Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-188, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 18 April 2017

Hydrology and Earth System Sciences
Discussions

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]Ghent University, Ghent, Belgium
[2]Forschungszentrum Jülich, Jülich, Germany
[3]NASA Goddard Space Flight Center, 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), improving soil moisture simulations over the Australian continent. 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 increases moderately from 0.61 to 0.68 when the root-zone is included in the updates. A slightly reduced improvement is achieved when restricting the assimilation to the upper soil layers. Furthermore, the long-term assimilation impact is analysed by looking at a set of quantiles computed 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 now still limited, longer L-band radiometer time series will become available and make model output improved by assimilating such data more usable for extreme event statistics.

1 Introduction

The potential to improve land surface simulations of soil moisture by assimilating satellite measurements is well known (Mohanty et al., 2017; Chen et al., 2014; Jia et al., 2009; De Lannoy et al., 2007; Parada and Liang, 2004). 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). The on-board passive Imaging Radiometer with Aperture Synthesis (MIRAS) instrument, sensitive at 1.4 GHz, measures multi-angular top of atmosphere brightness temperatures at horizontal (H) and vertical (V) polarisation influenced by, among others, surface soil moisture. 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 centimetres of the soil column, into the according model layers (Montzka et al., 2012; Reichle, 2008), the assimilation impact in deeper layers will depend on model physics (Montzka et al., 2011; Montaldo et al., 2001). Alternatively, deeper layers can be updated directly by making use of the relationship, i.e. covariance, between observable and unobservable layers by applying methods such as the various implementations of the Kalman Filter (Kalman et al., 1960). For plants, these deeper layers act as the root zone,

in which 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 an evapotranspiration model can benefit from assimilating these data over Australia and Lievens et al. (2015b), concluding that the positive assimilation impact on soil moisture can improve streamflow simulations, as shown for the Murray-Darling basin. The impact on both streamflow and evaporation is evaluated by Ridler et al. (2014) for western Denmark. More recently, Scholze et al. (2016) have assimilated SMOS retrievals together with CO2 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 and Level 3 soil moisture retrievals in essence solve minimisation problems between simulated brightness temperatures and the observed satellite signal (Kerr et al., 2012). The simulated top of atmosphere signal is thereby dependent on both static and dynamic ancillary data, which is based on input and output of a specific land surface model, e.g. in the case of 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 thus be beneficiary to directly assimilate the brightness temperatures as this allows for a consistent use of auxiliary information for the land surface model and the radiative transfer model (Han et al., 2013). Nevertheless, similar to when assimilating soil moisture retrievals and having to deal with potentially large biases between retrieved and modelled soil moisture, large biases are also common between modelled and observed brightness temperatures due to the many uncertainties involved. The source of such uncertainties, among others, may lie in the atmospheric forcings, the land surface representation or the land surface model itself (Drusch et al., 2009; Barella-Ortiz et al., 2015). Since assimilation is expected to correct random errors only and most assimilation algorithms rely on unbiased observations, i.e. bias-blind, it is 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 one 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). Case studies for assimilation and how to deal with the bias are given by Muñoz-Sabater (2015) and Muñoz-Sabater et al. (2012), both evaluating the assimilation of SMOS brightness temperatures into the ECMWF soil moisture analysis, or by Lievens et al. (2016) who assimilate the soil moisture products as well as brightness temperatures into the Variable Infiltration Capacity (VIC) model for the Murray-Darling basin. Similarly, De Lannoy and Reichle (2016b) compare brightness temperature and soil moisture assimilation over the U.S. and De Lannoy and Reichle (2016a) describe the assimilation of only brightness temperatures into the GEOS-5 Catchment Land Surface Model.

Taken as a whole, studies directly assimilating real brightness temperatures over large scales remain limited and the concept still needs to be further explored, particularly using different qualitative land surface models. Furthermore, the assimilation impact is often evaluated by solely comparing soil moisture time series to a limited number of in-situ measurements. One study looking more at the long-term assimilation effect, albeit using active microwave data, was carried out by Draper and

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-188, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 18 April 2017

Reichle (2015), highlighting the fact that despite the required unbiased nature of a bias-blind assimilation system, assimilation can correct for longer-term behaviour and thus be beneficial for monitoring extreme events, such as droughts. Within this study, we specifically look at model state biases that might be introduced into the model over longer time periods, e.g. due to the way model physics potentially react differently to positive or negative increments, model and atmospheric perturbations or unresolved seasonal discrepancies between model and observation. Model state biases can also be more complex in nature than a simple shift of the mean value away from the open loop run and introduce varying changes at different quantile levels within the cumulative distribution functions (CDFs). Such changes can be of great importance in the context of hydrological monitoring systems, since absolute soil moisture values are difficult to compare between grid cells due to e.g. differences in land cover and soil texture. Looking at the relative occurrence of a specific value is more useful, especially when trying to identify spatial patterns, such as areas suffering from extreme conditions like droughts or possible flooding. Examples for existing hydrological monitoring systems are for instance the US. drought monitor (Svoboda et al., 2002), the African Flood and Drought Monitor (Sheffield et al., 2014) or the German Drought Monitor (Samaniego et al., 2013), with all of them using soil moisture quantiles at grid cell level to characterise different levels of severity. With longer L-band time series becoming available, modelled soil moisture time series improved by assimilation will become sufficiently long for computing the relative occurrence of events and existing monitoring systems, now often relying on purely modelled data, might subsequently benefit from using such data.

Within this study we assimilate 6 full years of 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 as well in terms of induced model biases and quantile changes. The CLM is therefore coupled to the Community Microwave Emission Model (CMEM, 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 (Han et al., 2015b; Miyoshi and Yamane, 2007; Hunt et al., 2007). The observation bias problem 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 offline computed forward simulation anomalies at each grid point. The assimilation impact is evaluated by comparing open loop and assimilation results to in-situ measurements extracted from the International Soil Moisture Network (ISMN) (Dorigo et al., 2011) for two main assimilation experiments: In the first experiment (DA1) brightness temperature assimilation is restricted to the upper three CLM soil layers corresponding to a depth of 9 cm. The upper six model layers, reaching 50 cm, are updated in the second experiment (DA2). These two experiments enable us to examine to what extent CLM model physics are sufficient to propagate upper level increments to the root-zone in comparison to directly applying the increments in this depth. For these two experiments the soil texture perturbations applied for the ensemble generation were incrementally reduced with layer depth, avoiding to large updates in deep layers. Although the analysis is not focused on it, we have included a third experiment (DA0) using homogeneous soil texture perturbations across all layers, highlighting the problem of large increments in lower layers when the ensemble spread is too large. For the quantile analysis the quantiles at

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-188, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 18 April 2017

1 % steps are computed at each model grid-point, enabling the sufficient empirical estimation of the CDFs. As an example to highlight the effects of quantile changes, we show a very dry event defined at the 10 % quantile level and to what extent its classification changes for the open loop run and the data assimilation.

2 The Community Land Model

The Community Land Model (CLM) is the land surface component of the Community Earth System Model (CESM) and can be run either in coupled mode together with the other CESM components or offline using precomputed atmospheric forcings (Oleson et al., 2013). CLM provides global surface datasets which can be interpolated to pre-defined or custom resolutions and grid types both globally as well as regionally including single point simulations. For this continental scale study, we replace many of the surface datasets with more recent and higher resolution data, creating a consistent surface dataset at 0.25 degree resolution. The model is run at 30-minute time steps, with hourly outputs, allowing for the sufficiently correct temporal alignment of model and satellite observations.

2.1 Surface Datasets

Each grid cell within CLM is divided into land units covering a 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 over their percentage of coverage in 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 up-scaled 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 up-scaled 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 arc-second 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

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-188, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 18 April 2017

Hydrology and Earth System Sciences
Discussions

coarse scale soil fraction contains no organic matter. Bulk density was used to compute the organic matter content required by CLM assuming 0.58 g organic matter per kilogram. The rational for creating high-resolution datasets for CLM closely followed the approaches described in detail in Ke et al. (2012) and Han et al. (2012), similarly replacing the CLM standard datasets.

2.2 ERA-Interim atmospheric forcing

CLM provides forcings (CRUNCEP) not available for the required time period. However, the rolling release of ERA-Interim reanalysis data (Dee et al., 2011) with a time lag of only a few months enables the assimilation of relatively new satellite measurements, in this case SMOS, and thus ERA-Interim atmospheric data were used to force the CLM land surface model over Australia.

The variables 2 m air temperature, 2 m pressure, short-wave 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. 2 m wind speed was derived from the provided wind speed components in lateral and longitudinal direction. As ERA-Interim is 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 and the data needed to be handled respectively. Forecasts for flux variables are provided bi-daily at 0:00 and 12:00 UTC for 3, 6, 9 and 12 hour forecast periods and in accumulated form. For example, the precipitation forecast for a 6-hour time window is the accumulated precipitation over 6 hours. In order to obtain a precipitation estimate for the hours 3 - 6, the precipitation forecast for the first 3-hour 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 degrees spatial resolution to 0.25 degrees 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 timesteps is performed at CLM runtime applying an appropriate interpolation algorithm to each variable. Incoming radiation is interpolated using a cosine function simulating the position of the sun and 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 to couple it to different models, observation operators and assimilation algorithms. The version used within this study is coupled to the Community Land Model and the Community Microwave Emissions Model (CMEM, de Rosnay et al. (2009)) observation operator. Furthermore, the system uses the Local Ensemble Transform Kalman Filter (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 northwestern China (Han et al., 2012) and in the Rur catchment in Germany (Han et al., 2015c). The system allows for dual state parameter estimation as shown in Han et al. (2014).

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-188, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 18 April 2017

Hydrology and Earth System Sciences
Open Access
EGU
Discussions

DasPy has been developed with a focus on High-Performance Computing and 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. Although we are using Level 3 data with the antenna pattern already partially accounted for, the original SMOS footprint is 43 km and thus covers more than a single model grid cell, which would encourage assimilation in 3D. However, mostly for reasons of simplicity and the already performed inverse distance observation regridding, partially accounting for this, we here only assimilate observations directly covering a grid cell.

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

$$x^b_{n,k} = M_n(x^a_{n-1,k}) \tag{1}$$

The background ensemble perturbations X^b at the current time step can be computed as:

$$X^b = [x^b_1 - \bar{x}^b | ... | x^b_k - \bar{x}^b] \tag{2}$$

The individual ensemble states x^b are propagated into observation space using a forward operator H, in this case CMEM.

$$y^b_k = H(x^b_k) \tag{3}$$

and the forward simulation perturbations are defined as:

$$Y^b = [y^b_1 - \bar{y}^b | ... | y^b_k - \bar{y}^b] \tag{4}$$

Within the ensemble space the analysis error covariance $\tilde{P}^a$ is computed through

$$\tilde{P}^a = [(K-1)I + (Y^b)^T R^{-1} Y^b]^{-1} \tag{5}$$

allowing for the computation of $\bar{W}^a$ as the mean weighting vector

$$\bar{w}^a = \tilde{P}_a Y^{bT} R^{-1} (y^0 - \bar{y}^b) \tag{6}$$

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

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

The analysis perturbations are defined as X^a

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

with

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

where the analysis error covariance P^a is given by:

$$P^a = X^b \tilde{P}^a (X^b)^T \tag{10}$$

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, while 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 0.5. In order to avoid ensemble collapse during dry periods also soil texture is perturbed once at model startup with spatially correlated multiplicative noise and a standard deviation of 10 percent for clay and sand for the top two soil layers, while constraining the sum of soil and clay to a maximum of 98 percent. For the lower layers, the top layer multiplicative factor is rescaled by using the inverse relationship between each soil layers thickness and the summed soil layer thickness of the two top layers (see Table 1). This should insure 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. Because CLM derives hydraulic properties based on soil texture, it is to be noted that as a consequence each ensemble member runs with slightly modified model physics. As stated, for demonstration purpose a third experiment (DA0), was performed, using homogeneous texture perturbations for all soil layers.

3.3 Observation Operator

Forward simulations from the model space to the observation space are performed with the Community Microwave Emission Model (CMEM, version 5.1). Model output at each observation time, with the observation time rounded to the full hour, serve as input in order to simulate brightness temperatures as measured by the satellite. SMOS ascending and descending orbits have a local overpass time of approximately 6 am and 6 pm. 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

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-188, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 18 April 2017

of -10 hours for the entire Australian continent. This greatly decreases the number of analysis steps, since individual orbits within one day can be assimilated at once, and it is still assumed to provide a sufficiently correct temporal alignment between observations and model forward simulations.

CMEM requires time invariant information such as soil layer depth, sand, clay and water fractions, surface height as well as the dominant vegetation type covering the grid cell. Plant functional types are reclassified to ECOCLIMAP vegetation classes and the dominant type for low and high vegetation is then used by the CMEM (Champeaux et al., 2005). Based on this reclassification the LAI information is assigned to the dominant 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 for all defined soil layers and 2 m air temperature. CMEM supports different types of sub-modules for specific calculations. Within this study, the Mironov model (Mironov et al., 2004) is 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. 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 applying 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.

3.4 Observations and anomaly computation

SMOS daily brightness temperatures at horizontal H polarisation and 42.5 incidence angle provided by Centre Aval de Traitement des Donées (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 degree rectangular model grid using inverse-distance interpolation. Based on these data we compute the climatology for each day for the years 2010 - 2015 by averaging along a 7-day moving window across the 6 years, producing separate climatologies for ascending and descending orbits, and thus removing seasonal differences between forward simulations and observations (see e.g. De Lannoy and Reichle (2016a)). Anomalies are then computed between the climatologies and the original SMOS time series. Brightness temperature forward simulations based on an open loop run with 32 ensembles are performed and the ensemble mean climatology is derived the same way as the observation climatology. 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 taken in order to account for

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-188, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 18 April 2017

seasonal mean differences between simulations and observations and to remove the bias, without requiring more aggressive CDF-matching techniques at seasonal level. 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 K^2 was defined, accounting for 4 K instrument error and an assumed mean error of 3 K for the forward simulations and representativeness error. This can be seen as a low estimate and was based on the assumption that the brightness temperature binning around the 42.5 degree incidence angle should slightly reduce the instrument error. During assimilation at each time step the current forward simulation is subtracted from the precomputed forward simulation climatology to compute on-the-fly anomalies and compared to the precomputed SMOS anomaly. The difference between the two is the innovation between model and observations in observation space and 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 / descending orbits respectively.

4 Data assimilation and results

Two main data assimilation runs are carried out to assess the impact of the brightness temperature assimilation. The rescaled observed brightness temperatures are representative for the top layer and lower layers are updated making use of the covariances between the ensembles of the topmost CLM layer and the ensembles of all subsequent layers. Within the first experiment (DA1) the top 3 CLM layers, reaching a depth of 9 cm, are updated. The assimilation was not restricted to the top two layers, corresponding to the depth where SMOS is mainly sensitive, since the assimilation impact would have been further diminished, as discussed later on. The top 6 layers, later also referred to as the root zone, are updated within experiment two (DA2). Both experiments are validated by comparing the assimilation impact to an open loop run in respect to in-situ soil moisture observations. In-situ stations are available for a number of depths, mostly 5 cm and 8 cm, thus corresponding well to DA1, as well as 30 cm and a limited number of deeper depths, thus corresponding well to DA2. In addition to the in-situ validation, quantile shifts in respect to the open loop run are analysed, highlighting the long-term effects of the data assimilation. A set of quantiles is computed at each grid cell allowing for the empirical estimation of the cumulative distribution functions, since varying quantile shifts, both positive or negative, are possible at different quantile levels. For the quantile analysis, a third experiment (DA0) was added using homogeneous soil texture perturbations, as opposed to texture perturbations decreasing with layer depth, showing how sensitive the quantile analysis and the entire assimilation is from these perturbations. Finally, the implications of quantile changes were showcased for one specific dry event and placed within the context of hydrological monitoring systems.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-188, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 18 April 2017

Hydrology and Earth System Sciences
Open Access
EGU
Discussions

4.1 Correlation with in-situ observations

For validation purposes the CLM soil moisture model output is compared to in-situ measurements obtained from the International Soil Moisture Network (ISMN Dorigo et al. (2011)), which were additionally quality checked both visually and in an automatic way to remove erroneous soil moisture behaviour, e.g. identical values over long periods. For each in-situ measurement the weighted average of the corresponding CLM soil moisture is computed, using layer thickness as weights, prior to the comparison. When taking into account only measurements with at least one year of data, not necessarily consecutive, correlations improve from 0.613 for the open loop run to 0.640 and 0.681 for DA1 and DA2 respectively for top layer soil moisture (number of stations, n = 17). Bottom layer soil moisture improvements are lower with average correlation coefficients of 0.626 and 0.648 for DA1 and DA2, compared to 0.601 for the open loop run (n = 30). Upper soil moisture behaviour thus improves by also updating deeper layers and deeper layer soil moisture is slightly enhanced through only assimilating top level soil moisture with the assimilation effects only being propagated through model physics. Overall, updating the top 6 CLM layers results in the largest improvements. Figure 1 shows Taylor diagrams for the validation results of experiment DA2. The Taylor diagrams further show a slightly decreased normalised standard deviation of the assimilation results in respect to the in-situ measurements when compared to the open loop time series in respect to the in-situ data. In terms of standardised RMSE it is less conclusive, and RMSE is slightly reduced by assimilation for some stations and slightly increased for some others. These findings correspond well to experiment DA 1, not shown here.

4.2 Soil moisture increments

In a bias-blind assimilation system it can be expected that, especially over longer time periods, the mean increments should be very close to zero. Figure 2 and Figure 4 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. Distinctive areas of mean positive increments for the ascending orbit are located in the North, South-West and South-East of Australia, seemingly being linked to the occurrence of vegetation. Nevertheless, the positive bias does not exceed 0.5 % soil moisture and the remaining parts of Australia either show no mean increments bias or slightly negative values. These patterns are still visible for the descending orbit or the root-zone soil moisture but are on overall weaker, except for DA 0 with little differences between top layer and deep layer increments. Interesting to highlight, is the fact that top layer deviations from zero are strongest for the assimilation experiment DA1 compared to DA 0 and DA2, with the latter two updating both top layer soil as well as the root-zone. A possible interpretation might be that the assimilation into deeper layers results in a longer lasting effect, potentially moving the model closer to subsequent observations and thus reducing subsequent increments. Figure 3 and 5 show the increment standard deviations, which are substantially stronger for the ascending orbit. Top layer increments are larger for DA 1 followed by DA 2 and then DA 0. Deep layer increments are much larger for DA 0 when compared to DA 2. For the ensemble generation, precipitation is the most dominant dynamic factor and leads to immediate increased ensemble spread, and thus 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. The static soil texture perturbations,

although their effect are also dynamically dependent on the current soil moisture state, lead to model uncertainty not directly induced by the forcing perturbations. Soil texture perturbations decrease with layer depth for DA 1 and DA 2 (see perturbation scaling shown in Table 1) and lead to the reduced covariance between top layer and deep layers, which results in decreasing increments for lower layers. It is thereby important to note that in terms of absolute soil moisture, large increments in low layers remove or add vastly larger amounts of water than similar increments in shallow layers and will exhibit strong effects on all above layers (see soil thickness in Table 1). As an example, removal of water in a deep lying layer will lead to increased percolation in all above lying layers, resulting in a sudden drying out of the top soil layers. Small biases inherent in the data assimilation system thus might be largely exaggerated when allowing to large deep layer updates. This is visible for DA 0 where homogeneous soil texture perturbations were applied across all layers and increments are not significantly reduced for the root-zone (Figure 3). Concluding on the relationship between increment bias and increment standard deviation, it seems that there is either a decreased increment bias at the expense of larger deep layer increments, or vice versa. Finally, regarding the spatial patterns of the increments, areas with very low increments in all layers, especially along the east coast, are either based on a lack of observations, as these were filtered due to the active vegetation science flags, or the fact that the high LAI values for high vegetation prescribed within the forward operator mask any signal from soil moisture.

4.3 Soil moisture quantiles

In addition to 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 showed that small increment biases exist, potentially resulting in long term effects in the resulting analysis. Figures 6, 7 and 8 show the 10 % quantile, thus very dry conditions, for the top 9 CLM layers and to which extent it shifts for each grid cell when compared to the open loop run. For experiment DA 1 (Figure 7), assimilation has a small impact on the topmost soil layer with the quantile increasing by up to 5 % for spatially limited areas. Much smaller changes are visible for deeper layers but with changes nevertheless reaching layers that are not directly updated. For experiment DA 2 (Figure 8), whilst also updating layers 3 - 6, slightly larger impacts on the quantiles in deeper layers can be observed with more areas showing an increased 10 % quantile. For the largest part, the patterns, at least for upper layers, reflect well the patterns identified in Figure 4, showing the mean increments, although in deeper layers also some independent patterns emerge that cannot be directly explained by any increments bias. Figure 6 shows the impact on the 10 % quantile for the experiment DA 0, using homogeneous soil texture perturbations across all layers. Here, much larger effects are visible especially within the root-zone soil, increasing the quantiles over wide areas of the Australian inland by up to 5 %. Hard to pinpoint the exact mechanisms for this behaviour, it does highlight the potential implications of updating thick soil layers. The complex nature of these shifts at throughout the entire CDFs is shown in Figure 9. The continental average empirical cumulative distribution functions are plotted for the soil layers 1 - 6 for the open loop run as well as DA1 and DA2. As shown for the 10 % quantile, lower quantiles are increased on average through data assimilation, whereas a small decrease in the upper quantile values can be observed resulting in an overall decrease of the standard deviation of the soil moisture analysis. As stated, we want to place these findings within the context of possible applications of such

datasets, e.g. hydrological monitoring systems, which directly make use of grid cell quantiles and empirical CDFs. Draper and Reichle (2015) have shown that data assimilation is able to correct modelled soil moisture also at longer time intervals from sub-seasonal to seasonal scale. The correction of the short term behaviour, i.e. daily, of soil moisture and all connected fluxes is of importance for e.g. land-atmosphere feedbacks, but would have an negligible effect when analysing phenomena that spread across larger spatial scales and time intervals. These long term effects however, might be the result of the accumulated small updates during the assimilation run instead of large increments, that would drastically change the soil moisture regime. Further, when classifying an event, such as a drought defined at a specific quantile level, there will be a twofold impact from the assimilation, namely the change in soil moisture itself as well as the shifted quantile. The latter is expected to have the larger effect on such event statistics, since the soil moisture analysis nominally will fluctuate around the open loop simulations and not show a consistent bias. We highlight a sample drought event on the East coast to show to which extent it's classification changes through the assimilation impact. Figure 10 therefore 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 Figure 8 and 9, 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 visible 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 even when e.g. being affected by the same large scale precipitation event. An alternative to spatially correlated noise would be to further increase the ensemble size from 32 to a significantly higher number, which however will also require larger computational resources. Additionally, when trying to extract meaningful statistics on the occurrence events, such as droughts, by extracting these as clusters of grid cells over the spatial and temporal domain, it might be especially important to clean up the dataset in the case of data assimilation using simple filter algorithms, such as applied by Herrera-Estrada et al. (2017).

5 Conclusion and Discussion

The Community Land Model was set up for the Australian continent and we have substituted the surface datasets with higher resolution and more recent data. Further, we have replaced the offline forcing data with ERA-Interim reanalysis data. The assimilation over 6 full years, from 2010 – 2015, of SMOS brightness temperatures improved soil moisture simulations when compared to in-situ measurements in the order of 11 %, which is similar to the impact in other studies. The CLM model remained uncalibrated, as this is tremendous task for large areas and calibration towards soil moisture simulations is difficult for the lack of in-situ measurements, but given the results, we are confident that this specific CLM setup could be applied to larger areas or at global scale. CLM model physics alone did propagate assimilation effects into the root zone when restricting assimilation to the top three layers, improving the correlation with in-situ measurements both for the root zone and surface soil moisture. However, the improvements were largest when directly updating both surface soil moisture and root zone soil moisture. CLM layer thickness greatly increases with depth which translates into identical increments in relative soil moisture being very different in terms of absolute soil moisture. To thus restrict too large updates within the root zone we scaled the soil

texture perturbations. The scaling factor for each layer was based on the ratio between layer thickness and the layer thickness of the top two layers, corresponding to the assumed depth were SMOS is sensitive, namely ca. 5 centimetres.

Mean increments showed distinctive patterns with slight positive biases up to 1 % soil moisture in areas covered by vegetation. Although this might be due to a problem with the uncalibrated forward operator or with the fixed high vegetation class LAI values within the forward operator, a further possible cause is the use of climatological LAI data for CLM. The use of such data is common practise within current land data assimilation systems. However, due to the abundance of vegetation data we encourage future studies to look into possible improvements by using non-climatological LAI. In fact, the climatological LAI data is often aggregated from monthly or sub-monthly LAI values and could be easily replaced with the non-aggregated data. Climatological LAI might especially pose a problem for the monitoring of extreme events, such as droughts, since these would likely result in lower LAI values again influencing the forward simulations. With climatological LAI data these feedback processes will not be modelled. The potential of the direct brightness temperature assimilation, allowing for the consistent use of data across model and forward simulations, should be made full use of. Within assimilation systems where the forward operator is calibrated, it is also likely that the problem of equifinality for the forward operator parameters could thereby be reduced.

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, although the very high quantiles remained unchanged, overall resulting in a reduced analysis variability. This highlights the fact, that although in principle the assimilation experiment was set up using unbiased observations, analysis biases with non-linear behaviour may still be introduced. An additional experiment using homogeneous soil texture perturbations across all layers was carried out showing much larger analysis biases compared to the open loop run at the 10 % quantile level, highlighting the problem of too large updates within the root zone and the general sensitivity towards model perturbations.

The reduction of analysis variability by the assimilation might be partly attributed to the anomaly rescaling and to disentangle wished for systematic enhancements from erroneously introduced analysis is a challenge. There is still not the one perfect approach to rescale the observations to mach the model or to calibrate the forward observation and looking at long term CDF changes induced by the assimilation should be part of evaluating these different approaches, also keeping the final application of the analysis data in mind.

Longer L-band time series are becoming available through the continuation of existing missions as well as new ones, such as the Soil Moisture Active Passive Mission (SMAP, Entekhabi et al. (2010)). This in the long run opens up more application possibilities, since soil moisture datasets enhanced by assimilating will allow for the computation of more stable grid cell CDFs. This will be especially useful within the context of 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. Within this study we have shown that the impact on CDFs will have an effect on the quantile based classification of a drought event and change the spatial extent of the affected area.

Especially for droughts these datasets are beneficial, since the assimilation can update the root zone soil moisture, which is of vital importance for plant growth and thus for monitoring agricultural drought. Many drought monitoring systems currently rely on precipitation based indices, e.g. the Palmer Drought Severity Index (PSDI, Palmer (1965)), due to the high correlation of soil moisture with precipitation over longer time spans. However,this correlation can be much smaller on small time spans and soil moisture is influenced by a multitude of factors within the land surface complex, which a land surface model is better able to represent across user-required time steps and spatial resolutions (Sheffield et al., 2004). But also in the case of floods an improved root-zone representation will be beneficial for predicting how much precipitation is required to saturate top layer soil.

On overall, within this study we have attempted to show the possible complex behaviours induced by long-term assimilation and implications for the use of assimilation improved soil moisture simulations. Yet, it is clear that more studies should be carried out, bridging the gap between technical assimilation studies and the application. For instance, we have shown assimilation induced quantile changes for one specific data assimilation setup. The systematic analysis of various observation rescaling techniques and the impact on recorded drought and flood events should be part of future studies.

Author contributions. DR carried out all presented work and wrote the draft version of the publication. XH gave continuous support concerning the data assimilation system and the many code changes required for this work. HL gave input on the technical design of the data assimilation experiments as well as the evolving manuscript. CM and NV provided additional remarks and suggestions for the evolving manuscript.

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

Acknowledgements. The research presented in this paper is funded by BELSPO (Belgian Science Policy Office) in the frame of the STEREO III programme – project HYDRAS+ (SR/00/302). The assimilation experiments were run on the JURECA supercomputer at Jülich Forschungszentrum. The open-source assimilation system DasPy is available at https://github.com/daspy/daspy and was adapted to meet the requirements of this study. Hans Lievens is a postdoctoral research fellow of the Research Foundation Flanders (FWO).

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, Journal of hydrometeorology, 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, Hydrology and Earth System Sciences Discussions, 12, 13 019–13 067, doi:10.5194/hessd-12-13019-2015, http://www.hydrol-earth-syst-sci-discuss.net/12/13019/2015/, 2015.

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 Biogeochemical Cycles, 16, 5–1–5–23, doi:10.1029/2000GB001360, http://dx.doi.org/10.1029/2000GB001360, 2002.

Champeaux, J. L., Masson, V., and Chauvin, F.: ECOCLIMAP: a global database of land surface parameters at 1 km resolution, Meteorological Applications, 12, 29–32, doi:10.1017/S1350482705001519, http://dx.doi.org/10.1017/S1350482705001519, 2005.

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

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

De Lannoy, G. J., 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 Resources Research, 43, 2007.

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, Journal of Hydrometeorology, 17, 669–691, doi:10.1175/JHM-D-15-0037.1, http://dx.doi.org/10.1175/JHM-D-15-0037.1, 2016a.

De Lannoy, G. J. M. and Reichle, R. H.: Assimilation of SMOS brightness temperatures or soil moisture retrievals into a land surface model, Hydrology and Earth System Sciences, 20, 4895–4911, doi:10.5194/hess-20-4895-2016, http://www.hydrol-earth-syst-sci.net/20/4895/2016/, 2016b.

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, Journal of Hydrometeorology, 14, 765–785, doi:10.1175/JHM-D-12-092.1, http://dx.doi.org/10.1175/JHM-D-12-092.1, 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, Journal of Geophysical Research: Atmospheres, 114, n/a–n/a, doi:10.1029/2008JD010724, http://dx.doi.org/10.1029/2008JD010724, d05108, 2009.

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., et al.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Quarterly Journal of the royal meteorological society, 137, 553–597, 2011.

Dorigo, W., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., Oevelen, P. v., et al.: The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements, Hydrology and Earth System Sciences, 15, 1675–1698, 2011.

Draper, C. and Reichle, R.: The impact of near-surface soil moisture assimilation at subseasonal, seasonal, and inter-annual timescales, Hydrology and Earth System Sciences, 19, 4831–4844, doi:10.5194/hess-19-4831-2015, http://www.hydrol-earth-syst-sci.net/19/4831/2015/, 2015.

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, Journal of Hydrometeorology, 10, 213, doi:10.1175/2008JHM964.1, 2009.

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 Zyl, J. V.: The Soil Moisture Active Passive (SMAP) Mission, Proceedings of the IEEE, 98, 704–716, doi:10.1109/JPROC.2010.2043918, 2010.

Han, X., Li, X., Hendricks Franssen, H., Vereecken, H., and Montzka, C.: Spatial horizontal correlation characteristics in the land data assimilation of soil moisture, Hydrology and Earth System Sciences, 16, 1349–1363, 2012.

Han, X., Hendricks Franssen, H.-J., Li, X., Zhang, Y., Montzka, C., and Vereecken, H.: Joint assimilation of surface temperature and L-band microwave brightness temperature in land data assimilation, Vadose Zone Journal, 12, 2013.

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 resources research, 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, Geoscientific Model Development Discussions, 8, 7395–7444, doi:10.5194/gmdd-8-7395-2015, http://www.geosci-model-dev-discuss.net/8/7395/2015/, 2015a.

Han, X., Li, X., He, G., Kumbhar, P., Montzka, C., Kollet, S., Miyoshi, T., Rosolem, R., Zhang, Y., Vereecken, H., et al.: DasPy 1.0–the Open Source Multivariate Land Data Assimilation Framework in combination with the Community Land Model 4.5, Geoscientific Model Development Discussions, 8, 2015b.

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, e0116 435, 2015c.

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, doi:10.1371/journal.pone.0105992, http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0105992, 2014.

Herrera-Estrada, J. E., Satoh, Y., and Sheffield, J.: Spatiotemporal dynamics of global drought, Geophysical Research Letters, pp. n/a–n/a, doi:10.1002/2016GL071768, http://dx.doi.org/10.1002/2016GL071768, 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, International journal of climatology, 25, 1965–1978, 2005.

Hunt, B. R., Kostelich, E. J., and Szunyogh, I.: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter, Physica D: Nonlinear Phenomena, 230, 112 – 126, doi:http://dx.doi.org/10.1016/j.physd.2006.11.008, //www.sciencedirect.com/science/article/pii/S0167278906004647, data Assimilation, 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, Science in China Series D: Earth Sciences, 52, 1835–1848, 2009.

Kalman, R. E. et al.: A new approach to linear filtering and prediction problems, Journal of basic Engineering, 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, Geoscientific Model Development, 5, 1341–1362, doi:10.5194/gmd-5-1341-2012, http://www.geosci-model-dev.net/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 Transactions on Geoscience and Remote Sensing, 39, 1729–1735, doi:10.1109/36.942551, 2001.

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 Transactions on Geoscience and Remote Sensing, 50, 1384–1403, doi:10.1109/TGRS.2012.2184548, 2012.

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

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, Journal of Hydrometeorology, 16, 1109–1134, http://dx.doi.org/10.1175/JHM-D-14-0052.1, 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 Sensing of Environment, 168, 146–162, http://dx.doi.org/10.1016/j.rse.2015.06.025, 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 Sensing of Environment, 180, 292–304, http://dx.doi.org/10.1016/j.rse.2015.10.033, 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, International Journal of Applied Earth Observation and Geoinformation, 48, 146–162, http://dx.doi.org/10.1016/j.jag.2015.09.012, 2016.

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 Sensing of Environment, 180, 3–18, http://dx.doi.org/10.1016/j.rse.2015.12.025, 2016.

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 Transactions on Geoscience and Remote Sensing, 42, 773–785, doi:10.1109/TGRS.2003.823288, 2004.

Miyoshi, T. and Yamane, S.: Local Ensemble Transform Kalman Filtering with an AGCM at a T159/L48 Resolution, Monthly Weather Review, 135, 3841–3861, doi:10.1175/2007MWR1873.1, http://dx.doi.org/10.1175/2007MWR1873.1, 2007.

Mohanty, B. P., Cosh, M. H., Lakshmi, V., and Montzka, C.: Soil Moisture Remote Sensing: State-of-the-Science, Vadose Zone Journal, 16, 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 Resources Research, 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, Journal of Hydrology, 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, 16 291–16 333, 2012.

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

Muñoz-Sabater, J., Fouilloux, A., and de Rosnay, P.: Technical implementation of SMOS data in the ECMWF Integrated Forecasting System, IEEE Geoscience and Remote Sensing Letters, 9, 252–256, 2012.

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., 2013.

Palmer, W. C.: Meteorological drought, vol. 30, US Department of Commerce, Weather Bureau Washington, DC, 1965.

Parada, L. M. and Liang, X.: Optimal multiscale Kalman filter for assimilation of near-surface soil moisture into land surface models, Journal of Geophysical Research: Atmospheres, 109, 2004.

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 Transactions on Geoscience and Remote Sensing, 41, 2135–2139, doi:10.1109/TGRS.2003.815417, 2003.

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

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 Resources Research, 50, 8962–8981, doi:10.1002/2014WR015392, http://dx.doi.org/10.1002/2014WR015392, 2014.

Samaniego, L., Kumar, R., and Zink, M.: Implications of parameter uncertainty on soil moisture drought analysis in Germany, Journal of Hydrometeorology, 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 {CO2} in-situ observations to constrain the global terrestrial carbon cycle, Remote Sensing of Environment, 180, 334 – 345, doi:http://dx.doi.org/10.1016/j.rse.2016.02.058, //www.sciencedirect.com/science/article/pii/S0034425716300876, special Issue: ESA's Soil Moisture and Ocean Salinity Mission - Achievements and Applications, 2016.

Sheffield, J., Goteti, G., Wen, F., and Wood, E. F.: A simulated soil moisture based drought analysis for the United States, Journal of Geophysical Research: Atmospheres, 109, n/a–n/a, doi:10.1029/2004JD005182, http://dx.doi.org/10.1029/2004JD005182, d24108, 2004.

Sheffield, J., Wood, E. F., Chaney, N., Guan, K., Sadri, S., Yuan, X., Olang, L., Amani, A., Ali, A., Demuth, S., et al.: A drought monitoring and forecasting system for sub-Sahara African water resources and food security, Bulletin of the American Meteorological Society, 95, 861–882, 2014.

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., Rippey, B., Tinker, R., Palecki, M., Stooksbury, D., et al.: The drought monitor, Bulletin of the American Meteorological Society, 83, 1181–1190, 2002.

Vereecken, H., Schnepf, A., Hopmans, J., Javaux, M., Or, D., Roose, T., Vanderborght, J., Young, M., Amelung, W., Aitkenhead, M., et al.: Modeling soil processes: Review, key challenges, and new perspectives, Vadose zone journal, 15, 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 Resources Research, 50, 7505–7514, http://centaur.reading.ac.uk/38682/, 2014.

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

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-188, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 18 April 2017

Wigneron, J.-P., Kerr, Y., Waldteufel, P., Saleh, K., Escorihuela, M.-J., Richaume, P., Ferrazzoli, P., De Rosnay, P., Gurney, R., Calvet, J.-C., et al.: L-band microwave emission of the biosphere (L-MEB) model: Description and calibration against experimental data sets over crop fields, Remote Sensing of Environment, 107, 639–655, 2007.

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

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-188, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 18 April 2017

Hydrology and Earth System Sciences
Open Access
Discussions

Table 1. CLM soil layer depths and relative layer thickness in respect to 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

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-188, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 18 April 2017

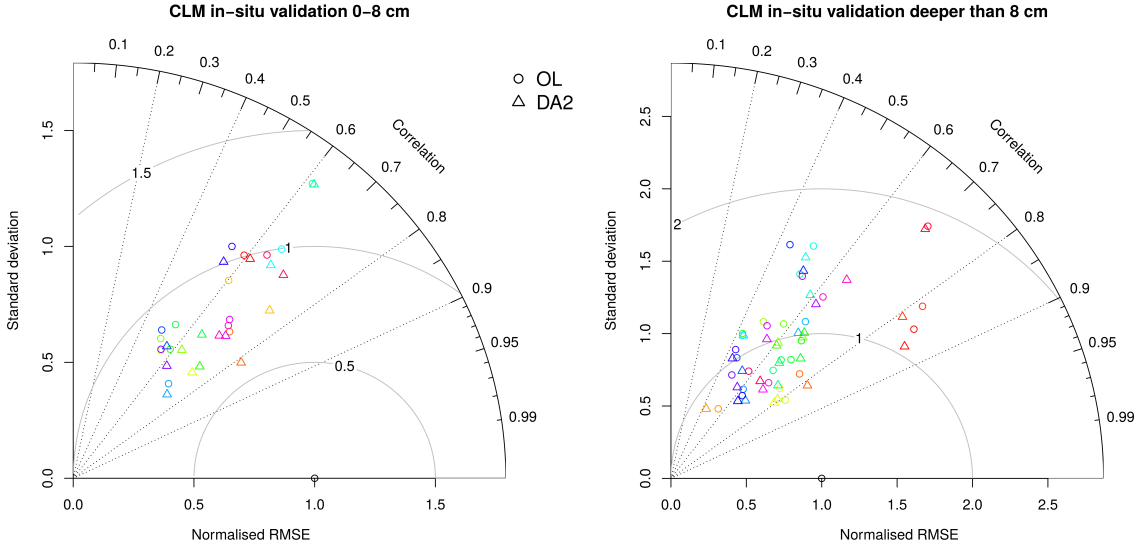

Figure 1. Taylor diagram showing assimilation impact on top layer soil moisture, defined as 8 cm soil depth, (left) and lower level soil moisture (right) in terms of correlation coefficient R, standard deviation and normalised RMSE. The colours correspond to the in-situ measurement sites but are not comparable between both panels.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-188, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 18 April 2017

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

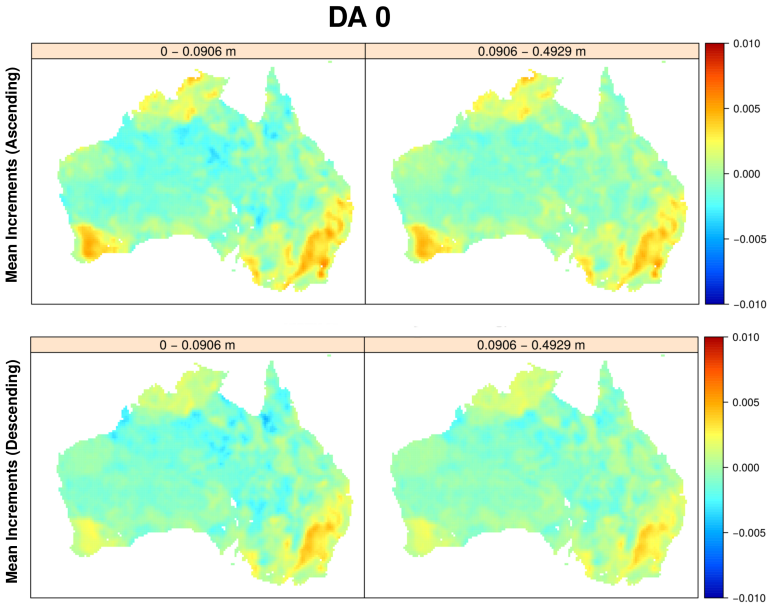

Figure 2. Mean of all increments for experiment DA 0 for top-layer soil moisture (left) and root zone soil moisture (right) for ascending (above) and descending orbits (below).

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-188, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 18 April 2017

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

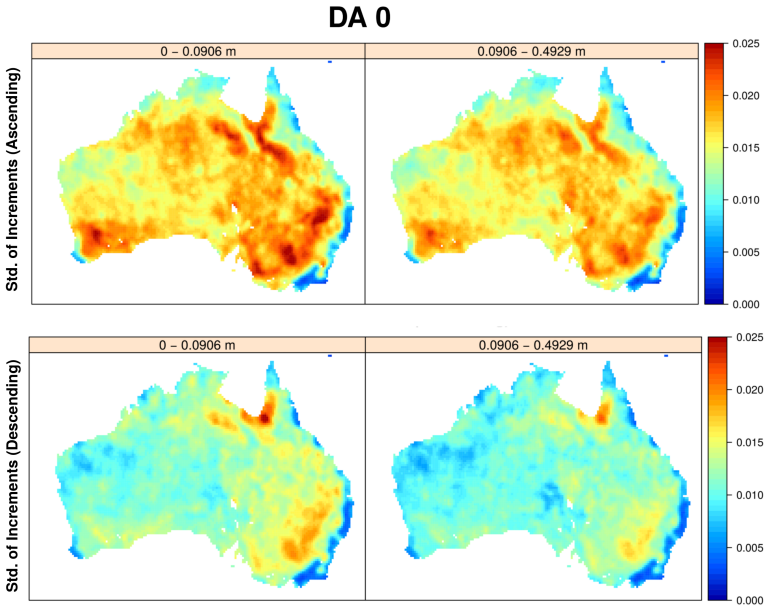

Figure 3. Standard deviation of all increments for experiment DA 0 for top-layer soil moisture (left) and root-zone soil moisture (right) for ascending (above) and descending orbits (below). Increments for the root-zone soil moisture are fairly similar to the top soil layers, due to the homogeneous texture perturbations applied across all layers.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-188, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 18 April 2017

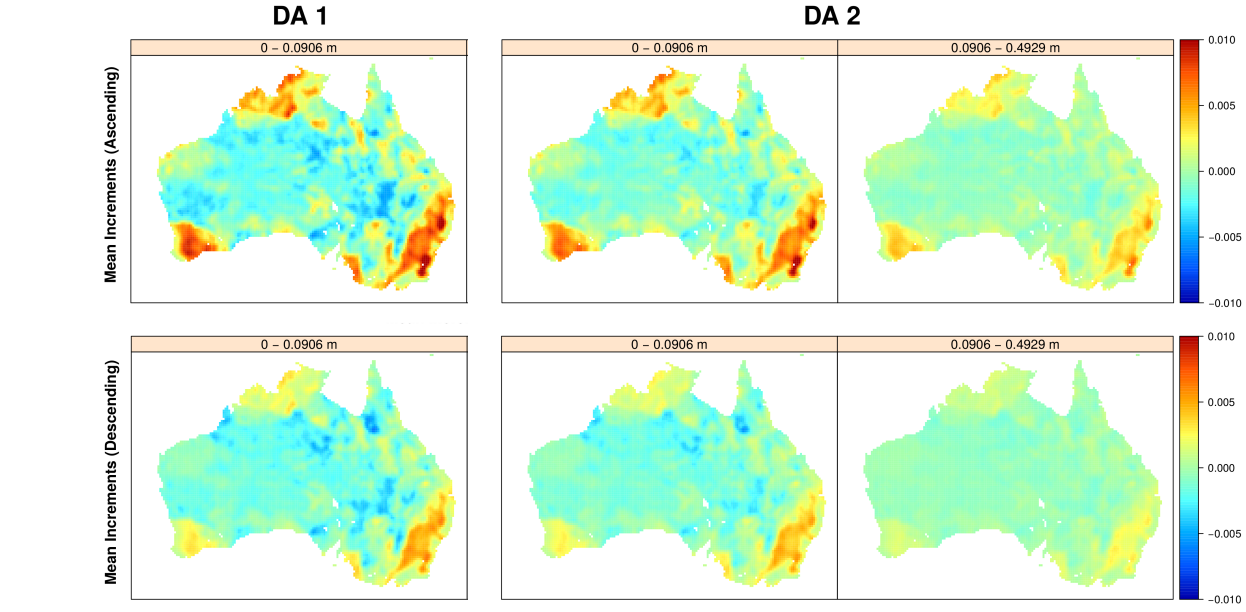

Figure 4. Mean of all increments for experiment DA 1 / DA 2 for top-layer soil moisture and root zone soil moisture for ascending (above) and descending orbits (below). Biases are strongest for the ascending orbit and distinctive spatial patterns are visible. Biases are strongly reduced both for deeper layers and descending orbits.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-188, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 18 April 2017

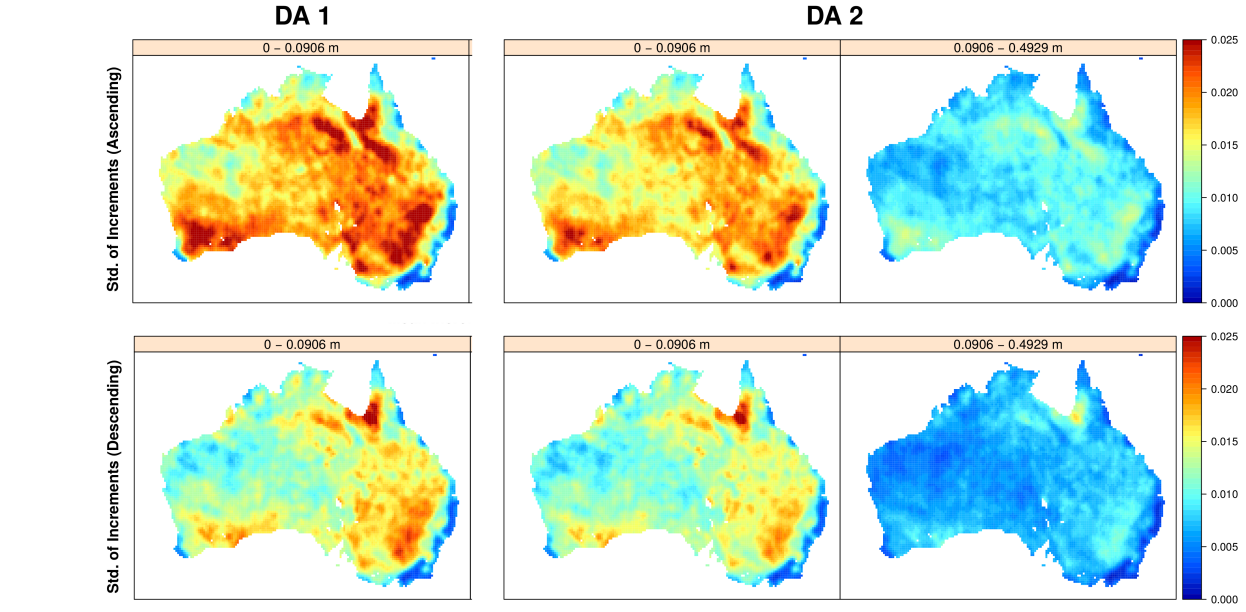

Figure 5. Standard deviation of all increments for experiment DA 1 / DA 2 for top-layer soil moisture and root zone soil moisture for ascending (above) and descending orbits (below). Increments are strongest for the ascending orbit and for top-layer soil moisture and even stronger when restricting assimilation 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.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-188, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 18 April 2017

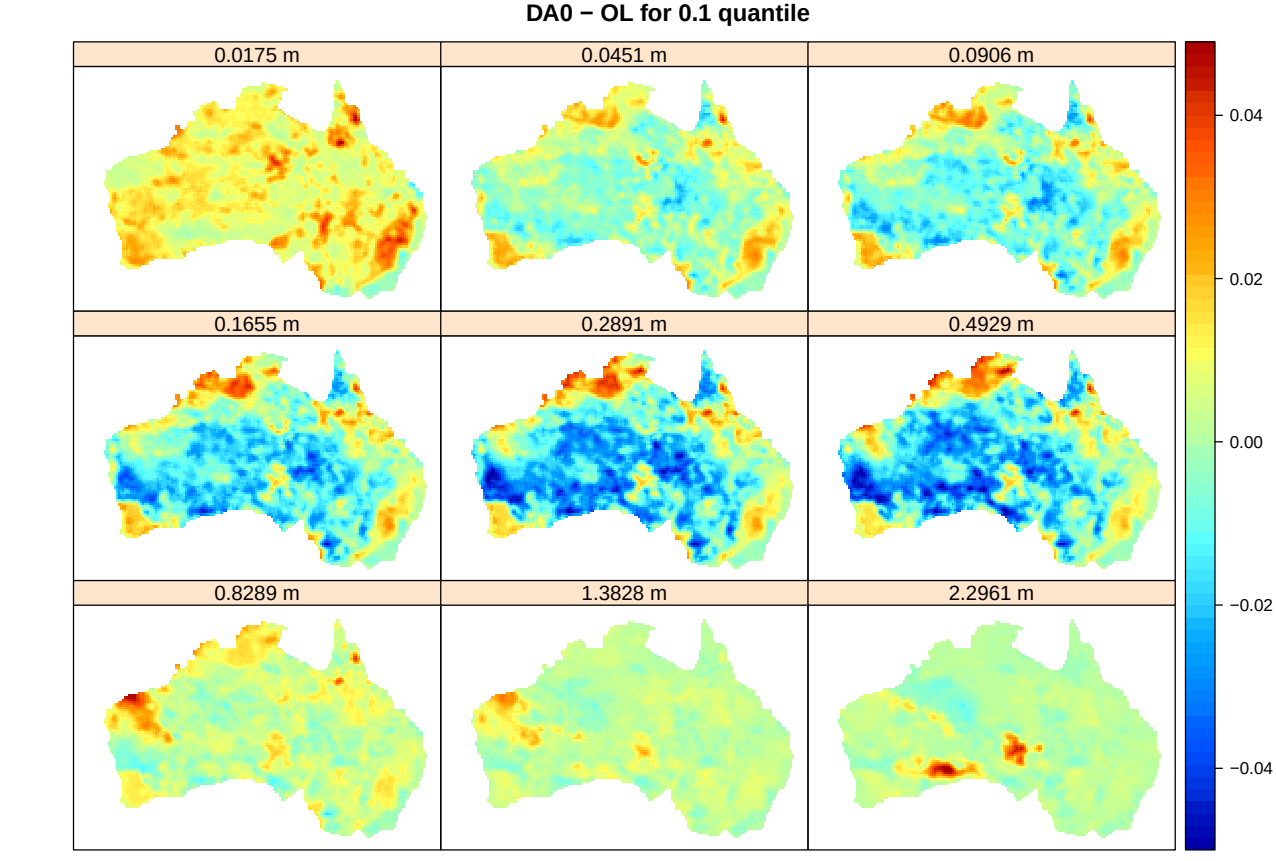

Figure 6. Differences in relative soil moisture [%/100] between open loop and DA0 experiment (DA 0 - OL) for the 10 % quantile. The individual panels correspond to the top 9 CLM soil layers.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-188, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 18 April 2017

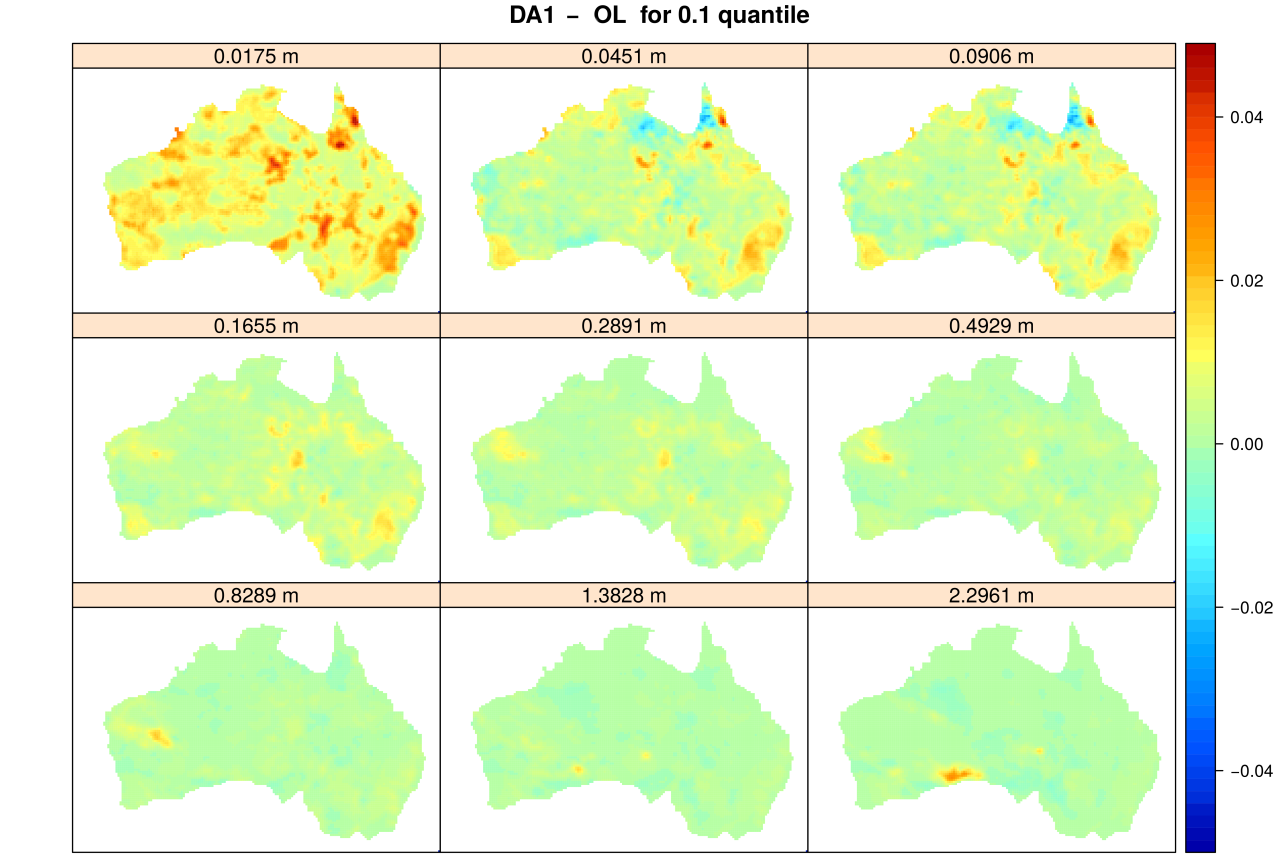

Figure 7. Differences in relative soil moisture [%/100] between open loop and DA1 experiment (DA 1 - OL) for 10 % quantile. The individual panels correspond to the top 9 CLM soil layers titled with their depth.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-188, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 18 April 2017

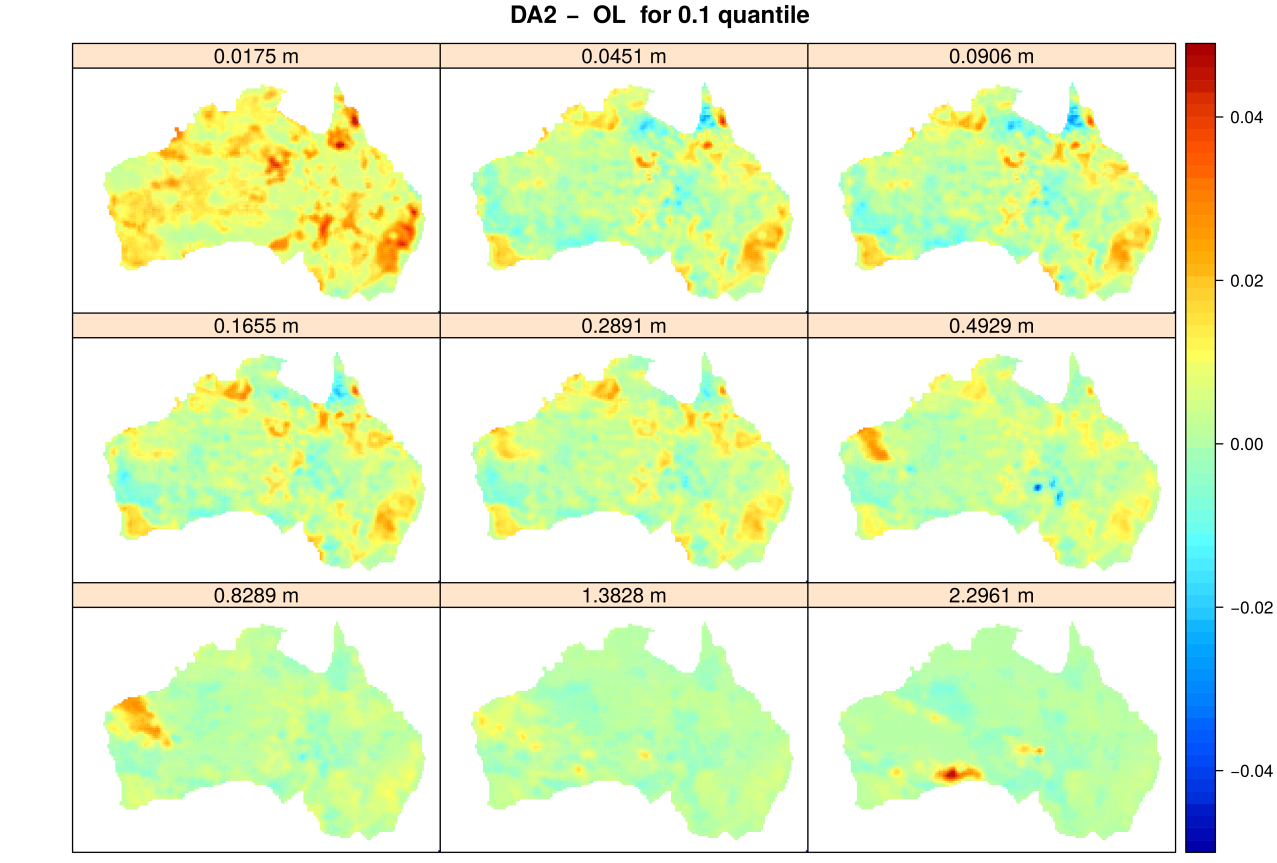

Figure 8. Differences in relative soil moisture [%/100] between open loop and DA2 experiment (DA2 - OL) for the 10 % quantile. The individual panels correspond to the top 9 CLM soil layers.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-188, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 18 April 2017

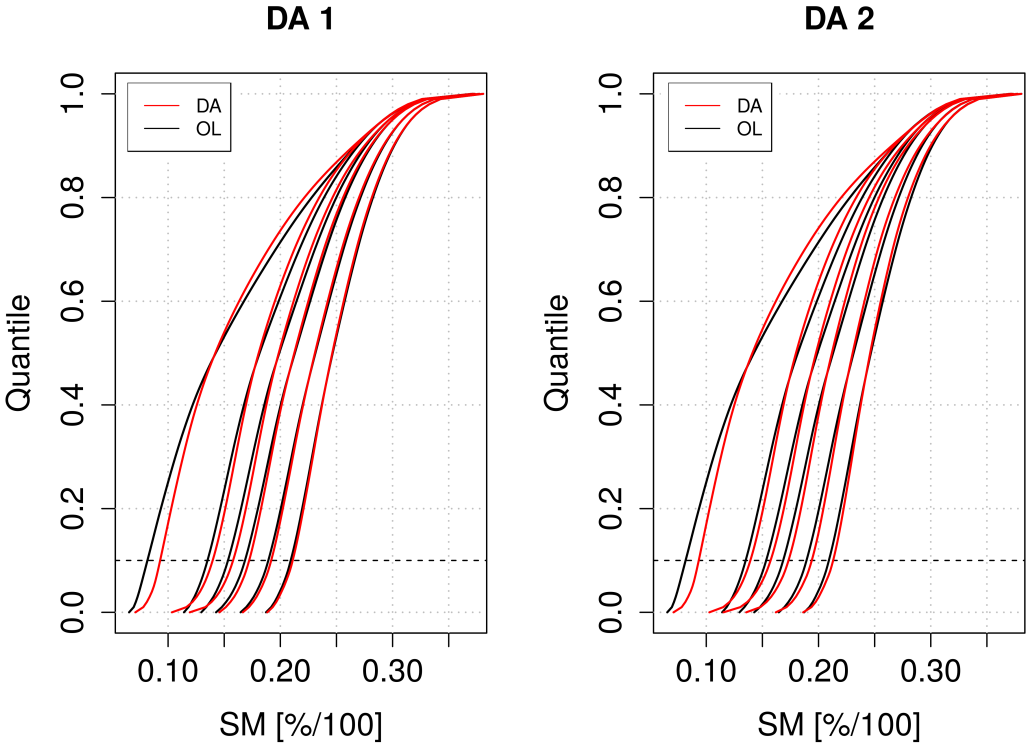

Figure 9. Cumulative distribution functions (CDFs) for the upper 6 CLM soil layers for experiments DA1 and DA2, based on quantiles computed for all data across the model domain. CDFs for open loop simulations are shown in black and assimilation results in red. Both panels show changes in CDF behaviour for the layers being updated in the respective experiments, i.e. layers 1-3 for DA1 and layers 1-6 for DA2. 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 figures 7 and 8.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-188, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 18 April 2017

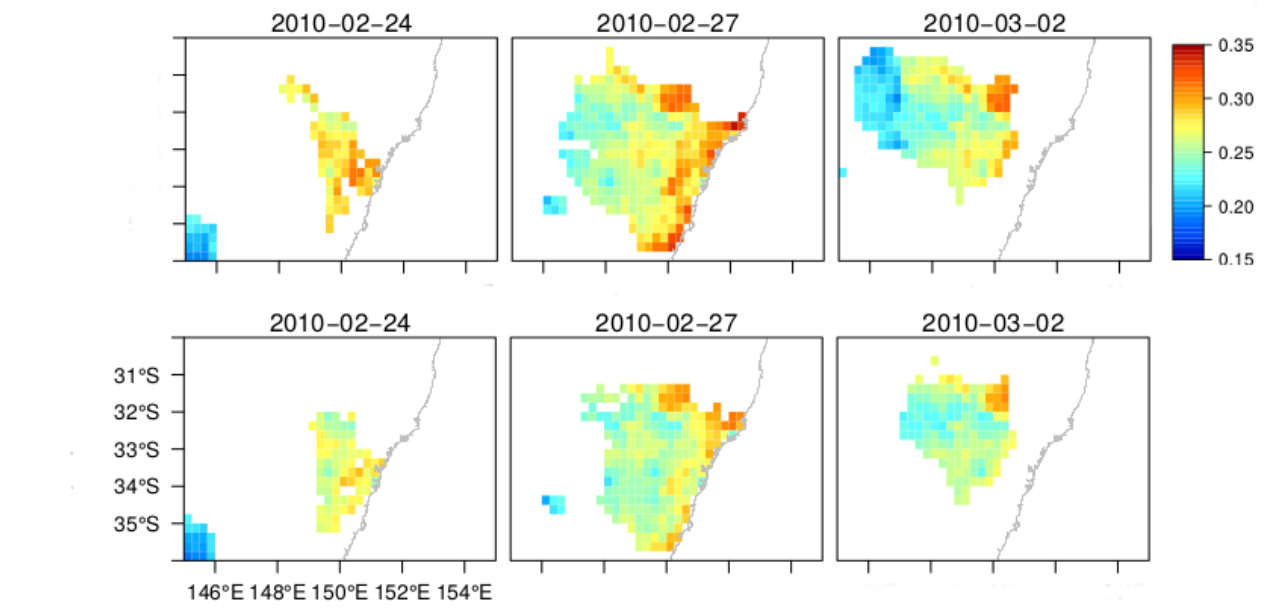

Figure 10. Sample drought event for February 2010, showing only the root zone soil moisture below the 10 % quantile level for the open loop (above) and experiment DA2 (below). The different spatial extent and differences in soil moisture itself, depending on the dataset used, at three different days are clearly visible.