Hydrology and
Earth System
Sciences

SMOS brightness temperature assimilation into the Community Land Model

Dominik Rains[1], **Xujun Han**[2], **Hans Lievens**[1,3], **Carsten Montzka**[2], and **Niko E. C. Verhoest**[1]

[1]Laboratory of Hydrology and Water Management, Ghent University, Ghent, Belgium
[2]Forschungszentrum Jülich GmbH, Institute of Bio- and Geosciences: Agrosphere (IBG-3), Jülich, Germany
[3]Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD, USA

Correspondence to: Dominik Rains (dominik.rains@ugent.be)

Abstract. SMOS (Soil Moisture and Ocean Salinity mission) brightness temperatures at a single incident angle are assimilated into the Community Land Model (CLM) across Australia to improve soil moisture simulations. Therefore, the data assimilation system DasPy is coupled to the local ensemble transform Kalman filter (LETKF) as well as to the Community Microwave Emission Model (CMEM). Brightness temperature climatologies are precomputed to enable the assimilation of brightness temperature anomalies, making use of 6 years of SMOS data (2010–2015). Mean correlation R with in situ measurements increases moderately from 0.61 to 0.68 (11 %) for upper soil layers if the root zone is included in the updates. A reduced improvement of 5 % is achieved if the assimilation is restricted to the upper soil layers. Root-zone simulations improve by 7 % when updating both the top layers and root zone, and by 4 % when only updating the top layers. Mean increments and increment standard deviations are compared for the experiments. The long-term assimilation impact is analysed by looking at a set of quantiles computed for soil moisture at each grid cell. Within hydrological monitoring systems, extreme dry or wet conditions are often defined via their relative occurrence, adding great importance to assimilation-induced quantile changes. Although still being limited now, longer L-band radiometer time series will become available and make model output improved by assimilating such data that are more usable for extreme event statistics.

1 Introduction

The potential to improve land surface simulations of soil moisture by assimilating information derived from satellite measurements is well known (Parada and Liang, 2004; De Lannoy et al., 2007; Jia et al., 2009; Chen et al., 2014; Mohanty et al., 2017). Soil moisture products based on data from a number of missions have been used, e.g. ASCAT (Brocca et al., 2010, 2012; Dharssi et al., 2011; Draper et al., 2011), AMSR-E (Reichle et al., 2007; Yang et al., 2007; Draper et al., 2009a) or a combination of both (Draper et al., 2012; Renzullo et al., 2014). Launched in November 2009, the Soil Moisture and Ocean Salinity (SMOS) spacecraft is the first mission specifically designed to map soil moisture from space (Kerr et al., 2001; Mecklenburg et al., 2016), with the second one being the similar SMAP mission launched in 2015 (Entekhabi et al., 2010). The passive Microwave Imaging Radiometer with Aperture Synthesis (MIRAS) instrument aboard SMOS, sensitive to 1.4 GHz electromagnetic emissions, measures multi-angular top-of-atmosphere brightness temperatures at horizontal (H) and vertical (V) polarisation. These brightness temperatures are ingested into a complex retrieval algorithm, resulting in soil moisture estimates (Kerr et al., 2012) readily usable for analysis, input for higher-level products or data assimilation. When assimilating these products, which roughly represent the top 5 cm of the soil column, into the according model layers (Reichle, 2008; Montzka et al., 2012), the assimilation impact in deeper layers will depend on model physics (Montaldo et al., 2001; Kumar et al., 2009; Montzka et al., 2011). Alternatively, by making use of one of the key advantages of the various implementations of the Kalman filter (Kalman, 1960), deeper

and unobserved layers can be updated directly. For plants, these deeper layers act as the root zone, where soil moisture has a profound effect on biochemical processes, thus limiting the effect of data assimilation not only to soil moisture (Vereecken et al., 2016). Examples for assimilating SMOS soil moisture retrievals are, among others, given by Martens et al. (2016), showing that the GLEAM evapotranspiration model can benefit from assimilating these data over Australia, or Lievens et al. (2015b), who conclude that the positive assimilation impact on soil moisture can improve streamflow simulations for the variable infiltration capacity (VIC) model, as shown in the Murray–Darling Basin. The impact on both streamflow and evaporation is evaluated by Ridler et al. (2014) for western Denmark. Leroux et al. (2016) assimilate SMOS soil moisture products into the Distributed Hydrology Soil Vegetation Model (DHSVM), improving water table depth and streamflow simulations, thereby greatly reducing the uncertainties introduced by the use of uncorrected near-real-time precipitation forcings. Scholze et al. (2016) have assimilated SMOS retrievals together with CO_2 measurements to constrain the global carbon cycle.

Apart from assimilating the retrieved soil moisture products, it is also possible to directly assimilate the brightness temperatures, which should, in theory, eliminate a number of problems. For instance, the SMOS Level 2 soil moisture retrievals represent the optimum fits between simulated brightness temperatures and the observed satellite signal (Kerr et al., 2012). The simulated top-of-atmosphere signal thereby depends on both static and dynamic ancillary data based on input and output of a specific land surface model, e.g. for SMOS retrievals the European Centre for Medium-Range Weather Forecasts HTESSEL land surface model (Balsamo et al., 2009). When using a modified or different land surface model, it can be beneficial to directly assimilate the brightness temperatures in order to use consistent auxiliary information for the land surface model and the radiative transfer model. In the case of assimilating soil moisture retrievals, the auxiliary data used by the model are likely to be correlated with the data used by the retrievals. This inevitably leads to cross-correlated errors between the model and the retrievals, which may have a negative impact on the assimilation performance (De Lannoy and Reichle, 2016a). Some examples of brightness temperature assimilation studies are given by Jia et al. (2009), Muñoz-Sabater (2015), De Lannoy and Reichle (2016a) and Lievens et al. (2016). Taken as a whole, assimilating L-band brightness temperatures in practical terms is quite a new concept which still needs further exploring. The assimilation impact is mostly evaluated by comparing soil moisture time series to a limited number of in situ measurements. Given the proven positive impact and the increased availability of longer time series of satellite observations, hydrological monitoring systems, such as for droughts or floods, are likely to benefit from these data. However, little is still known about long-term assimilation impacts, e.g. on quantiles, which are often used for applications such as drought monitoring.

Within this study, we assimilate SMOS brightness temperatures at H polarisation over Australia from January 2010 until December 2015 into the Community Land Model (version 4.5, Oleson et al., 2013) and evaluate the assimilation impact both in terms of correlation improvements towards in situ measurements and in terms of long-term induced model biases, i.e. changes in quantiles, for the state variable soil moisture. We place the findings within the context of hydrological monitoring systems, which mostly use cumulative distribution functions (CDFs) as a basis to classify areas of interest. A good overview on the evolution of such hydrological monitoring systems is given by Van Dijk and Renzullo (2011).

We have selected Australia as a study site as we consider it as an ideal test domain for the long-term brightness temperature assimilation. It is quite heterogeneous in terms of climate and largely uninfluenced by human activity and therefore mostly unaffected by radio frequency interference (Leroux et al., 2013). Although large parts are covered by drylands, the land cover varies along the coastline and includes some densely forested areas in the Australian Alps as well as pasture and areas of intense agricultural activity in the south-east and south-west. The lack of large, densely vegetated areas, which mask out the L-band emissions sensitive to soil moisture, is beneficial. Furthermore, soil moisture information based on satellite data is often advertised as being especially useful for monitoring hydrological extremes such as floods and droughts, both of which Australia is susceptible to (van Dijk et al., 2013; Johnson et al., 2016; Kiem et al., 2016). In addition to the ones already mentioned, a number of studies have focused on Australia, covering soil moisture retrieval (Van der Schalie et al., 2015), assimilation studies (Lievens et al., 2015b), validation studies and field campaigns for SMOS (Peischl et al., 2009; Panciera et al., 2008), as well as SMAP (Panciera et al., 2014) and soil moisture downscaling experiments (Piles et al., 2011; Merlin et al., 2012; Dumedah et al., 2015). More recently, SMOS soil moisture and GRACE water storage data have been jointly assimilated by Tian et al. (2017). A comparison of SMOS satellite soil moisture retrievals with products based on other sensors is given by Su et al. (2013). The potential of AMSR-E soil moisture retrievals has been shown by Draper et al. (2009b) and the joint assimilation of ASCAT and AMSR-E data has been tested by Renzullo et al. (2014). Downscaled AMSR-E soil moisture observations were assimilated within the Murrumbidgee Basin by López López et al. (2016).

The Community Land Model (CLM) provides all outputs required for the brightness temperature forward simulations, which further motivates the direct assimilation of brightness temperatures. Being part of the fully coupled Community Earth System Model (CESM), it can be used for future coupled land–atmosphere studies using a similar setup as for the

brightness temperature assimilation. A full description of the CLM surface data used for modelling the Australian continent will be given in Sect. 2.

In order to obtain the brightness temperature forward simulations, the CLM is coupled to the Community Microwave Emission Model (version 5.1, Drusch et al., 2009) forward operator within the data assimilation system DasPy (Han et al., 2015a). The increments are computed with the local ensemble transform Kalman filter (LETKF) (Miyoshi and Yamane, 2007; Hunt et al., 2007). The observation bias between forward simulations and observed brightness temperatures is encountered by assimilating anomalies. Remaining differences in mean and variance are resolved by quantile mapping the entire observation anomaly time series towards the open-loop forward simulation anomalies at each grid point. Details on the implementation of the assimilation system, the forward simulations and the observation treatment will be given in Sect. 3.

The in situ data used for the validation are from the OzNet and CosmOz measurement networks (Smith et al., 2012; Hawdon et al., 2014) and were obtained through the International Soil Moisture Network (ISMN) (Dorigo et al., 2011). For the quantile analysis, quantiles at 1 % steps are computed at each model grid point, allowing the empirical estimation of the cumulative distribution functions. To exemplify the effects of quantile changes, we show a dry event defined at the 10 % quantile level and to what extent its spatial extent changes when comparing the open-loop run to the data assimilation results. Part of the experiments is also to show how the CLM translates assimilation updates restricted to the upper soil layers into the root zone purely through model physics as compared to directly updating both the upper soil layers as well as the root zone, with the findings being set in relation to the quantile analysis. The results of the experiments will be given in Sect. 4, followed by the discussion and conclusion in Sect. 5.

2 The Community Land Model

The Community Land Model is the land surface component of the Community Earth System Model and can be run offline with precomputed atmospheric forcings (Oleson et al., 2013). CLM provides global surface datasets which can be interpolated to predefined or custom resolutions and grid types both globally as well as regionally, including single point simulations. Interpolating the included surface datasets resulted in artefacts for elevation and grid cell elevation variance as well as plant functional types, with one plant functional type clearly linked to latitudinal borders. We replaced these, but also other surface datasets, with suitable alternatives. For the choice of datasets, we kept possible future global applications in mind, which the results of this study could be compared with. At the same time, we believe that the Australian continent is well represented by the chosen

datasets or that no better-suited alternatives were available for the requirements of this study. A description of these datasets follows in the next section. The model resolution was defined at 0.25°, which agrees well with the Level 3 observations provided in the EASE 25 km grid. The model is run at 30 min time steps, with hourly outputs, allowing for a correct temporal alignment of model and satellite observations.

2.1 Surface datasets

Each grid cell within CLM is divided into land units covering a certain percentage of the total grid cell area. Possible land units consist of vegetation, wetlands, lakes, glaciers and urban areas. Vegetated land units have a single set of soil properties but can be populated by several plant functional types (PFTs), again defined by their percentage of coverage with respect to the entire grid cell (Bonan et al., 2002). We have updated the model PFTs with information from the Moderate Resolution Imaging Spectroradiometer (MODIS) MCD12Q1 (version 5) land cover products, provided at 500 m resolution in sinusoidal projection and containing a classification of each grid cell describing the dominant plant functional type. On the basis of WorldClim climate data (Hijmans et al., 2005), these plant functional types are reclassified to the CLM-compatible PFTs (Bonan et al., 2002). PFTs were then aggregated to the model resolution, computing the percentage of 500 m pixels of each plant functional type per grid cell. Monthly leaf area index (LAI) values for each PFT within a grid cell were computed by averaging the MODIS 8-daily MCD15A3H (version 6) LAI product, also provided at 500 m resolution in sinusoidal projection, over the assimilation period (2010–2015) to derive the monthly climatology and to replace the original climatological LAI values of CLM. The high-resolution LAI values were upscaled to model resolution by mapping the 500 m pixels to the 500 m reclassified PFT values within each grid cell and subsequently averaging these per PFT. Stem area index (SAI) values were also computed on the basis of the high-resolution MODIS LAI data and likewise upscaled to model resolution, replacing the standard CLM values. Urban and lake areas were extracted from the MODIS land cover information MCD12Q1. Mean topographic height and standard deviation for each grid cell were downscaled from the 3 arcsec HydroSHEDS digital elevation model (Lehner et al., 2008). Soil texture, namely clay and sand fractions, as well as organic matter content, were obtained from the global International Soil Reference and Information System (ISRIC) soil database (Hengl et al., 2014) and mapped to the 10 CLM soil layers by nearest-neighbour interpolation according to their respective depths. The ISRIC database provided information on organic matter as the gravimetric percentage of the fine-scale soil fraction and we assumed that the coarse-scale soil fraction contains no organic matter. Bulk density was used to compute the organic matter content required by CLM, as-

suming 0.58 g organic matter per kilogram. The rationale for creating high-resolution datasets for CLM closely followed the approaches described in detail in Ke et al. (2012) and Han et al. (2012), who similarly replaced the CLM standard datasets.

2.2 ERA-Interim atmospheric forcing

CLM provides forcings (CRUNCEP) which do not cover the required time period. Therefore, due to the release of ERA-Interim reanalysis data (Dee et al., 2011) with a time lag of only a few months, these data were used to force the CLM land surface model over Australia.

The variables 2 m air temperature, 2 m pressure, shortwave incoming radiation and total precipitation were extracted, and specific humidity was computed from the ERA-Interim 2 m dew point temperature and 2 m air temperature. The 2 m wind speed was derived from the provided wind speed components in lateral and longitudinal directions. With ERA-Interim being produced by assimilating a multitude of observations into an atmospheric model, some of these variables are the result of the analysis step and others of the forecast step; thus, the data needed to be handled respectively. Forecasts for flux variables are provided bi-daily at 00:00 and 12:00 UTC for 3, 6, 9 and 12 h forecast periods and in accumulated form. For example, the precipitation forecast for a 6 h time window is the accumulated precipitation over 6 h. In order to obtain a precipitation estimate for 03:00–06:00 UTC, the precipitation forecast for the first 3 h window needs to be subtracted. This disaggregation was performed for all flux variables to obtain 3-hourly forcing estimates. Analysis variables are valid as instantaneous estimates and no disaggregation had to be performed in their case. The atmospheric forcings were bi-linearly interpolated from 0.75° spatial resolution to 0.25° model resolution. A similar approach for creating atmospheric forcing data based on ERA-Interim, but with additional corrections through ancillary data, is described in Weedon et al. (2014). Time interpolation from 3-hourly to 1-hourly time steps is performed at CLM runtime with an appropriate interpolation algorithm applied to each variable. Incoming radiation is interpolated by using a cosine function simulating the position of the Sun; for precipitation, a nearest-neighbour interpolation is used. For the remaining variables, linear interpolation is applied.

3 Assimilation system

The assimilation experiments are performed with the open-source multivariate data assimilation system DasPy. Mainly coded in Python, its modular design, in principle, allows the coupling of different models, observation operators and assimilation algorithms. The version used within this study is coupled to the Community Land Model and the Community Microwave Emission Model (CMEM; de Rosnay et al.,

2009) observation operator. Furthermore, the system uses the LETKF implementation by Miyoshi and Yamane (2007) for computing the actual increments. Several studies have been performed using DasPy, including the assimilation of synthetic brightness temperatures within the Babaohe River basin in north-western China (Han et al., 2012) and in the Rur catchment in Germany (Han et al., 2015b). The system allows for dual state parameter estimation, as shown in Han et al. (2014).

DasPy has been developed with a focus on high-performance computing. Parallelism is achieved through ParallelPython, OpenMP, the Message Parsing Interface (MPI) and MPI4Python. Ensemble members can be distributed across different nodes with the core assimilation system, including the LETKF, being confined to one node. Some of the operations are implemented in C++ within the Python environment, using Weave, to further optimise performance. The LETKF itself is a fully parallel Fortran implementation called through F2PY (Fortran2Python).

3.1 Local ensemble transform Kalman filter

The local ensemble transform Kalman filter (Hunt et al., 2007) is one of the implementations of the ensemble Kalman square root filter and is deterministic as opposed to stochastic, thus not introducing random noise into the observations. The LETKF has the advantage over other non-localised implementations that the analysis performed for each grid point is limited to a local domain, which makes it computationally more efficient and less susceptible to long-range spurious correlations. The original SMOS footprint is 43 km across and thus covers more than a single model grid cell, which would justify the assimilation in 3-D. However, mostly for reasons of simplicity, and also due to the previously performed inverse distance observation regridding partially accounting for this, we only use observations directly covering a grid cell.

Mathematically, the LETKF can be described as follows. Model states for each ensemble member k from a total of K ensemble members are propagated over time by the model M, starting at a previous time step, $n-1$, e.g. a previous analysis step within the data assimilation scheme, x_{n-1}^{a}. This results in a new background estimate of the state vector $\boldsymbol{x}^b$ consisting of the soil moisture states for all ensemble members at the current time step n.

$$\boldsymbol{x}_{n,k}^{b} = M_n\left(\boldsymbol{x}_{n-1,k}^{a}\right) \tag{1}$$

The background ensemble perturbations $\mathbf{X}^b$ at the current time step can be computed as

$$\mathbf{X}^b = \left[\boldsymbol{x}_1^b - \overline{\boldsymbol{x}}^b | \ldots | \boldsymbol{x}_K^b - \overline{\boldsymbol{x}}^b\right]. \tag{2}$$

The individual ensemble states $\boldsymbol{x}^b$ are mapped into observation space using a forward operator $\mathbf{H}$, in this case CMEM:

$$\boldsymbol{y}_k^b = \mathbf{H}\left(\boldsymbol{x}_k^b\right), \tag{3}$$

and the forward simulation perturbations are defined as

$$\mathbf{Y}^b = \left[\boldsymbol{y}_1^b - \overline{\boldsymbol{y}}^b | \dots | \boldsymbol{y}_K^b - \overline{\boldsymbol{y}}^b \right]. \tag{4}$$

Within the ensemble space, with $\mathbf{R}$ being the observation variance, the analysis error covariance $\widetilde{\mathbf{P}}^a$ is computed through

$$\widetilde{\mathbf{P}}^a = \left[(K-1)\mathbf{I} + (\mathbf{Y}^b)^T \mathbf{R}^{-1} \mathbf{Y}^b \right]^{-1}, \tag{5}$$

allowing for the computation of $\overline{\boldsymbol{w}}^a$ as the mean weighting vector, with y^o being the observations,

$$\overline{\boldsymbol{w}}^a = \widetilde{\mathbf{P}}_a (\mathbf{Y}^b)^T \mathbf{R}^{-1} \left(\boldsymbol{y}^o - \overline{\boldsymbol{y}}^b \right), \tag{6}$$

resulting in the analysis mean $\overline{\boldsymbol{x}}_a$.

$$\overline{\boldsymbol{x}}_a = \overline{\boldsymbol{x}}_b + \mathbf{X}^b \overline{\boldsymbol{w}}^a \tag{7}$$

The analysis perturbations are defined as $\mathbf{X}^a$:

$$\mathbf{X}^a = \mathbf{X}^b + \mathbf{W}^a, \tag{8}$$

with

$$\mathbf{W}^a = \sqrt{(K-1)\widetilde{\mathbf{P}}^a}, \tag{9}$$

where the analysis error covariance P^a is given by

$$\mathbf{P}^a = \mathbf{X}^b \widetilde{\mathbf{P}}^a (\mathbf{X}^b)^T. \tag{10}$$

3.2   Ensemble generation

Model uncertainty is simulated by running the model in ensembles with perturbations applied either to the atmospheric forcings, surface dataset, model parameters or possible combinations of these. In order to account for the model uncertainty in this study, CLM is run with 32 ensembles with spatially uncorrelated perturbations added to some of the ERA-Interim forcing data, namely air temperature, shortwave radiation and precipitation. Shortwave radiation is perturbed with multiplicative noise with a standard deviation of 0.3, whereas for temperature additive noise with a standard deviation of 2.5 K is applied. Finally, precipitation is perturbed with multiplicative log-normal noise with a standard deviation of 0.3. The perturbation factors are the same as used by Reichle et al. (2007) and Han et al. (2014). No spatially correlated noise was added, as the experiments are carried out in 1-D, only using one observation per grid cell. To avoid ensemble collapse during dry periods, soil texture is also perturbed once at model startup. Here, multiplicative noise with a standard deviation of 10 % for clay and sand for the top two soil layers is applied. For lower layers, the top-layer multiplicative factor is rescaled by using the inverse relationship

between the thickness of each soil layer and the summed soil layer thickness of the two top layers (see Table 1). This is to ensure that increments in lower soil layers do not result in very large changes in soil water in absolute terms, since soil layer thickness greatly increases towards lower layers. With CLM deriving hydraulic properties based on soil texture, it is to be noted that as a consequence each ensemble member runs with slightly modified model physics.

3.3   Observation operator

Forward simulations from the model space to the observation space are performed with CMEM version 5.1. Model output at each observation time, with the observation time rounded to the full hour, serves as input in order to simulate brightness temperatures as measured by the satellite. SMOS ascending and descending orbits have local overpass times of approximately 06:00 and 18:00 UTC. Forward simulations are thus computed at 08:00 UTC on the same day and 20:00 UTC on the previous day for descending and ascending acquisitions, respectively, assuming an average time shift of $-10\,\text{h}$ for the entire Australian continent. This greatly decreases the number of analysis steps, since individual orbits within one day can be assimilated at once, assuming that a sufficiently correct temporal alignment between observations and model forward simulations is provided. With the western parts of Australia deviating by 2 h and the ERA-Interim forcings being interpolated from 3-hourly to 1-hourly data, we consider this approach to be acceptable for the purpose of the study.

CMEM requires time-invariant information such as soil layer depth, sand, clay and water fractions, surface height as well as the vegetation types covering the grid cell. The CLM plant functional types need to be reclassified to ECO-CLIMAP vegetation classes (Champeaux et al., 2005) and the percentage of the low and high vegetation types is then used by the CMEM. Based on this reclassification, the LAI information is assigned to the ECOCLIMAP low vegetation classes accordingly. For the offline forward simulations, CLM was run with LAI as daily output in order to make use of the model-internal LAI interpolation, creating a smooth LAI time series based on the monthly surface dataset. This also ensures that the LAI values used for the CMEM forward simulation are the same as those used within CLM during assimilation. LAI values for high vegetation classes are fixed within CMEM and not taken from the CLM input data. Other dynamic fields used in the forward simulations are soil moisture and soil temperature and 2 m air temperature. CMEM supports different types of submodules for specific calculations. Within this study, the Mironov model (Mironov et al., 2004) has been chosen for the dielectric constant computation. Vegetation temperature is computed directly by CLM and used as an input without the need of an approximation, e.g. through air temperature. Effective temperature is obtained through the Wigneron model (Wigneron et al., 2001) and applied in the dielectric model.

Table 1. CLM soil layer depths and relative layer thickness with respect to the sum of the two top layers. The relative thickness was used as a scaling factor for the soil perturbations, effectively decreasing ensemble spread and error covariance for lower levels.

Layer depth (m)	0.018	0.045	0.09	0.17	0.290	0.493	0.829	1.383	2.296	3.802
Perturb. scaling	1	1	1	0.60	0.36	0.22	0.13	0.08	0.05	0.03

For smooth surface emissivity, soil roughness and vegetation opacity, the Fresnel, Choudhury (Choudhury et al., 1979) and Wigneron (Wigneron et al., 2007) models are used, respectively. Finally, atmospheric contributions are estimated with the Pellarin methodology (Pellarin et al., 2003). For all modules, the standard parameters for CMEM 5.1 remained unchanged and the forward observation model was not calibrated. Although the standard parameters are very unlikely to be perfect for the different land cover classes, we argue that this approach is not necessarily worse than the alternative of calibrating the radiative transfer model. By modifying parameters, such as surface roughness, the bias between forward simulations and observations can be removed, but in some cases at the expense of a reduced sensitivity towards soil moisture. Therefore, we remove the bias between simulations and observations through the computation of forward simulation and observation anomalies.

3.4 Observations and anomaly preparation

Large biases are common between modelled and observed brightness temperatures due to the many uncertainties involved, such as in the atmospheric forcing, the land surface representation, the land surface model itself as well as the radiative transfer model and its parameterisation (Drusch et al., 2009; Barella-Ortiz et al., 2017), with this study being no exception. The assimilation is expected to correct random errors only, i.e. bias blind, and it is therefore necessary to remove the bias prior to assimilation (Yilmaz and Crow, 2013). Calibrating the radiative transfer model to closely match the observed time series is a possible solution, as shown by Drusch et al. (2009), De Lannoy et al. (2013) and Lievens et al. (2015a), with the alternative being the rescaling of the measurements to mimic more closely the forward simulations (Lievens et al., 2015b), as mentioned above. The details of preparing the observations prior to assimilation are given here.

SMOS Level 3 daily brightness temperatures at horizontal H polarisation and 42.5 incidence angle provided by Centre Aval de Traitement des Données (CATDS) are used in the study and processed for the years 2010–2015 (version 310). The data are rigorously filtered by using ancillary data from the corresponding Level 3 soil moisture products (version 300), excluding measurements with a probability of radio frequency interference (RFI) greater than 0.2 and a Data Quality Index (DQX) value greater than 0.07. Measurements with a number of activated science flags, namely strong topography, snow, flooding, urban areas, coastal zone

and precipitation, are not considered either. The filtered observation data are regridded from the Equal-Area Scalable Earth Grid 2 (EASE2) 25 km grid to the 0.25° rectangular model grid by using inverse-distance interpolation.

On the basis of these data, we compute the climatology for each day and at each grid point for the years 2010–2015 by averaging along a 14-day moving window across the 6 years, producing separate climatologies for ascending and descending orbits. Anomalies are then computed by subtracting the original SMOS time series from the climatologies. Brightness temperature forward simulations based on an open-loop run with 32 ensembles are performed and the ensemble mean climatology and forward simulation anomalies are derived in the same way as for the observations. The SMOS anomalies are then quantile matched to the ensemble average forward simulation anomalies to account for the differences in variance. The full approach of anomaly computation and quantile matching is to account for seasonal differences between simulations and observations (see, e.g. De Lannoy and Reichle, 2016a) and to remove the bias without more aggressive CDF-matching techniques at seasonal level being required. The original brightness temperature simulations over the entire period exhibited a mean warm bias of 21 K for the ascending orbit and 26 K for the descending orbit. Anomaly correlations prior to quantile matching are 0.21 and 0.39, and after quantile matching 0.38 and 0.60 for ascending and descending orbits, respectively. Based on the scaling factor between the standard deviation of the original and CDF-matched SMOS anomalies, the observation variance is recomputed. The unscaled observation variance $R = 5\,\text{K}^2$ was defined, accounting for a standard instrument error of 3 K and an assumed combined standard mean error of 4 K for the forward simulations and representativeness error. The instrument error can be seen as a low estimate and is based on the assumption that the brightness temperature binning around the 42.5° incidence angle results in a slight reduction, when compared to the 4 K instrument error usually applied for Level 1 data.

During assimilation at each time step, the current forward simulation is subtracted from the precomputed forward simulation climatology to compute on-the-fly anomalies. The difference between this simulated anomaly and the SMOS anomaly is the innovation, which is used within the LETKF algorithm. The assumption is made that the forward simulation climatology does not significantly change during the assimilation run. In total, there are 2063 and 2044 observations for the ascending and descending orbits, respectively.

4 Data assimilation and results

In total, three assimilation experiments are carried out, updating either top-layer soil moisture or both top-layer and root-zone soil moisture. Only updating the upper soil moisture allows for testing the ability of the model to feed the assimilation effects into the root zone through model physics only. Updating the root zone is carried out with two sets of soil texture perturbations, which largely influence the modelled background error. The objective is to validate the assimilation impact by comparing the time series before and after assimilation to a number of in situ measurements. In addition, shifts in the soil moisture quantiles with respect to the open-loop run are analysed to highlight some long-term effects of the data assimilation. A set of quantiles is computed at each grid cell to allow the empirical estimation of the cumulative distribution functions, since shifts, both positive or negative, are possible at different quantile levels. Both the impact on correlation and the long-term effects are set in relation to which layers are updated and to the model background error in the root zone. The experiments and their results are described in the following and set into the context of their potential effect on hydrological monitoring systems, as shown for the exemplary classification of a dry event. We believe this to be relevant, since L-band data, or data from other sources, are in the long run likely to be incorporated into more and more operational systems.

The spatial patterns of the open-loop soil moisture simulations at different depths were compared to the locally optimised AWRA-L land surface model (http://www.bom.gov.au/water/landscape) to ensure that the CLM simulations are plausible.

In the first experiment (DA 1), only the upper three CLM soil layers, corresponding to a depth of 9 cm, are updated. Although the brightness temperatures are only sensitive to soil moisture in up to 5 cm depth, DA 1 was defined as updating the top three layers, since a number of in situ measurements are taken from a depth of up to 8 cm. For these in situ sites, measurements are also available for deeper layers, and we thus define top-layer soil moisture as the soil moisture updated in DA 1. The upper six model layers, reaching 50 cm soil depth, are updated in the second experiment (DA 2). We refer to the lower three of these soil layers as the root zone. These two experiments enable us to examine to what extent CLM model physics alone are sufficient to update the root zone through the effects of the assimilation on the upper layers, as in comparison to directly applying the increments in this depth. For the experiments DA 1 and DA 2, soil texture perturbations were incrementally reduced with layer depth, minimising the impact of potentially large updates in deep layers. Since increments are computed in relative soil moisture, identical increments affect absolute soil water very differently, greatly exaggerating the assimilation impact for deeper layers. Perturbations for the two top layers remain unchanged, thus not decreasing the ensemble spread for the layers where SMOS is sensitive to soil moisture. The soil texture of the subsequent layers is perturbed by decreasing the perturbation factor by the inverse ratio between the respective layer thickness and the layer thickness of the two top layers (see ensemble generation under Sect. 3 and Table 1). Within a third experiment (DA 0), homogeneous soil texture perturbations are applied across all layers, highlighting the problem of large increments in lower layers. As will be shown, the larger ensemble spread in DA 0 actually further improves the correlation with in situ measurements but at the expense of introducing strong long-term effects. For all experiments, the brightness temperature forward simulations are computed by using the CLM output of all layers. The L-band simulations are thereby mostly affected by the output of up to 5 cm depth, which corresponds to the sensitivity of the SMOS sensor.

4.1 Correlation with in situ observations

For validation, hourly CLM soil moisture output is compared to in situ measurements obtained from the International Soil Moisture Network (ISMN) (Dorigo et al., 2011). OzNet in situ measurement probes are located within the Murrumbidgee catchment in south-east Australia, a limited spatial domain which does, however, cover a range of different land cover classes representative of Australia (Smith et al., 2012). The Murrumbidgee catchment was also chosen as a site for a SMOS validation campaign (Peischl et al., 2012). Measurements within the OzNet network are taken with time-domain reflectometry (TDR) probes at shallow levels (5 or 8 cm) and at deeper layers (30, 60 and 90 cm). The in situ measurements that are part of the CosmOz network are taken by using cosmic-ray neutron probes and are therefore representative of a larger horizontal footprint than the more traditional measurements. CosmOz measurement sites are located within the Murrumbidgee catchment as well as at selected locations close to the Australian coast. In addition to the original description of the measurement networks, Renzullo et al. (2014) and Holgate et al. (2016), for instance, offer an extensive overview of the CosmOz and OzNet measurement sites. Su et al. (2013) give more details on the Murrumbidgee catchment and the locations of the OzNet measurement sites. For all in situ measurements sites, the weighted average of the corresponding CLM soil moisture layers is taken, with the layer thickness being used as the respective weights. Figure 1 shows where the Murrumbidgee catchment is situated, as well as the land cover data used for the CLM simulations.

Taking into account only measurements with at least 1 year of data, not necessarily consecutive, correlations improve from 0.613 for the open-loop run to 0.640, 0.678 and 0.681 for DA 1, DA 2 and DA 0 for top-layer soil moisture (number of stations $n = 17$). Root-zone soil moisture improvements are smaller, with average correlation coefficients of 0.626, 0.644 and 0.648 for DA 1, DA 2 and DA 0 compared to 0.601
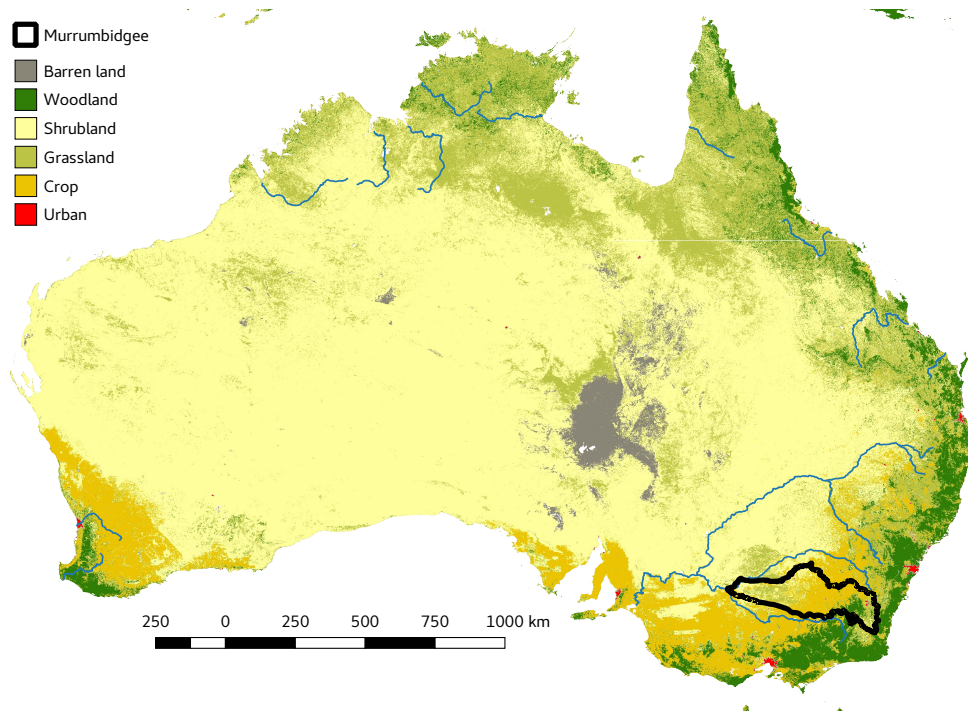
Figure 1. CLM plant functional types based on MODIS MCDQ12 land cover classification and ECOCLIMAP climate zones at 500 m resolution prior to the aggregation to 0.25°. Some classes are here aggregated for visualisation purposes, e.g. evergreen temperate and evergreen tropical forests are both shown as woodland. The boundary of the Murrumbidgee catchment, which is the site of the OzNet in situ measurements, is outlined.

for the open-loop run ($n = 31$). For the upper-level soil moisture, correlation improves for all in situ measurement stations, whereas for the root-zone soil moisture a single in situ station shows a deterioration of correlation for DA 1. In the case of DA 2 and DA 0, the correlation at two stations, albeit at different ones, slightly deteriorates. On average, upper moisture behaviour thus improves by additionally updating deeper layers, whereas deeper-layer soil moisture is slightly enhanced through only updating top-level soil moisture, with the assimilation effects only being applied through model physics. All in all, updating the top six CLM layers results in the largest improvements, even more so if the identical soil texture perturbations are applied to all soil layers within experiment DA 0, thereby increasing the assimilation impact through an increased ensemble spread and background error. The individual in situ measurements around the area of the Murrumbidgee catchment and the respective correlation changes for all three experiments towards the open-loop run are shown in Fig. 2. For top-layer soil moisture, the largest improvements are visible for the sites located in the centre of the catchment (Yanco site) with clear improvements for DA 2 and DA 0 when compared to DA 1. In the case of the root zone, multiple measurements at different depths were averaged using the measurement depth as weights and compared to the corresponding weighted average soil moisture simulations. Here, improvements are also highest for the Yanco site,

except for one measurement location showing a deterioration of correlation for DA 0. The area around the Yanco site is flat and semi-arid with mostly low vegetation and thus more ideal for L-band soil moisture sensitivity. The lesser improvements for the other in situ sites towards the east therefore could be explained by the more complex terrain, less homogeneous soil texture and higher vegetation influencing the L-band signal, as discussed by Su et al. (2013).

Figure 3 shows the Taylor diagrams for the in situ validation of experiment DA 2. As opposed to Fig. 2, all original measurements are included with no vertical aggregation performed. The Taylor diagrams reveal a slightly decreased normalised standard deviation when compared to the open-loop time series. In terms of standardised root mean square error (RMSE), it is less conclusive, with RMSE being slightly reduced for some stations and slightly increased for others, but never significantly. These findings correspond well to experiments DA 0 and DA 1 (not shown).

When comparing the average changes in correlation for the non-vertically aggregated sites, only for the 10 used CosmOz sites, correlation increases by about 0.016 for DA 1 (three sites with +0.03 and two very close to zero). For DA 2, the average correlation increases by 0.02 (two sites showing +0.05 and one −0.02) and for DA 3 by 0.13 (three sites slightly deteriorating). This highlights that the assimilation improvements are stronger for the OzNet sites located in the
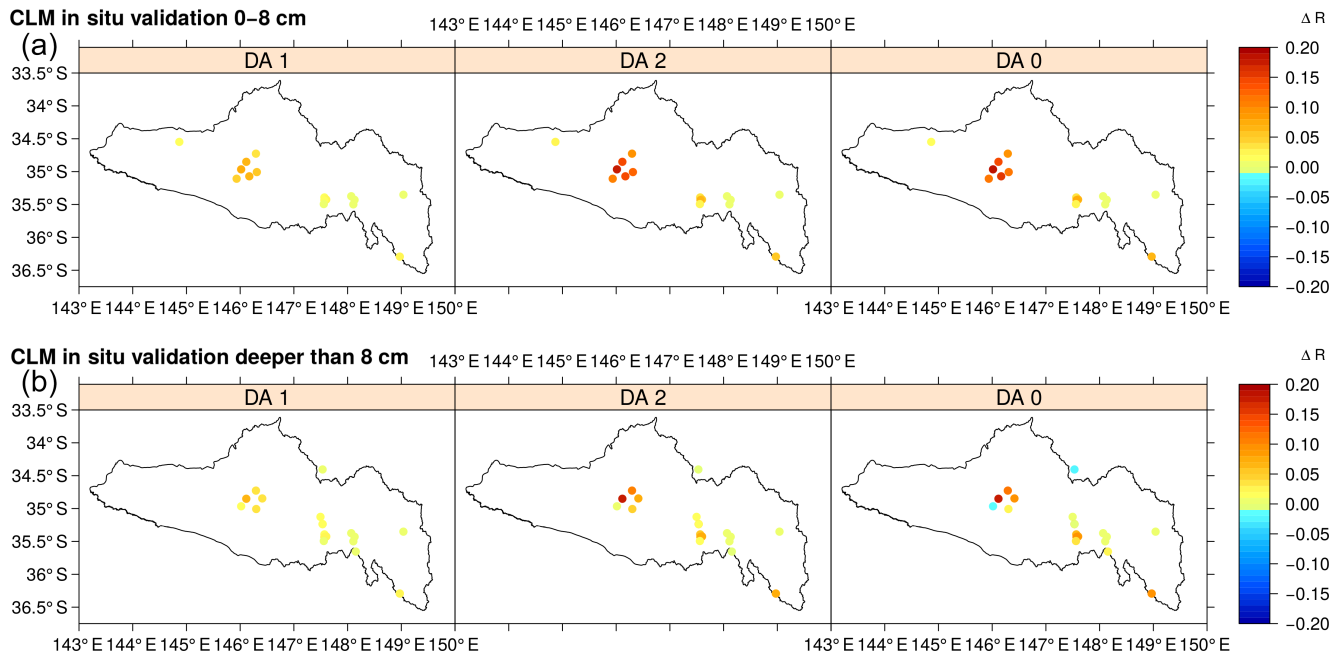
Figure 2. Change in correlation R for experiments DA 1, DA 2 and DA 0 both for top-layer soil moisture **(a)** as well as the root-zone soil moisture **(b)** within the Murrumbidgee catchment. In the case of multiple measurements at the same location, the weighted average of the measured as well as modelled soil moisture was computed in accordance with the corresponding CLM layer thickness.
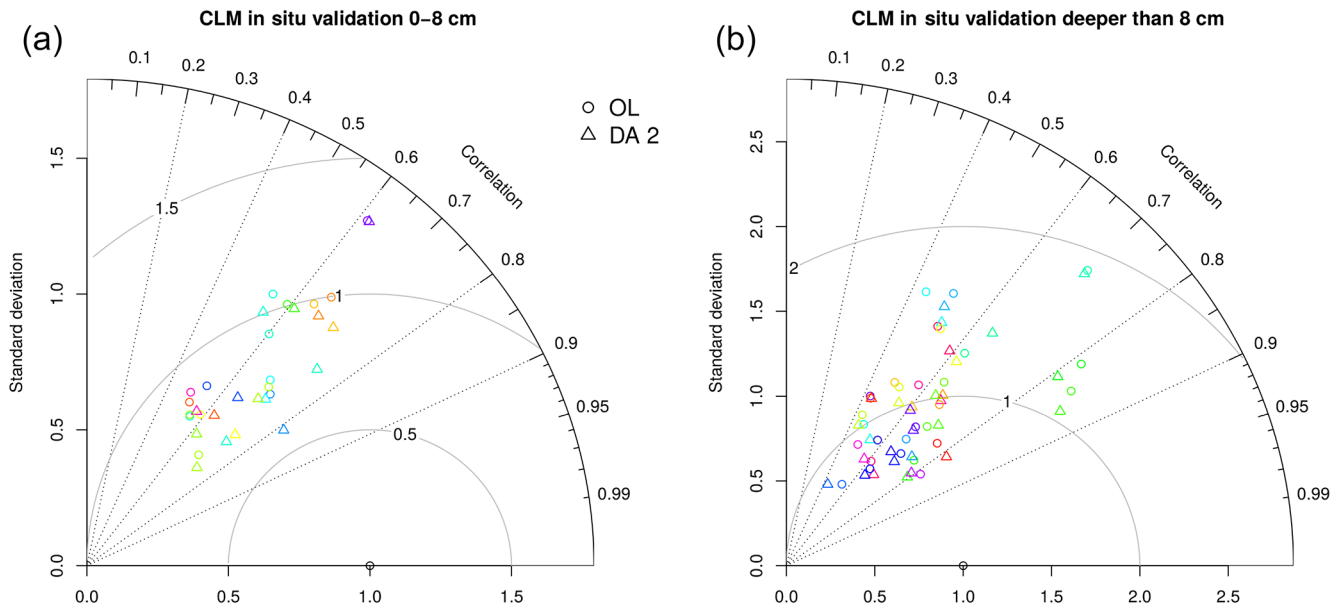

Figure 3. Taylor diagrams showing assimilation impact on top-layer soil moisture, defined as 8 cm soil depth **(a)**, and lower-level soil moisture **(b)** in terms of correlation coefficient R, standard deviation and normalised root mean square error (RMSE) for all in situ measurement sites. Measurements at multiple depths are not aggregated.

Murrumbidgee catchment. Partly, this might be attributed to the fact that the CosmOz measurements are valid for a variable soil depth, depending on the current soil moisture conditions. For validation, a single soil depth was used, which is reported to the ISMN network. Also, CosmOz sites are partly situated along the coast or close to water bodies and within areas of higher vegetation, making improvements through data assimilation more challenging, as reported by Renzullo et al. (2014). When only considering the CosmOz sites, correlation decreases for DA 3 with respect to DA 2, which con-

www.hydrol-earth-syst-sci.net/21/5929/2017/

Hydrol. Earth Syst. Sci., 21, 5929–5951, 2017

tradicts the findings when taking all measurements into account.

Altogether, the results demonstrate that the assimilation system is capable of improving soil moisture simulations across the Australian continent, both for top-layer soil moisture and the root zone. However, as with most assimilation studies, validation sites are sparse and do not cover the manifold combinations of soil texture, land cover, climate, etc., which might all have an impact on the assimilation performance. The representativeness error of the in situ measurement equally remains a problem, with the spatial support of the measurements (in the case of TDR probes, only point measurements) being smaller than the area covered by satellite. The assimilation system is therefore not designed to remove the relative bias between soil moisture simulations and observations, since the exact truth remains unknown, but to improve the temporal behaviour of the simulations, which has for the most part been achieved.

4.2 Soil moisture increments

Especially over long time periods, the mean increments in a bias-blind assimilation system can be expected to be very close to zero. Figures 4 and 6 show the mean soil moisture increments over the assimilation period (2010–2015) for the experiments DA 0, DA 1 and DA 2, separately for the ascending and descending orbits. Both the top three soil layers and the root-zone soil layers were averaged. Distinctive areas of mean positive increments for the ascending orbit are visible in the north, south-west and south-east of Australia, seemingly being linked to the occurrence of vegetation (see Fig. 1). The areas in the south-west and south-east as well as in the north correspond well to the subtropical, temperate and tropical climate zones, respectively, and are subject to higher precipitation than the dry areas inland, although other areas along the coast have similar precipitation. The areas in the south-west and south-east correspond to the wheat-growing areas of Australia. The misrepresentation of these areas through the CLM surface datasets, such as the use of climatological LAI instead of actual LAI, might well be the source of such patterns. Also, satellite-based estimates of leaf area index are not error-free, as has been shown specifically for the Murrumbidgee catchment (McColl et al., 2011). Irrigation, which is predominantly applied within the south-east, could be an additional source of error for limited areas, since the forward simulations will be based on seasonally too-low soil moisture, causing an incorrect estimation of the brightness temperature seasonality and the subsequent anomaly computation.

Nevertheless, for the ascending overpasses, the positive biases hardly exceed 0.5 % soil moisture and the remaining parts of Australia either show no mean increment bias or slightly negative values, both in areas covered by mostly sparse vegetation and the inner drylands. For the descending orbits, the patterns are still visible but are weaker, both for top

layers and the root zone. The exception is DA 0, where little differences between top-layer and deep-layer increments are noticeable. Interesting to highlight is the fact that top-layer deviations from zero are strongest for the assimilation experiment DA 1, compared to DA 0 and DA 2, which both update top-layer soil as well as root zone. The reason may be that updating deeper layers results in a more lasting effect, moving the model closer to subsequent observations and thus reducing subsequent increments.

Figures 5 and 7 show the increment standard deviations. Spatial patterns of the assimilation impact are very distinctive but do not necessarily correspond to the patterns seen in the mean increments, although they do partly match, as for the south-west and south-east. The relatively large increments in the western Wheatbelt and the Murray–Darling Basin, some areas close to the western coast of Queensland and the eastern coast of the Northern Territory show a standard deviation of 2.5 %. The areas in the north seem to be consistent with the occurrence of tussock grasses, as shown by the Australian National land cover map (http://www.ga.gov.au/scientific-topics/earth-obs/). Minimal or zero increments in all layers, especially along the eastern coast, are due to a lack of observations, as these were removed due to the active vegetation science flags or the fact that the high LAI values for high vegetation prescribed within the forward operator mask all signals from soil moisture.

Concerning the different assimilation experiments, top-layer increments are largest for DA 1, followed by DA 2 and then DA 0. Being the most dominant dynamic factor for the ensemble generation, precipitation leads to an immediate increase in ensemble spread and, as a consequence, to a larger background error for the very shallow soil layers, thereby also increasing the observation impact. This impact on the ensemble spread will, however, be dampened and temporally lagged for deeper layers. With increasing layer depth, the soil texture perturbations play a more important role in determining the background error. This is visible for DA 0, where homogeneous soil texture perturbations were applied across all layers, and increments in the root zone are not significantly smaller than for the layers above (Fig. 5). In contrast, root-zone increments applied within experiment DA 2 are far smaller than for the upper layers. For the validation with in situ measurements, we showed that these larger increments for DA 0 actually result in a slightly increased correlation over DA 1.

Concluding on the behaviour of increment bias and increment standard deviation, it seems that there is a relationship to root-zone updates. Both increment bias and increment standard deviation are largest for DA 1, where the root zone is not updated at all. The top-layer increment standard deviation decreases for DA 2, whilst also updating the root zone, with a slight decrease of the increment bias. Compared to DA 2, the increment standard deviation is larger in the root zone for DA 0 and the top-layer increment bias decreases substantially.
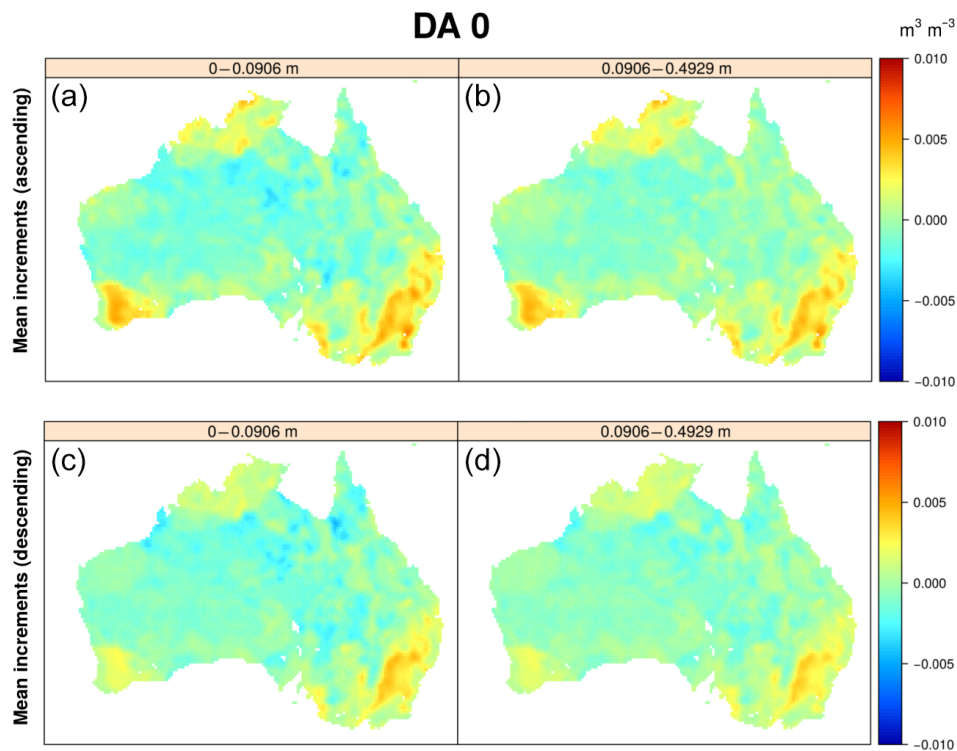
Figure 4. Mean of all increments for experiment DA 0 for top-layer soil moisture **(a, c)** and root-zone soil moisture **(b, d)** for ascending **(a, b)** and descending orbits **(c, d)**.
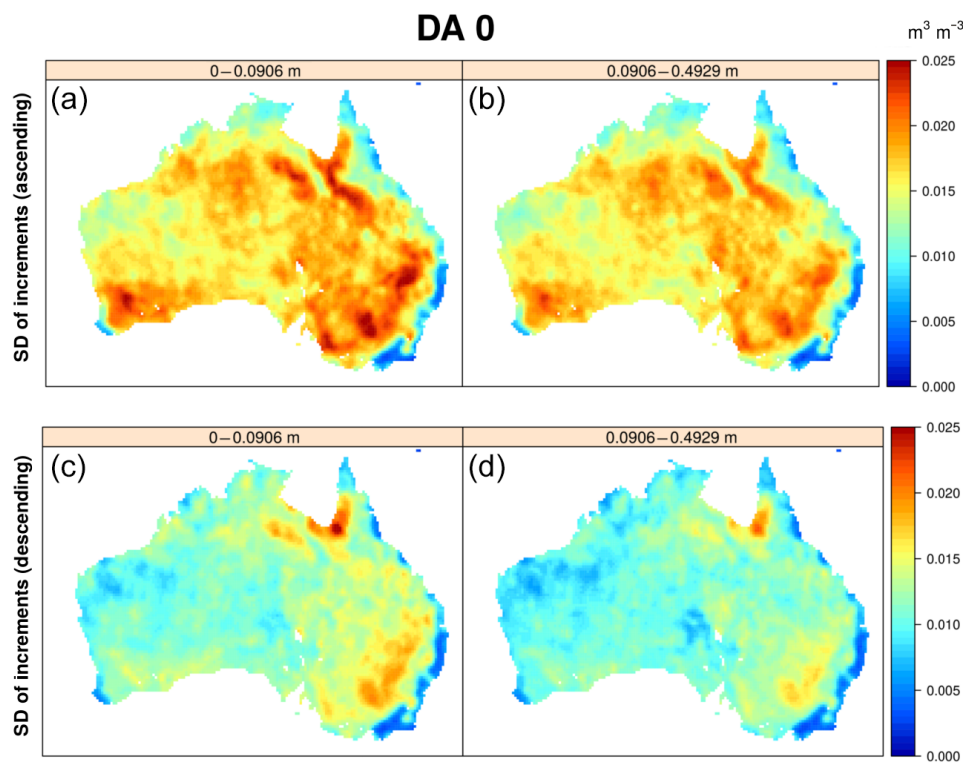

Figure 5. Increments standard deviation of all increments for experiment DA 0 for top-layer soil moisture **(a, c)** and root-zone soil moisture **(b, d)** for ascending **(a, b)** and descending orbits **(c, d)**. Increments for the root-zone soil moisture are fairly similar to the top soil layers, due to the homogeneous texture perturbations applied across all layers.
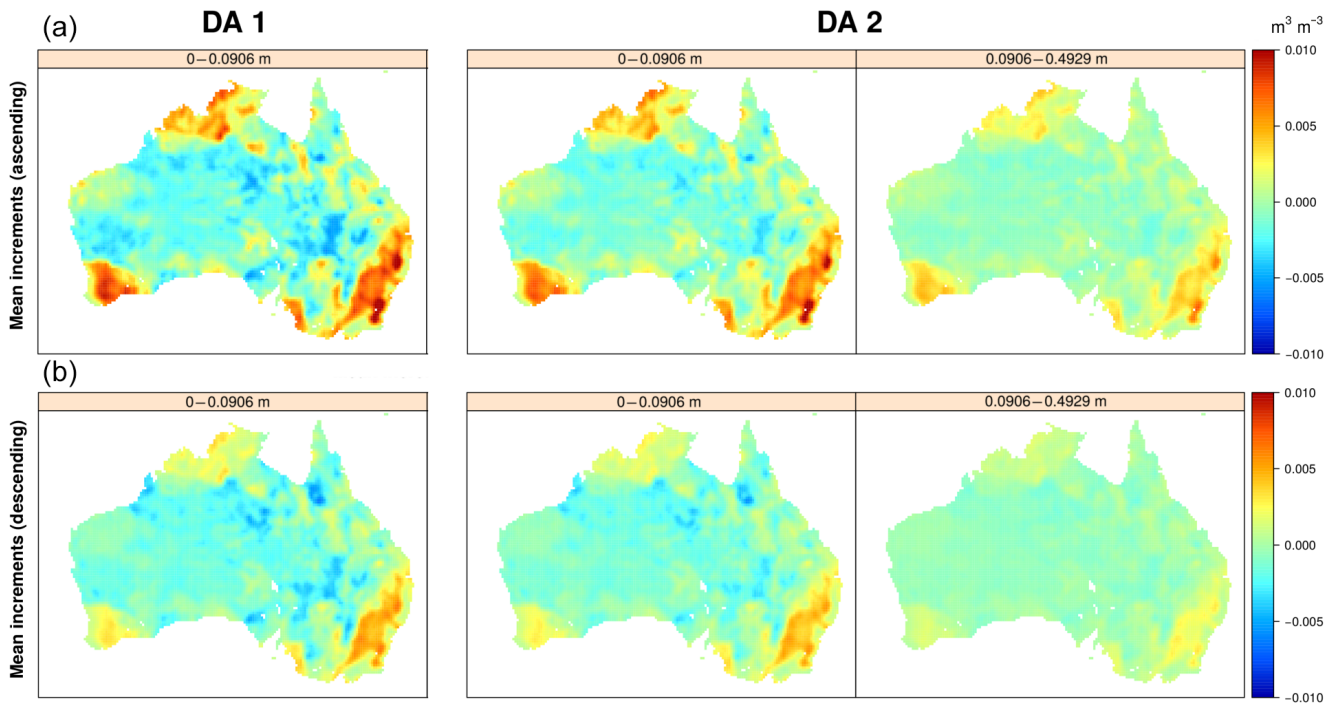
Figure 6. Mean of all increments for experiments DA 1 and DA 2 for top-layer soil moisture and root-zone soil moisture for ascending **(a)** and descending orbits **(b)**. Biases are strongest for the ascending orbit, and distinctive spatial patterns are visible. Biases are strongly reduced both for deeper layers and descending orbits.
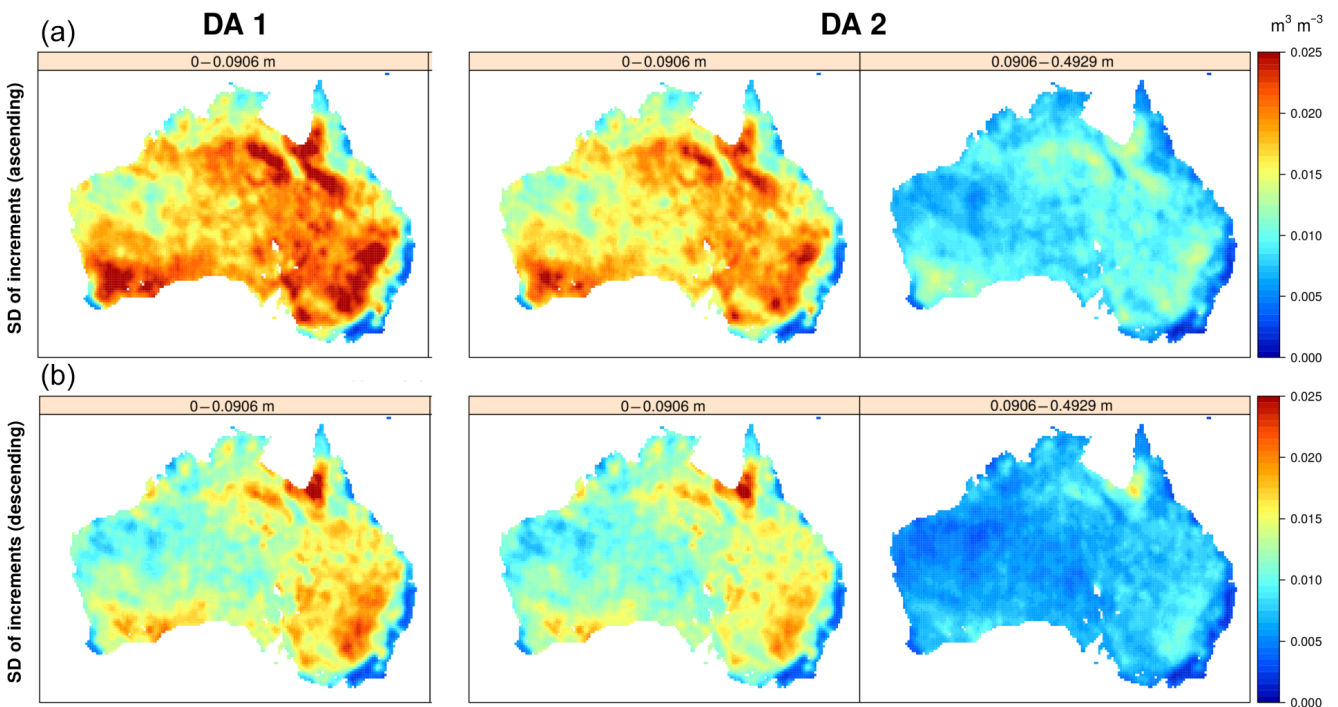
Figure 7. Standard deviation of increments for experiments DA 1 and DA 2 for top-layer soil moisture and root-zone soil moisture for ascending **(a)** and descending orbits **(b)**. Increments are strongest for the ascending orbit and for top-layer soil moisture and even stronger when restricting assimilation to these layers, as in DA 1. Increments are very low or zero for the forested areas along the coastline, either due to the absence of observations or the high LAI values masking any soil moisture signal within the forward operator.
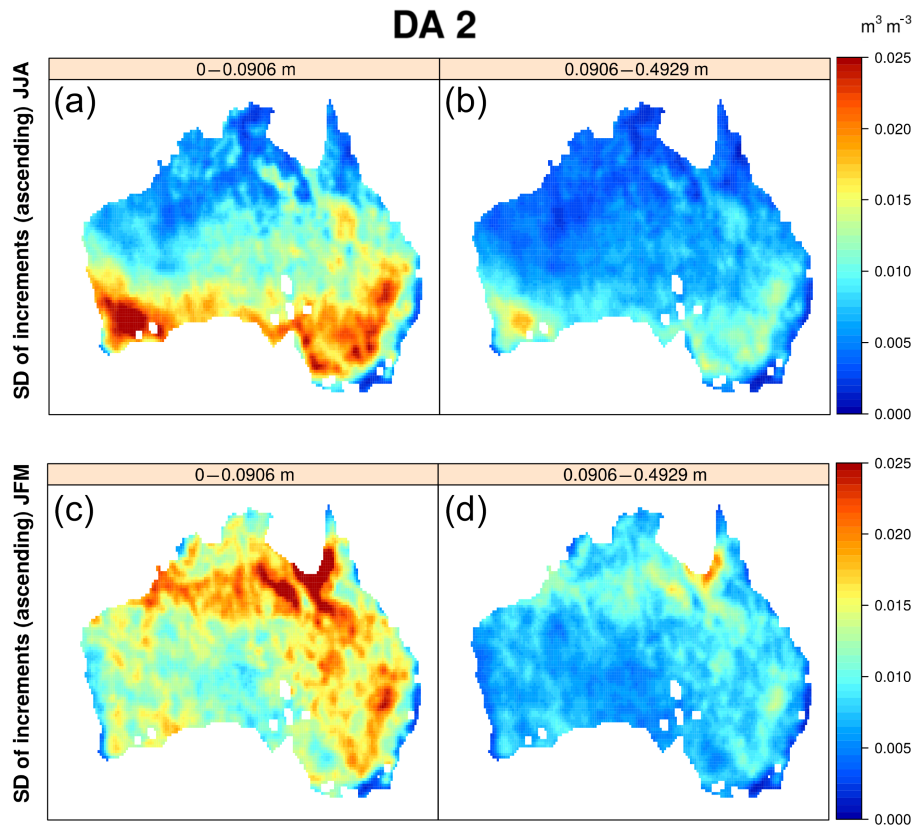
Figure 8. Increment standard deviation for ascending orbits for June–August **(a, b)** and January–March **(c, d)** for experiment DA 2. For the austral winter, increments are strongest for the south of Australia, especially the agricultural areas. During austral summer, the increments are strongest for the northern grasslands.

As for the differences between the ascending and descending orbits, we conclude that they can be partly explained by referring back to the fact that soil moisture retrievals are expected to be of a higher quality for the ascending orbits (Hornbuckle and England, 2005; Kerr et al., 2010).

To highlight some of the seasonal effects, Fig. 8 shows the increment standard deviation exemplary for DA 2 and the ascending orbit both for the months January to March and June to August. For the austral winter, increments are largest for the agricultural areas in the south-west and south-east, now in the growing season, and these seasonal effects clearly dominate the average of the increment standard deviation (compare to Fig. 7). Similarly, the patterns in the north, mostly linked to grassland, are visible in the yearly average and the shown months contribute the most to their existence. Differing seasonal effects of the assimilation impact were also observed by Martens et al. (2016) and Tian et al. (2017), although the observed patterns are distinctively different.

When comparing the winter patterns to areas where irrigation takes place, as shown by van Dijk et al. (2013), irrigated areas within the Murray–Darling Basin can be identified through an increased increment standard deviation. Here, the SMOS observations correct soil moisture dynam-

ics which are not explicitly modelled. Kumar et al. (2015), Escorihuela and Quintana-Seguí (2016) and De Lannoy and Reichle (2016b) have similarly reported on the potential of SMOS to observe irrigation.

4.3 Soil moisture quantiles

Apart from looking at the increments, we compute a set of quantiles at 1 % intervals for each CLM soil layer and each grid point, both for the assimilation experiments and the open-loop run. Although in principle the assimilation system should be designed bias-free with similar positive and negative increments, the previous section has revealed that small increment biases do exist, potentially causing long-term effects in the resulting analysis. Figures 9, 10 and 11 show the 10 % quantile changes, thus very dry conditions, in relation to the open-loop run for the top nine CLM layers. For experiment DA 1 (Fig. 10), assimilation has an impact on the topmost soil layer with the quantile increasing by a maximum of approximately 1 % for large areas and by up to 4 % for spatially very limited areas. Much smaller changes are visible for the second and third layers, with some areas also showing a negative impact by up to 2 %. CLM model physics result in changes being also visible within the root zone, CLM layers
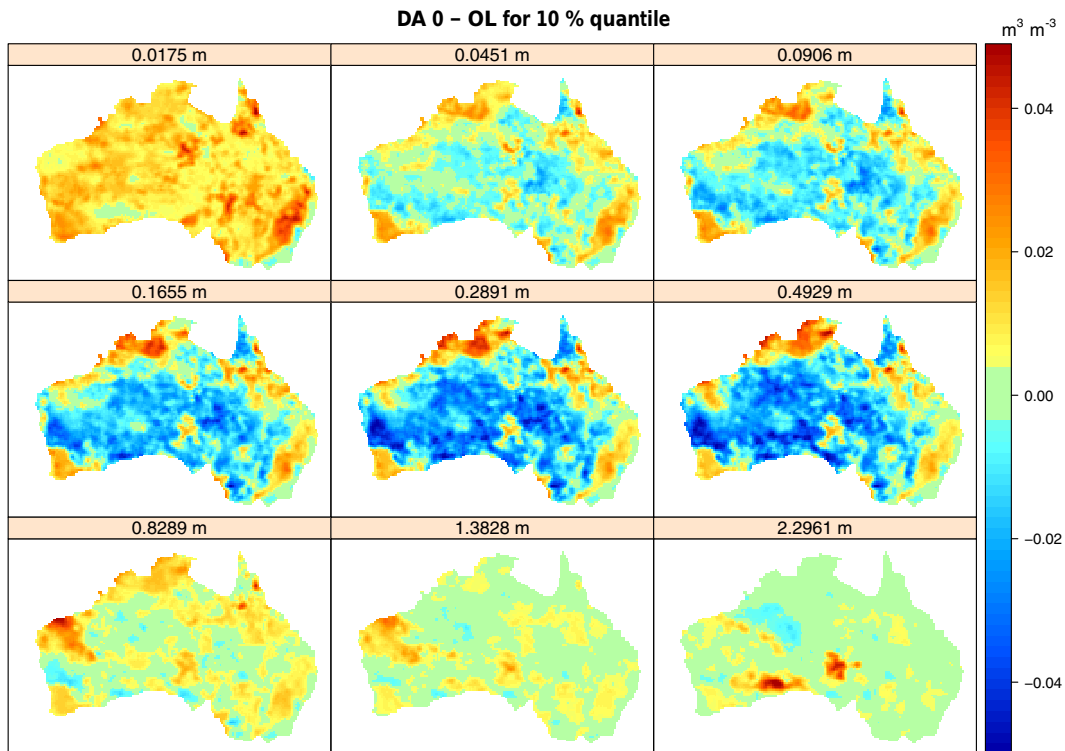
DA 0 − OL for 10 % quantile

Figure 9. Differences in relative soil moisture between open-loop and DA 0 experiments (DA 0 − OL) for the 10 % quantile. The individual panels correspond to the top nine CLM soil layers.
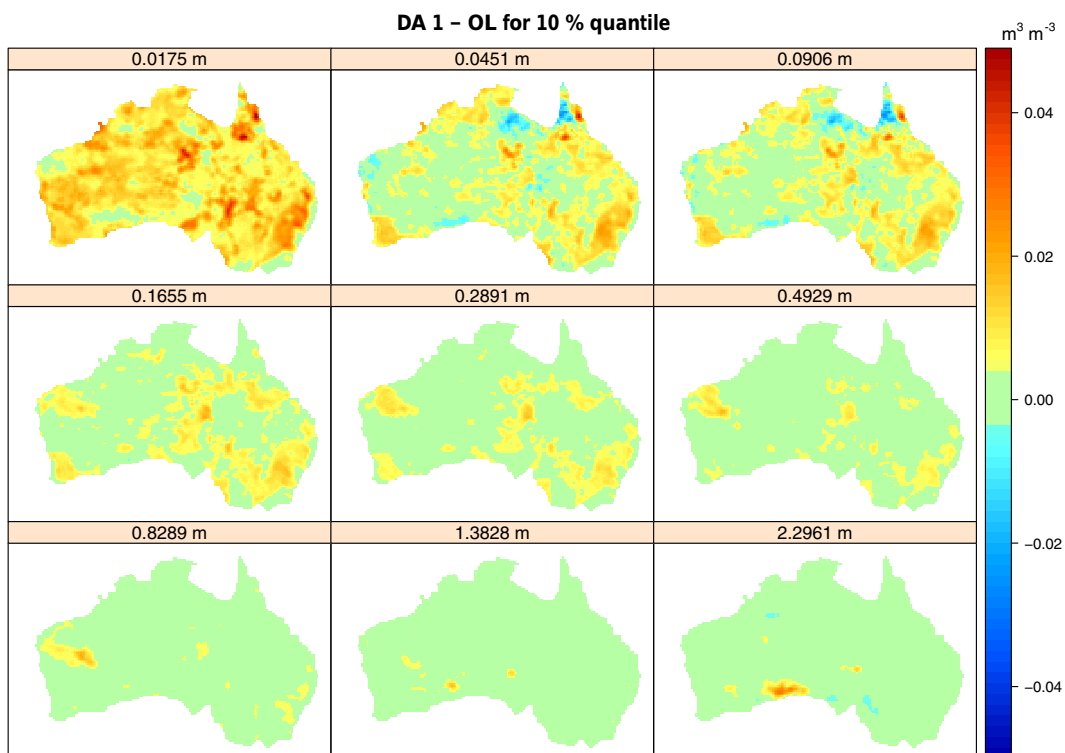
DA 1 − OL for 10 % quantile

Figure 10. Differences in relative soil moisture between open-loop and DA 1 experiments (DA 1 − OL) for 10 % quantile. The individual panels correspond to the top nine CLM soil layers titled with their depth.
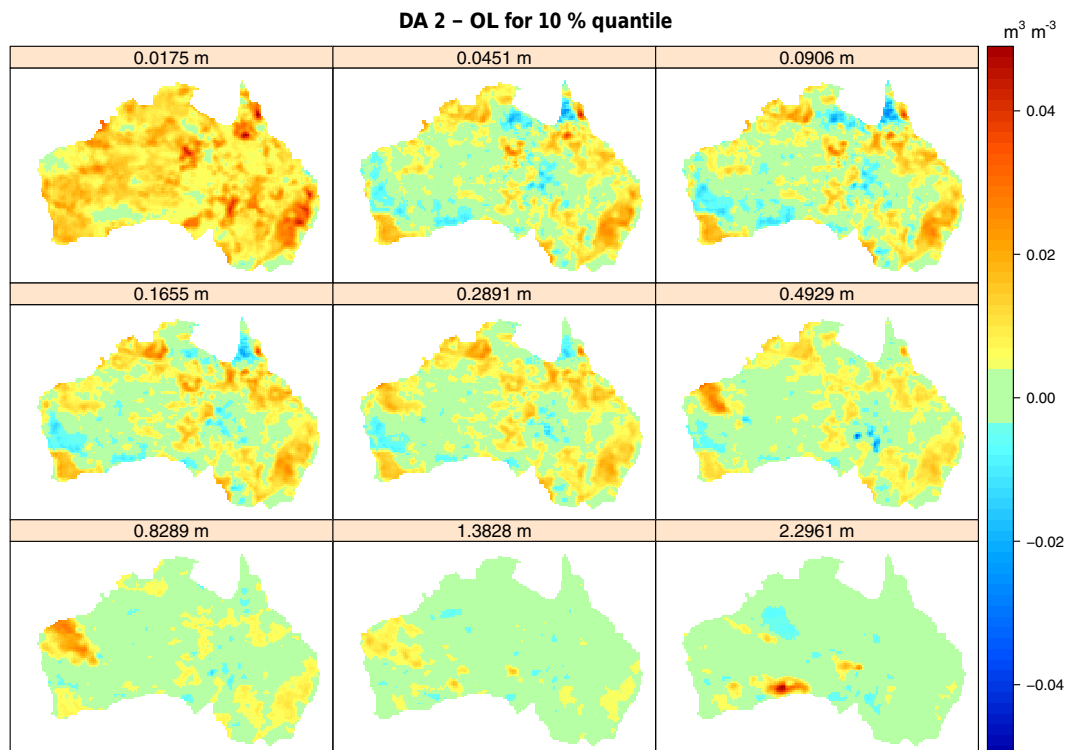
Figure 11. Differences in relative soil moisture between open-loop and DA 2 experiments (DA 2 − OL) for the 10 % quantile. The individual panels correspond to the top nine CLM soil layers.

3–6, and below. One of the visible patterns is again south-east Australia. For the very deep layers, some independent patterns emerge which are not visible for the above layers. Most notably in the Nullarbor Plain, on the south coast of Australia, where the 10 % quantile increases by up to 2 %. Such patterns are related to strong singular increments in very dry areas which accumulate in the deepest layers. Due to the low temporal dynamics in these lower layers, any added water will have a lasting effect especially on lower quantiles. For experiment DA 2 (Fig. 11), with the root zone also being updated, larger impacts on the quantiles in deeper layers can be observed. For the most part, the patterns reflect well the ones identified for DA 1. Figure 9 shows the impact on the 10 % quantile for experiment DA 0 with homogeneous soil texture perturbations being used. As expected, significant effects are visible especially within the root zone, with quantiles being decreased over wide areas of the Australian inland by up to 5 %. This is the result of the mean increments being slightly negative for inland Australia, which has a large effect when allowing large updates. Also, since absolute soil moisture increases with layer depth due to increasing layer thickness, removal of water in low layers increases drainage in the above layers, resulting in these drying out. This is especially visible for layers 2 and 3, where inland Australia to a far greater extent shows a lowered 10 % quantile in comparison to DA 1 and DA 2. For the lowest layer, a clear positive quantile shift

is visible in the area of Lake Eyre. The land cover map in Fig. 1 shows this as the only area that is classified as bare soil, although it is mostly a salt plain with water levels of the lake itself being strongly seasonal. A number of observations were therefore flagged, making the computation of a stable brightness temperature climatology challenging.

Figures 10–11 focus on the changes of the 10 % quantile. However, the spatial patterns identified do not necessarily reflect changes at other quantile levels. The complex nature of these shifts throughout the entire CDFs is shown in Fig. 12. The continental average empirical cumulative distribution functions are plotted for soil layers 1–6 for the open-loop run as well as for DA 1 and DA 2. Lower quantiles are increased on average through data assimilation, although at extremely low levels the behaviour tends to reverse again. Here, the quantiles decrease when compared to the open loop. For the upper quantiles, a small decrease can also be observed. The point where the decrease turns into an increase, with the assimilation having an on average neutral impact, is roughly the 50 % quantile for the top layer. For the subsequent layers, this point decreases towards the 40 % quantile. Although DA 0 resulted in the best correlation with the in situ measurements, it was disregarded at this point since the assimilation impact was too disruptive by strongly drying out the model across many layers.
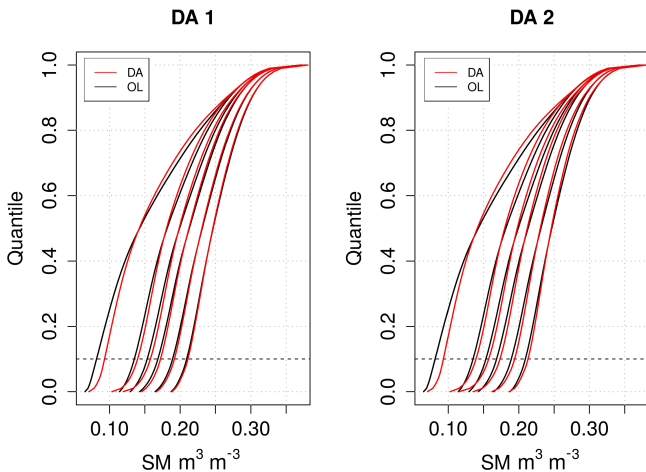
Figure 12. Cumulative distribution functions (CDFs) for the upper six CLM soil layers for experiments DA 1 and DA 2, based on quantiles computed for all data across the model domain. CDFs for open-loop simulations are shown in black and assimilation results in red. Both panels show changes in CDF behaviour for the layers being updated in the respective experiments, i.e. layers 1–3 for DA 1 and layers 1–6 for DA 2. Soil moisture increases systematically with soil depth allowing for the easy identification of the layers within the plot. The dashed vertical line marks the 10 % quantile, corresponding to Figs. 10 and 11.

For DA 1 and DA 2, the interplay of the quantile changes at the various levels results in an average decrease of the standard deviation of the soil moisture analysis, which to a certain extent could be attributed to the anomaly rescaling. First of all, due to the low sample count, any quantile-mapping procedure tends to be a challenge around the extremes of the distributions. Additionally, the exact observation error for each observation is unknown and although expected to be zero on average, with a small sample size, the observation error might on average deviate from zero, affecting the rescaling, since the true limits of the observation anomaly CDFs are unknown. However, many other reasons will equally play a role and, as it has been shown, increment bias and standard deviation are linked to certain geographic areas. The disentanglement of the desired systematic enhancements from erroneously introduced effects remains a challenge. There is still no perfect approach for rescaling the observations to match the model or for calibrating the forward observation model. Looking at long-term CDF changes induced by assimilation can be part of evaluating these different approaches with the final application of analysis data to be kept in mind.

Finally, as an example, we want to place quantile behaviour within the context of possible hydrological monitoring systems which directly make use of grid cell quantiles and empirical CDFs. The correction of short-term behaviour alone, i.e. hourly or daily, has a minor effect when analysing phenomena that spread across larger spatial scales

and time intervals, although large increments, e.g. due to corrected precipitation during a storm, can have an effect on the start and end points of an observed phenomenon, such as a drought. When classifying such an event defined at a specific quantile level, there will be a 2-fold impact from the assimilation: the changes in the quantile of interest as well as the change in soil moisture itself. Here, we highlight a sample dry event on the east coast to show to what extent its classification changes through the assimilation impact. Figure 13 shows root-zone soil moisture at or below the 10 % quantile level for the open-loop run as well as the data assimilation experiment DA 2 for soil moisture conditions in early 2010, thus at the beginning of the assimilation period. Due to the higher 10 % quantile for DA 2, as seen in Figs. 11 and 12, the spatial extent of the cluster for DA 2 is reduced, but the spatial patterns of soil moisture remain largely the same. At some time periods, not shown here, a higher degree of noise is noticeable within the assimilation dataset. This is likely due to the fact that non-spatially correlated noise was applied to the meteorological forcings, resulting in a heterogeneous background error field for grid points. We thus conclude that despite having carried out the assimilation in 1-D, spatially correlated noise is recommended for such applications. An alternative would be to further increase the ensemble size but at the expense of higher computational resources. Additionally, when trying to extract meaningful statistics on the occurrence of events, such as droughts, it might be particularly important to clean up the dataset in the case of data assimilation using simple filter algorithms, such as those applied by Herrera-Estrada et al. (2017). We want to highlight the fact that the shown event is for demonstration purposes and not linked to any major drought event, which would require a more in-depth analysis and references to independent data sources.

5 Discussion and conclusion

The Community Land Model was set up for the Australian continent and coupled to the Community Microwave Emission Model. We have substituted the surface datasets with higher resolution and more recent data. Additionally, we have replaced the offline forcings with the ERA-Interim reanalysis. The assimilation over 6 full years, from 2010 to 2015, of SMOS brightness temperature anomalies with the LETKF improved soil moisture simulations when compared to in situ measurements on the order of up to 11 % for top soil moisture. Both the CLM model and the forward observation model were not calibrated, therefore implying that the assimilation system could be applied to other areas.

In detail, three data assimilation experiments were carried out. Within the first experiment, the top three layers were updated, which mostly correspond to the depth where SMOS is sensitive to changes in soil moisture and top-layer in situ measurements are available. The correlation with top-layer
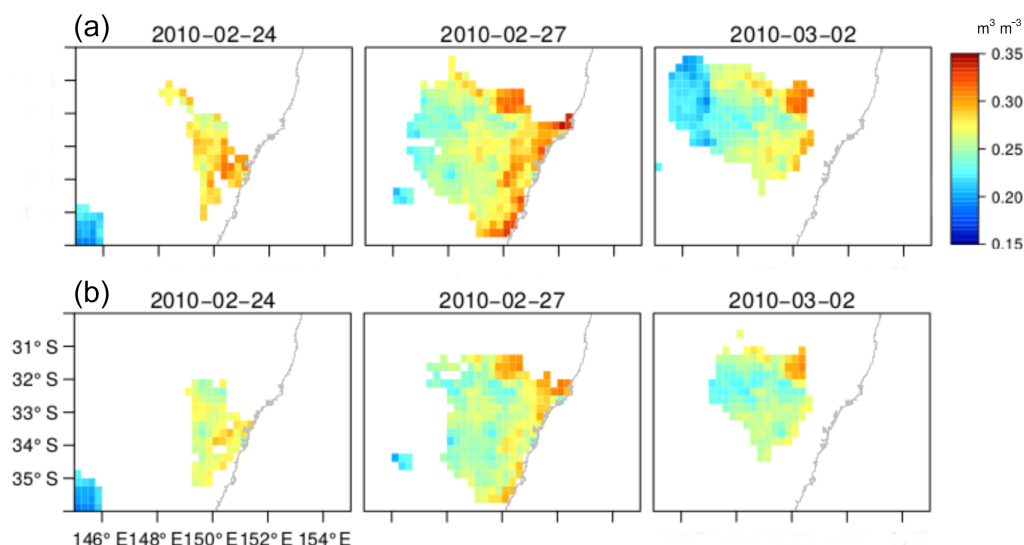
Figure 13. Dry period in February 2010, showing only the root-zone soil moisture below the 10 % quantile level for the open loop **(a)** and experiment DA 2 **(b)**. The different spatial extent and differences in soil moisture itself, depending on the dataset used, on three different days are clearly visible. The figure is centred around the central coast of New South Wales.

soil moisture measurements increased by 5 %; root-zone soil moisture increased by 4 %. Within the second experiment, both top soil moisture and the root zone were updated, resulting in correlation improvements of 11 and 7 %, respectively. The CLM is therefore able to translate top-layer updates into deeper-layer soils. Greater improvements can be achieved by additionally updating the root zone directly. For these two experiments, soil texture perturbations were reduced with increasing layer depth. With CLM layer thickness vastly increasing with depth, homogeneous soil perturbations across all layers result in large deep-layer updates in terms of absolute soil moisture. This was demonstrated in a third experiment, where correlation with in situ measurements was highest compared to the first two experiments, namely 11 and 8 % for top and root-zone soil moisture, respectively. This coincides with the findings by Kumar et al. (2009), who report that soil moisture simulations profit more from assimilation with an exaggerated coupling between top-layer and root-zone soil moisture than vice versa. Within this context, we interpret the overly large root-zone updates for the third experiment, based on the large covariance between root-zone and top-layer ensembles, as an artificially exaggerated coupling. Kumar et al. (2009) also state that when compared to other land surface models, CLM actually shows an overall lower coupling strength. Larger improvements in root-zone soil moisture simulations therefore might be possible when using the identical assimilation setup with a different land surface model.

Mean increments showed distinctive patterns with slight positive biases up to 1 % soil moisture in areas covered by denser vegetation and neutral to slightly negative impact for areas mostly covered by sparse vegetation. A possible cause

could be the use of climatological LAI data, which is common practice within current land data assimilation systems. Due to the abundance of operational available vegetation data, we would like to encourage future studies to look into possible improvements by using non-climatological LAI, where cloud cover permits. Climatological LAI might especially pose a problem for the monitoring of extreme events, such as droughts, since these tend to result in lower LAI values again influencing the forward simulations. Further, it might also be useful to add perturbations to the LAI data in order to better simulate the uncertainties of the forward simulations. It is known that remotely sensed LAI data are erroneous. McColl et al. (2011), for example, show for the Murrumbidgee Basin that MODIS MOD15A2 estimates are too high for lightly vegetated areas and too low for densely vegetated areas. This coincides well with the contradicting patterns of this study with the increments showing a slight positive bias for densely vegetated areas and vice versa. McColl et al. (2011) further describe the quasi-Gaussian distribution of the LAI uncertainties with a bias of −0.35 and a standard deviation of 0.82.

The dependence of SMOS soil moisture retrievals on landscape features, such as vegetation, within the Murrumbidgee Basin was also shown by Su et al. (2013). The ability of explicitly accounting for these effects within the forward simulations and to avoid cross correlations with ancillary data used within the soil moisture retrieval is one of the advantages of the brightness temperature assimilation when compared to the assimilation of retrievals. Draper et al. (2009b) have evaluated AMSR-E soil moisture over Australia which correlates well with in situ measurements. The seasonal soil moisture patterns also well reflect the ones observed within

this study, most notably in the northern tropical regions as well as in the east. Draper et al. (2009b) argue that although vegetated areas mostly correspond to higher soil moisture, the retrievals might also be contaminated by the vegetation signal.

Areas of denser vegetation are also mostly linked to higher precipitation. Seasonal variations for the increments are clearly linked to the seasonal precipitation patterns and thus to the vegetation growing season. The multiplicative precipitation perturbations applied here have a large impact on the total simulated model uncertainty and the impact of the observations across different geographic areas. Renzullo et al. (2014) applied an average multiplicative error of 60 % over Australia, closely matching the 50 % applied in this study, for the BAWAP rain-gauge-based precipitation data based on the analysis by Jones et al. (2009). Strong spatial variations exist for these precipitation estimates, largely influenced by the amount of locally available gauge stations. The ERA-Interim analysis data used within this study, produced by assimilating a multitude of both satellite as well as in situ data, equally have errors linked to geographic areas. Estimates on these are, however, not provided with the product. The updated ERA-5 reanalysis data, which will include information on ensemble mean and spread, could therefore be a significant step forward in characterising the background error induced by the meteorological forcings when using global, non-locally optimised data. López López et al. (2016) have assimilated AMSR-E soil moisture data across the Murrumbidgee Basin using global precipitation data as well as locally optimised forcing data. Since the latter increases the open-loop accuracy, the positive assimilation effect here is actually reduced.

Within the broader context of Australian soil moisture analysis, a study comparing soil moisture output from several models as well as satellite products to in situ data carried out by Holgate et al. (2016) showed that SMOS retrievals are favourable across measurement sites, except for one with dense tropical vegetation. This is attributed to the likely advantage of their deeper penetration capability. The same study also shows that retrievals from ascending orbits perform best. These findings relate to our study, where soil moisture simulations were mostly improved across all measurement sites and the unscaled ascending brightness temperature acquisitions both showed a smaller bias towards the forward simulations and had a larger effect on the soil moisture analysis. However, Su et al. (2013) report that their comparison of SMOS retrievals to in situ measurements within the Murrumbidgee Basin showed that the descending orbits performed better. Holgate et al. (2016) also show that similarities are largest within the group of the satellite retrievals and within the group of the different model outputs, on average, with larger differences between models and retrievals. This further motivates us to combine observations and model output in an optimal way through data assimilation specifically for Australia, as performed here. Kumar et al. (2017)

compare soil moisture from simple model output to complex land surface simulations, arguing that the latter perform better within Australia. The CLM land surface model is a fully physical-based land surface scheme solving the energy and mass balance and provides all data required for the forward simulation of the brightness temperatures, allowing the full use of L-band brightness temperature observations.

This enables the correct temporal alignment between observations and forward simulations is especially important to achieve a high accuracy for the forward simulations. Here, we slightly simplified the forward simulations by merging all ascending or descending daily overpasses and computing the forward simulations at one common time. The temporal offset is thereby maximum 3 h, which is within the temporal resolution of the forcing data. No artefacts were identified, although we encourage a more precise temporal alignment within operational systems, as is mostly done. Soil moisture retrievals are only valid for one specific time instance and inter-daily variation of soil moisture can be considerable due to precipitation events. The assimilation of observations into a model thus is advantageous, since it allows for more correct daily estimates by averaging the model time steps, which is 30 min here.

CLM uses fixed soil layer depths, which is likely the most beneficial model structure for comparing the assimilation effects spatially, since the covariance between observations and state variables does not vary depending on spatially non-uniform layer depths, as is the case with some other land surface models. The validity of updating very deep layers with information derived from surface observations is, however, questionable. Therefore, joint assimilation schemes also assimilating data from satellites such as GRACE are preferable, as was done by Tian et al. (2017).

Long-term assimilation effects were analysed by estimating the cumulative distribution functions for each grid cell prior to and after assimilation. On average, lower quantiles are shifted towards wetter conditions and higher quantiles are slightly shifted to drier conditions, resulting in reduced analysis variability. Spatial patterns in the quantiles do, however, change significantly at different quantile levels. We have shown these exemplary for the 10 % quantile. Here, for the experiment using homogeneous soil texture perturbations across all layers, the root-zone soil moisture showed a strong reduction compared to the open-loop run at the 10 % quantile level. Patterns visible in the increment bias were strongly exaggerated, highlighting the problem of too-large updates within the root zone and the general sensitivity towards model perturbations.

The uncalibrated forward operator and the therefore necessary rescaling of the observations might be one possible cause for the reduced analysis variability as well as the spatial patterns. The observation rescaling is especially a challenge around the very low or very high values. The number of samples within these regions of the CDFs is small and the observations are contaminated with errors, which might not be

zero on average. In this case, the tails of the observation distribution will not represent the true maximum and minimum values. Furthermore, the standard parameterisation of the forward operator is certainly not perfect. The spatiotemporal patterns linked to geographic areas or specific land cover classes, such as areas of higher vegetation, could likely be reduced by calibrating the forward observation model towards the observations at a grid cell level. This, however, comes with its own problems, for instance, a possible reduction of the sensitivity towards soil moisture or several parameter sets achieving the same result (i.e. equifinality).

Hydrological monitoring systems, where it is important to identify the relative occurrence of certain soil moisture levels and to monitor patterns both over space and time, are more and more likely to incorporate the assimilation of brightness temperatures with sufficiently long data records becoming available. Draper and Reichle (2015) have shown that data assimilation is able to correct modelled soil moisture also at longer time intervals from subseasonal to seasonal scale and seasonal differences in the assimilation effect are reported across many studies, as also shown here. Existing hydrological monitoring systems, such as the US Drought Monitor (Svoboda et al., 2002), the African Flood and Drought Monitor (Sheffield et al., 2014), the German Drought Monitor (Samaniego et al., 2013) or the Australian Water Resource Assessment (Van Dijk et al., 2011; Vaze et al., 2013) all use soil moisture quantiles at grid cell level to characterise different levels of severity and facilitate the comparison of soil moisture levels between grid cells. We have shown that the assimilation-induced quantile changes will have an effect on the spatiotemporal classification of areas above or below a certain quantile level, although the characteristics of these changes will be highly dependent on the model and data assimilation system. Hopefully, assimilation will benefit the monitoring and analysis of future severe events, such as the Millennium Drought in Australia (van Dijk et al., 2013). Modelling the complex feedback processes between soil moisture and vegetation is likely best performed using raw brightness temperatures and therefore the use of consistent data between the land surface model and the forward simulations.

In this paper, a relatively long time series of SMOS brightness temperatures has been assimilated into the Community Land Model across the Australian continent and soil moisture simulations are improved for the very largest part of in situ measurements, both for top-layer and root-zone soil moisture. Finally, the Community Land Model is part of the Community Earth System Model and the here-presented data assimilation system will in future also enable the analysis of the long-term impact of L-band brightness temperature assimilation within coupled land–atmosphere experiments.

References

Balsamo, G., Beljaars, A., Scipal, K., Viterbo, P., van den Hurk, B., Hirschi, M., and Betts, A. K.: A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System, J. Hydrometeorol., 10, 623–643, 2009.

Barella-Ortiz, A., Polcher, J., de Rosnay, P., Piles, M., and Gelati, E.: Comparison of measured brightness temperatures from SMOS with modelled ones from ORCHIDEE and H-TESSEL over the Iberian Peninsula, Hydrol. Earth Syst. Sci., 21, 357–375, https://doi.org/10.5194/hess-21-357-2017, 2017.

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, 5-1–5-23, 2002.

Brocca, L., Melone, F., Moramarco, T., Wagner, W., Naeimi, V., Bartalis, Z., and Hasenauer, S.: Improving runoff pre-

diction through the assimilation of the ASCAT soil moisture product, Hydrol. Earth Syst. Sci., 14, 1881–1893, https://doi.org/10.5194/hess-14-1881-2010, 2010.

Brocca, L., Moramarco, T., Melone, F., Wagner, W., Hasenauer, S., and Hahn, S.: Assimilation of surface-and root-zone ASCAT soil moisture products into rainfall–runoff modeling, IEEE T. Geosci. Remote, 50, 2542–2555, 2012.

Champeaux, J. L., Masson, V., and Chauvin, F.: ECOCLIMAP: a global database of land surface parameters at 1 km resolution, Meteorol. Appl., 12, 29–32, 2005.

Chen, F., Crow, W. T., and Ryu, D.: Dual forcing and state correction via soil moisture assimilation for improved rainfall–runoff modeling, J. Hydrometeorol., 15, 1832–1848, 2014.

Choudhury, B., Schmugge, T. J., Chang, A., and Newton, R.: Effect of surface roughness on the microwave emission from soils, J. Geophys. Res.-Oceans, 84, 5699–5706, 1979.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553–597, 2011.

De Lannoy, G. J. M. and Reichle, R. H.: Global assimilation of multiangle and multipolarization SMOS brightness temperature observations into the GEOS-5 Catchment Land Surface Model for soil moisture estimation, J. Hydrometeorol., 17, 669–691, 2016a.

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, https://doi.org/10.5194/hess-20-4895-2016, 2016b.

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

De Lannoy, G. J. M., Reichle, R. H., and Pauwels, V. R. N.: Global calibration of the GEOS-5 L-band microwave radiative transfer model over non frozen land using SMOS observations, J. Hydrometeorol., 14, 765–785, 2013.

de Rosnay, P., Drusch, M., Boone, A., Balsamo, G., Decharme, B., Harris, P., Kerr, Y., Pellarin, T., Polcher, J., and Wigneron, J.-P.: AMMA Land Surface Model Intercomparison Experiment coupled to the Community Microwave Emission Model: ALMIP-MEM, J. Geophys. Res.-Atmos., 114, D05108, https://doi.org/10.1029/2008JD010724, 2009.

Dharssi, I., Bovis, K. J., Macpherson, B., and Jones, C. P.: Operational assimilation of ASCAT surface soil wetness at the Met Office, Hydrol. Earth Syst. Sci., 15, 2729–2746, https://doi.org/10.5194/hess-15-2729-2011, 2011.

Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., van Oevelen, P., Robock, A., and Jackson, T.: The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements, Hydrol. Earth Syst. Sci., 15, 1675–1698, https://doi.org/10.5194/hess-15-1675-2011, 2011.

Draper, C. and Reichle, R.: The impact of near-surface soil moisture assimilation at subseasonal, seasonal, and interannual timescales, Hydrol. Earth Syst. Sci., 19, 4831–4844, https://doi.org/10.5194/hess-19-4831-2015, 2015.

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

Draper, C. S., Walker, J. P., Steinle, P. J., De Jeu, R. A., and Holmes, T. R.: An evaluation of AMSR-E derived soil moisture over Australia, Remote Sens. Environ., 113, 703–710, 2009b.

Draper, C., Mahfouf, J.-F., Calvet, J.-C., Martin, E., and Wagner, W.: Assimilation of ASCAT near-surface soil moisture into the SIM hydrological model over France, Hydrol. Earth Syst. Sci., 15, 3829–3841, https://doi.org/10.5194/hess-15-3829-2011, 2011.

Draper, C. S., Reichle, R. H., De Lannoy, G. J. M., and Liu, Q.: Assimilation of passive and active microwave soil moisture retrievals, Geophys. Res. Lett., 39, L04401, https://doi.org/10.1029/2011GL050655, 2012.

Drusch, M., Holmes, T., de Rosnay, P., and Balsamo, G.: Comparing ERA-40-based L-band brightness temperatures with Skylab observations: A calibration/validation study using the Community Microwave Emission Model, J. Hydrometeorol., 10, 213–226, https://doi.org/10.1175/2008JHM964.1, 2009.

Dumedah, G., Walker, J. P., and Merlin, O.: Root-zone soil moisture estimation from assimilation of downscaled Soil Moisture and Ocean Salinity data, Adv. Water Resour., 84, 14–22, 2015.

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., Entin, J. K., Goodman, S. D., Jackson, T. J., Johnson, J., Kimball, J., Piepmeier, J. R., Koster, R. D., Martin, N., McDonald, K. C., Moghaddam, M., Moran, S., Reichle, R., Shi, J. C., Spencer, M. W., Thurman, S. W., Tsang, L., and Van Zyl, J.: The soil moisture active passive (SMAP) mission, P. IEEE, 98, 704–716, 2010.

Escorihuela, M. J. and Quintana-Seguí, P.: Comparison of remote sensing and simulated soil moisture datasets in Mediterranean landscapes, Remote Sens. Environ., 180, 99–114, 2016.

Han, X., Li, X., Hendricks Franssen, H. J., Vereecken, H., and Montzka, C.: Spatial horizontal correlation characteristics in the land data assimilation of soil moisture, Hydrol. Earth Syst. Sci., 16, 1349–1363, https://doi.org/10.5194/hess-16-1349-2012, 2012.

Han, X., 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, 2014.

Han, X., Li, X., He, G., Kumbhar, P., Montzka, C., Kollet, S., Miyoshi, T., Rosolem, R., Zhang, Y., Vereecken, H., and Franssen, H.-J. H.: DasPy 1.0 – the Open Source Multivariate Land Data Assimilation Framework in combination with the Community Land Model 4.5, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmdd-8-7395-2015, 2015a.

Han, X., Li, X., Rigon, R., Jin, R., and Endrizzi, S.: Soil moisture estimation by assimilating L-band Microwave brightness temperature with geostatistics and observation localization, PloS one, 10, e0116435, https://doi.org/10.1371/journal.pone.0116435, 2015b.

Han, X., Li, X., He, G., Kumbhar, P., Montzka, C., Kollet, S., Miyoshi, T., Rosolem, R., Zhang, Y., Vereecken, H., and

Franssen, H.-J. H.: Python Multivariate Land Data Assimilation System with High Performance Computing for CLM 4.5, available at: https://github.com/daspy/daspy (last access: 25 September 2017), 2015c.

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, 2014.

Hengl, T., de Jesus, J. M., MacMillan, R., Batjes, N. H., Heuvelink, G. B., Ribeiro, E., Samuel-Rosa, A., Kempen, B., Leenaars, J. G., Walsh, M. G., and Gonzalez, M. R.: SoilGrids1km – Global Soil Information Based on Automated Mapping, PLoS ONE, 9, e105992, https://doi.org/10.1371/journal.pone.0105992, 2014.

Herrera-Estrada, J. E., Satoh, Y., and Sheffield, J.: Spatiotemporal dynamics of global drought, Geophys. Res. Lett., 44, 2254–2263, https://doi.org/10.1002/2016GL071768, 2017.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A.: Very high resolution interpolated climate surfaces for global land areas, Int. J. Climatol., 25, 1965–1978, 2005.

Holgate, C., Jeu, R. D., van Dijk, A., Liu, Y., Renzullo, L., Vinodkumar, Dharssi, I., Parinussa, R., Schalie, R. V. D., Gevaert, A., Walker, J., McJannet, D., Cleverly, J., Haverd, V., Trudinger, C., and Briggs, P.: Comparison of remotely sensed and modelled soil moisture data sets across Australia, Remote Sens. Environ., 186, 479–500, https://doi.org/10.1016/j.rse.2016.09.015, 2016.

Hornbuckle, B. K. and England, A. W.: Diurnal variation of vertical temperature gradients within a field of maize: Implications for satellite microwave radiometry, IEEE Geosci. Remote S., 2, 74–77, 2005.

Hunt, B. R., Kostelich, E. J., and Szunyogh, I.: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter, data Assimilation, Physica D, 230, 112–126, 2007.

Jia, B., Xie, Z., Tian, X., and Shi, C.: A soil moisture assimilation scheme based on the ensemble Kalman filter using microwave brightness temperature, Sci. China Ser. D, 52, 1835–1848, 2009.

Johnson, F., White, C. J., van Dijk, A., Ekstrom, M., Evans, J. P., Jakob, D., Kiem, A. S., Leonard, M., Rouillard, A., and Westra, S.: Natural hazards in Australia: floods, Climatic Change, 139, 21–35, 2016.

Jones, D. A., Wang, W., and Fawcett, R.: High-quality spatial climate data-sets for Australia, Aust. Meteorol. Ocean. J., 58, 233–248, 2009.

Kalman, R. E.: A new approach to linear filtering and prediction problems, J. Basic Eng., 82, 35–45, 1960.

Ke, Y., Leung, L. R., Huang, M., Coleman, A. M., Li, H., and Wigmosta, M. S.: Development of high resolution land surface parameters for the Community Land Model, Geosci. Model Dev., 5, 1341–1362, https://doi.org/10.5194/gmd-5-1341-2012, 2012.

Kerr, Y. H., Waldteufel, P., Wigneron, J. P., Martinuzzi, J., Font, J., and Berger, M.: Soil moisture retrieval from space: the Soil Moisture and Ocean Salinity (SMOS) mission, IEEE T. Geosci. Remote, 39, 1729–1735, 2001.

Kerr, Y. H., Waldteufel, P., Wigneron, J.-P., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M.-J., Font, J., Reul, N., Gruhier, C., Juglea, S. E., Drinkwater, M. R., Hahne, A., Martin-Neira, M., and Mecklenburg, S.: The SMOS mission: New tool for monitoring key elements ofthe global water cycle, P. IEEE, 98, 666–687, 2010.

Kerr, Y. H., Waldteufel, P., Richaume, P., Wigneron, J. P., Ferrazzoli, P., Mahmoodi, A., Bitar, A. 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, 2012.

Kiem, A. S., Johnson, F., Westra, S., van Dijk, A., Evans, J. P., O'Donnell, A., Rouillard, A., Barr, C., Tyler, J., Thyer, M., Jakob, D., Woldemeskel, F., Sivakumar, B., and Mehrotra, R.: Natural hazards in Australia: droughts, Climatic Change, 139, 37–54, 2016.

Kumar, S. V., Reichle, R. H., Koster, R. D., Crow, W. T., and Peters-Lidard, C. D.: Role of Subsurface Physics in the Assimilation of Surface Soil Moisture Observations, J. Hydrometeorol., 10, 1534–1547, 2009.

Kumar, S. V., Peters-Lidard, C. D., Santanello, J. A., Reichle, R. H., Draper, C. S., Koster, R. D., Nearing, G., and Jasinski, M. F.: Evaluating the utility of satellite soil moisture retrievals over irrigated areas and the ability of land data assimilation methods to correct for unmodeled processes, Hydrol. Earth Syst. Sci., 19, 4463–4478, https://doi.org/10.5194/hess-19-4463-2015, 2015.

Kumar, V., Dharssi, I., Bally, J., Steinle, P., McJannet, D., and Walker, J.: Comparison of soil wetness from multiple models over Australia with observations, Water Resour. Res., 53, 633–646, https://doi.org/10.1002/2015WR017738, 2017.

Lehner, B., Verdin, K., and Jarvis, A.: New global hydrography derived from spaceborne elevation data, Eos, 89, 93–94, 2008.

Leroux, D. J., Kerr, Y. H., Richaume, P., and Fieuzal, R.: Spatial distribution and possible sources of SMOS errors at the global scale, Remote Sens. Environ., 133, 240–250, 2013.

Leroux, D. J., Pellarin, T., Vischel, T., Cohard, J.-M., Gascon, T., Gibon, F., Mialon, A., Galle, S., Peugeot, C., and Seguis, L.: Assimilation of SMOS soil moisture into a distributed hydrological model and impacts on the water cycle variables over the Ouémé catchment in Benin, Hydrol. Earth Syst. Sci., 20, 2827–2840, https://doi.org/10.5194/hess-20-2827-2016, 2016.

Lievens, H., Al Bitar, A., Verhoest, N., Cabot, F., De Lannoy, G., Drusch, M., Dumedah, G., Franssen, H., Kerr, Y., Tomer, S., Martens, B., Merlin, O., Pan, M., van den Berg, M., Vereecken, H., Walker, J., Wood, E., and Pauwels, V.: Optimization of a radiative transfer forward operator for simulating SMOS brightness temperatures over the Upper Mississippi Basin, J. Hydrometeorol., 16, 1109–1134, 2015a.

Lievens, H., Tomer, S., Al Bitar, A., De Lannoy, G., Drusch, M., Dumedah, G., Hendricks Franssen, H.-J., Kerr, Y., Martens, B., Pan, M., Roundy, J., Vereecken, H., Walker, J., Wood, E., Verhoest, N., and Pauwels, V.: SMOS soil moisture assimilation for improved hydrologic simulation in the Murray Darling Basin, Australia, Remote Sens. Environ., 168, 146–162, 2015b.

Lievens, H., De Lannoy, G., Al Bitar, A., Drusch, M., Dumedah, G., Hendricks Franssen, H.-J., Kerr, Y., Tomer, S., Martens, B., Merlin, O., Pan, M., Roundy, J., Vereecken, H., Walker, J., Wood, E., Verhoest, N., and Pauwels, V.: Assimilation of SMOS soil moisture and brightness temperature products into a land surface model, Remote Sens. Environ., 180, 292–304, 2016.

López López, P., Wanders, N., Schellekens, J., Renzullo, L. J., Sutanudjaja, E. H., and Bierkens, M. F. P.: Improved large-scale hydrological modelling through the assimilation of streamflow and downscaled satellite soil moisture observations, Hydrol. Earth Syst. Sci., 20, 3059–3076, https://doi.org/10.5194/hess-20-3059-2016, 2016.

Martens, B., Miralles, D., Lievens, H., Fernández-Prieto, D., and Verhoest, N.: Improving terrestrial evaporation estimates over continental Australia through assimilation of SMOS soil moisture, Int. J. Appl. Earth Obs., 48, 146–162, 2016.

McColl, K., Pipunic, R., Ryu, D., and Walker, J.: Validation of the MODIS LAI product in the Murrumbidgee catchment, Australia, in: Proc. 19th Int. Congress on Modelling and Simulation, 12–16 December 2011, Perth, Australia, 1973–1979, 2011.

Mecklenburg, S., Drusch, M., Kaleschke, L., Rodriguez-Fernandez, N., Reul, N., Kerr, Y., Font, J., Martin-Neira, M., Oliva, R., Daganzo-Eusebio, E., Grant, J., Sabia, R., Macelloni, G., Rautiainen, K., Fauste, J., de Rosnay, P., Munoz-Sabater, J., Verhoest, N., Lievens, H., Delwart, S., Crapolicchio, R., de la Fuente, A., and Kornberg, M.: ESA's Soil Moisture and Ocean Salinity mission: from science to operational applications, Remote Sens. Environ., 180, 3–18, 2016.

Merlin, O., Rudiger, C., Al Bitar, A., Richaume, P., Walker, J. P., and Kerr, Y. H.: Disaggregation of SMOS soil moisture in Southeastern Australia, IEEE T. Geosci. Remote, 50, 1556–1571, 2012.

Mironov, V. L., Dobson, M. C., Kaupp, V. H., Komarov, S. A., and Kleshchenko, V. N.: Generalized refractive mixing dielectric model for moist soils, IEEE T. Geosci. Remote, 42, 773–785, 2004.

Miyoshi, T. and Yamane, S.: Local Ensemble Transform Kalman Filtering with an AGCM at a T159/L48 Resolution, Mon. Weather Rev., 135, 3841–3861, 2007.

Mohanty, B. P., Cosh, M. H., Lakshmi, V., and Montzka, C.: Soil Moisture Remote Sensing: State-of-the-Science, Vadose Zone J., 16, https://doi.org/10.2136/vzj2016.10.0105, 2017.

Montaldo, N., Albertson, J. D., Mancini, M., and Kiely, G.: Robust simulation of root zone soil moisture with assimilation of surface soil moisture data, Water Resour. Res., 37, 2889–2900, 2001.

Montzka, C., Moradkhani, H., Weihermüller, 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, 2011.

Montzka, C., Pauwels, V., Franssen, H.-J. H., Han, X., and Vereecken, H.: Multivariate and multiscale data assimilation in terrestrial systems: A review, Sensors, 12, 16291–16333, 2012.

Muñoz-Sabater, J.: Incorporation of passive microwave Brightness Temperatures in the ECMWF soil moisture analysis, Remote Sensing, 7, 5758–5784, 2015.

Oleson, K. W., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Koven, C. D., Levis, S., Li, F., Riley, W. J., Subin, Z. M., Swenson, S. C., and Thornton, P. E.: Technical Description of version 4.5 of the Community Land Model (CLM), Tech. rep., NCAR, Boulder, Colorado, USA, 2013.

Panciera, R., Walker, J. P., Kalma, J. D., Kim, E. J., Hacker, J. M., Merlin, O., Berger, M., and Skou, N.: The NAFE'05/CoSMOS Data Set: Toward SMOS Soil Moisture Retrieval, Downscaling, and Assimilation, IEEE T. Geosci. Remote, 46, 736–745, 2008.

Panciera, R., Walker, J. P., Jackson, T. J., Gray, D. A., Tanase, M. A., Ryu, D., Monerris, A., Yardley, H., Rüdiger, C., Wu, X., Gao, Y., and Hacker, J. M.: The Soil Moisture Active Passive Experiments (SMAPEx): Toward Soil Moisture Retrieval From the SMAP Mission, IEEE T. Geosci. Remote, 52, 490–507, 2014.

Parada, L. M. and Liang, X.: Optimal multiscale Kalman filter for assimilation of near-surface soil moisture into land surface models, J. Geophys. Res.-Atmos., 109, D24109, https://doi.org/10.1029/2004JD004745, 2004.

Peischl, S., Walker, J., Allahmoradi, M., Barrett, D., Gurney, R., Kerr, Y., Kim, E., Le Marshall, J., Rüdiger, C., Ryu, D., and Ye, N.: Towards validation of SMOS using airborne and ground data over the Murrumbidgee catchment, in: Proc. MODSIM, 18th World IMACS/MODSIM Congress, 13–17 July 2009, Cairns, Australia, 3733–3739, 2009.

Peischl, S., Walker, J. P., Rüdiger, C., Ye, N., Kerr, Y. H., Kim, E., Bandara, R., and Allahmoradi, M.: The AACES field experiments: SMOS calibration and validation across the Murrumbidgee River catchment, Hydrol. Earth Syst. Sci., 16, 1697–1708, https://doi.org/10.5194/hess-16-1697-2012, 2012.

Pellarin, T., Wigneron, J. P., Calvet, J. C., Berger, M., Douville, H., Ferrazzoli, P., Kerr, Y. H., Lopez-Baeza, E., Pulliainen, J., Simmonds, L. P., and Waldteufel, P.: Two-year global simulation of L-band brightness temperatures over land, IEEE T. Geosci. Remote, 41, 2135–2139, 2003.

Piles, M., Camps, A., Vall-Llossera, M., Corbella, I., Panciera, R., Rudiger, C., Kerr, Y. H., and Walker, J.: Downscaling SMOS-derived soil moisture using MODIS visible/infrared data, IEEE T. Geosci. Remote, 49, 3156–3166, 2011.

Reichle, R. H.: Data assimilation methods in the Earth sciences, Adv. Water Resour., 31, 1411–1418, 2008.

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

Renzullo, L. J., Van Dijk, A., Perraud, J.-M., Collins, D., Henderson, B., Jin, H., Smith, A., and McJannet, D.: Continental satellite soil moisture data assimilation improves root-zone moisture analysis for water resources assessment, J. Hydrol., 519, 2747–2762, 2014.

Ridler, M.-E., Madsen, H., Stisen, S., Bircher, S., and Fensholt, R.: Assimilation of SMOS-derived soil moisture in a fully integrated hydrological and soil-vegetation-atmosphere transfer model in Western Denmark, Water Resour. Res., 50, 8962–8981, 2014.

Samaniego, L., Kumar, R., and Zink, M.: Implications of parameter uncertainty on soil moisture drought analysis in Germany, J. Hydrometeorol., 14, 47–68, 2013.

Scholze, M., Kaminski, T., Knorr, W., Blessing, S., Vossbeck, M., Grant, J., and Scipal, K.: Simultaneous assimilation of SMOS soil moisture and atmospheric CO_2 in situ observations to constrain the global terrestrial carbon cycle, Remote Sens. Environ., 180, 334–345, 2016.

Sheffield, J., Wood, E. F., Chaney, N., Guan, K., Sadri, S., Yuan, X., Olang, L., Amani, A., Ali, A., Demuth, S., and Ogallo, L.: A drought monitoring and forecasting system for sub-Sahara African water resources and food security, B. Am. Meteorol. Soc., 95, 861–882, 2014.

Smith, A. B., Walker, J. P., Western, A. W., Young, R. I., Ellett, K. M., Pipunic, R. C., Grayson, R. B., Siriwardena, L., Chiew, F. H. S., and Richter, H.: The Murrumbidgee soil moisture monitoring network data set, Water Resour. Res., 48, W07701, https://doi.org/10.1029/2012WR011976, 2012.

Su, C.-H., Ryu, D., Young, R. I., Western, A. W., and Wagner, W.: Inter-comparison of microwave satellite soil moisture retrievals over the Murrumbidgee Basin, southeast Australia, Remote Sens. Environ., 134, 1–11, 2013.

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., Rippey, B., Tinker, R., Palecki, M., Stooksbury, D., Miskus, D., and Stephens, S.: The drought monitor, B. Am. Meteorol. Soc., 83, 1181–1190, 2002.

Tian, S., Tregoning, P., Renzullo, L. J., van Dijk, A. I. J. M., Walker, J. P., Pauwels, V. R. N., and Allgeyer, S.: Improved water balance component estimates through joint assimilation of GRACE water storage and SMOS soil moisture retrievals, Water Resour. Res., 53, 1820–1840, 2017.

Van der Schalie, R., Parinussa, R., Renzullo, L. J., Van Dijk, A., Su, C.-H., and de Jeu, R. A.: SMOS soil moisture retrievals using the land parameter retrieval model: Evaluation over the Murrumbidgee Catchment, southeast Australia, Remote Sens. Environ., 163, 70–79, 2015.

van Dijk, A. I. J. M. and Renzullo, L. J.: Water resource monitoring systems and the role of satellite observations, Hydrol. Earth Syst. Sci., 15, 39–55, https://doi.org/10.5194/hess-15-39-2011, 2011.

Van Dijk, A., Bacon, D., Barratt, D., Crosbie, R., Daamen, C., Fitch, P., Frost, A., Guerschman, J., Henderson, B., King, E., McVicar, T., Renzullo, L., Stenson, M., and Viney, N.: Design and development of the Australian Water Resources Assessment system, in: Water Information Research and Development Alliance: Science Symposium Proceedings, edited by: Sims, J., Merrin, L., Ackland, R., and Herron, N., 1–5 August 2011, Melbourne, Australia, CSIRO, 17–27, 2011.

van Dijk, A. I. J. M., Beck, H. E., Crosbie, R. S., de Jeu, R. A. M., Liu, Y. Y., Podger, G. M., Timbal, B., and Viney, N. R.: The Millennium Drought in southeast Australia (2001–2009): Natural and human causes and implications for water resources, ecosystems, economy, and society, Water Resour. Res., 49, 1040–1057, https://doi.org/10.1002/wrcr.20123, 2013.

Vaze, J., Viney, N., Stenson, M., Renzullo, L., Van Dijk, A., Dutta, D., Crosbie, R., Lerat, J., Penton, D., Vleeshouwer, J., Peeters, L., Teng, J., Kim, S., Hughes, J., Dawes, W., Zhang, Y., Leighton, B., Perraud, J.-M., Joehnk, K., Yang, A., Wang, B., Frost, A., Elmahdi, A., Smith, A., and Daamen, C.: The Australian Water Resource Assessment Modelling System (AWRA), in: 20th International Congress on Modelling and Simulation, 1–6 December 2013, Adelaide, Australia, vol. 16, 2013.

Vereecken, H., Schnepf, A., Hopmans, J., Javaux, M., Or, D., Roose, T., Vanderborght, J., Young, M., Amelung, W., Aitkenhead, M., Allison, S. D., Assouline, S., Baveye, P., Berli, M., Brüggemann, N., Finke, P., Flury, M., Gaiser, T., Govers, G., Ghezzehei, T., Hallett, P., Hendricks Franssen, H. J., 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., Šimu̇nek, J., Tiktak, A., Van Dam, J., van der Zee, S. E. A. T. M., Vogel, H. J., Vrugt, J. A., Wöhling, T., and Young, I. M.: Modeling soil processes: Review, key challenges, and new perspectives, Vadose Zone J., 15, https://doi.org/10.2136/vzj2015.09.0131, 2016.

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., and Viterbo, P.: The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data, Water Resour. Res., 50, 7505–7514, 2014.

Wigneron, J.-P., Laguerre, L., and Kerr, Y. H.: A simple parameterization of the L-band microwave emission from rough agricultural soils, IEEE T. Geosci. Remote Sens., 39, 1697–1707, 2001.

Wigneron, J.-P., Kerr, Y., Waldteufel, P., Saleh, K., Escorihuela, M.-J., Richaume, P., Ferrazzoli, P., De Rosnay, P., Gurney, R., Calvet, J.-C., Grant, J. P., Guglielmetti, M., Hornbuckle, B., Mätzler, C., Pellarin, T., and Schwank, M.: L-band microwave emission of the biosphere (L-MEB) model: Description and calibration against experimental data sets over crop fields, Remote Sens. Environ., 107, 639–655, 2007.

Yang, K., Watanabe, T., Koike, T., Li, X., Fujii, H., Tamagawa, K., Ma, Y., and Ishikawa, H.: Auto-calibration system developed to assimilate AMSR-E data into a land surface model for estimating soil moisture and the surface energy budget, J. Meteorol. Soc. Jpn. Ser. II, 85, 229–242, 2007.

Yilmaz, M. T. and Crow, W. T.: The optimality of potential rescaling approaches in land data assimilation, J. Hydrometeorol., 14, 650–660, 2013.