{m}^3 \text{m}^{-3}$ ) and the empirical exponent  $B$  (Oleson et al., 2013):

$$k(z_i) = \begin{cases} \phi_{\text{ice}} k_{\text{sat}}(z_i) \left[ \frac{(\theta_i + \theta_{i+1})}{(\theta_{\text{sat},i} + \theta_{\text{sat},i+1})} \right]^{2B_i+3}, & 1 \leq i \leq N_{\text{levsoil}} - 1 \\ \phi_{\text{ice}} k_{\text{sat}}(z_i) \left( \frac{\theta_i}{\theta_{\text{sat},i}} \right)^{2B_i+3}, & i = N_{\text{levsoil}} \end{cases}, \quad (1)$$

where  $\phi_{\text{ice}}$  is the ice impedance factor. The ice impedance factor was implemented to simplify an increased tortuosity of water flow in a partly frozen pore space. It is calculated with  $\phi_{\text{ice}} = 10^{-\Omega F_{\text{ice}}}$  using the resistance factor  $\Omega = 6$  and the frozen fraction of soil porosity  $F_{\text{ice}} = \theta_{\text{ice}}/\theta_{\text{sat},i}$ . Soil hydraulic properties are calculated separately for the mineral (min) and organic matter (om) soil components. Total porosity  $\theta_{\text{sat},i}$  is calculated using the fraction of organic matter ( $f_{\text{om},i}$ ) with the following:

$$\theta_{\text{sat},i} = (1 - f_{\text{om},i}) \theta_{\text{sat},\text{min},i} + f_{\text{om},i} \theta_{\text{sat},\text{om}}, \quad (2)$$

where the organic matter porosity is  $\theta_{\text{sat},\text{om}} = 0.9$  and sand content in percentage (%) determines the mineral soil porosity  $\theta_{\text{sat},\text{min}}$  as follows:

$$\theta_{\text{sat},\text{min}} = 0.489 - 0.00126 \times \% \text{sand}. \quad (3)$$

Analogous, the exponent  $B$  is calculated with

$$B_i = (1 - f_{\text{om},i}) B_{\text{min},i} + f_{\text{om},i} B_{\text{om}}, \quad (4)$$

where  $B_{\text{om}} = 2.7$  is the organic exponent and the mineral exponent  $B_{\text{min},i}$  is determined by clay content in percentage (%) with the following:

$$B_{\text{min},i} = 2.91 + 0.159 \times \% \text{clay}. \quad (5)$$

Saturated hydraulic conductivity is calculated for a connected and an unconnected fraction of the grid cell with the following:

$$k_{\text{sat}}(z_i) = (1 - f_{\text{perc}}) k_{\text{sat},\text{uncon}}(z_i) + f_{\text{perc},i} k_{\text{sat},\text{om}}(z_i), \quad (6)$$

where  $f_{\text{perc},i}$  is the fraction of a grid cell where water flows with saturated hydraulic conductivity of the organic matter ( $k_{\text{sat},\text{om}}(z_i)$  in  $\text{mm s}^{-1}$ ) through the organic material only, the so-called connected flow pathway. The saturated hydraulic conductivity of the unconnected part ( $k_{\text{sat},\text{uncon}}(z_i)$ ,  $\text{mm s}^{-1}$ ) depends on organic and saturated mineral soil hydraulic conductivity:

$$k_{\text{sat},\text{uncon}} = (1 - f_{\text{perc}}) \left( \frac{1 - f_{\text{om}}}{k_{\text{sat},\text{min}}} + \frac{f_{\text{om}} - f_{\text{perc}}}{k_{\text{sat},\text{om}}} \right)^{-1}, \quad (7)$$

where saturated hydraulic conductivity for mineral soil is calculated from the grid cell sand content as follows:

$$k_{\text{sat},\text{min}}[z_i] = 0.0070556 \times 10^{-0.884+0.0153 \times \% \text{sand}}. \quad (8)$$

The fraction  $f_{\text{perc}}$  is calculated with the following:

$$f_{\text{perc}} = \begin{cases} 0.908 \times (f_{\text{om}} - 0.5)^{0.139}, & f_{\text{om}} \geq 0.5 \\ 0, & f_{\text{om}} < 0.5 \end{cases}. \quad (9)$$

Soil matric potential (mm) is defined as a function of saturated soil matric potential (mm) with the following:

$$\psi_i = \psi_{\text{sat},i} \left( \frac{\theta_i}{\theta_{\text{sat},i}} \right)^{-B_i} = [(1 - f_{\text{om},i}) \psi_{\text{sat},\text{min},i} + f_{\text{om},i} \psi_{\text{sat},\text{om}}] \left( \frac{\theta_i}{\theta_{\text{sat},i}} \right)^{-B_i}, \quad (10)$$

where saturated organic matter matric potential is  $\psi_{\text{sat},\text{om}} = -10.3$  mm and saturated mineral soil matric potential is calculated from sand content as follows:

$$\psi_{\text{sat},\text{min},i} = -10.0 \times 10^{1.88-0.0131 \times \% \text{sand}} \quad (11)$$

### 2.3 Cosmic-ray forward model

SWC retrievals were calculated from neutron intensity observations with COSMIC (Shuttleworth et al., 2013) following calibration results and the procedure of Baatz et al. (2014). COSMIC parameterizes neutron transport within the soil subsurface and was calibrated against the more complex Monte Carlo Neutron Particle model MCNPx (Pelowitz, 2005). COSMIC needs considerably less CPU time than the MCNPx model. The code was tested at multiple sites for soil moisture determination (Baatz et al., 2014; Rosolem et al., 2014) and analysed in detail by Rosolem et al. (2014).

COSMIC assumes that a number of high-energy neutrons enter the soil. In the soil, the number of high-energy neutrons is reduced by interactions within the soil, leading to generation of fast neutrons in each soil layer. Before resurfacing, the number of fast neutrons is reduced again by their interaction with nuclei of elements within soil (Shuttleworth et al., 2013). The number of neutrons,  $N_{\text{CRP}}$ , that reaches the CRNS can be summarized in a single integral as follows:

$$N_{\text{CRP}} = N_{\text{COSMIC}} \int_0^{\infty} \left\{ A(z) [\alpha \rho_{\text{s}}(z) + \rho_{\text{w}}(z)] \exp \left( - \left[ \frac{m_{\text{s}}(z)}{L_1} + \frac{m_{\text{w}}(z)}{L_2} \right] \right) \right\} \cdot dz, \quad (12)$$

where  $N_{\text{COSMIC}}$  is an empirical coefficient that is CRNS-specific and needs to be estimated by calibration,

$A(z)$  is the integrated average attenuation of fast neutrons,  $\alpha = 0.404 - 0.101 \times \rho_S$  is the site-specific empirical coefficient for the creation of fast neutrons by soil,  $\rho_S$  is the dry soil bulk density ( $\text{g cm}^{-3}$ ),  $\rho_w$  is the total soil water density ( $\text{g cm}^{-3}$ ), and  $m_s$  and  $m_w$  are the masses of soil and water, respectively, per area ( $\text{g cm}^{-2}$ )  $L_1 = 162.0 \text{ g cm}^{-2}$  and  $L_2 = 129.1 \text{ g cm}^{-2}$  are scattering lengths for fast neutrons in solids and water, respectively, that were estimated using the MCNPx code (Shuttleworth et al., 2013). The integrated average attenuation of fast neutrons  $A(z)$  can be found numerically by solving

$$A(z) = \left(\frac{2}{\pi}\right) \int_0^\pi /2 \exp\left(\frac{-1}{\cos(\gamma)} \left[\frac{m_s(z)}{L_3} + \frac{m_w(z)}{L_4}\right]\right) \cdot d\gamma, \quad (13)$$

where  $\gamma$  is the angle along a vertical line below the CRNS detector to the element that contributes to the attenuation of fast neutrons, and  $L_3 = -31.65 + 99.29 \times \rho_S$  and  $L_4 = 3.16 \text{ g cm}^{-2}$  are the scattering lengths for fast neutrons in soil and water, respectively, determined using the MCNPx code (Shuttleworth et al., 2013). The COSMIC operator is discretized into 300 layers of 1 cm thickness up to a depth of 3 m. For each CLM grid cell in the model domain, simulated SWC in all CLM layers is used to generate a weighted SWC retrieval using the COSMIC code. Simulated SWC is handed from the CLM simulation history files to the COSMIC operator. Given the vertical SWC distribution of the individual CLM grid cell, COSMIC internally calculates the contribution of each layer to the simulated neutron intensity signal at the soil surface in COSMIC. In this study, the contribution of each CLM soil layer was used to calculate the weighted CLM SWC retrieval corresponding to the vertical distribution of simulated SWC in each grid cell.

Measured neutron intensity of CRNS was used to inversely determine a CRNS SWC retrieval, as by Baatz et al. (2014) assuming a homogeneous vertical SWC distribution. Then, the weighted CLM SWC retrieval is used in the data assimilation scheme to relate the CRNS SWC retrieval to the model state. Alternatively, neutron flux data could be assimilated directly within the catchment. This would require calibration data throughout the catchment, which is only feasible using spatially distributed data sets (e.g. Avery et al., 2016). However, high stands of biomass are a major factor for calibration in the Rur catchment (Baatz et al., 2015), and estimates of biomass come along with high uncertainties. To circumvent the introduction of these additional uncertainties, SWC retrievals are assimilated in this study. Changes in on-site biomass were assumed to be negligible.

## 2.4 Data assimilation

To further expand the work of Han et al. (2016), this study uses the LETKF (Hunt et al., 2007) to assimilate SWC retrievals by CRNSs into the land surface model CLM. Updates were calculated either for SWC states or jointly for SWC states and soil parameters, depending on the experiment setup. For state updates only, the LETKF was used as proposed by Hunt et al. (2007). Calculations were made for an ensemble of model simulations which differed depending on variations in model forcings and input parameters. The states of the different ensemble members are indicated by  $\mathbf{x}_i^f$  where  $i = 1, \dots, N$  and  $N$  is the number of ensemble members; “f” marks the model prediction or forecast before the update. The individual state vectors  $\mathbf{x}_i^f$  contain the CLM-simulated SWC of the 10 soil layers and the vertically weighted SWC retrieval obtained with the COSMIC operator. For each grid cell, a matrix  $\mathbf{X}^f$  can be constructed which contains the deviations of the simulated states with respect to the ensemble mean  $\bar{\mathbf{x}}^f$ :

$$\mathbf{X}^f = \left[ \mathbf{x}_1^f - \bar{\mathbf{x}}^f, \dots, \mathbf{x}_N^f - \bar{\mathbf{x}}^f \right]. \quad (14)$$

In the case of joint state–parameter updates, a state augmentation approach was followed (Hendricks Franssen and Kinzelbach, 2008; Han et al., 2014). In this case, the augmented model state matrix  $\mathbf{X}^f$  is constructed from the simulated SWC of the 10 soil layers, weighted SWC, and the grid cell’s sand, clay and organic matter content.

In order to relate the measured neutron intensity with the simulated SWC of CLM, the observation operator  $\mathbf{H}$  (COSMIC) is applied on the measured neutron intensity in order to obtain the expected weighted SWC retrieval at each of the observation locations for each of the stochastic realizations:

$$\mathbf{y}_i^f = \mathbf{H} \left( \mathbf{x}_i^f \right). \quad (15)$$

The ensemble realizations of the modelled SWC retrievals at the measurement locations  $\mathbf{y}_1^f$  to  $\mathbf{y}_N^f$  with respect to the ensemble mean  $\bar{\mathbf{y}}^f$  are stored in the matrix  $\mathbf{Y}^f$ :

$$\mathbf{Y}^f = \left[ \mathbf{y}_1^f - \bar{\mathbf{y}}^f, \dots, \mathbf{y}_N^f - \bar{\mathbf{y}}^f \right]. \quad (16)$$

The observation error correlation was reduced in space by the factor  $f_{\text{red}}$  using the spherical model:

$$f_{\text{red}} = 1 - (1.5 \times d/d_{\text{max}}) + \left(0.5 \times [d/d_{\text{max}}]^3\right), \quad (17)$$

where  $d$  is the distance to the observation and  $d_{\text{max}} = 40 \text{ km}$  is the maximum observation correlation length, about half the size of the catchment. Only SWC retrievals within the maximum observation correlation length were used for assimilation. This leads to a “localized” size of  $\mathbf{Y}^f$  and the observation error covariance matrix  $\mathbf{R}$ . The intermediate covariance matrix  $\mathbf{P}^a$  (also called analysis error covariance matrix) is calculated according to the following:

$$\mathbf{P}^a = \left[ (N - 1)\mathbf{I} + \mathbf{Y}^f \mathbf{R}^{-1} \mathbf{Y}^f \right]. \quad (18)$$

In addition, the mean weight vector  $\bar{\mathbf{w}}^a$  is obtained as follows:

$$\bar{\mathbf{w}}^a = \mathbf{P}^a \mathbf{Y}^{fT} \mathbf{R}^{-1} (\mathbf{y}^0 - \bar{\mathbf{y}}^f), \quad (19)$$

where  $\mathbf{y}^0$  contains the CRNS SWC retrievals at the measurement locations. In the ensemble space, a perturbation matrix  $\mathbf{W}^a$  is calculated from the symmetric square root of  $\mathbf{P}^a$ :

$$\mathbf{W}^a = \left[ (N - 1)\mathbf{P}^a \right]^{1/2}. \quad (20)$$

The final analysis  $\mathbf{X}^a$  is obtained from the following:

$$\mathbf{X}^a = \bar{\mathbf{x}}^f + \mathbf{X}^f [\bar{\mathbf{w}}^a + \mathbf{W}^a]. \quad (21)$$

A more detailed description of the LETKF can be found in Hunt et al., (2007), and details on the implementation of the LETKF in combination with CLM are given by Han et al. (2015).

### 3 Model and experiment setup

#### 3.1 Model setup

In this study, discretization and parameterization of the hydrological catchment was done on the basis of high-resolution data. The model of the Rur catchment was spatially discretized by rectangular grid cells of  $0.008^\circ$  size ( $\sim 750$  m). The model time step was set to hourly. Land cover was assumed to consist of vegetated land units only, and a single plant functional type (PFT) for each grid cell was defined. The plant functional types were derived from a remotely sensed land use map using RapidEye and ASTER data with 15 m resolution (Waldhoff, 2012). Contents of sand, clay and organic matter were derived from the high-resolution regional soil map BK50 (Geologischer Dienst Nordrhein-Westfalen, 2009). Alternative simulations were also performed with the FAO soil map of the global Harmonized World Soil Database (FAO, 2012) and with a biased soil texture with a fixed sand content of 80 % and clay content of 10 % (S80 soil map). Average sand and clay content are 22.5 and 21.4 % for the BK50 soil map and 39 and 22 % for the FAO soil map, respectively. The FAO soil map and the biased soil map represent large error with respect to the soil properties of the BK50 soil map. The FAO soil map and S80 soil map simulations allow the evaluation of the joint state-parameter estimation approach because, given the expected bias, we can evaluate to what extent the soil properties are modified by the data assimilation. This is important because in many regions across the Earth a high-resolution soil map is not available. Land surface models are applied for those regions, for example in the context of global simulations, and

hence might be strongly affected by the error in soil properties.

Maximum saturated fraction, a surface parameter which is used for runoff generation, was calculated from a 10 m digital elevation model (scilands GmbH, 2010). Leaf area index data were derived from monthly-averaged Moderate Resolution Imaging Spectrometer data (MODIS). CLM was forced with hourly atmospheric data from the COSMO\_DE reanalysis data set for the years 2010 to 2013 from the German Weather Service (Deutscher Wetterdienst, DWD). The data were downscaled from a resolution of  $2.8 \text{ km} \times 2.8 \text{ km}$  to the CLM resolution using linear interpolation based on Delaunay triangulation. Forcing data include precipitation, incident solar and longwave radiation, air temperature, air pressure, wind speed, and relative humidity at the lowest atmospheric level.

#### 3.2 Model ensemble

Uncertainty was introduced into the regional CLM model by perturbed soil parameters and external model forcings. Contents of sand, clay and organic matter were perturbed with spatially correlated noise from a uniform sampling distribution with mean zero and standard deviations 10 and 30 % (Han et al., 2015). Soil texture perturbation considers that in CLM a single set of pedotransfer functions is assumed to be valid throughout the globe while pedotransfer functions are usually specific to regions (e.g. Patil and Singh, 2016). In other words, the perturbation of soil texture also covers the uncertainty in the pedotransfer function itself. By perturbing texture, soil parameters are also perturbed through the pedotransfer functions used in CLM as specified in Sect. 2.2. Precipitation ( $\sigma = 0.5$  or  $1.0$ ; lognormal distribution) and shortwave radiation ( $\sigma = 0.3$ ; lognormal distribution) were perturbed with multiplicative noise with mean equal to 1. Longwave radiation ( $\sigma = 20 \text{ W m}^{-2}$ ) and air temperature ( $\sigma = 1 \text{ K}$ ) were perturbed with additive noise. The forcing perturbations were imposed with correlations in space (5 km) using a fast Fourier transform. Correlation in time was introduced with an AR(1) model with autoregressive parameter 0.33. These correlations and standard deviations were chosen based on previous data assimilation experiments (Reichle et al., 2010; Kumar et al., 2012; De Lannoy et al., 2012; Han et al., 2015). In this work, only results for precipitation perturbation with  $\sigma = 0.5$  will be shown, as results for  $\sigma = 1.0$  were similar. An ensemble size of 95 realizations was used in the simulations. Based on previous work (Baatz et al., 2015), the SWC retrieval uncertainty for CRNS was estimated to be  $0.03 \text{ cm}^3 \text{ cm}^{-3}$  while fluctuations in the measurement standard deviation, related to the nonlinear relation between observed neutron intensity and SWC, were assumed to be negligible.

**Table 2.** Overview of simulation scenarios: open loop (OL-\*) with variation in the soil maps BK50, FAO and S80, data assimilation run with state update (Stt) or joint state and parameter update (PAR) with variation in the soil map perturbation (−10 and −30), and jackknife evaluation runs (jk8-S80-1 to 9, jk8-BK50-1 to 9 and jk4-S80-A to C).

Simulation code	Update		Sand content			Soil perturbation	
	State	Parameter	BK50	80 % fix	FAO	10	30
OL-BK50			+				+
OL-S80				+			+
OL-FAO					+		+
Stt-BK50	+		+				+
Stt-S80	+			+			+
PAR-BK50-30	+	+	+				+
PAR-BK50-10	+	+	+			+	
PAR-S80-30	+	+		+			+
PAR-S80-10	+	+		+		+	
PAR-FAO-30	+	+			+		+
jk8-BK50-1 to 9	+	+	+				+
jk8-S80-1 to 9	+	+		+			+
jk4-S80-A to C	+	+		+			+

### 3.3 Experiment set-up

All simulation experiments in this study used initial conditions from a single 5-year spin-up run in which a single forcing data set of the year 2010 was repeatedly used as atmospheric input. The soil moisture regime became stable after the 5-year spin-up period, and additional spin-up simulations would not affect soil moisture in the consecutive years. After this 5-year spin-up, soil parameters and forcing data of the consecutive years were perturbed. From 1 January 2011 onwards, CLM was propagated forward with an ensemble of 95 realizations. On 20 March 2011, the first SWC retrieval was assimilated, and assimilation of SWC retrievals continued until 31 December 2012. In the data assimilation period, soil properties were estimated at every time step when observations were made available. For the year 2013, the model was propagated forward without data assimilation but with an ensemble of 95 realizations. The year 2013 was used exclusively as the evaluation period for data assimilation experiments.

In total, 31 simulation experiments were carried out using different setups (Table 2). The present setups are intended to cover three different initial soil maps, three different sizes of a CRNS network and two different parameter perturbations. Three open loop simulations were run without data assimilation and soil parameter perturbation of 30 % for the BK50 soil map (OL-BK50), the FAO soil map (OL-FAO) and the S80 soil map (OL-S80). These simulations are referred to as reference runs for the respective soil map. Simulation results of data assimilation runs were compared to the reference runs for quantification of data assimilation benefits. Simulations were done with joint state–parameter estimation (PAR-\*), two for the BK50 soil map (PAR-BK50-\*), one for the FAO soil map (PAR-FAO-30), and two for the

S80 soil map (PAR-S80-\*). Soil texture was perturbed by 10 or 30 % as indicated by the experiment name (Table 2). Two simulations were done with state updates only for the BK50 soil map (Stt-BK50) and the S80 soil map (Stt-S80). These 10 simulations form the basic set of experiments.

Besides the data assimilation experiments, a larger number of jackknifing simulations were also conducted to evaluate the impact of the CRNS data assimilation on SWC at unobserved locations in the model domain. In nine jackknife experiments, data from eight CRNS locations were assimilated (jk8-\* simulations) and data of the one remaining CRNSs were not assimilated but kept for evaluation. In addition, three simulations were conducted where data of four CRNSs were assimilated (jk4-\* simulations), and data of the five remaining CRNSs were used for evaluation. These three simulations represent a CRNS network with much less than the existing nine CRNSs. At the evaluation locations, simulated SWC (which is affected by the assimilation of the other eight probes) was compared to CRNS SWC retrievals. For jackknife simulations, the perturbation of soil texture was set to 30 %. States and parameters at these sites were jointly updated, and simulations were made using either the BK50 or the S80 soil maps as initial parameterization. Therefore, a total of 21 jackknife simulations were performed.

Simulation results were evaluated with the root mean square error ( $E_{\text{RMS}}$ ):

$$E_{\text{RMS}} = \sqrt{\frac{\sum_{t=1}^n (\theta_{t,\text{CLM}} - \theta_{t,\text{CRNS}})^2}{n}}, \quad (22)$$

where  $n$  is the total number of time steps,  $\theta_{t,\text{CLM}}$  is SWC simulated by CLM at time step  $t$  and  $\theta_{t,\text{CRNS}}$  is the CRNS SWC retrieval at time step  $t$ . In case SWC was assimilated at

the corresponding time step,  $\theta_{t,CLM}$  is SWC prior to assimilation. In case the  $E_{RMS}$  is estimated at a single point in time over all CRNSs available, the number of time steps  $n$  can be replaced by the number of CRNSs available. The second evaluation measurement in this study is the bias which is, in contrast to the  $E_{RMS}$ , a measure for systematic deviation:

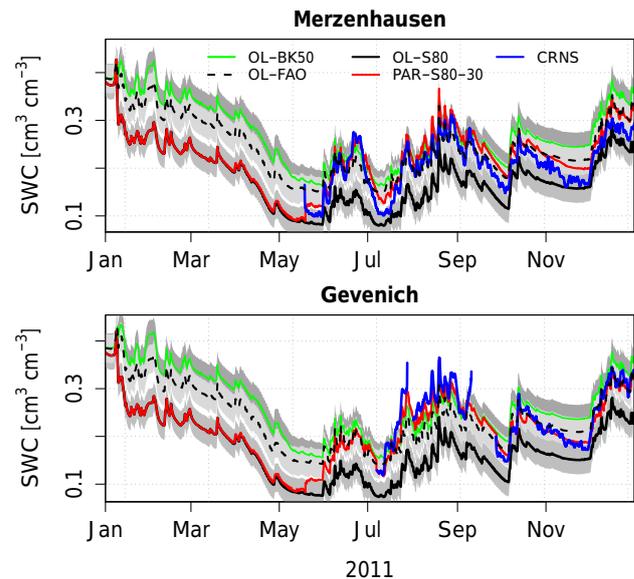
$$\text{bias} = \frac{\sum_{t=1}^n (\theta_{t,CLM} - \theta_{t,CRNS})}{n}. \quad (23)$$

## 4 Results and discussion

### 4.1 General results

Table 3 summarizes the performance statistics in terms of  $E_{RMS}$  and bias for the assimilation period (2011 and 2012) and evaluation period (2013). Presented are results for the open loop scenarios with the BK50, FAO and S80, and data assimilation scenarios. Errors of open loop simulations were highest for the S80 simulation ( $0.11 \text{ cm}^3 \text{ cm}^{-3}$ ), followed by the FAO simulation ( $0.07 \text{ cm}^3 \text{ cm}^{-3}$ ) and the BK50 simulation ( $0.04 \text{ cm}^3 \text{ cm}^{-3}$ ). Mean absolute bias was highest for the S80 soil map ( $0.11 \text{ cm}^3 \text{ cm}^{-3}$ ), now as high for the FAO soil map ( $0.06 \text{ cm}^3 \text{ cm}^{-3}$ ) and lowest for the BK50 soil map ( $0.02 \text{ cm}^3 \text{ cm}^{-3}$ ). Data assimilation improved simulations more for the S80 soil map ( $E_{RMS}$  reduced by  $0.08 \text{ cm}^3 \text{ cm}^{-3}$ ) than for the FAO soil map ( $E_{RMS}$  reduced by  $0.04 \text{ cm}^3 \text{ cm}^{-3}$ ) or the BK50 soil map ( $E_{RMS}$  reduced by  $0.01 \text{ cm}^3 \text{ cm}^{-3}$ ). The BK50 soil map led to  $E_{RMS}$  values in open loop simulations lower than  $0.05 \text{ cm}^3 \text{ cm}^{-3}$ , which left little room for error reduction considering a measurement error of  $0.03 \text{ cm}^3 \text{ cm}^{-3}$ . However, slight improvements by  $0.01 \text{ cm}^3 \text{ cm}^{-3}$  were possible at monitored locations in the data assimilation period but not in the evaluation period. Joint state–parameter estimation improved simulation results, as shown by the reduced  $E_{RMS}$  and bias for the S80 and the FAO soil maps. The verification period (2013) with the updated soil hydraulic parameters for the FAO soil map resulted in an  $E_{RMS}$  value of  $0.05 \text{ cm}^3 \text{ cm}^{-3}$ , also clearly an improvement compared to the open loop run with an  $E_{RMS}$  of  $0.07 \text{ cm}^3 \text{ cm}^{-3}$  (Table 3). Joint state–parameter updating resulted in similar  $E_{RMS}$  values for all three initial soil maps: the BK50, FAO and S80 soil maps (each  $0.03 \text{ cm}^3 \text{ cm}^{-3}$ ). State updates (Stt-S80) improved  $E_{RMS}$  and bias for the S80 soil map ( $E_{RMS} = 0.06 \text{ cm}^3 \text{ cm}^{-3}$  for assimilation period) but much less compared to the joint state–parameter updates (PAR-S80-30). The  $E_{RMS}$  and bias for simulations with 10 and 30 % perturbation of soil texture only showed very small differences (smaller than  $0.01 \text{ cm}^3 \text{ cm}^{-3}$ ).

The temporal course of simulated soil moisture in 2011 at the two sites, Merzenhausen and Gevenich, is shown in Fig. 2. The figure illustrates that simulated SWC at both sites



**Figure 2.** Temporal evolution of simulated soil water content (SWC) retrievals, calculated with open loop (OL-\*) simulations and data assimilation including parameter updating (PAR-S80-30), together with the CRNS SWC retrieval during the first year of simulation at the sites Merzenhausen and Gevenich. Simulated SWC was vertically weighted using the COSMIC operator to obtain the appropriate SWC corresponding to the CRNS SWC retrieval.

was lowest with the S80 soil map (OL-S80) and highest with the BK50 soil map (OL-BK50), and the FAO soil map resulted in intermediate soil moisture (OL-FAO). Mean open loop SWC in 2011 was  $0.17 \text{ cm}^3 \text{ cm}^{-3}$  for the S80 soil map,  $0.24 \text{ cm}^3 \text{ cm}^{-3}$  for the FAO soil map and  $0.27 \text{ cm}^3 \text{ cm}^{-3}$  for the BK50 soil map at both sites. Measurements with CRNS started in May 2011 at Merzenhausen. At Gevenich, the first observation was recorded on 7 July 2011. In the data assimilation run shown (PAR-S80-30), modelled SWC was immediately affected at both sites, Merzenhausen and Gevenich, as soon as data at Merzenhausen were assimilated. By July, simulated SWC with the biased soil map and data assimilation (PAR-S80-30) was already close to the CRNS SWC retrieval at the Gevenich site (Fig. 2). This demonstrates the beneficial impact of data availability for assimilation at one site and the information brought into space by the data assimilation scheme. Figure 2 also shows that the BK50 open loop run was close to the observed SWC at both sites, even without data assimilation.

Figure 3 shows the temporal course of SWC from January 2011 to December 2013 at Heinsberg and Wildenrath. Assimilation and evaluation results are shown for the open loop (OL-S80 and OL-FAO) simulations, only state updates (Stt-S80), joint state–parameter updates (PAR-S80-30) and CRNS SWC retrievals. At Heinsberg, results show that simulated SWC with assimilation was closer to the CRNS when both states and parameters were updated (PAR-S80-

**Table 3.** Root mean square error ( $E_{\text{RMS}}$ ,  $\text{cm}^3 \text{cm}^{-3}$ ) and mean absolute bias ( $\text{cm}^3 \text{cm}^{-3}$ ) for open loop simulations (OL-\*), data assimilation with state updates (Stt-\*) and joint state–parameter updates (PAR-\*) for both the assimilation period (2011 and 2012) and the evaluation period (2013). Error and bias was averaged over all sites with observations. Site-specific errors and biases are provided in Appendix A1 to A4. The best cases are marked in bold.

Soil map	Simulation	Site average			
		Data assimilation 2011 and 2012		Evaluation period 2013	
		$E_{\text{RMS}}$	Absolute bias	$E_{\text{RMS}}$	Absolute bias
BK50	OL-BK50	0.04	0.02	0.04	0.02
	<b>Stt-BK50</b>	<b>0.03</b>	<b>0.01</b>	<b>0.04</b>	<b>0.01</b>
	PAR-BK50-10	0.03	0.01	0.05	0.03
	PAR-BK50-30	0.03	0.01	0.05	0.03
FAO	OL-FAO	0.07	0.06	0.07	0.06
	<b>PAR-FAO-30</b>	<b>0.03</b>	<b>0.02</b>	<b>0.05</b>	<b>0.03</b>
Biased (S80)	OL-S80	0.12	0.11	0.12	0.11
	Stt-S80	0.06	0.05	0.10	0.09
	<b>PAR-S80-10</b>	<b>0.03</b>	<b>0.01</b>	<b>0.05</b>	<b>0.03</b>
	<b>PAR-S80-30</b>	<b>0.03</b>	<b>0.02</b>	<b>0.04</b>	<b>0.02</b>

30) than if only states were updated (Stt-S80). This is the case for both the assimilation and the evaluation periods. At the beginning of the evaluation period (first few days of 2013), the Stt-S80 simulation shows an increase in bias between modelled SWC and CRNS. The bias of Stt-S80 remained throughout the evaluation period. In contrast, if parameters were previously updated (PAR-S80-30), modelled SWC was close to the CRNS during the evaluation period. Open loop SWC modelled with the FAO soil map is lower than the CRNS SWC retrievals at Heinsberg and higher than CRNS SWC retrievals at Wildenrath. At Wildenrath, results of the OL-S80 run suggest that the initial sand content of the biased soil map is closer to the optimal sand content than the sand content of the FAO soil map. Consequently, the OL-FAO bias was  $-0.05$  and  $0.05 \text{ cm}^3 \text{cm}^{-3}$  for Heinsberg and Wildenrath, respectively (Tables A1 and A4 in Appendix). At both sites, absolute bias was reduced with joint state–parameter updates to equal or less than  $0.01 \text{ cm}^3 \text{cm}^{-3}$  (S80 and FAO soil map). The reduced bias is also well reflected in the temporal course of modelled SWC with joint state–parameter updates (PAR-S80-30).

It is interesting to notice that the error values for the verification period are very similar if soil hydraulic parameters were estimated in the assimilation period, independent of the initial soil map (Table 3).  $E_{\text{RMS}}$  values for the 2013 simulations with state updates only (Stt-BK50 and Stt-S80) show that in the evaluation period the improvements by state updates (without parameter updates) were small (reduction by  $0.02$  and  $0.00 \text{ cm}^3 \text{cm}^{-3}$  for S80 and BK50, respectively) compared to the improvements obtained by joint state–parameter updates (reduction by  $0.07 \text{ cm}^3 \text{cm}^{-3}$  for

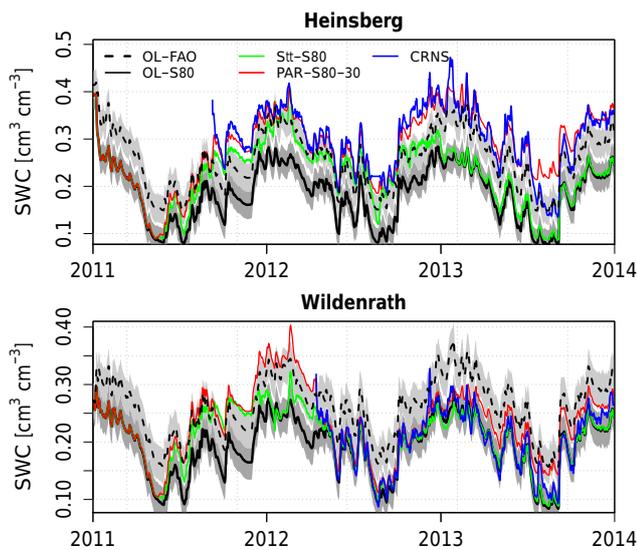
S80). This illustrates the benefits of joint state–parameter updates compared to state updates only, and that soil moisture states are strongly determined by soil hydraulic parameters. The case of only state updates also illustrates that the improved characterization of soil moisture states in the assimilation period results in improved initial states for the verification period (Table 3) but in the verification period these improvements lose their influence quickly over time (Fig. 3).

#### 4.2 Temporal evolution of mean $E_{\text{RMS}}$

Figure 4 shows the temporal evolution of the hourly  $E_{\text{RMS}}$  calculated for all nine CRNSs.  $E_{\text{RMS}}$  was highest for the S80 open loop run and lowest for the PAR-S80-30 simulation. The FAO soil map resulted in errors mostly between  $0.05$  and  $0.1 \text{ cm}^3 \text{cm}^{-3}$ , which are lower than the S80 soil map but not as good as simulation results with joint state–parameter updates (PAR-S80-30) or with the BK50 soil map (OL-BK50). State updates did not improve modelled SWC as much as joint state–parameter updates. For most of the time, the  $E_{\text{RMS}}$  of the Stt-S80 run is larger than the  $E_{\text{RMS}}$  of the OL-BK50 run. During the evaluation period, the open loop run with the FAO soil map (OL-FAO) also performs better than the Stt-S80 run. In contrast, joint state–parameter updates to the S80 soil map improved the  $E_{\text{RMS}}$  most of the time compared to open loop simulations (OL-BK50, OL-FAO and OL-S80). As shown in Fig. 4, the PAR-S80-30 simulation performed best out of the four simulations during the assimilation period 2011–2012. During the evaluation period 2013, OL-BK50 and PAR-S80-30 performed equally well, except in summer 2013 when the PAR-S80-30 simula-

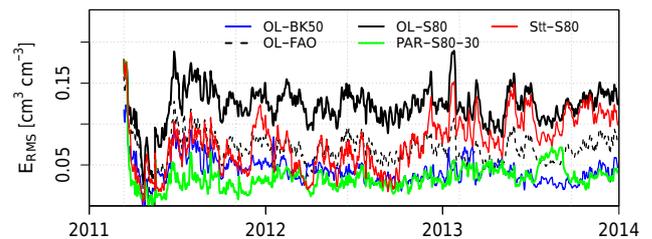
**Table 4.** Root mean square error ( $E_{RMS}$ ,  $\text{cm}^3 \text{cm}^{-3}$ ) and mean absolute bias ( $\text{cm}^3 \text{cm}^{-3}$ ) for open loop (OL-\*), jackknife simulations with eight CRNSs (simulations jk8-S80-1 to 9 were averaged) and with four CRNSs (simulations jk4-S80-A to C). Results were averaged over the omitted sites only. Data at omitted sites were not assimilated, while at the other sites data were assimilated. At sites where data were assimilated,  $E_{RMS}$  and bias were equal to the values found in simulation PAR-S80-30. Site-specific errors and biases are provided in the Appendix A1 to A4. The best cases are marked in bold.

Soil map	Simulation	Site average			
		Data assimilation 2011 and 2012		Evaluation period 2013	
		$E_{RMS}$	Absolute bias	$E_{RMS}$	Absolute bias
BK50	<b>OL-BK50</b>	<b>0.04</b>	<b>0.02</b>	<b>0.04</b>	<b>0.02</b>
	jk8-BK50-1 to 9	0.06	0.04	0.05	0.04
Biased (S80)	OL-S80	0.12	0.11	0.12	0.11
	<b>jk8-S80-1 to 9</b>	<b>0.06</b>	<b>0.05</b>	<b>0.06</b>	<b>0.04</b>
	jk4-S80-A	0.08	0.06	0.06	0.04
	jk4-S80-B	0.06	0.05	0.06	0.05
	jk4-S80-C	0.07	0.05	0.07	0.06



**Figure 3.** Temporal evolution of simulated soil water content (SWC) retrievals, calculated with open loop (OL-\*), data assimilation with state update only (Stt-S80), and data assimilation including parameter updating (PAR-S80-30), together with the CRNS SWC retrieval at the sites Heinsberg and Wildenrath for the data assimilation period 2011 and 2012 and the evaluation period 2013. Simulated SWC was vertically weighted to obtain the appropriate SWC corresponding to the CRNS SWC retrieval.

tion yielded much higher  $E_{RMS}$  values than the BK50 open loop run.



**Figure 4.** Temporal evolution of root mean square error ( $E_{RMS}$ ) for hourly SWC retrievals.  $E_{RMS}$  is calculated hourly for all nine CRNSs for open loop (OL-\*) runs for soil maps BK50, FAO and S80; joint state–parameter updates (PAR-S80-30); and state updates only (Stt-S80) during the assimilation period with joint state–parameter updates (2011 and 2012) and the verification period (2013).

### 4.3 Jackknife simulations

The jackknife simulations investigated the impact of CRNS data on improving simulated SWC at locations beyond the CRNS stations. Spatial improvements are possible by spatial correlation structures of atmospheric forcings, soil hydraulic parameters and soil moisture which are taken into account by the local ensemble transform Kalman filter. The error and bias shown in Table 4 refer to jackknife simulations with the BK50 and the S80 soil map. On average, over the three runs where only data of four CRNSs were assimilated (jk4-S80-\*), the  $E_{RMS}$  was  $0.07 \text{ m}^3 \text{ m}^{-3}$ , which is much lower than the  $E_{RMS}$  for the open loop run ( $0.12 \text{ m}^3 \text{ m}^{-3}$ ) and only a bit higher than the case where eight CRNSs were assimilated ( $E_{RMS} = 0.06 \text{ m}^3 \text{ m}^{-3}$  for jk8-S80-\*). The improved simulation results were also due to the bias reduction from 0.11 to  $0.05 \text{ m}^3 \text{ m}^{-3}$  in the case of four and  $0.04 \text{ m}^3 \text{ m}^{-3}$  in

the case of eight assimilated CRNSs. However, for the BK50 soil map where  $E_{\text{RMS}}$  ( $0.04 \text{ m}^3 \text{ m}^{-3}$ ) and bias ( $0.02 \text{ m}^3 \text{ m}^{-3}$ ) of the open loop run were already good, the jackknife simulations led to slightly higher  $E_{\text{RMS}}$  ( $0.05 \text{ m}^3 \text{ m}^{-3}$ ) and bias ( $0.04 \text{ m}^3 \text{ m}^{-3}$ ). More detailed site statistics (Tables A1 to A4 of the Appendix) demonstrate that all jackknife simulations with the S80 soil map resulted in an improved  $E_{\text{RMS}}$  at the omitted locations compared to the open loop simulation, except for Wildenrath. At sites with large open loop  $E_{\text{RMS}}$ , the assimilation could reduce the  $E_{\text{RMS}}$  by 50 % or more.

The jackknife simulations illustrate that a network of CRNSs can improve modelled SWC if the soil map information is not sufficient. This suggests that assimilation of CRNS data is particularly useful for regions with little information on subsurface parameters. A trade-off can be expected between the initial uncertainty on soil moisture and parameters, and the density of a CRNS network. In the case of a large uncertainty, like in regions with limited information about soils or a strongly biased soil map (e.g. FAO or S80 soil map) and a low density of meteorological stations, a sparse network of probes can already be helpful for improving soil moisture characterization. The results of the real-world jackknife experiments demonstrated that four CRNSs are already beneficial, but it is desirable to have more CRNSs for improved parameter estimates. The results also suggest that the additional information gain for an extra CRNS is reduced for a denser network, because the soil moisture characterization did not improve so much more when eight instead of four CRNSs were used for assimilation. However, in regions with a high density of meteorological stations and a high-resolution soil map, it can be expected that a denser CRNS network than that in this study is needed to further lower the error of soil moisture characterization. Further potentially synthetic experiments in other regions with networks of CRNSs are needed to obtain more quantitative information about this.

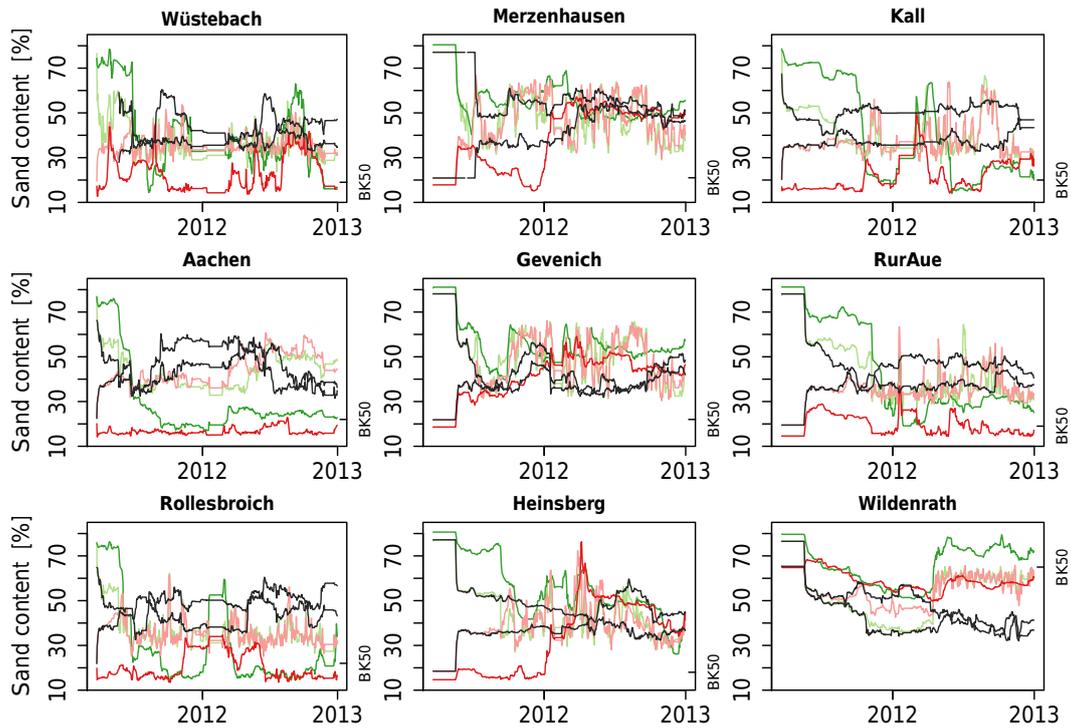
#### 4.4 Temporal evolution of parameter estimates and parameter uncertainty

The temporal evolution of sand content estimates during the assimilation period for the nine sites with CRNSs is shown in Fig. 5 for PAR-S80-30, PAR-S80-10, PAR-BK50-30, PAR-BK50-10, jk8-S80-\* and jk8-BK50-\*. Time series start on 20 March 2011, the date of the first assimilated CRNS SWC retrieval at Wüstebach. At Wüstebach and sites within the influence sphere of Wüstebach (Aachen, Kall and Rollesbroich), sand content estimates were updated from 20 March 2011 onwards. Because of the localization, all other sites show a first update in sand content in May 2012 when Rollesbroich and Merzenhausen start operating, and their data were assimilated. During the data assimilation period with joint state–parameter updates, all sites show variability in sand content estimates over time with differences in magnitude. Values and spread in sand content estimates amongst the experiments is smaller at the sites Merzen-

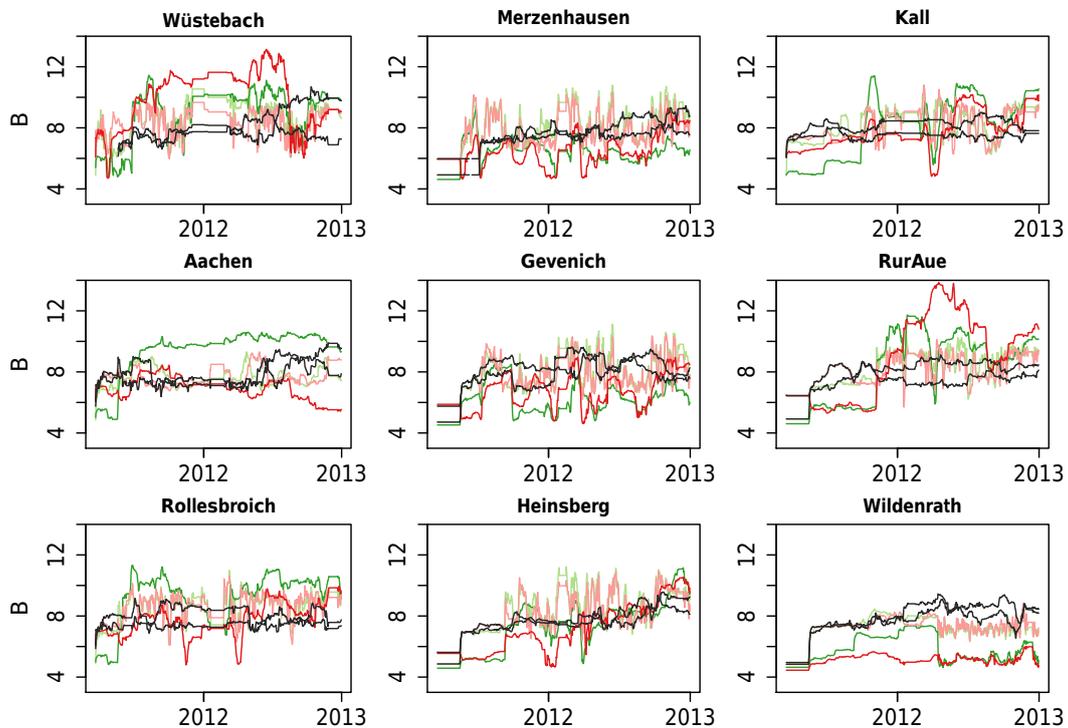
hausen, Gevenich, RurAue, Heinsberg and Wildenrath, compared to the sites Wüstebach, Aachen and Rollesbroich where spread is considerably larger. At the sites Merzenhausen, Kall, Gevenich, RurAue and Heinsberg, sand content estimates of the jackknife simulations were close to the sand content of the other data assimilation experiments with joint state–parameter estimation. A comparison of parameter estimates at the end of the assimilation period indicates that initial soil parameterization has a limited effect on the resulting parameter estimates. Parameter estimates of jk8-BK50-\* and jk8-S80-\* are close together at the end of the assimilation period.

Estimates of the soil hydraulic parameter  $B$  and saturated hydraulic conductivity are shown in Figs. 6 and 7 for PAR-S80-30, PAR-S80-10, PAR-BK50-30, PAR-BK50-10, jk8-S80-\* and jk8-BK50-\*. Updates of soil hydraulic parameters start in March and May 2011 with the assimilation of CRNS SWC retrievals, depending on the location. The  $B$  parameter estimates increase for all simulations. Throughout the whole assimilation period the empirical  $B$  parameter varies considerably within short time intervals. The total range of the  $B$  parameter estimates is between 2.7 and 14 at all sites. At the sites Merzenhausen, Kall, Aachen, Gevenich and Rollesbroich,  $B$  generally ranges between 6 and 10. At Wüstebach, Heinsberg and RurAue, estimates of  $B$  range most of the time between 8 and 12, and at Wildenrath,  $B$  is below 8. Initial saturated hydraulic conductivity  $k_{\text{sat}}$  is rather high ( $k_{\text{sat}} > 0.015 \text{ mm s}^{-1}$ ) in the case of high sand content, i.e. for the S80 soil map, and rather low ( $k_{\text{sat}} < 0.005 \text{ mm s}^{-1}$ ) in the case of low sand content, i.e. for the BK50 soil map. In the case of the S80 soil map, at all sites except Wildenrath, high initial  $k_{\text{sat}}$  estimates decrease quickly through joint state–parameter updates to values below  $0.01 \text{ mm s}^{-1}$ . The initial spread in  $k_{\text{sat}}$  estimates amongst the simulation scenarios decreases at most sites. At Wüstebach, Merzenhausen, Aachen, Gevenich, RurAue and Heinsberg, the spread is rather small, particularly at the end of the assimilation period, while at Wildenrath  $k_{\text{sat}}$  ranges from 0.005 to 0.015 for individual experiments at the end of the assimilation period.

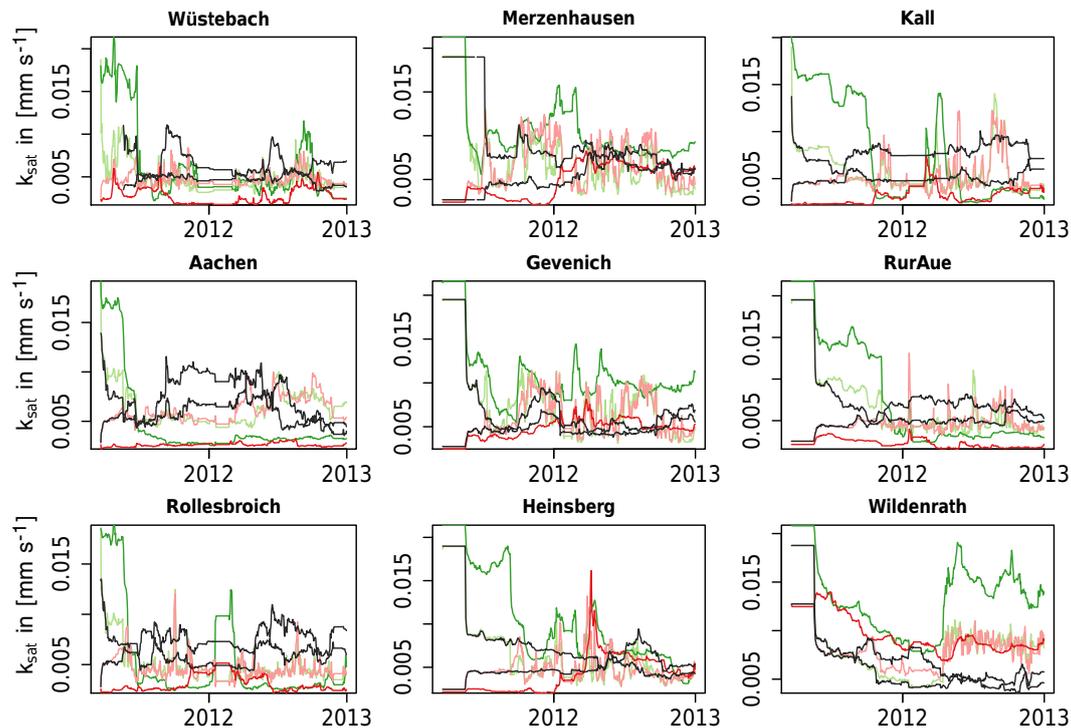
Temporally unstable parameter estimates imply that there may be multiple or seasonal optimal parameter values. This is also supported by the findings of the temporal behaviour of site-average  $E_{\text{RMS}}$  (Fig. 4), e.g. during the evaluation period when, in the dry summer of 2013, the  $E_{\text{RMS}}$  peaks for the PAR-S80-30 simulation. In this context, it is important to mention that many possible error sources were not subject to calibration in this study but could be crucial for an even better modelled soil moisture and more reliable soil parameter estimation. In this study we only considered uncertainty of soil parameters, but vegetation parameters are also uncertain. Also, a number of other CLM-specific hydrologic parameters (e.g. decay factor for subsurface runoff and maximum subsurface drainage) strongly influence state variables in CLM and hence show potential for optimization (Sun et al., 2013). Considering this uncertainty from multiple param-



**Figure 5.** At nine sites, estimates of percentage sand content are shown for simulations with parameter update: PAR-S80-30 (green), PAR-S80-10 (light green), PAR-BK50-30 (red), PAR-BK50-10 (light red), jk8-S80-\* (black) and jk8-BK50-\* (black). The value of the BK50 soil map is marked at the second y axis.



**Figure 6.** At nine sites, estimates of the  $B$  parameter (top 15 cm) are shown for simulations with parameter update: PAR-S80-30 (green), PAR-S80-10 (light green), PAR-BK50-30 (red), PAR-BK50-10 (light red), jk8-S80-\* (black) and jk8-BK50-\* (black).



**Figure 7.** At nine sites, estimates of saturated hydraulic conductivity (top 15 cm) are shown for simulations with parameter update: PAR-S80-30 (green), PAR-S80-10 (light green), PAR-BK50-30 (red), PAR-BK50-10 (light red), jk8-S80-\* (black) and jk8-BK50-\* (black).

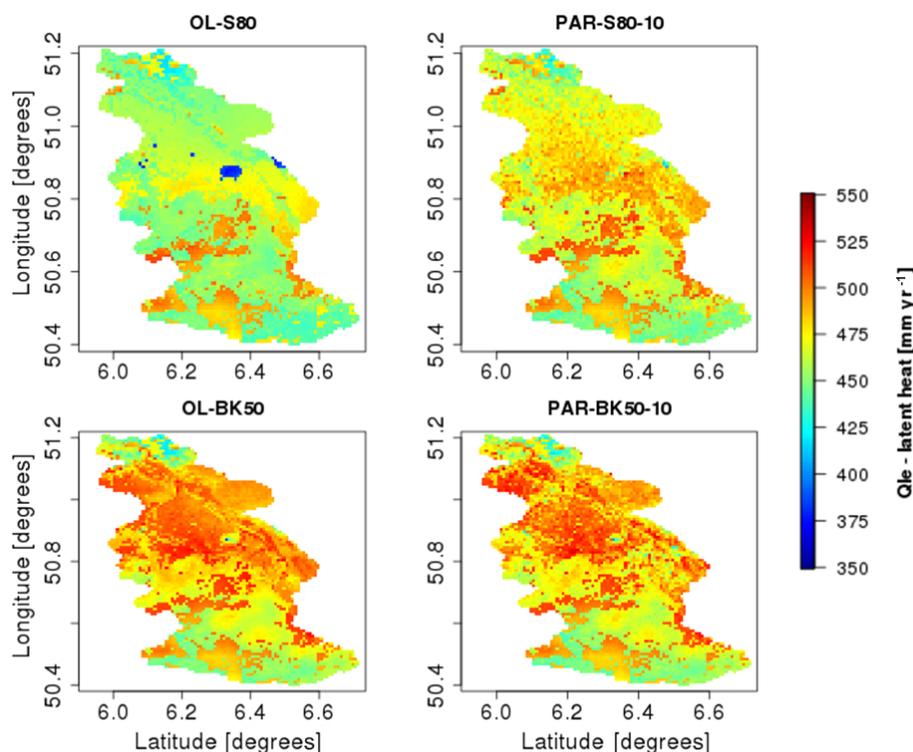
eters could give a better parameter-uncertainty characterization (Shi et al., 2014). Precipitation is also an important forcing for hydrologic modelling. For this study, precipitation data from the COSMO\_DE re-analysis were used. A product which optimally combines precipitation estimates from radar and gauge measurements is expected to give better precipitation estimates than the reanalysis. This could improve the soil moisture characterization and also potentially lead to better parameter estimates. Further improvements and constraining of parameter uncertainty is also possible using multivariate data assimilation with observations such as latent heat flux (e.g. Shi et al., 2014). Also, other error sources related to the model structure play a significant role. These options should be subject to future investigations.

#### 4.5 Latent heat flux

Latent heat flux, or evapotranspiration (ET), is another important diagnostic variable of land surface models (e.g. Best et al., 2015) and is of importance for atmospheric models. Results of the data assimilation experiments showed that soil texture updates altered soil moisture states significantly. In Fig. 8 it is shown that joint state–parameter estimation also altered ET during the evaluation period. Figure 8 shows ET within the evaluation period 2013 across the whole catchment for four simulation experiments. On the one hand, ET was similar for both open loop simulations (OL-S80 and OL-BK50) in the south of the catchment. On the other hand, ET

in the north was up to 80 mm lower per year for the S80 open loop run compared to the BK50 open loop run. The differences can be linked to the drier soil conditions for OL-S80 compared to OL-BK50 simulation results. The differences in ET between the runs with and without parameter updates were larger for the S80 soil map than for the BK50 soil map. For PAR-S80-10, ET increased by up to 40 mm yr<sup>-1</sup> in the northern part of the catchment through data assimilation while the change in ET from OL-BK50 to PAR-BK50-10 is rather small. This is linked to the comparatively larger updates made to soil hydraulic parameters.

Additionally, the impact of soil parameter estimates on ET is different in the north of the catchment compared to the south. While ET in the north of the catchment was impacted by the estimated soil properties during the evaluation period in 2013 for PAR-S80-10, ET in the south was not impacted as much by estimated soil properties. This is related to the fact that in the north ET is moisture-limited in summer, whereas in the south this is not moisture-limited but energy-limited. Therefore, ET in the north is sensitive to variations in soil hydraulic parameter values, whereas in the south this is not the case. In the south, ET is sensitive to model forcings such as incoming shortwave radiation. Nearing et al. (2016) came to the conclusion that soil parameter uncertainty dominates soil moisture uncertainty and forcing uncertainty dominates ET uncertainty. Our findings in the southern part of the catchment support their conclusion, but in the northern part of



**Figure 8.** Annual evapotranspiration (ET) is shown in the year 2013 (evaluation period, no assimilation). This figure demonstrates the impact of parameter updates (PAR-S80-10 and PAR-BK50-10) in comparison to open loop (OL-S80) and the reference soil map (OL-BK50). ET changes in the north but not as much in the south.

the catchment soil parameter uncertainty strongly affects ET. Hence, particularly in the northern part of the catchment, further observations such as ET measurements are desirable for further improving the land surface model. These additional observations could be used for future land surface model benchmarking (Best et al., 2015) or for more constrained parameter estimates (Shi et al., 2015).

## 5 Conclusions and outlook

This real-world case study on assimilating CRNS SWC retrievals into a land surface model shows the potential of CRNS networks to improve subsurface parameterization in regional land surface models, especially if prior information on soil properties is limited. CRNS SWC retrievals were assimilated into the land surface model CLM version 4.5 using the LETKF. SWC and subsurface parameters were updated with the LETKF at unmonitored locations in the catchment considering model and observation uncertainties. Joint state–parameter estimates improved soil moisture estimates during the assimilation and during the evaluation period. The error and bias for the soil moisture characterization was strongly reduced for simulations initialized with a biased soil map and similarly well if initialized with the FAO soil map. Simulations initialized with a biased or global soil map approached

similar error statistics with joint state–parameter updates to the ones obtained when the regional soil map was used as input to the simulations. Error values in simulations with the regional soil map were not improved during the evaluation period, because open loop simulation results were already close to the observations. The beneficial results of joint state–parameter updates were confirmed by additional jackknife experiments with eight and four CRNSs for assimilation. In many areas of the world, only global soil maps (e.g. the FAO soil map) are available but there are no detailed high-resolution regional soil maps. This study has shown that in these areas a more advanced subsurface characterization is possible using CRNS measurements and the data assimilation framework presented in this study.

For now, neutron intensity observations by CRNSs were not assimilated directly. In future studies it would be desirable to use the COSMIC operator for assimilating neutron intensity observations directly. However, in this case the impact of biomass on the CRNS measurement signal would have to be taken into account. Therefore, it is desirable to further develop the COSMIC operator to include the impact of biomass on neutron intensities. Using the biogeochemical module of CLM would then allow for modelling of local vegetation states as input for the measurement operator. Remotely sensed vegetation states are another option to characterize vegetation states as input for the measurement op-

erator. Both methods require additional field measurements for the verification of vegetation state estimates. The further extension of the data assimilation framework would also enable the estimation of additional land surface parameters. In addition, the impact of other subsurface parameters, such as subsurface drainage parameters and the surface drainage decay factor, on SWC states and radiative surface fluxes has already been shown (Sun et al., 2013). Estimation of these parameters is desirable because of the inherent uncertainty of these globally tuned parameters. However, estimation of soil texture and organic matter content was demonstrated to already be beneficial for improved SWC modelling. Hence, this study represents a way forward towards the integration of CRNS information in the calibration or real-time updating of land surface models.

*Data availability.* Most data presented in this study are freely available via the TERENO data portal TEODOOR (<http://teodoor.icg.kfa-juelich.de/>). Atmospheric data were licensed by the German Weather Service (Deutscher Wetterdienst, DWD), and the BK50 soil map was licensed by the Geologischer Dienst Nordrhein-Westfalen.

## Appendix A

**Table A1.**  $E_{\text{RMS}}$  ( $\text{cm}^3 \text{cm}^{-3}$ ) at CRNS sites for open loop runs and different data assimilation scenarios, for the assimilation period (2011 and 2012). For jackknife experiments (21 in total) only the error of the omitted sites is reported. The best cases are marked in bold.

Soil map	2011 and 2012	Rollesbroich	Merzenhausen	Gevenich	Heinsberg	Kall	RurAue	Wüstebach	Aachen	Wildenrath	Average $E_{\text{RMS}}$
BK50	OL-BK50	0.058	0.060	0.039	0.039	0.046	0.034	0.056	0.032	0.017	0.042
	Stt-BK50	0.031	0.039	0.021	0.021	0.030	0.024	0.039	0.023	0.017	<b>0.027</b>
	PAR-BK50-10	0.033	0.036	0.020	0.019	0.032	0.025	0.035	0.045	0.015	<b>0.029</b>
	PAR-BK50-30	0.030	0.032	0.018	0.018	0.028	0.024	0.040	0.044	0.016	<b>0.028</b>
	jk8-BK50-1 to 9	0.067	0.056	0.065	0.033	0.047	0.051	0.062	0.050	0.091	0.058
FAO	OL-FAO	0.097	0.033	0.029	0.056	0.082	0.096	0.079	0.098	0.056	0.070
	PAR-FAO-30	0.029	0.033	0.018	0.019	0.028	0.025	0.042	0.056	0.017	<b>0.030</b>
Biased (S80)	OL-S80	0.169	0.054	0.082	0.119	0.152	0.161	0.110	0.169	0.020	0.115
	Stt-S80	0.098	0.019	0.036	0.050	0.082	0.054	0.083	0.086	0.018	0.058
	PAR-S80-10	0.031	0.035	0.023	0.023	0.033	0.024	0.041	0.048	0.015	<b>0.030</b>
	PAR-S80-30	0.029	0.032	0.018	0.019	0.028	0.024	0.042	0.068	0.016	<b>0.031</b>
	jk8-S80-1 to 9	0.081	0.038	0.060	0.035	0.068	0.043	0.057	0.073	0.095	0.061
	jk4-S80-A	0.064	0.038	0.059	0.076	–	0.157	–	–	–	0.079
	jk4-S80-B	0.077	0.041	–	0.051	0.062	0.079	–	–	–	0.062
jk4-S80-C	–	0.073	0.056	–	0.051	–	–	0.078	0.109	0.073	

**Table A2.**  $E_{\text{RMS}}$  ( $\text{cm}^3 \text{cm}^{-3}$ ) at CRNS sites for open loop, data assimilation and jackknife simulations on the basis of a comparison with CRNS SWC retrievals for the verification period (2013). For jackknife experiments (21 in total) only the error of the omitted sites is reported. The best cases are marked in bold.

Soil map	2013	Rollesbroich	Merzenhausen	Gevenich	Heinsberg	Kall	RurAue	Wüstebach	Aachen	Wildenrath	Average $E_{\text{RMS}}$
BK50	OL-BK50	0.04	0.07	0.04	0.03	0.05	0.04	0.05	0.04	0.02	<b>0.040</b>
	Stt-BK50	0.04	0.05	0.03	0.03	0.05	0.04	0.05	0.04	0.02	<b>0.039</b>
	PAR-BK50-10	0.07	0.06	0.04	0.04	0.06	0.06	0.04	0.06	0.02	0.048
	PAR-BK50-30	0.05	0.06	0.04	0.03	0.07	0.05	0.04	0.05	0.04	0.047
	jk8-BK50-1 to 9	0.04	0.05	0.04	0.03	0.05	0.04	0.05	0.06	0.11	0.052
FAO	OL-FAO	0.08	0.04	0.04	0.05	0.09	0.09	0.07	0.09	0.07	0.068
	PAR-FAO-30	0.06	0.06	0.04	0.04	0.06	0.03	0.05	0.07	0.04	<b>0.049</b>
Biased (S80)	OL-S80	0.16	0.06	0.11	0.12	0.16	0.15	0.10	0.17	0.02	0.115
	Stt-S80	0.10	0.06	0.11	0.11	0.10	0.15	0.10	0.16	0.02	0.100
	PAR-S80-10	0.06	0.04	0.04	0.04	0.06	0.04	0.05	0.06	0.02	<b>0.047</b>
	PAR-S80-30	0.05	0.06	0.04	0.04	0.05	0.03	0.05	0.05	0.04	<b>0.044</b>
	jk8-S80-1 to 9	0.08	0.05	0.04	0.04	0.059	0.04	0.06	0.04	0.11	0.057
	jk4-S80-A	0.05	0.03	0.05	0.04	–	0.16	–	–	–	0.065
	jk4-S80-B	0.05	0.07	–	0.04	0.07	0.07	–	–	–	0.061
jk4-S80-C	–	0.05	0.04	–	0.07	–	–	0.06	0.13	0.069	

**Table A3.** Bias ( $\text{cm}^3 \text{cm}^{-3}$ ) at CRNS sites for open loop, data assimilation and jackknife simulations compared to CRNS SWC retrievals for the data assimilation period (2011 and 2012). For jackknife experiments (21 in total) only the bias of the omitted sites is reported. The best cases are marked in bold.

Soil map	2011 and 2012	Rollesbroich	Merzenhausen	Gevenich	Heinsberg	Kall	RurAue	Wüstabach	Aachen	Wildenrath	Mean absolute bias
BK50	OL-BK50	-0.05	0.05	0.02	-0.01	-0.02	-0.01	-0.03	0.00	0.00	<b>0.02</b>
	Stt-BK50	-0.02	0.03	0.00	-0.01	-0.01	-0.01	-0.02	0.00	-0.01	<b>0.01</b>
	PAR-BK50-10	0.00	0.03	0.00	0.00	0.00	0.00	-0.02	-0.03	0.00	<b>0.01</b>
	PAR-BK50-30	-0.02	0.03	0.00	-0.01	-0.01	-0.01	-0.03	-0.03	0.00	<b>0.01</b>
	jk8-BK50-1 to 9	-0.05	0.05	0.05	-0.01	0.00	-0.04	-0.04	-0.04	0.09	0.04
FAO	OL-FAO	-0.09	0.02	-0.01	-0.05	-0.07	-0.09	-0.06	-0.09	0.05	0.06
	PAR-FAO-30	-0.01	0.03	0.00	-0.01	-0.01	-0.01	-0.03	-0.04	0.01	<b>0.02</b>
Biased (S80)	OL-S80	-0.17	-0.05	-0.08	-0.12	-0.15	-0.16	-0.09	-0.17	-0.01	0.11
	Stt-S80	-0.09	0.00	-0.03	-0.04	-0.07	-0.04	-0.07	-0.07	-0.01	0.05
	PAR-S80-10	0.00	0.03	0.00	-0.01	0.00	0.00	-0.03	-0.04	0.00	<b>0.01</b>
	PAR-S80-30	-0.02	0.03	0.00	-0.01	-0.01	-0.01	-0.03	-0.06	0.00	<b>0.02</b>
	jk8-S80-1 to 9	-0.07	0.02	0.05	-0.01	-0.06	-0.03	-0.03	-0.07	0.09	0.05
	jk4-S80-A	-0.05	-0.02	-0.03	-0.06	-	-0.16	-	-	-	0.06
	jk4-S80-B	-0.07	0.02	-	-0.04	-0.05	-0.07	-	-	-	0.05
	jk4-S80-C	-	0.04	0.02	-	-0.02	-	-	-0.07	0.11	0.05

**Table A4.** Bias ( $\text{cm}^3 \text{cm}^{-3}$ ) at CRNS sites for open loop, data assimilation and jackknife simulations compared to CRNS SWC retrievals for the data assimilation period (2011 and 2012). For jackknife experiments (21 in total) only the bias of the omitted sites is reported. The best cases are marked in bold.

Soil map	2013	Rollesbroich	Merzenhausen	Gevenich	Heinsberg	Kall	RurAue	Wüstabach	Aachen	Wildenrath	Mean absolute bias
BK50	OL-BK50	-0.03	0.06	0.01	0.00	-0.02	0.00	-0.02	0.01	0.00	<b>0.02</b>
	Stt-BK50	-0.01	0.04	0.00	0.00	-0.01	-0.01	-0.02	0.00	-0.01	<b>0.01</b>
	PAR-BK50-10	0.06	0.05	0.01	0.02	0.04	0.04	0.02	-0.04	0.00	0.03
	PAR-BK50-30	0.03	0.05	0.00	0.02	0.04	0.03	-0.01	-0.03	0.03	0.03
	jk8-BK50-1 to 9	-0.02	0.04	0.01	-0.01	-0.03	-0.02	-0.04	-0.05	0.11	0.04
FAO	OL-FAO	-0.08	0.02	-0.02	-0.04	-0.08	-0.08	-0.05	-0.08	0.06	0.06
	PAR-FAO-30	0.04	0.05	0.00	0.02	0.03	0.00	-0.02	-0.06	0.03	<b>0.03</b>
Biased (S80)	OL-S80	-0.15	-0.05	-0.10	-0.11	-0.16	-0.15	-0.09	-0.16	-0.01	0.11
	Stt-S80	-0.09	-0.05	-0.10	-0.10	-0.08	-0.14	-0.09	-0.15	-0.01	0.09
	PAR-S80-10	0.04	0.03	-0.03	0.03	0.05	0.03	0.03	-0.04	-0.01	<b>0.03</b>
	PAR-S80-30	0.03	0.05	-0.01	0.02	0.03	0.01	-0.01	-0.03	0.03	<b>0.02</b>
	jk8-S80-1 to 9	-0.07	0.03	0.02	0.02	-0.05	-0.02	-0.04	-0.03	0.10	0.04
	jk4-S80-A	0.00	0.01	0.03	-0.03	-	-0.15	-	-	-	0.04
	jk4-S80-B	-0.03	0.06	-	-0.03	-0.06	-0.06	-	-	-	0.05
	jk4-S80-C	-	0.04	0.02	-	-0.05	-	-	-0.05	0.13	0.06

*Competing interests.* The authors declare that they have no conflict of interest.

*Acknowledgements.* The authors gratefully acknowledge the support by the SFB-TR32 “Pattern in Soil–Vegetation–Atmosphere Systems: Monitoring, Modelling and Data Assimilation”, funded by the Deutsche Forschungsgemeinschaft (DFG), and TERENO (Terrestrial Environmental Observatories), funded by the Helmholtz-Gemeinschaft. The authors also gratefully acknowledge the computing time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JURECA at Jülich Supercomputing Centre (JSC). Finally, the authors acknowledge and thank four anonymous referees for providing constructive comments and the Editor, Nunzio Romano, for guiding the revision process.

The article processing charges for this open-access publication were covered by a Research Centre of the Helmholtz Association.

Edited by: N. Romano

Reviewed by: four anonymous referees

## References

- Ajami, H., McCabe, M. F., Evans, J. P., and Stisen, S.: Assessing the impact of model spin-up on surface water-groundwater interactions using an integrated hydrologic model, *Water Resour. Res.*, 50, 2636–2656, doi:10.1002/2013wr014258, 2014.
- Anderson, J. L.: An ensemble adjustment Kalman filter for data assimilation, *Mon. Weather Rev.*, 129, 2884–2903, doi:10.1175/1520-0493(2001)129<2884: Aeakff>2.0.Co;2, 2001.
- Avery, W. A., Finkenbinder, C., Franz, T. E., Wang, T. J., Nguyen-Robertson, A. L., Suyker, A., Arkebauer, T., and Munoz-Arriola, F.: Incorporation of globally available datasets into the roving cosmic-ray neutron probe method for estimating field-scale soil water content, *Hydrol. Earth Syst. Sci.*, 20, 3859–3872, doi:10.5194/hess-20-3859-2016, 2016.
- Baatz, R., Bogena, H. R., Hendricks Franssen, H. J., Huisman, J. A., Qu, W., Montzka, C., and Vereecken, H.: Calibration of a catchment scale cosmic-ray probe network: A comparison of three parameterization methods, *J. Hydrol.*, 516, 231–244, doi:10.1016/j.jhydrol.2014.02.026, 2014.
- Baatz, R., Bogena, H. R., Hendricks Franssen, H. J., Huisman, J. A., Montzka, C., and Vereecken, H.: An empirical vegetation correction for soil water content quantification using cosmic ray probes, *Water Resour. Res.*, 51, 2030–2046, doi:10.1002/2014WR016443, 2015.
- Bateni, S. M. and Entekhabi, D.: Surface heat flux estimation with the ensemble Kalman smoother: Joint estimation of state and parameters, *Water Resour. Res.*, 48, W08521, doi:10.1029/2011wr011542, 2012.
- Best, M. J., Abramowitz, G., Johnson, H. R., Pitman, A. J., Balsamo, G., Boone, A., Cuntz, M., Decharme, B., Dirmeyer, P. A., Dong, J., Ek, M., Guo, Z., Haverd, V., Van den Hurk, B. J. J., Nearing, G. S., Pak, B., Peters-Lidard, C., Santanello, J. A., Stevens, L., and Vuichard, N.: The Plumbing of Land Surface Models: Benchmarking Model Performance, *J. Hydrometeorol.*, 16, 1425–1442, doi:10.1175/Jhm-D-14-0158.1, 2015.
- Bogena, H. R., Herbst, M., Hake, J. F., Kunkel, R., Montzka, C., Pütz, T., Vereecken, H., and Wendland, F.: MOSYRUR – Water balance analysis in the Rur basin, in: *Schriften des Forschungszentrums Jülich, Reihe Umwelt/Environment*, Forschungszentrum Jülich, Jülich, 2005.
- Bogena, H. R., Huisman, J. A., Baatz, R., Hendricks-Franssen, H. J., and Vereecken, H.: Accuracy of the cosmic-ray soil water content probe in humid forest ecosystems: The worst case scenario, *Water Resour. Res.*, 49, 5778–5791, doi:10.1002/wrcr.20463, 2013.
- Bonan, G. B., Levis, S., Kergoat, L., and Oleson, K. W.: Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models, *Global Biogeochem. Cy.*, 16, 1021, doi:10.1029/2000GB001360, 2002.
- Brutsaert, W.: *Hydrology: an introduction*, Cambridge University Press, Cambridge, New York, 605 pp., 2005.
- Burgers, G., van Leeuwen, P. J., and Evensen, G.: Analysis scheme in the ensemble Kalman filter, *Mon. Weather Rev.*, 126, 1719–1724, doi:10.1175/1520-0493(1998)126<1719:Asitek>2.0.Co;2, 1998.
- Chen, F., Manning, K. W., LeMone, M. A., Trier, S. B., Alfieri, J. G., Roberts, R., Tewari, M., Niyogi, D., Horst, T. W., Oncley, S. P., Basara, J. B., and Blanken, P. D.: Description and evaluation of the characteristics of the NCAR high-resolution land data assimilation system, *J. Appl. Meteorol. Clim.*, 46, 694–713, doi:10.1175/Jam2463.1, 2007.
- Chen, Y. and Zhang, D. X.: Data assimilation for transient flow in geologic formations via ensemble Kalman filter, *Adv. Water Resour.*, 29, 1107–1122, doi:10.1016/j.advwatres.2005.09.007, 2006.
- Clapp, R. B. and Hornberger, G. M.: Empirical Equations for Some Soil Hydraulic-Properties, *Water Resour. Res.*, 14, 601–604, doi:10.1029/Wr014i004p00601, 1978.
- 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, doi:10.1029/Wr020i006p00682, 1984.
- Crow, W. T.: Correcting land surface model predictions for the impact of temporally sparse rainfall rate measurements using an ensemble Kalman filter and surface brightness temperature observations, *J. Hydrometeorol.*, 4, 960–973, 2003.
- Crow, W. T., Berg, A. A., Cosh, M. H., Loew, A., Mohanty, B. P., Panciera, R., de Rosnay, P., Ryu, D., and Walker, J. P.: Upscaling Sparse Ground-Based Soil Moisture Observations for the Validation of Coarse-Resolution Satellite Soil Moisture Products, *Rev. Geophys.*, 50, Rg2002, doi:10.1029/2011rg000372, 2012.
- De Lannoy, G. J. M. and Reichle, R. H.: Assimilation of SMOS brightness temperatures or soil moisture retrievals into a land surface model, *Hydrol. Earth Syst. Sci.*, 20, 4895–4911, doi:10.5194/hess-20-4895-2016, 2016.
- De Lannoy, G. J. M., Reichle, R. H., Arsenault, K. R., Houser, P. R., Kumar, S., Verhoest, N. E. C., and Pauwels, V. R. N.: 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, 2012.

- 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.-Atmos.*, 114, D20104, doi:10.1029/2008JD011650, 2009.
- Dunne, S. and Entekhabi, D.: An ensemble-based reanalysis approach to land data assimilation, *Water Resour. Res.*, 41, W02013, doi:10.1029/2004wr003449, 2005.
- Eltahir, E. A. B.: A soil moisture rainfall feedback mechanism 1. Theory and observations, *Water Resour. Res.*, 34, 765–776, doi:10.1029/97wr03499, 1998.
- Erdal, D., Neuweiler, I., and Wollschlager, U.: Using a bias aware EnKF to account for unresolved structure in an unsaturated zone model, *Water Resour. Res.*, 50, 132–147, doi:10.1002/2012wr013443, 2014.
- Erdal, D., Rahman, M. A., and Neuweiler, I.: The importance of state transformations when using the ensemble Kalman filter for unsaturated flow modeling: Dealing with strong nonlinearities, *Adv. Water Resour.*, 86, 354–365, doi:10.1016/j.advwatres.2015.09.008, 2015.
- Evans, J. G., Ward, H. C., Blake, J. R., Hewitt, E. J., Morrison, R., Fry, M., Ball, L. A., Doughty, L. C., Libre, J. W., Hitt, O. E., Rylett, D., Ellis, R. J., Warwick, A. C., Brooks, M., Parkes, M. A., Wright, G. M. H., Singer, A. C., Boorman, D. B., and Jenkins, A.: Soil water content in southern England derived from a cosmic-ray soil moisture observing system – COSMOS-UK, *Hydrol. Process.*, 30, 4987–4999, doi:10.1002/hyp.10929, 2016.
- Evensen, G.: Sequential Data Assimilation with a Nonlinear Quasi-Geostrophic Model Using Monte-Carlo Methods to Forecast Error Statistics, *J. Geophys. Res.-Oceans.*, 99, 10143–10162, doi:10.1029/94jc00572, 1994.
- FAO: ISRIC ISSCAS Harmonized World Soil Database v1.2, Rome, Italy, 2012.
- Franz, T. E., Zreda, M., Ferre, T. P. A., Rosolem, R., Zweck, C., Stillman, S., Zeng, X., and Shuttleworth, W. J.: Measurement depth of the cosmic ray soil moisture probe affected by hydrogen from various sources, *Water Resour. Res.*, 48, W08515, doi:10.1029/2012wr011871, 2012.
- Franz, T. E., Zreda, M., Ferre, T. P. A., and Rosolem, R.: An assessment of the effect of horizontal soil moisture heterogeneity on the area-average measurement of cosmic-ray neutrons, *Water Resour. Res.*, 49, 6450–6458, doi:10.1002/Wrcr.20530, 2013.
- Geologischer Dienst Nordrhein-Westfalen: Informationssystem Bodenkarte 50, 1:50 000, Geologischer Dienst Nordrhein-Westfalen, Krefeld, Germany, 2009.
- Han, X., Li, X., Franssen, H. J. H., Vereecken, H., and Montzka, C.: Spatial horizontal correlation characteristics in the land data assimilation of soil moisture, *Hydrol. Earth Syst. Sci.*, 16, 1349–1363, doi:10.5194/hess-16-1349-2012, 2012.
- Han, X., Franssen, H. J. H., Rosolem, R., Jin, R., Li, X., and Vereecken, H.: Correction of systematic model forcing bias of CLM using assimilation of cosmic-ray Neutrons and land surface temperature: a study in the Heihe Catchment, China, *Hydrol. Earth Syst. Sci.*, 19, 615–629, doi:10.5194/hess-19-615-2015, 2015.
- Han, X., Franssen, H. J. H., Bello, M. A. J., Rosolem, R., Bogena, H., Alzamora, F. M., Chanzy, A., and Vereecken, H.: Simultaneous Soil Moisture and Properties Estimation for a Drip Irrigated Field by Assimilating Cosmic-ray Neutron Intensity, *J. Hydrol.*, 539, 611–624, doi:10.1016/j.jhydrol.2016.05.050, 2016.
- Han, X. J., Franssen, H. J. H., Li, X., Zhang, Y. L., Montzka, C., and Vereecken, H.: Joint Assimilation of Surface Temperature and L-Band Microwave Brightness Temperature in Land Data Assimilation, *Vadose Zone J.*, 12, 3, doi:10.2136/vzj2012.0072, 2013.
- Han, X. J., Franssen, H. J. H., Montzka, C., and Vereecken, H.: 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, 2014.
- Hawdon, A., McJannet, D., and Wallace, J.: Calibration and correction procedures for cosmic-ray neutron soil moisture probes located across Australia, *Water Resour. Res.*, 50, 5029–5043, doi:10.1002/2013WR015138, 2014.
- Hendricks Franssen, H. J. and Kinzelbach, W.: 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, 2008.
- Hunt, B. R., Kostelich, E. J., and Szunyogh, I.: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter, *Physica D*, 230, 112–126, doi:10.1016/j.physd.2006.11.008, 2007.
- Iwema, J., Rosolem, R., Baatz, R., Wagener, T., and Bogena, H. R.: Investigating temporal field sampling strategies for site-specific calibration of three soil moisture-neutron intensity parameterisation methods, *Hydrol. Earth Syst. Sci.*, 19, 3203–3216, doi:10.5194/hess-19-3203-2015, 2015.
- Jackson, T. J., Le Vine, D. M., Hsu, A. Y., Oldak, A., Starks, P. J., Swift, C. T., Isham, J. D., and Haken, M.: Soil moisture mapping at regional scales using microwave radiometry: The Southern Great Plains Hydrology Experiment, *IEEE T. Geosci. Remote*, 37, 2136–2151, doi:10.1109/36.789610, 1999.
- Kerr, Y. H., Waldteufel, P., Richaume, P., Wigneron, J. P., Ferrazoli, P., Mahmoodi, A., Al Bitar, A., Cabot, F., Gruhier, C., Juglea, S. E., Leroux, D., Mialon, A., and Delwart, S.: The SMOS Soil Moisture Retrieval Algorithm, *IEEE T. Geosci. Remote*, 50, 1384–1403, doi:10.1109/Tgrs.2012.2184548, 2012.
- Kirkpatrick, J. B., Green, K., Bridle, K. L., and Venn, S. E.: Patterns of variation in Australian alpine soils and their relationships to parent material, vegetation formation, climate and topography, *Catena*, 121, 186–194, doi:10.1016/j.catena.2014.05.005, 2014.
- Kohli, M., Schron, M., Zreda, M., Schmidt, U., Dietrich, P., and Zacharias, S.: Footprint characteristics revised for field-scale soil moisture monitoring with cosmic-ray neutrons, *Water Resour. Res.*, 51, 5772–5790, doi:10.1002/2015WR017169, 2015.
- Korres, W., Reichenau, T. G., Fiener, P., Koyama, C. N., Bogena, H. R., Comelissen, T., Baatz, R., Herbst, M., Diekkruger, B., Vereecken, H., and Schneider, K.: Spatio-temporal soil moisture patterns – A meta-analysis using plot to catchment scale data, *J. Hydrol.*, 520, 326–341, doi:10.1016/j.jhydrol.2014.11.042, 2015.
- Koster, R. D., Dirmeyer, P. A., Guo, Z. C., Bonan, G., Chan, E., Cox, P., Gordon, C. T., Kanae, S., Kowalczyk, E., Lawrence, D., Liu, P., Lu, C. H., Malyshev, S., McAvaney, B., Mitchell, K., Mocko, D., Oki, T., Oleson, K., Pitman, A., Sud, Y. C., Taylor, C. M., Verseghy, D., Vasic, R., Xue, Y. K., Yamada, T., and Team, G.: Regions of strong coupling between soil moisture and precipi-

- tation, *Science*, 305, 1138–1140, doi:10.1126/science.1100217, 2004.
- Kumar, S. V., Reichle, R. H., Harrison, K. W., Peters-Lidard, C. D., Yatheendradas, S., and Santanello, J. A.: A comparison of methods for a priori bias correction in soil moisture data assimilation, *Water Resour. Res.*, 48, W03515, doi:10.1029/2010wr010261, 2012.
- Kurtz, W., He, G. W., Kollet, S. J., Maxwell, R. M., Vereecken, H., and Franssen, H. J. H.: TerrSysMP-PDAF (version 1.0): a modular high-performance data assimilation framework for an integrated land surface-subsurface model, *Geosci. Model Dev.*, 9, 1341–1360, doi:10.5194/gmd-9-1341-2016, 2016.
- Lawrence, D. M. and Slater, A. G.: Incorporating organic soil into a global climate model, *Clim. Dynam.*, 30, 145–160, doi:10.1007/s00382-007-0278-1, 2008.
- Lee, J. H.: Spatial-Scale Prediction of the SVAT Soil Hydraulic Variables Characterizing Stratified Soils on the Tibetan Plateau from an EnKF Analysis of SAR Soil Moisture, *Vadose Zone J.*, 13, 11, doi:10.2136/vzj2014.06.0060, 2014.
- Montzka, C., Canty, M., Kunkel, R., Menz, G., Vereecken, H., and Wendland, F.: Modelling the water balance of a mesoscale catchment basin using remotely sensed land cover data, *J. Hydrol.*, 353, 322–334, doi:10.1016/j.jhydrol.2008.02.018, 2008.
- Montzka, C., Moradkhani, H., Weihermuller, L., Franssen, H. J. H., Canty, M., and Vereecken, H.: Hydraulic parameter estimation by remotely-sensed top soil moisture observations with the particle filter, *J. Hydrol.*, 399, 410–421, doi:10.1016/j.jhydrol.2011.01.020, 2011.
- Montzka, C., Pauwels, V. R. N., Franssen, H. J. H., Han, X. J., and Vereecken, H.: Multivariate and Multiscale Data Assimilation in Terrestrial Systems: A Review, *Sensors-Basel*, 12, 16291–16333, doi:10.3390/s121216291, 2012.
- Montzka, C., Grant, J. P., Moradkhani, H., Franssen, H. J. H., Weihermuller, L., Drusch, M., and Vereecken, H.: 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, 2013.
- Moradkhani, H., Sorooshian, S., Gupta, H. V., and Houser, P. R.: Dual state-parameter estimation of hydrological models using ensemble Kalman filter, *Adv. Water Resour.*, 28, 135–147, doi:10.1016/j.advwatres.2004.09.002, 2005.
- Nearing, G. S., Mocko, D. M., Peters-Lidard, C. D., Kumar, S. V., and Xia, Y. L.: Benchmarking NLDAS-2 Soil Moisture and Evapotranspiration to Separate Uncertainty Contributions, *J. Hydrometeorol.*, 17, 745–759, doi:10.1175/Jhm-D-15-0063.1, 2016.
- Oglesby, R. J. and Erickson, D. J.: Soil-Moisture and the Persistence of North-American Drought, *J. Climate*, 2, 1362–1380, doi:10.1175/1520-0442(1989)002<1362:Smatpo>2.0.Co;2, 1989.
- Oleson, K., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Koven, C. D., Levis, S., Li, F., Riley, J. M., Subin, Z. M., Swenson, S., Thornton, P. E., Bozbiyik, A., Fisher, R., Heald, C. L., Kluzek, E., Lamarque, J.-F., Lawrence, P. J., Leung, L. R., Lipscomb, W., Muszala, S. P., Ricciuto, D. M., Sacks, W. J., Sun, Y., Tang, J., and Yang, Z.-L.: Technical description of version 4.5 of the Community Land Model (CLM), NCAR Technical Note NCAR/TN-503+STR, NCAR, Boulder, Colorado, 1–435, doi:10.5065/D6RR1W7M, 2013.
- Patil, N. G. and Singh, S. K.: Pedotransfer Functions for Estimating Soil Hydraulic Properties: A Review, *Pedosphere*, 26, 417–430, doi:10.1016/S1002-0160(15)60054-6, 2016.
- Pauwels, V. R. N., Balenzano, A., Satalino, G., Skriver, H., Verhoest, N. E. C., and Mattia, F.: Optimization of Soil Hydraulic Model Parameters Using Synthetic Aperture Radar Data: An Integrated Multidisciplinary Approach, *IEEE T. Geosci. Remote*, 47, 455–467, doi:10.1109/Tgrs.2008.2007849, 2009.
- Pelowitz, D. B.: MCNPX user's manual, version 5, Rep. LA-CP-05-0369, Los Alamos National Laboratory, Los Alamos, 2005.
- Ray, N. H.: Gas Chromatography. 1. The Separation and Estimation of Volatile Organic Compounds by Gas-Liquid Partition Chromatography, *J. Appl. Chem.*, 4, 21–25, 1954.
- Reichle, R. H.: Data assimilation methods in the Earth sciences, *Adv. Water Resour.*, 31, 1411–1418, doi:10.1016/j.advwatres.2008.01.001, 2008.
- Reichle, R. H., McLaughlin, D. B., and Entekhabi, D.: Hydrologic data assimilation with the ensemble Kalman filter, *Mon. Weather Rev.*, 130, 103–114, doi:10.1175/1520-0493(2002)130<0103:Hdawte>2.0.Co;2, 2002a.
- Reichle, R. H., Walker, J. P., Koster, R. D., and Houser, P. R.: Extended versus ensemble Kalman filtering for land data assimilation, *J. Hydrometeorol.*, 3, 728–740, doi:10.1175/1525-7541(2002)003<0728:Evekff>2.0.Co;2, 2002b.
- Reichle, R. H., Kumar, S. V., Mahanama, S. P. P., Koster, R. D., and Liu, Q.: Assimilation of Satellite-Derived Skin Temperature Observations into Land Surface Models, *J. Hydrometeorol.*, 11, 1103–1122, doi:10.1175/2010jhm1262.1, 2010.
- Robinson, D. A., Campbell, C. S., Hopmans, J. W., Hornbuckle, B. K., Jones, S. B., Knight, R., Ogden, F., Selker, J., and Wendroth, O.: Soil moisture measurement for ecological and hydrological watershed-scale observatories: A review, *Vadose Zone J.*, 7, 358–389, doi:10.2136/Vzj2007.0143, 2008.
- Rosolem, R., Hoar, T., Arellano, A., Anderson, J. L., Shuttleworth, W. J., Zeng, X., and Franz, T. E.: Translating aboveground cosmic-ray neutron intensity to high-frequency soil moisture profiles at sub-kilometer scale, *Hydrol. Earth Syst. Sci.*, 18, 4363–4379, doi:10.5194/hess-18-4363-2014, 2014.
- scilands GmbH: Digital Elevation Model 10 without anthropogenic landforms, Göttingen, 2010.
- Seneviratne, S. I., Luthi, D., Litschi, M., and Schar, C.: Land-atmosphere coupling and climate change in Europe, *Nature*, 443, 205–209, doi:10.1038/Nature05095, 2006.
- Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling, A. J.: Investigating soil moisture–climate interactions in a changing climate: A review, *Earth-Sci. Rev.*, 99, 125–161, doi:10.1016/j.earscirev.2010.02.004, 2010.
- Shi, Y. N., Davis, K. J., Zhang, F. Q., Duffy, C. J., and Yu, X.: Parameter estimation of a physically based land surface hydrologic model using the ensemble Kalman filter: A synthetic experiment, *Water Resour. Res.*, 50, 706–724, doi:10.1002/2013wr014070, 2014.
- Shi, Y. N., Davis, K. J., Zhang, F. Q., Duffy, C. J., and Yu, X.: Parameter estimation of a physically-based land surface hydrologic model using an ensemble Kalman filter: A multivariate real-data experiment, *Adv. Water Resour.*, 83, 421–427, doi:10.1016/j.advwatres.2015.06.009, 2015.

- Shrestha, P., Sulis, M., Masbou, M., Kollet, S., and Simmer, C.: A Scale-Consistent Terrestrial Systems Modeling Platform Based on COSMO, CLM, and ParFlow, *Mon. Weather Rev.*, 142, 3466–3483, doi:10.1175/Mwr-D-14-00029.1, 2014.
- Shuttleworth, J., Rosolem, R., Zreda, M., and Franz, T.: The COsmic-ray Soil Moisture Interaction Code (COSMIC) for use in data assimilation, *Hydrol. Earth Syst. Sci.*, 17, 3205–3217, doi:10.5194/hess-17-3205-2013, 2013.
- Simmer, C., Thiele-Eich, I., Masbou, M., Amelung, W., Bogena, H., Crewell, S., Diekkruiger, B., Ewert, F., Franssen, H. J. H., Huisman, J. A., Kemna, A., Klitzsch, N., Kollet, S., Langensiepen, M., Lohnert, U., Rahman, A. S. M. M., Rascher, U., Schneider, K., Schween, J., Shao, Y. P., Shrestha, P., Stiebler, M., Sulis, M., Vanderborght, J., Vereecken, H., van der Kruk, J., Waldhoff, G., and Zerenner, T.: Monitoring and modeling the terrestrial system from pores to catchment, The Transregional Collaborative Research Center on Patterns in the Soil–Vegetation–Atmosphere System, *B. Am. Meteorol. Soc.*, 96, 1765–1787, doi:10.1175/Bams-D-13-00134.1, 2015.
- Song, X. H., Shi, L. S., Ye, M., Yang, J. Z., and Navon, I. M.: Numerical Comparison of Iterative Ensemble Kalman Filters for Unsaturated Flow Inverse Modeling, *Vadose Zone J.*, 13, 2, doi:10.2136/vzj2013.05.0083, 2014.
- Sun, Y., Hou, Z., Huang, M., Tian, F., and Leung, L. R.: Inverse modeling of hydrologic parameters using surface flux and runoff observations in the Community Land Model, *Hydrol. Earth Syst. Sci.*, 17, 4995–5011, doi:10.5194/hess-17-4995-2013, 2013.
- Temimi, M., Lakhankar, T., Zhan, X. W., Cosh, M. H., Krakauer, N., Fares, A., Kelly, V., Khanbilvardi, R., and Kumassi, L.: Soil Moisture Retrieval Using Ground-Based L-Band Passive Microwave Observations in Northeastern USA, *Vadose Zone J.*, 13, 3, doi:10.2136/vzj2013.06.0101, 2014.
- Vereecken, H., Huisman, J. A., Bogena, H., Vanderborght, J., Vrugt, J. A., and Hopmans, J. W.: On the value of soil moisture measurements in vadose zone hydrology: A review, *Water Resour. Res.*, 44, W00d06, doi:10.1029/2008wr006829, 2008.
- Vereecken, H., Schnepf, A., Hopmans, J. W., Javaux, M., Or, D., Roose, D. O. T., Vanderborght, J., Young, M. H., Amelung, W., Aitkenhead, M., Allison, S. D., Assouline, S., Baveye, P., Berli, M., Bruggemann, N., Finke, P., Flury, M., Gaiser, T., Govers, G., Ghezzehei, T., Hallett, P., Franssen, H. J. H., Heppell, J., Horn, R., Huisman, J. A., Jacques, D., Jonard, F., Kollet, S., Lafolie, F., Lamorski, K., Leitner, D., McBratney, A., Minasny, B., Montzka, C., Nowak, W., Pachepsky, Y., Padarian, J., Romano, N., Roth, K., Rothfuss, Y., Rowe, E. C., Schwen, A., Simunek, J., Tiktak, A., Van Dam, J., van der Zee, S. E. A. T. M., Vogel, H. J., Vrugt, J. A., Wohling, T., and Young, I. M.: Modeling Soil Processes: Review, Key Challenges, and New Perspectives, *Vadose Zone J.*, 15, 5, doi:10.2136/vzj2015.09.0131, 2016.
- Villarreyes, C. A. R., Baroni, G., and Oswald, S. E.: Inverse modelling of cosmic-ray soil moisture for field-scale soil hydraulic parameters, *Eur. J. Soil Sci.*, 65, 876–886, doi:10.1111/ejss.12162, 2014.
- Vrugt, J. A., Diks, C. G. H., Gupta, H. V., Bouten, W., and Verstraten, J. M.: 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, 2005.
- Waldhoff, G.: Enhanced Land Use Classification of 2009 for the Rur catchment, CRC/TR32 Database (TR32DB), Cologne, doi:10.5880/TR32DB.2, 2012.
- Whan, K., Zscheischler, J., Orth, R., Shongwe, M., Rahimi, M., Asare, E. O., and Seneviratne, S. I.: Impact of soil moisture on extreme maximum temperatures in Europe, *Weather Clim. Extr.*, 9, 57–67, doi:10.1016/j.wace.2015.05.001, 2015.
- Wu, C. C. and Margulis, S. A.: Feasibility of real-time soil state and flux characterization for wastewater reuse using an embedded sensor network data assimilation approach, *J. Hydrol.*, 399, 313–325, doi:10.1016/j.jhydrol.2011.01.011, 2011.
- Wu, C. C. and Margulis, S. A.: Real-Time Soil Moisture and Salinity Profile Estimation Using Assimilation of Embedded Sensor Datastreams, *Vadose Zone J.*, 12, 1, doi:10.2136/vzj2011.0176, 2013.
- Zacharias, S., Bogena, H., Samaniego, L., Mauder, M., Fuss, R., Putz, T., Frenzel, M., Schwank, M., Baessler, C., Butterbach-Bahl, K., Bens, O., Borg, E., Brauer, A., Dietrich, P., Hajnsek, I., Helle, G., Kiese, R., Kunstmann, H., Klotz, S., Munch, J. C., Papen, H., Priesack, E., Schmid, H. P., Steinbrecher, R., Rosenbaum, U., Teutsch, G., and Vereecken, H.: A Network of Terrestrial Environmental Observatories in Germany, *Vadose Zone J.*, 10, 955–973, doi:10.2136/Vzj2010.0139, 2011.
- Zreda, M., Desilets, D., Ferre, T. P. A., and Scott, R. L.: Measuring soil moisture content non-invasively at intermediate spatial scale using cosmic-ray neutrons, *Geophys. Res. Lett.*, 35, L21402, doi:10.1029/2008GL035655, 2008.
- Zreda, M., Shuttleworth, W. J., Zeng, X., Zweck, C., Desilets, D., Franz, T., and Rosolem, R.: COSMOS: the COsmic-ray Soil Moisture Observing System (vol. 16, pg. 4079, 2012), *Hydrol. Earth Syst. Sci.*, 17, 1065–1066, doi:10.5194/hess-17-1065-2013, 2012.