Responses to comments from Reviewer 1

We thank the referee for the helpful comments. Due to the numerous changes in the manuscript, the new version is uploaded as a separate file.

Major Comments

This paper presents a data assimilation (DA) study where the SMOS brightness temperature is assimilated into the CLM model, forced with ERA-Interim surface meteorological fields, over the Australia area. The CMEM model is taken as the observation operator to simulate the 42.5 o incidence angle brightness temperature in H polarization and the LETKF algorithm from the DasPy package is used to perform the filter update.

The model ensemble is generated by perturbing both model parameters and forcing inputs. Three sets of DA experiments are carried out (DA1, DA2, DA0) with different numbers of soil layers included in the filter update and different ways to perturb the soil parameters. The filter updates are performed over brightness temperature anomalies (with seasonal cycle removed), which is different from most other studies. CDF matching is performed on the anomalies. Validations are carried out against ISMN in-situ observations. The results and analysis are focused on the soil moisture increments during the filter update and low soil moisture quantiles

This is a very carefully designed and carried out data assimilation study with its main novelty in assimilating brightness temperature anomalies. The investigation and results are significant and the quality of both the research and its presentation is very good – I see no major issues with the choices of the processing methods along the entire chain of DA procedures. The DA improvement, as measured by soil moisture skills (against ISMN), is reported as moderate, which is consistent with similar studies. The discussions are relatively weak, especially on the effects of DA at different temporal scales. Draper and Reichle, 2015 decomposes the soil moisture time series into dynamics at different time scales (long-term, seasonal, and short-term) for the analysis. It is not exactly clear how (and why) the anomaly assimilation (which has the seasonal signals removed) changes the way the DA behaves at seasonal to longer time scales. Some time series plots and related analysis are needed to help on this. Also, the study area is very large and heterogeneous in terms of soil and vegetation – should there be any stratification on the analysis of the results, e.g., statistics over different types of soil/vegetation? I think the paper can be published in HESS with minor revisions.

We have discussed this quite a bit and in the end did not include details on the temporal effects of the data assimilation. The paper is already quite extensive. However, we can gladly do this if the referee further suggest to do so. In that case we would suggest something on the increments, as shown in the two example figures below. We did not include any time series of soil moisture itself, since this would be for one specific location and thus not convey too much information. A land cover map has been added to the publication and patterns seen in the increments and quantiles are related to some features in Australia.

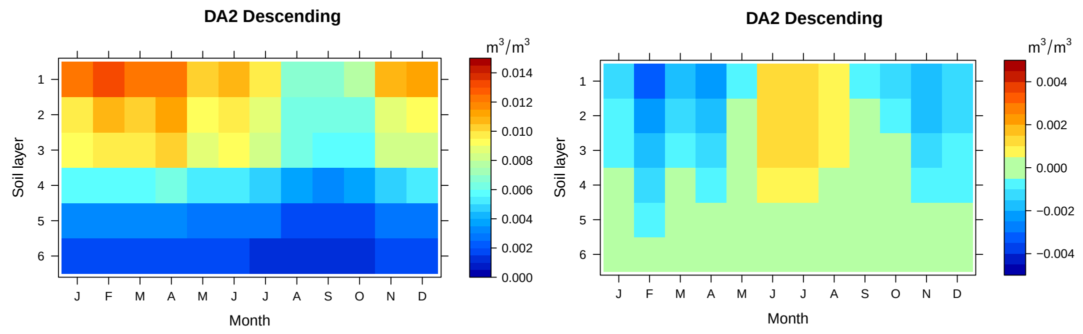

Figure 1. Standard deviation of increments (left) and increment bias (right).

Addition from 29th August:
We have included a land cover map based on the MODIS plant functional types used for the CLM model. Some of the results and patterns are set into relationship with land cover / vegetation. We have further added a map showing the seasonal differences in the assimilation impact.

Details: Page 9, line 6-7: the unites for observation errors are confusing – should they all be K^2 if they are all variances? Or they should all be in K if they are the standard deviation? My guess is that they are all in K because $4^2 + 3^2 = 5^2$.
Thank you for pointing this out. Yes, we have clarified this.

Figures 2, 4, 6, 7, 8: Maps here contain both negative and positive values and the sign of the data also matters. So it'll be much easier for the readers if a particular color (e.g. white) is used for the 0 values and two different sets of color shades (e.g. one set of warm shades and one set of cool shades) are used for positive and negative values. We have adjusted the colourbars and expanded the mid green as the neutral zone.

Figures 6, 7, 8: What is [%/100]? Should it be just [%]? Change "0.1 quantile" to "10 % quantile".
We have changed the figures accordingly.

Responses to comments from Reviewer 2

We thank Luigi Renzullo for the very in-depth review, providing feedback both on major points as well as numerous details. For clarity our responses are added to his original comments and highlighted in colour.

The changes to the manuscript are attached in an extra file, since changes are pretty extensive.

GENERAL COMMENTS

The paper explores a very interesting idea of assimilating brightness temperature observations, as opposed to derived soil moisture products, into a land surface model. However, no compelling argument is provide as to why this might be 'better' than assimilating the derived soil moisture product. Nor is there any real examination to the improvements, or otherwise, to the model performance. Most troubling however is the lack of evaluation against local information about the continental water balance to see if the patterns the authors have identified may be corroborated with either independent data or research. Why choose Australia as a case study but ignore the very many articles about data assimilation for water balance over the country? More detail critique is provided in the following. I recommend major revision and another review.

We have reworked the relevant text passages within the introduction, highlighting the advantages of the assimilation of brightness temperature as opposed to soil moisture products. The CLM land surface model used in this study provides all necessary dynamic information required for the brightness temperature forward simulations, e.g. soil temperature, vegetation temperature etc. Furthermore, the static surface datasets within of the model are also used within the forward simulations. In contrast to this, soil moisture products are based on retrievals using output and surface datasets from other models, thus introducing inconsistencies. However, brightness temperature assimilation does have its own issues, e.g. shortcomings in the forward simulations and biases between simulated and observed brightness temperatures.

We have compared the modelled soil moisture to in-situ stations over Australia and evaluated the assimilation performance in terms of the correlation coefficient R and the Root Mean Square Error (RMSE), as done in other studies. The validation has been expanded, giving some more detail on the in-situ measurements and citing the relevant publications. We have included maps showing the Murrumbidgee basin where most of the in-situ measurements are located as well as the Australian land cover classification used by the model. Spatial patterns for the increments and also quantile evaluation are put into context of the Australian landscape.

We have added some lines in the introduction about why Australia was chosen as a study area and highlighted the main soil moisture assimilation studies that we have found.

MAJOR ITEMS: * Most concerning is that there appears to be no interest in gaining new insight about Australian hydrology or indeed assessing the validity of the model estimates beyond soil moisture. No papers referencing Australian sources on the continent's hydrology or water cycle, so how do you know if the results are any good. There are clearly patterns in the results that may or may not be known to Australian research community. A simple first check is to see if the results accord with those from http://www.bom.gov.au/water/landscape/. The limited evaluation against in situ data in the Murrumbidgee catchment is weak, including no mention of the sites locations, the depth measublue, nor what is measublue (e.g. volumetric water content or wetness, neutron count etc).

Please see above response. Further we would like to add that the study strongly focuses on soil moisture, and therefore the wider Australian water balance has not been discussed. We have compared the spatial patterns of the soil moisture simulations to the above link and now mention this. A quick check did reveal that for instance evaporation fluxes change by roughly 5 %,

* Should mention in the introduction that while L-band on SMOS may be the first 'ded- icated' mission for soil moisture, there has been a long history of data assimilation development in C-band soil moisture retrievals (AMSR-E,-2 and ASCAT

for example) and SMAP is yet another L-band mission that is providing global coverage. Moreover you would be wise to cite work from research who have performed assimilation with an Australian focus (you are not the first) and you cannot ignore the rich legacy of work conducted in understanding Australian hydrology.

We have included example studies on the use of ASCAT and AMSR-E in the introduction as well as studies focusing on Australia.

* Simulations appear to be made for layers 0-9 cm, however the L-band sees emissions from at (at best) 0-5cm. Comment on this disparity and the impact, if any, on simulated brightness temperatures.

For all experiments the forward simulations use model output from the 10 CLM layers. These reach far deeper than where L-band emissions mostly originate from, as stated up to roughly 5 cm. The forward operator accounts for this and the simulations therefore are also only sensitive to model output of 0-5cm. We have clarified this within the description of the experiments.

SPECIFIC ITEMS: P1,L17: Change to ' . . . sensitive to 1.4 GHz electromagnetic emissions, measures . . . multi angular top-of-atmosphere . . .'
We have changed this.

P1,L18: Delete 'influenced by, among others, surface soil moisture.' Sentence is too long and 'among others' doesn't make sense. Among other what?
We have shortened the sentence accordingly.

P1,L22: I suggest including the key reference by Kumar et al. 2009 in the list which describes the mechanisms how top layer soil moisture assimilation can improve root-zone estimates in LSM's. [Kumar, S. V., Reichle, R. H., Koster, R. D., Crow, W. T., & Peters-Lidard, C. D. (2009). Role of Subsurface Physics in the Assimilation of Surface Soil Moisture Observations. Journal of Hydrometeorology, 10(6), 1534–1547. https://doi.org/10.1175/2009JHM1134.1]
The citation has been included.

P2,L11: Suggest rewording the sentence to: ". . . retrievals represent the optimum fits between simulated brightness temperatures and the . . . "
The sentence has been changed.

P2,L19: Modify: "Sources of uncertainty include atmospheric forcing, . . ."
We have changed this.

P2,L32-33: Not clear what is meant here. Elaborate on the link between brightness temperature and 'qualitative' models. Why would this even be a consideration?
We have removed the term 'qualitative' as we agree it was not clear. We simply meant that brightness temperature assimilation should be tested with different land surface models

P3,L2: "Within this" should be the start of a new paragraph.
A new paragraph has been inserted.

P3,L2-35: Very lengthy introduction to what this paper is about. Strongly suggest restructuring to be more clear in the lead up to Section 2 about what the objectives of this paper are. The paragraph should start: "In this paper we . . " and itemise the objectives. This will help the reader link the findings with the objective of the work.
We have revised the introduction and moved some parts to the conclusion part of the paper (relevance of findings for drought monitoring systems, ellaborations on the use of CDFs and quantiles for extreme event characterisations in a shortened form).

P3,L19: Why Australia? This needs to be clearly articulated.
We have added some sentences on the motivation of choosing Australia as a study area.

P3,L26-28: If the results will be evaluated over Australia then you should cite "Smith et al. (2012)". Yes, these data are part of ISMN, but should cite the official source. [Smith, A. B., Walker, J. P., Western, A. W., Young, R. I., Ellett, K. M., Pipunic, R. C., ... Richter, H. (2012). The Murrumbidgee soil moisture monitoring network data set. Water Resources Research, 48(7), 1–6. https://doi.org/10.1029/2012WR011976]
The citation has been included.

P3,L33: Change "avoiding to large" to "minimising the impact of potential large".
The sentence has been altered.

P4,L6: So did you use coupled or uncoupled mode? Why mention these if you're not going to specify here why.
We do not mention the coupled mode anymore

P4,L9: What are these more recent higher resolution data sets? Explain that these will be described in Section 2.1 and 2.2. Why 0.25 degree? What is CLM normally run at? 0.25 degrees is quite coarse for continental studies, makes me think why not extend to the whole world. That way t=you can use the whole ISMN and not just the tiny little southeast corner of Australia?
We have included the reference to the relevant section. The 0.25 degree resolution matches the SMOS observatiosn well, which is now stated in the text. CLM itself can be run at many resolutions, although usually coupled global simulations are quite a bit coarser than 0.25 degrees, e.g. 0.5 or 0.75 degrees. The motivation of using Australia has been included, see above.

P4,L13-32: How do you know if the derived surface information is accurate for Australia? What local information/expertise have you consulted? There is A LOT of research work (none of which are cite here) that shows these MODIS products are not representative of truth in Australia, (let alone the soils information). I would accept that a global study may use inferior information because it is the only data available with global coverage, but because this investigation focuses on Australia, it must be addressed! If accuracy is not an issue for this investigation (because assimilation compensates for the model deficiencies, including parameterisation) than you should state it explicitly here.
We agree we should have consulted more local expertise. However, despite focusing on Australia we did have future global applications in mind choosing the datasets. We have included a study on the validation of LAI within the Murrumbidgee area in the Assimilation and results section, linking visible patterns to possible LAI errors. Selecting MODIS data was also motivated by the fact that a clear rational exists in using these data as CLM Plant Functional Types and LAI values. The soil data used incorporates local information, albeit into a global product. Concerning the forcing datasets, we have consulted local expertise but did not find forcings at the requried spatial and temporal resolution.

P5,L26: Change to '. . . allows the coupling of different . . .'
We have changed this.

P6,L15: Is that "K ensembles" or "K ensemble members"? Clarify.
Changed to "K ensemble members."

P6,L22: I recommend "mapped" instead of "propagated". Propagated is only relevant to mapping through time (or space).
We have changed this.

P7,L29: The UTC to local time conversion may work for eastern Australia but not central or western Australia. How big an impact do you think a 2 hour error in timing will make on simulations?
We have included that the assumed error is justifiable, since the 2 hour mismatch is smaller than the temporal resolution of the forcings. We do agree the approach is not optimal.

P8,L16-17: Modelled brightness temperature can be extremely sensitive to choices in h, the roughness parameter. How have you dealt with this? Perhaps through the bias correction? Explain.

We have included that the roughness parameter is important, but calibrating the forward simulations towards the observations might actually deteriorate the sensitivity towards soil moiture. We therefore keep the original parameters for good variability and remove the bias through CDF-matching.

P8,L24: Is RFI an issue over Australia? If so, where will it be most likely. If not, then say so.
Australia is largely unaffected by RFI, we have added this information

P10,L2-3: You need to be more specific. These are the OzNet network in the southeast of Australian in a catchment called the Murrumbidgee, I presume. If so, confirm and cite the relevant work (Smith et al, 2012). If not, then you need to explaining where the in situ data are located, how deep they measure, etc.
We have added the reference and also added a brief description of the sites location and depths. The OzNet sites location is now shown in a map.

P10,L18-19: I would have thought the innovations would be close to zero on average (in fact that is one of the tests to see if your filter is operating optimally). Do you mean innovation or increment? Clarify. Also, what are the units on the increment? They appear to highlight dryland agricultural areas, e.g. western Australian wheat belt. Can you comment on the patterns and their connection to the surface parameterisation?
We argue that for most parts the increments are close to zero but that deviations do exist. They are given in vol % soil moisture, which has been added to the graphics. Within the text we now refer to the possible error in the LAI values or other possible reasons, such as irrigation for limited areas.

P11,L19: Please comment on the strong positive features in Fig. 6 in the 0.8 - 2.3 m layers. They are clearly linked to features in the landscape. What can you say about them?
We have linked the patterns to Lake Eyre and the Nullarbor plain, they are the result of strong increments accumulating in the lower layers. The Nullarbor plain for instance is very dry, and adding water in the deep layers with low temporal variability will lead to strong quantile changes.

P11,L34-P12,L1: A more relevant way to "place these findings in the context of" hydrological monitoring systems is to compare with actual modelling system output. A simple web search shows you can gain a lot of information about water balance in Australia from http://www.bom.gov.au/water/landscape/ I strongly urge you to consider locally relevant information to assess your results.
We have included the site, stating that CLM output was compared to the AWRA-L simulations to check for consistency.

P12,L10-13: How do you know it was a drought event? What other independent corroborating evidence supports this?
We now refer to the event as being relatively dry, it is to highlight the influence of quantile changes without making any quantitative evaluations on real droughts. This could be interesting for future research.

P12,L23-24: You should mention the coupling to CMEM, as CLM does not estimate brightness temperatures.
We now mention CMEM.

P13,L5-7: Agree with revisiting the use of LAI climatology. Recommend further than you examine the usefulness in Australia.
Quite likely a next study, restricted to a more local area. We will consult local expertise upfront.

Figure 1: Why cant the two panels be compablue? They should be able to be compablue. The point need to be identified, otherwise why have them a s separate shapes and colours?

This has been corrected for.

Figure2 2-8,10: Why no label on the colour bar? Insert units. ('Unitless' is acceptable)
Done

Figure 10: Where are we looking. Consider a location diagram/inset or mention: "central coast of New South Wales." for example.
The proposed text has been added.

Responses to comments from reviewer 3

We thank the referee for his comments and hope to have answered them as best possible. Since the changes to the manuscript are extensive, the new version is uploaded as a separate file.

This study investigated the benefit of integrating SMOS brightness temperature and the Community Land Model over Australia. Three different scenarios were performed to update different layers of soil moisture by the LETKF method. The results were evaluated using ground soil moisture measurements. Personally, I think this paper was well written. The organization was reasonable and the experimental design was clear. However, there were still some major issues need to be addressed before it can be considered for publication. A more systematic literature review on remote sensing data - land surface model assimilation need to be conducted. There are two groups of remotely sensed soil moisture (or brightness temperature) assimilation studies, one for soil moisture estimation typically through land surface models, and the other for runof and streamflow prediction normally through catchment hydrologic models. The current introduction mixed these two together, with a lack of detailed review on remote sensing constrained land surface modelling. The contribution of this study should be better articulated based on the review of the current progress on this topic. The authors discussed extensively on bias issue in the Introduction and Results sections, which I agree is an important issue; however, I did not see what is new in this study in addressing this issue. The CDF matching is a traditional approach with the advantage of removing relative bias. However, the problem is that it does not estimate and disaggregate the relative bias into model one and observational one. I did not see how this study addressed this issue. The design of the different DA experiments were not well justified. Technically, there is no problem to update all soil moisture layers through cross covariance, which should maximize the benefit of assimilating remotely sensed surface soil moisture by addressing the gross error accumulated in the deep soil moisture. So what was the point of just update the first 9 cm? It may be argued that updating only surface soil moisture could test the ability of the CLM to update the deep soil moisture by the model dynamics itself; however, I do not think a Kalman filter is the best choice to answer this question. The error in deep soil moisture is an accumulation of the error from the surface soil moisture and a smoother to assimilate the RS data to update both current and past surface soil moisture will have a better capacity on testing the capability of the model to update deep soil moisture through model dynamics. Besides, more in-depth analysis and discussions need to be added. For instance, what is the implication of the results from this study on the issues such bias? Whether the results is reasonable (and being improved after data assimilation) for the whole Australia? Also, I would suggest the authors to be careful in using the words "assimilate" and "update". It should be very clear through the paper that RS surface soil moisture was "assimilated" while different layers in the model were "updated". P2L31: Based on the review above, I cannot get to the conclusion that TB assimilation is under researched compared with soil moisture retrieval assimilation. P7L10-15: Why 32 ensembles? Why no spatial correlation was considered while most of the errors are known to be spatially correlated? How these error parameters are estimated/determined? 50% of rainfall is a lot, I reckon. P10L1-5: A bit of details on the soil moisture measurements quality control.

Literature review: We have added some assimilation case studies, some specifically focusing on the study area. Also, while referencing these we now mention the model that has been used. We agree that catchment models vs land surface models can be quite different but don't see a big problem in referencing these together, since the assimilation steps are usually quite similar.

CDF-matching: We agree that the introduction was too long (as also pointed out by referee 2). We have therefore shortened and streamlined it towards the objectives of the study. Quite some detail on observation rescaling has been removed, or when appropriate moved to other parts, since it might have caused the false impression of the issue being resolved within this study or that this study heavily focuses on it.

Experiment design: We hope to have clarified the design of the experiments within the introduction as well as the assimilation and results section. Concerning the udating into different layers, we argue that errors in the upper layers are actually best fed into the deeper layers through model physics. The experiment DA 2 however shows, that directly updating the root-zone leads to further improvements. Due to the high temporal variability of upper soil layers we believe the Kalman filter is the method of choice for close to surface soil moisture. We agree that for lower layers a smoother might be an interesting option, such as is used for the assimilation of GRACE data.

Analysis and Discussion: We have added more in-depth discussions on the in-situ validation, patterns in the increments as well as the quantile analysis. A land cover map as well as maps to show the in-situ validation within the Murrumbidgee catchment have been added. We have clarified that the consistent improvement of correlation with in-situ measurements makes us believe that the results are valid for all Australia, although the problem of sparse in-situ measurement sites remains.

Assimilate vs Update: We have substituted the wording "assimilate" with "updating" where appropriate.

We have changed the sentence stating that TB assimilation is under researched to that it is relatively new in practical terms.

Number of Ensembles: We have added that around 30 Ensembles is common for land data assimilation studies.

Spatial noise: The assimilation was performed in 1D, thus not requiring spatially correlated perturbations. This has been added in the text. The relevant references have been added from which the perturbation factors were taken, including the rainfall perturbations.

Quality control: We have removed the sentence, the quality control was carried out globally and actually no sites in Australia were affected.

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, Laboratory of Hydrology and Water Management, 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) ~~, improving soil moisture simulations over the Australian continent~~across Australia. 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~~ (11%) for upper soil layers if the root-zone is included in the updates. A ~~slightly reduced improvement is achieved when restricting the assimilation~~ reduced improvement of 5% is achieved if the assimilation is restricted to the upper soil layers. ~~Furthermore, the~~ 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 deviation are compared for the experiments. 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~~ assimilation-induced quantile changes. Although ~~now still limited~~still being limited now, 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 information derived from 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).~~ (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)~~. The on-board~~ , the second one being the similar SMAP mission launched in 2015 (Entekhabi et al., 2010b). The passive Imaging Radiometer with Aperture Synthesis (MIRAS) instrument ~~, sensitive at~~ on-board SMOS, sensitive to 1.4 GHz electromagnetic emissions, 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)~~(Reichle, 2008; Montzka et al., 2012), the assimilation impact in deeper layers will depend on model physics ~~(Montzka et al., 2011; Montaldo et al., 2001)~~(Montaldo et al., 2001; Kumar et al., 2009; Montzka et al., 2011). Alternatively, ~~deeper layers can be updated directly~~ by making use of one of the key advantages 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), deeper and unobserved layers can be updated directly. For plants, these deeper layers act as the root zone, ~~in which~~ 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. (2016a), showing that ~~an~~ the GLEAM evapotranspiration model can benefit from assimilating these data over Australia~~and Lievens et al. (2015b), concluding~~ , or Lievens et al. (2015b), who conclude that the positive assimilation impact on soil moisture can improve streamflow simulations for the VIC model, as shown ~~for~~ in the Murray-Darling basin. The impact on both streamflow and evaporation is evaluated by Ridler et al. (2014) for western Denmark. ~~More recently,~~ Leroux et al. (2016) assimilate SMOS soil moisture products into the DHSVM model, 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 ~~CO2~~ 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 ~~and Level 3~~ soil moisture retrievals ~~in essence solve minimisation problems~~ represent the optimum fits between simulated brightness temperatures and the observed satellite signal (Kerr et al., 2012). The simulated top of atmosphere signal ~~is thereby dependent~~ thereby depends 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~~ 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 thus~~ , it can be beneficiary to directly assimilate the brightness temperatures ~~as this allows for a consistent use of~~ in order to use consistent auxiliary information for the land surface model and the radiative transfer model~~(Han et al., 2013). Nevertheless, similar to when~~ . In the case of 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~~ , 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 ~~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.~~

Jia et al. (2009), Muñoz-Sabater (2015), De Lannoy and Reichle (2016a) and Lievens et al. (2016). 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~~ assimilating L-band brightness temperatures in practical terms is quite a new concept which still needs further exploring. The assimilation impact is ~~often evaluated by solely~~ mostly evaluated by 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 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~~Given the proven positive impact and the increased availability of longer time series of satellite observations, hydrological monitoring systems, such as ~~droughts.~~ 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 ~~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 biasescan 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~~ 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, ~~since absolute soil moisture values are difficult to compare between grid cells due to e. g.differences in land cover~~ which mostly use 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, 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 ~~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 droughtsor 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., 2 with all of them using soil moisture quantiles at grid cell level to characterise different levels of severity. With longer~~ south-west. The lack of large densely vegetated areas, which mask out the 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.~~ 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, which Australia is both 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 L-band specific studies have focused on Australia, covering soil moisture retrieval (Van der Schalie et al., 2015), assimilation studies (Lievens et al., 2015c), 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). The potential of AMSR-E soil moisture retrievals has been shown by Draper et al. (2009b). A comparison of SMOS satellite soil moisture retrievals with products based on other sensors is given by Su et al. (2013). The joint assimilation of ASCAT and AMSR-E data has been tested by Renzullo et al. (2014). More recently, SMOS soil moisture and GRACE water storage have been jointly assimilated by Tian et al. (2017). Downscaled AMSR-E soil moisture observations were assimilated within the Murrumbidgee basin by López López et al. (2016).

~~Within this study we assimilate 6 full years of SMOS brightness temperaturesat 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~~ 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 section 2.

In order to obtain the brightness temperature forward simulations, the CLM is 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; H~~ 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~~ quantile-mapping the entire observation anomaly time series towards the ~~offline computed~~ open-loop forward simulation anomalies at each grid point. ~~The assimilation impact is evaluated by comparing open loop and assimilation results to~~ Details on the implementation of the assimilation system, the forward simulations and the observation treatment will be given in section 3.

The 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~~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 ~~the~~ quantiles at 1 % steps are computed at each model grid-point, ~~enabling the sufficient~~ allowing a sufficiently precise empirical estimation of the ~~CDFs. As an example to highlight~~ cumulative distribution functions. To exemplify 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~~ 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 into relationship to the quantile analysis. The results of the experiments will be given in section 4, followed up by the discussion and conclusion in section 5.

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~~ offline with pre-computed 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~~ 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 is follows in the next section. The model resolution was defined at 0.25 ~~degree resolution~~degrees, which agrees well with the Level 3 observations provided in the EASE 25 km grid. The model is run at 30-minute time steps, with hourly outputs, allowing for ~~the~~ a 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 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~~ plant functional types (PFTs), again defined ~~over~~ by 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~~ 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~~ 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 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~~ who similarly replaced the CLM standard datasets.

2.2 ERA-Interim atmospheric forcing

CLM provides forcings (CRUNCEP) ~~not available for~~ which do not cover the required time period. ~~However, the rolling~~ Therefore, due to the release of ERA-Interim reanalysis data (Dee et al., 2011) with a ~~time lag~~ time-lag of only a few months~~enables the assimilation of relatively new satellite measurements, in this case SMOS, and thus ERA-Interim atmospheric~~ , these 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~~ With ERA-Interim ~~is~~ 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~~and~~ , thus 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~~ 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~~ with an appropriate interpolation algorithm applied to each variable. Incoming radiation is interpolated by 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~~ 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 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. (2014b).

DasPy has been developed with a focus on High-Performance Computing~~and parallelism~~ . 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~~ The original SMOS footprint is 43 km across and thus covers more than a single model grid cell, which would ~~encourage~~ justify the assimilation in 3D. However, mostly for reasons of simplicity~~and the already~~ , and also due to the previously performed inverse distance observation regridding ~~,~~ partially accounting for this, we ~~here only assimilate~~ only use observations directly covering a grid cell. Also, about 90 % of the signal observed by SMOS does originate from a footprint closely matching the model grid.

Mathematically, the LETKF can be described as follows~~.~~: Model states for each ensemble member k from a total of K ~~ensembles~~ ensemble members 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~~. This results 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~~ mapped 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~~ 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 ~~0.5. In order to~~ 0.3. The perturbation factors are the same as used by Reichle et al. (2007) and Han et al. (2014a). No spatially correlated noise was added, as the experiments are carried out in 1D, only using one observation per grid cell. To avoid ensemble collapse during dry periods~~also~~ , soil texture is also perturbed once at model startup~~with spatially correlated multiplicative noise and~~ . Here, multiplicative noise with 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 ,~~ is applied. For lower layers the top layer multiplicative factor is rescaled by using the inverse relationship between ~~each soil layers thickness~~ the thickness of each soil layer and the summed soil layer thickness of the two top layers (see Table 1). This ~~should insure~~ 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. ~~Because CLM derives~~ 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. ~~As stated, for demonstration purpose a third experiment (DA0), was performed, using homogeneous texture perturbations for all soil layers~~Concerning the number of ensembles, an amount of around 30 is common in brightness temperature assimilation studies and should allow sufficient error estimation.

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 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~~ assuming that a sufficiently correct temporal alignment between observations and model forward simulations ~~.~~ is provided. With the western parts of Australia deviating by 2 hours and the ERA-Interim forcings being interpolated from 3-hourly to 1-hourly data, we consider this approach to be acceptable.

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~~ (Champeaux et al., 2005) and the 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~~ 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. 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~~ 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 static bias between simulations and observations through observation rescaling.

3.4 Observations and anomaly ~~computationSMOS~~ 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 parametrization (Drusch et al., 2009; Barella-Ortiz et al., 2015), 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 ~~Donées~~ 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~~ 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 by using inverse-distance interpolation. ~~Based on these data~~

On the basis of 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~~ open-loop run with 32 ensembles are performed and the ensemble mean climatology is derived in the same way as the observation climatology. SMOS anomalies are then ~~quantile matched~~ 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 seasonal mean differences between simulations and observations and to remove the bias ~~, without requiring~~ 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 K^2 was defined, accounting for ~~4 K instrument error~~ a standard instrument error of 3 K and an assumed combined standard mean error of ~~3~~4 K for the forward simulations and representativeness error. ~~This~~ The instrument error can be seen as a low estimate and ~~was~~ is based on the assumption that the brightness temperature binning around the 42.5 degree incidence angle ~~should slightly reduce the instrument error.~~ 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~~and compared to the precomputed SMOS anomaly~~. The difference between ~~the two~~ this simulated anomaly and the SMOS anomaly is the innovation~~between model and observations in observation space and~~, 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 / descending orbits respectively.

4 Data assimilation and results

~~Two main data assimilation runs~~ In total, three assimilation experiments 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~~ , 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 ~~assimilation impact to an open loop run in respect to in-situ soil moisture observations. In-situ stations are available for~~ time series before and after assimilation to 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-situ measurements. In addition~~to the in-situ validation, quantile shifts in~~ , shifts in the soil moisture quantiles in respect to the ~~open loop~~ open-loop run are analysed ~~, highlighting the~~ to highlight some long-term effects of the data assimilation. A set of quantiles is computed at each grid cell ~~allowing for~~ to allow 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~~ Both the impact on correlation and the long-term effects are set into relationship 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 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, ~~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.~~ 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 section 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~~purposes the~~, hourly 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~~ ISMN (Dorigo et al., 2011). OzNet in-situ ~~measurement~~ 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 for 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 TDR-probes at shallow levels, mostly 5 cm or 8 cm, and at deeper layers, mostly 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 for 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 ~~is computed, using layer thickness as weights, prior to the comparison.When taking~~ 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 one year of data, not necessarily consecutive, correlations improve from 0.613 for the ~~open loop~~ open-loop run to 0.640, 0.678 and 0.681 for ~~DA1 and DA2 respectively~~ DA 1, DA 2 and DA 0 for top layer soil moisture (number of stations ~~,~~ n = 17). ~~Bottom layer~~ Root-zone soil moisture improvements are ~~lower~~ smaller, with average correlation coefficients of 0.626, 0.644 and 0.648 for ~~DA1 and DA2,~~ DA 1, DA 2 and DA 0 compared to 0.601 for the ~~open loop~~ open-loop run (n = ~~30). Upper soil moisture~~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 soil moisture behaviour thus improves by ~~also~~ additionally updating deeper layers~~and~~ , whereas deeper layer soil moisture is slightly enhanced through only ~~assimilating~~ updating top level soil moisture, with the assimilation effects only being ~~propagated~~ applied through model physics. ~~Overall~~All in all, updating the top ~~6~~ six CLM layers results in the largest improvements~~. Figure 3 shows~~ , 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 Figure 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 CLM layer thickness as weights. 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 ~~validation results of experiment DA2~~in-situ validation of experiment DA 2. As opposed to Figure 2, all original measurements are included with no vertical aggregation performed. The Taylor diagrams ~~further show~~ reveal a slightly decreased normalised standard deviation ~~of the assimilation results in respect to the in-situ measurements~~ when compared to the ~~open loop time seriesin respect to the in-situ data~~open-loop time series. In terms of standardised RMSE it is less conclusive, ~~and RMSE is slightly reduced by assimilation~~ with RMSE being slightly reduced for some stations and slightly increased for ~~some others~~others, but never significantly. These findings correspond well to ~~experiment DA~~ experiments DA 0 and DA 1 ~~, not shownhere~~(not shown).

When comparing the average changes in correlation for the non-vertically aggregated sites only for the ten used CosmOz sites, correlation increases by about 0.016 for DA1 (three sites with +0.03 and two very close to zero). For DA2 the average correlation increases by 0.02 (two sites showing +0.05 and one -0.02) and for DA3 by 0.13 (three sites slightly deteriorating). This highlights that the assimilation improvements are stronger for the OzNet sites located in the 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 DA3 in respect to DA2 which contradicts the findings when taking all measurements into account.

Altogether, the results demonstrate that the assimilation system has been sufficiently well designed to improve the modelling of soil moisture for 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 many-fold 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

~~In a bias-blind assimilation system it can be expected that , especially over longer time periods ,~~ Especially over long time periods the mean increments ~~should be~~ in a bias-blind assimilation system can be expected to be very close to zero. Figure 4 and Figure 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 ~~located in the North, South-West and South-East~~ visible in the north, south-west and south-east of Australia, seemingly being linked to the occurrence of vegetation ~~. Nevertheless, the positive bias does not~~ (see Figure 1). The areas in the south-west and south-east as well as in the north correspond well to the subtropical, temperature 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 increments bias or slightly negative values~~. These~~ , both in areas covered by mostly sparse vegetation and the inner drylands. For the descending orbits the patterns are still visible ~~for the descending orbit or the~~ but are weaker, both for top layers and the root-zone~~soil moisture but are on overall weaker, except for~~ . The exception is DA 0~~with~~ , 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 ~~DA1~~ DA 1, compared to DA 0 and ~~DA2, with the latter two updating both~~ DA 2, which both update top layer soil as well as ~~the root-zone. A possible interpretation might be that the assimilation into~~ root zone. The reason may be that updating deeper layers results in a ~~longer~~ more lasting effect, ~~potentially~~ moving the model closer to subsequent observations and thus reducing subsequent increments.

Figure 5 and 7 show the increment standard deviations~~, which are substantially stronger for the ascending orbit. Top~~ . 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 ~~larger~~ largest 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~~ Being the most dominant dynamic factor for the ensemble generation, precipitation ~~is the most dominant dynamic factor and leads to immediate increased ensemble spread , andthus a~~ 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. ~~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.~~ 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 ~~reduced for the root-zone~~ smaller than for the layers above (Figure 5). As a 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 ~~relationship between~~ behaviour of increment bias and increment standard deviation, it seems that there is ~~either a decreased increment bias at~~ 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.

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 The thermal equilibrium within the soils, which is more pronounced at 6 a.m. local time, reduces the error in the forward simulations. The mean bias between the forward simulations and observations is 5 K less for the ascending orbits, which supports this explanation.

To highlight some of the seasonal effects, Figure 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 ~~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 .~~ 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 Figure 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. (2016b) 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 dynamics which are not explicitly modelled. Kumar et al. (2015); Escorihuela and Quintana-Seguí (2016 De Lannoy and Reichle (2016b) have similarly reported on the potential of SMOS to observe irrigation.

4.3 Soil moisture quantiles

~~In addition to~~ 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~~ open-loop run. Although in principle the assimilation system should be designed bias-free with similar positive and negative increments, the previous section ~~showed~~ has revealed that small increment biases do exist, potentially ~~resulting in long term~~ 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 ~~9 CLM layersand to which extent it shifts for each grid cell when compared to the open loop run~~nine CLM layers. For experiment DA 1 (Figure 10), assimilation has ~~a small~~ an impact on the topmost soil layer with the quantile increasing by ~~up to 5~~ by a maximum of ca. 1 % for large areas and by up to 4 % for spatially very limited areas. Much smaller changes are visible for ~~deeper layers but with changes nevertheless reaching layers that are not directly updated.~~ the second and third layer, 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 three to six, 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 (Figure 11), ~~whilst also updating layers 3 - 6, slightly~~ with the root zone also being updated, 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,~~ most part the patterns reflect well the ~~patterns identified in Figure 6, showing the mean increments, although in deeper layers also some independent patterns emerge that cannot be directly explained by any increments bias. Figure~~ ones identified for DA 1. Figure 9 shows the impact on the 10 % quantile for ~~the~~ experiment DA 0 ~~, using~~ with homogeneous soil texture perturbations ~~across all layers. Here, much larger~~ being used. As expected, significant effects are visible especially within the ~~root-zone soil, increasing the quantiles~~ root zone, with quantiles being decreased 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.~~ 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 to dry out. This is especially visible for layers two and three, 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 Figure 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 ~~at~~ throughout the entire CDFs is shown in Figure 12. The continental average empirical cumulative distribution functions are plotted for ~~the~~ soil layers 1 - 6 for the ~~open loop~~ open-loop run as well as ~~DA1 and DA2. As shown for the 10 % quantile, lower~~ for DA 1 and DA 2. 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~~ although at extremely low levels the behaviour tends to reverses 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.

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~~. As stated,~~ , 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 to rescaling the observations to match the model or to 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 ~~these findings~~ quantile behaviour 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~~ short-term behaviour alone, i.e. ~~daily, of soil moisture and all connected fluxes is of importance for e.g. land-atmosphere feedbacks, but would have an negligible~~ hourly or daily, has a minor 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~~ , although large increments, ~~that would drastically change the soil moisture regime.Further, when classifying an event~~ e.g. due to corrected precipitation during a storm, can have an effect on the start and end point of an observed phenomenon, such as a drought. When classifying such an event defined at a specific quantile level, there will be a twofold impact from the assimilation~~, namely the change in soil moisture itself~~ : the changes in the quantile of interest 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~~ change in soil moisture itself. Here we highlight a sample ~~drought~~ dry event on the ~~East~~ east coast to show to ~~which extent it's~~ what extent its classification changes through the assimilation impact. Figure 13 ~~therefore shows root zone~~ shows root-zone soil moisture at or below the 10 % quantile level for the ~~open loop~~ 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 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 ~~visible~~ 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~~even when e. g. being affected by the same large scale precipitation event. An alternative to~~ . We thus conclude that despite having carried out the assimilation in 1D, spatially correlated noise is recommended for such applications. An alternative would be to further increase the ensemble size~~from 32 to a significantly higher number, which however will also require larger~~ , but at the expense of higher computational resources. Additionally, when trying to extract meaningful statistics on the occurrence of events, such as droughts, ~~by extracting these as clusters of grid cells over the spatial and temporal domain,~~ it might be ~~especially~~ particularly 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). We want to highlight the fact that the shown event is for demonstration purpose 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 ~~we~~ coupled to the Community Microwave Emission Model. We have substituted the surface datasets with higher resolution and more recent data. ~~Further~~Additionally, we have replaced the offline ~~forcing data with~~ forcings with the ERA-Interim reanalysis~~data~~. The assimilation over 6 full years, from 2010 – 2015, of SMOS brightness ~~temperatures~~ temperature anomalies with the LETKF improved soil moisture simulations when compared to in-situ measurements in the order of up to 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~~ 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 ~~larger areasor at global scale.CLM model physics alone did propagate assimilation effects into the root zone when restricting assimilation to~~ other areas.

In detail, three data assimilation experiments were carried out: Within the first experiment 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~~ 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 soil moisture measurements increased by 5 %, root-zone soil moisture increased by 4 %. Within the second experiment both top soil moisture and ~~root zone soil moisture.~~ 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 ~~greatly increases with depth which translates into identical increments in relative soil moisture being very different in~~ vastly increasing with depth homogeneous soil perturbations across all layers result in large deep layer updates 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.~~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 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 ~~vegetation. Although this might be due to a problem with the uncalibrated forward operator or with the fixed high vegetationclass LAI values within the forward operator, a further possible cause is~~ 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~~for CLM. The use of such data is common practise~~ , which is common practice within current land data assimilation systems. ~~However, due~~ 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~~. 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~~, where cloud cover permits. Climatological LAI might especially pose a problem for the monitoring of extreme events, such as droughts, since these ~~would likely~~ tend to 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~~ 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 is 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 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 -0.82 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 ~~, allowing for the consistent use of data across model and forward simulations, should be made~~ 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 has 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, with on average larger differences between models and retrievals. This further motivates 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 ~~. Within assimilation systems where the forward operator is calibrated, it is also likely that the problem of equifinality for~~ 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 hours, 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, here 30 minutes.

CLM uses fixed soil layer-depths which is likely the ~~forward operator parameters could thereby be reduced.~~ most beneficial model structure for comparing the assimilation effects spatially, since the covariance between observations and state variable 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, ~~although the very high quantiles remained unchanged, overall resulting in a~~ resulting in 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~~ 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~~was carried out showing much larger analysis biases~~ , the root-zone soil moisture showed a strong reduction compared to the ~~open loop~~ 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 ~~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~~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 parameterization of the forward operator is certainly not perfect. The spatio-temporal patterns linked to geographic areas or specific land cover classes, such as ~~the Soil Moisture Active Passive Mission (SMAP, Entekhabi et al. (2010a)). 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~~ 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~~. 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 systemscurrently rely on precipitation based indices, e. g. the Palmer Drought Severity Index (PSDI, Palmer (1965)), due to the high correlation~~, 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 sub-seasonal 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 ~~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~~ levels between grid cells. We have shown that the assimilation induced quantile changes will have an effect on the spatio-temporal 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). To model the complex feedback processes between soil moisture and vegetation is likely best performed using raw brightness temperatures and the therefore use of consistent data between 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.~~ model and the forward simulations.

~~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 moisturesimulations. 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~~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.

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). The authors thank everyone involved in the OzNET and CosmOZ soil moisture networks as well as the maintainers of the International Soil Moisture Network.

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, 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, 2002.

Brocca, L., Melone, F., Moramarco, T., Wagner, W., Naeimi, V., Bartalis, Z., and Hasenauer, S.: Improving runoff prediction through the assimilation of the ASCAT soil moisture product, Hydrology and Earth System Sciences, 14, 1881, 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 Transactions on Geoscience and Remote Sensing, 50, 2542–2555, 2012.

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, 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, 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, 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, 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, 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.

Dharssi, I., Bovis, K., Macpherson, B., and Jones, C.: Operational assimilation of ASCAT surface soil wetness at the Met Office, Hydrology and Earth System Sciences, 15, 2729–2746, 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, 2015.

Draper, C., Mahfouf, J.-F., and Walker, J.: An EKF assimilation of AMSR-E soil moisture into the ISBA land surface scheme, Journal of Geophysical Research: Atmospheres, 114, 2009a.

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, Hydrology and Earth System Sciences, 15, 3829, 2011.

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 Sensing of Environment, 113, 703–710, 2009b.

Draper, C. S., Reichle, R. H., De Lannoy, G. J. M., and Liu, Q.: Assimilation of passive and active microwave soil moisture retrievals, Geophysical Research Letters, 39, n/a–n/a, doi:10.1029/2011GL050655, http://dx.doi.org/10.1029/2011GL050655, l04401, 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, Journal of Hydrometeorology, 10, 213, 2009.

Dumedah, G., Walker, J. P., and Merlin, O.: Root-zone soil moisture estimation from assimilation of downscaled Soil Moisture and Ocean Salinity data, Advances in Water Resources, 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 Zyl, J. V.: The Soil Moisture Active Passive (SMAP) Mission, Proceedings of the IEEE, 98, 704–716, 2010a.

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., et al.: The soil moisture active passive (SMAP) mission, Proceedings of the IEEE, 98, 704–716, 2010b.

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

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, 2014a.

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, 2014b.

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, 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.

Hawdon, A., McJannet, D., and Wallace, J.: Calibration and correction procedures for cosmic-ray neutron soil moisture probes located across Australia, Water Resources Research, 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, 2014.

Herrera-Estrada, J. E., Satoh, Y., and Sheffield, J.: Spatiotemporal dynamics of global drought, Geophysical Research Letters, pp. n/a–n/a, 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.

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 Sensing of Environment, 186, 479 – 500, doi:http://dx.doi.org/10.1016/j.rse.2016.09.015, http://www.sciencedirect.com/science/article/pii/S0034425716303583, 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 Geoscience and Remote Sensing Letters, 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, Physica D: Nonlinear Phenomena, 230, 112 – 126, 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.

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, Australian Meteorological and Oceanographic Journal, 58, 233, 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, 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, 2001.

Kerr, Y. H., Waldteufel, P., Wigneron, J.-P., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M.-J., Font, J., Reul, N., Gruhier, C., et al.: The SMOS mission: New tool for monitoring key elements ofthe global water cycle, Proceedings of the 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 Transactions on Geoscience and Remote Sensing, 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., Peters-Lidard, C., Santanello, J., Reichle, R., Draper, C., Koster, R., Nearing, G., and Jasinski, M.: Evaluating the utility of satellite soil moisture retrievals over irrigated areas and the ability of land data assimilation methods to correct for unmodeled processes, Hydrology and Earth System Sciences, 19, 4463, 2015.

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, Journal of Hydrometeorology, 10, 1534–1547, 2009.

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 Resources Research, 53, 633–646, doi:10.1002/2015WR017738, http://dx.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 Sensing of Environment, 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, Hydrology and Earth System Sciences, 20, 2827–2840, 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, Journal of Hydrometeorology, 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 Sensing of Environment, 168, 146–162, 2015b.

Lievens, H., Tomer, S. K., Al Bitar, A., De Lannoy, G., Drusch, M., Dumedah, G., Franssen, H.-J. H., Kerr, Y., Martens, B., Pan, M., et al.: SMOS soil moisture assimilation for improved hydrologic simulation in the Murray Darling Basin, Australia, Remote Sensing of Environment, 168, 146–162, 2015c.

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, 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, Hydrology and Earth System Sciences, 20, 3059–3076, doi:10.5194/hess-20-3059-2016, https://www.hydrol-earth-syst-sci.net/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, International Journal of Applied Earth Observation and Geoinformation, 48, 146–162, 2016a.

Martens, B., Miralles, D., Lievens, H., Fernández-Prieto, D., and Verhoest, N. E.: Improving terrestrial evaporation estimates over continental Australia through assimilation of SMOS soil moisture, International Journal of Applied Earth Observation and Geoinformation, 48, 146–162, 2016b.

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, pp. 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 Sensing of Environment, 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 Transactions on Geoscience and Remote Sensing, 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 Transactions on Geoscience and Remote Sensing, 42, 773–785, 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, 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.

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 Transactions on Geoscience and Remote Sensing, 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 Transactions on Geoscience and Remote Sensing, 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, Journal of Geophysical Research: Atmospheres, 109, 2004.

Peischl, S., Walker, J., Allahmoradi, M., Barrett, D., Gurney, R., Kerr, Y., Kim, E., Le Marshall, J., Rüdiger, C., Ryu, D., et al.: Towards validation of SMOS using airborne and ground data over the Murrumbidgee catchment, in: Proc. MODSIM, pp. 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, Hydrology and Earth System Sciences, 16, 1697–1708, 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 Transactions on Geoscience and Remote Sensing, 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 Transactions on Geoscience and Remote Sensing, 49, 3156–3166, 2011.

Reichle, R. H.: Data assimilation methods in the Earth sciences, Advances in Water Resources, 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), Journal of Geophysical Research: Atmospheres, 112, 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, Journal of hydrology, 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 Resources Research, 50, 8962–8981, 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, 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, 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.

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 Resources Research, 48, 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 Sensing of Environment, 134, 1 – 11, 2013.

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.

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 Resources Research, 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 Sensing of Environment, 163, 70–79, 2015.

Van Dijk, A. and Renzullo, L. J.: Water resource monitoring systems and the role of satellite observations, Hydrology and Earth System Sciences, 15, 39–55, 2011.

Van Dijk, A., Bacon, D., Barratt, D., Crosbie, R., Daamen, C., Fitch, P., Frost, A., Guerschman, J., Henderson, B., King, E., et al.: Design and development of the Australian Water Resources Assessment system, in: Proceedings, Water Information Research and Development Alliance Science Symposium, 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 Resources Research, 49, 1040–1057, doi:10.1002/wrcr.20123, http://dx.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., et al.: The Australian Water Resource Assessment Modelling System (AWRA), in: 20th International Congress on Modelling and Simulation, 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., 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, 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.

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.

Yang, K., Watanabe, T., Koike, T., Li, X., Fujii, H., Tamagawa, K., Ishikawa, H., et al.: Auto-calibration system developed to assimilate AMSR-E data into a land surface model for estimating soil moisture and the surface energy budget, Journal of the Meteorological Society of Japan. Ser. II, 85, 229–242, 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.

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

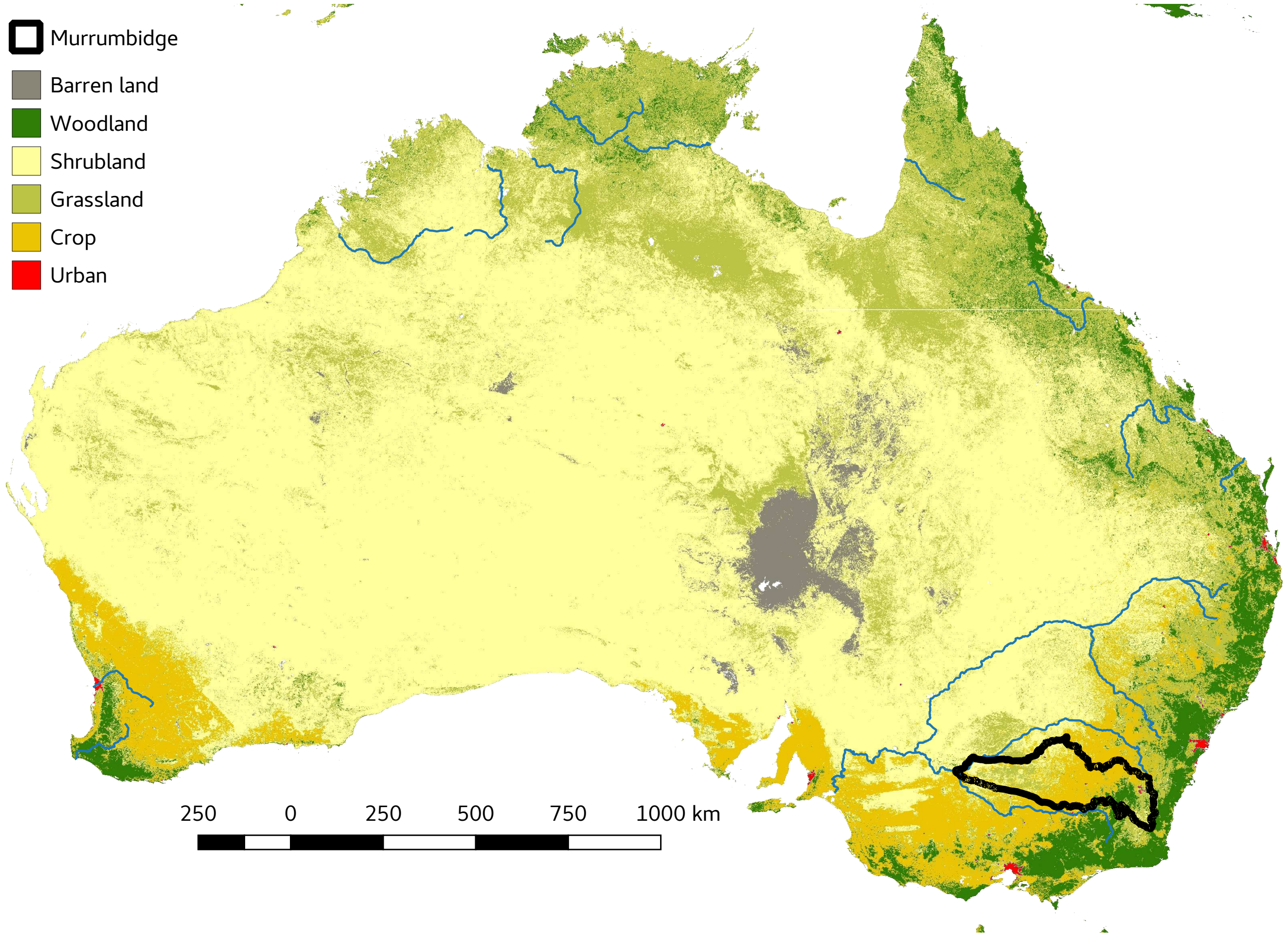

Figure 1. ~~Taylor diagram showing assimilation impact~~ CLM plant functional types based on ~~top layer soil moisture, defined as 8 cm soil depth, (left)~~ MODIS MCDQ12 land cover classification and ~~lower level soil moisture (right) in terms of correlation coefficient R~~ECOCLIMAP climate zones at 500 m resolution prior to the aggregation to 0.25 degrees. Some classes are here aggregated for visualisation purpose, ~~standard deviation~~ e.g. evergreen temperate and ~~normalised RMSE~~evergreen tropical forests are both shown as Woodland. The ~~colours correspond to~~ boundary of the Murrumbidgee catchment, which is the site of the OzNet in-situ ~~measurement sites but are not comparable between both panels~~measurements, is shown.

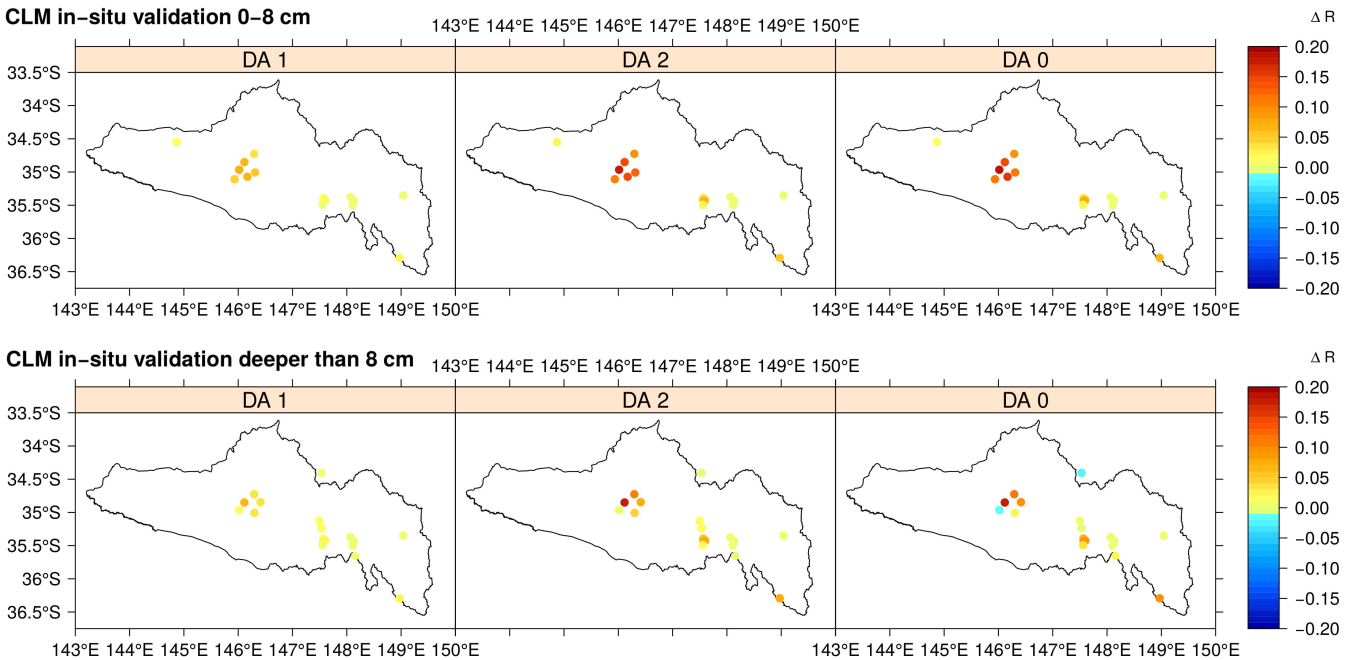

Figure 2. Change in correlation R for experiments DA 1, DA 2 and DA 0 both for top layer soil moisture (top panel) as well as the root-zone soil moisture (bottom panel) within the Murrumbidgee catchment. In the case of multiple measurements at the same location, the weighted average of the measured as well as modelled soil moisture was computed in accordance to the corresponding CLM layer thickness.

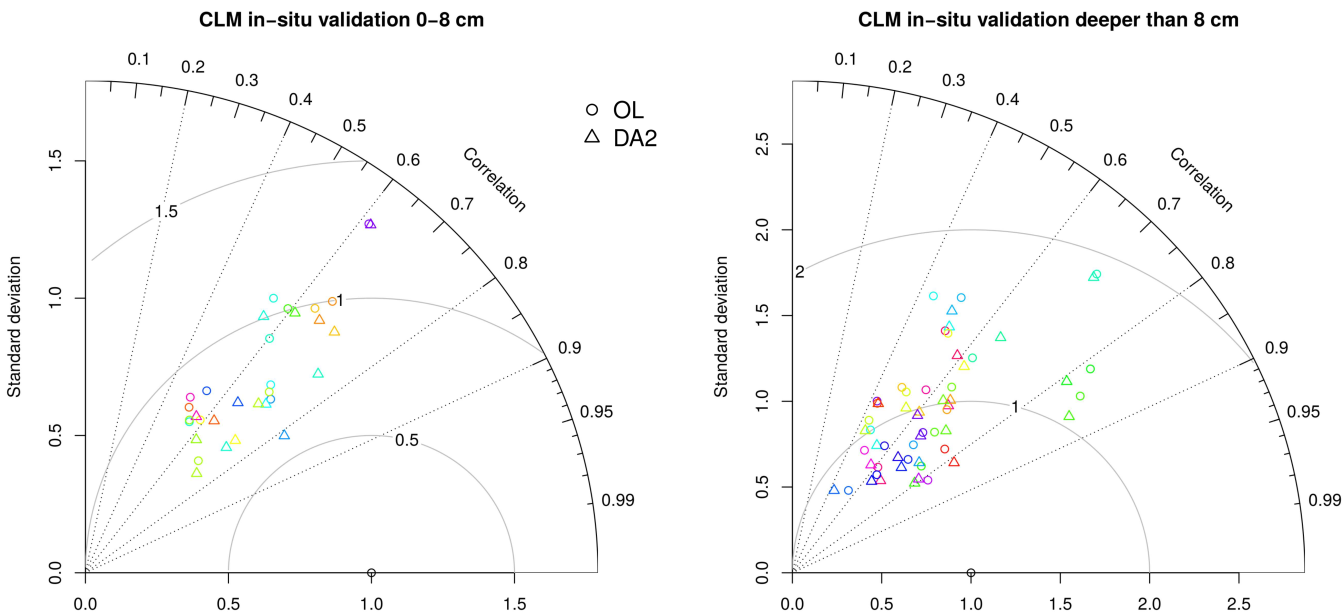

Figure 3. Taylor diagram showing assimilation impact on top layer soil moisture, defined as 8 cm soil depth, (left) and lower level soil moisture (right) in terms of correlation coefficient R, standard deviation and normalised RMSE for all in-situ measurement sites. Measurements at multiple depths are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

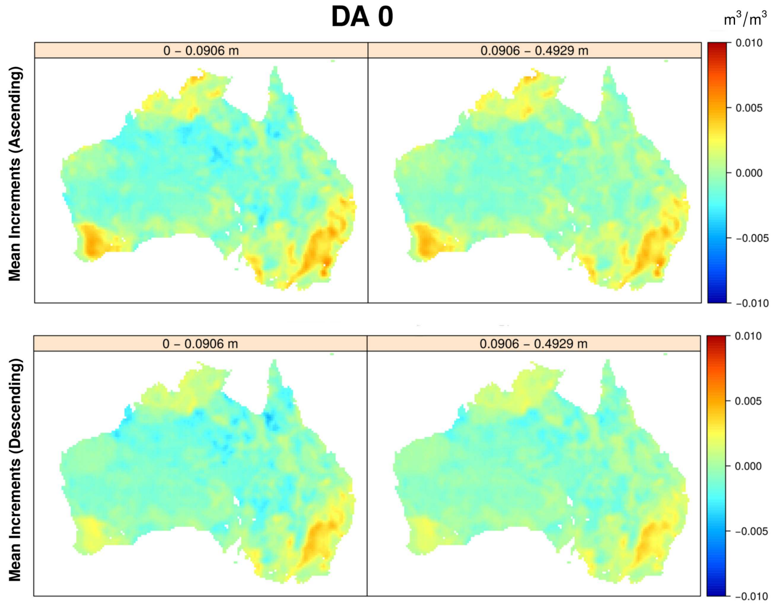

Figure 4. Mean of all increments for experiment DA 0 for top-layer soil moisture (left) and root zone soil moisture (right) for ascending (above) and descending orbits (below).

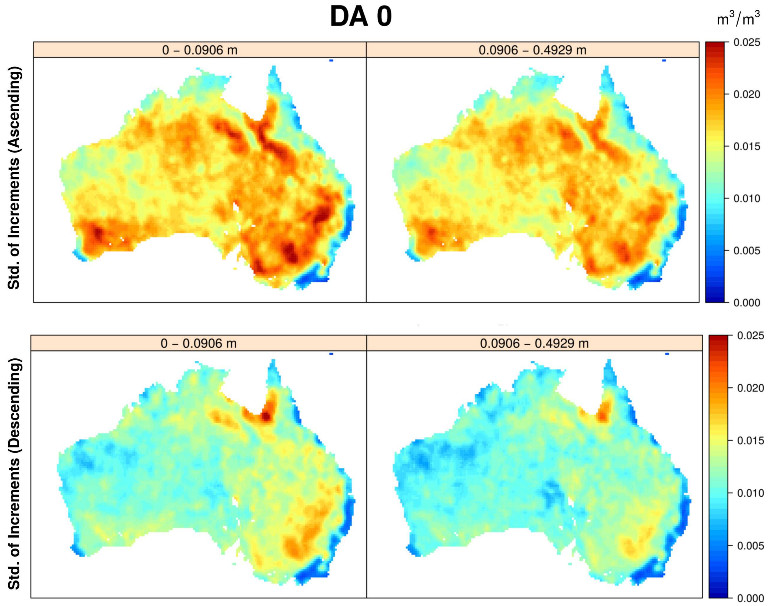

Figure 5. ~~Standard~~ Increments standard deviation of all increments for experiment DA 0 for top-layer soil moisture (left) and root-zone soil moisture (right) for ascending (above) and descending orbits (below). Increments for the root-zone soil moisture are fairly similar to the top soil layers, due to the homogeneous texture perturbations applied across all layers.

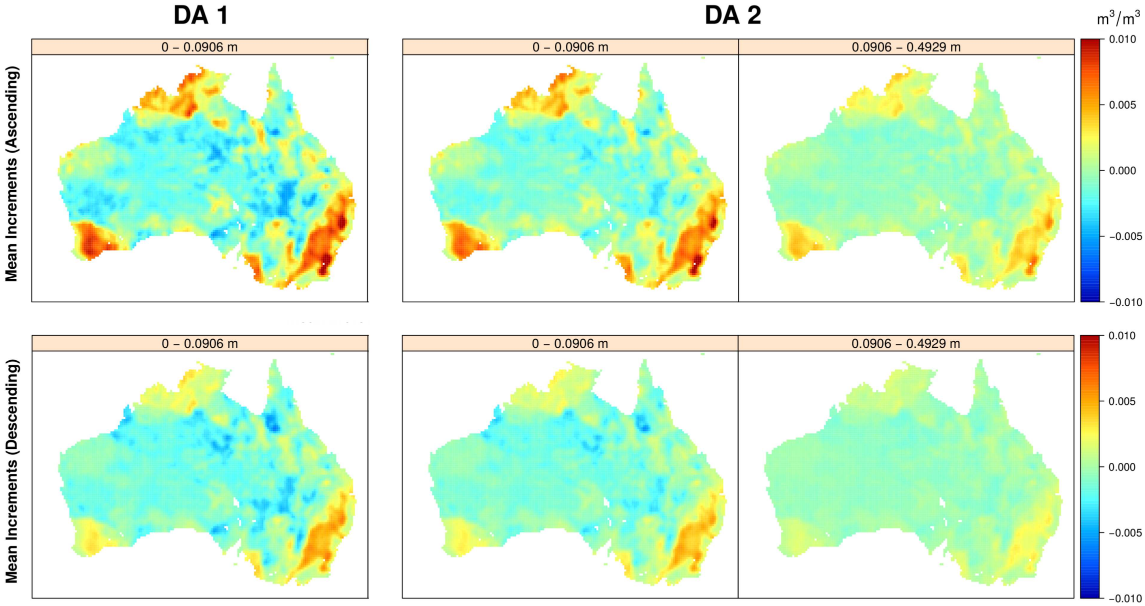

Figure 6. Mean of all increments for experiment DA 1 / DA 2 for top-layer soil moisture and root zone soil moisture for ascending (above) and descending orbits (below). Biases are strongest for the ascending orbit and distinctive spatial patterns are visible. Biases are strongly reduced both for deeper layers and descending orbits.

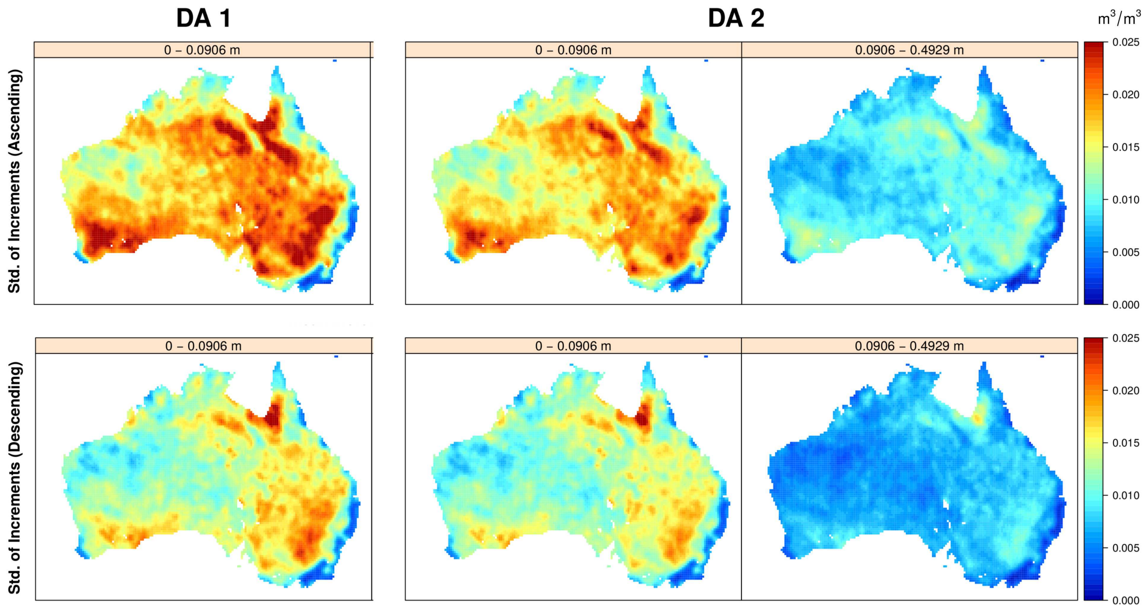

Figure 7. Standard deviation of ~~all~~ increments for experiment DA 1 / DA 2 for top-layer soil moisture and root zone soil moisture for ascending (above) and descending orbits (below). Increments are strongest for the ascending orbit and for top-layer soil moisture and even stronger when restricting 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.

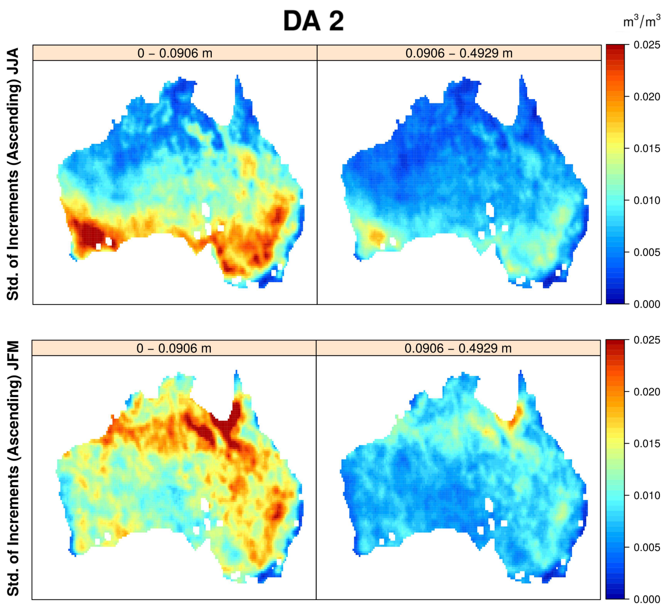

Figure 8. Increment standard deviation for ascending orbits for July - August (top) and January - March (bottom) for experiment DA 2. For the austral winter increments are strongest for the south of Australia, especially the agricultural areas. During austral summer the increments are strongest for the northern grasslands.

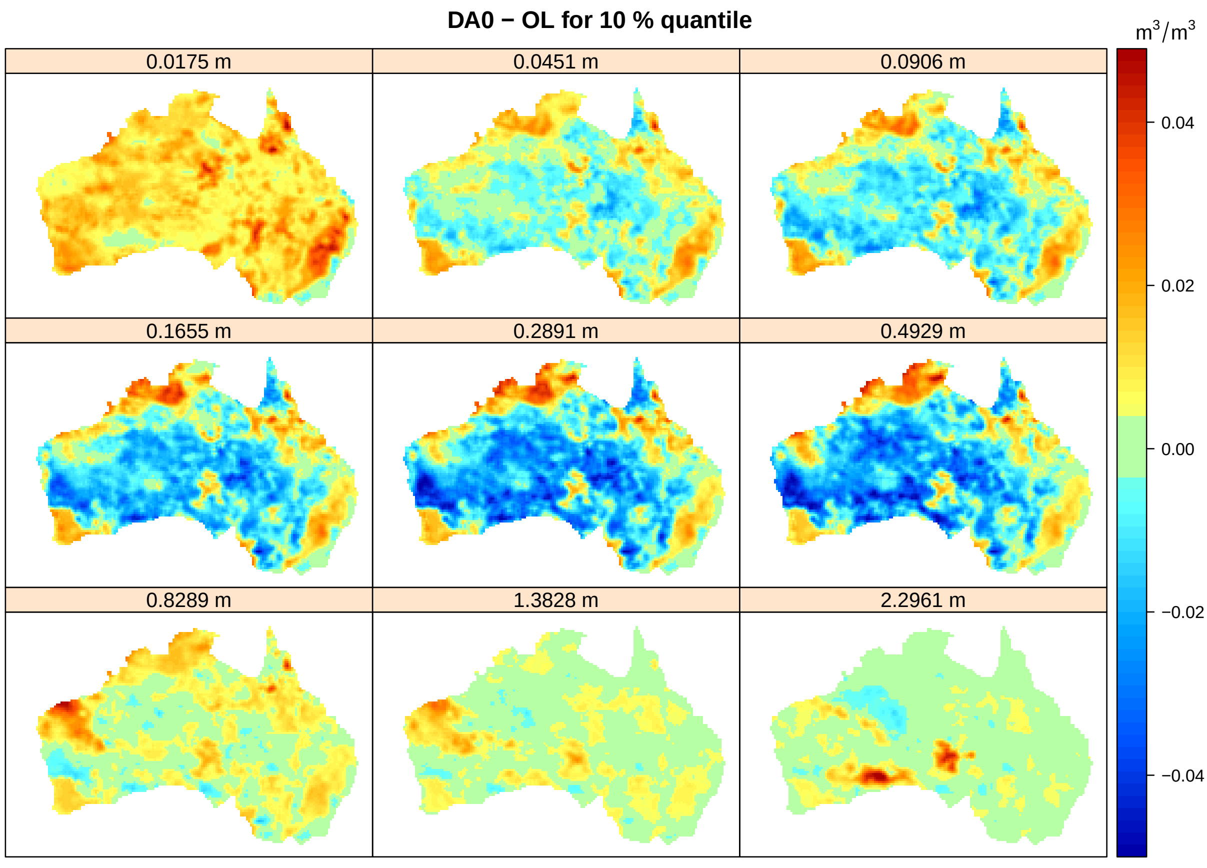

Figure 9. Differences in relative soil moisture ~~%/100~~between ~~open loop~~ open-loop and ~~DA0~~ DA 0 experiment (DA 0 - OL) for the 10 % quantile. The individual panels correspond to the top 9 CLM soil layers.

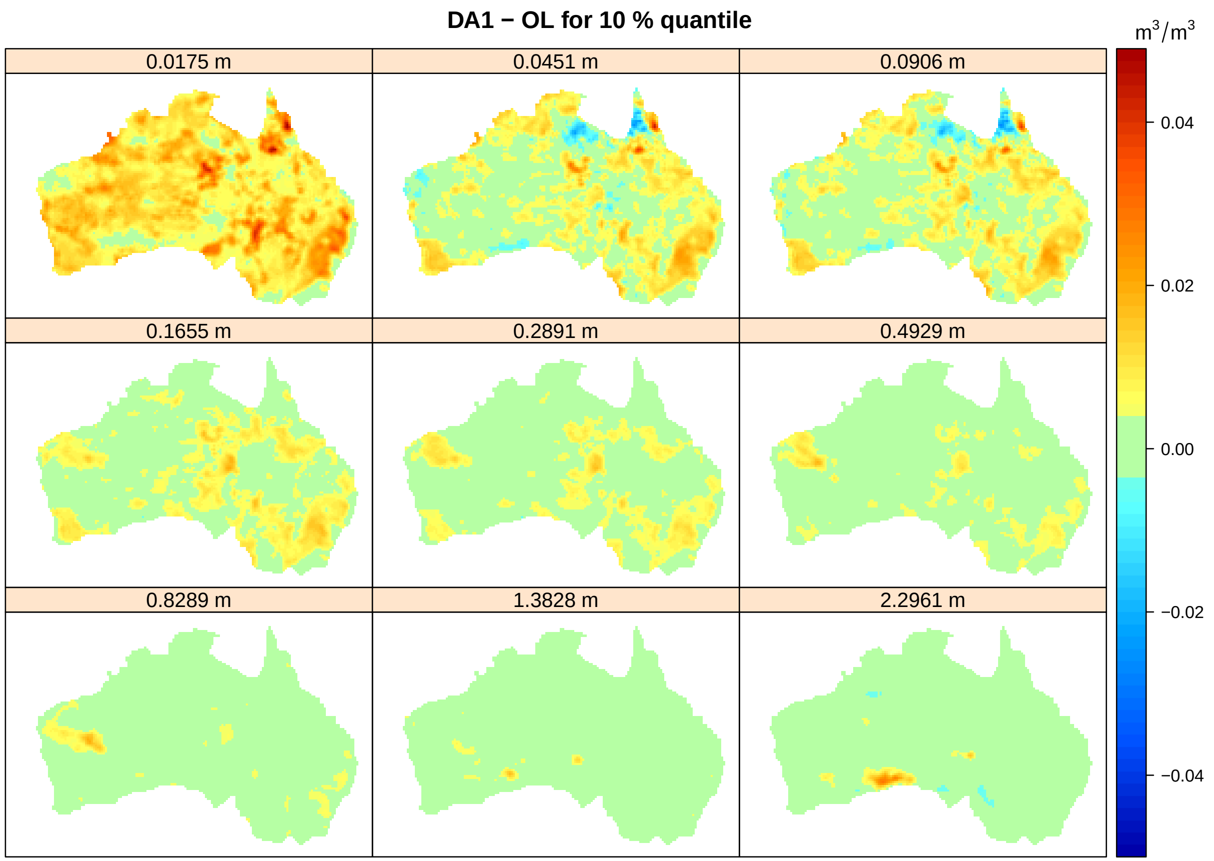

Figure 10. Differences in relative soil moisture ~~%/100~~between ~~open loop~~ open-loop and ~~DA1~~ DA 1 experiment (DA 1 - OL) for 10 % quantile. The individual panels correspond to the top 9 CLM soil layers titled with their depth.

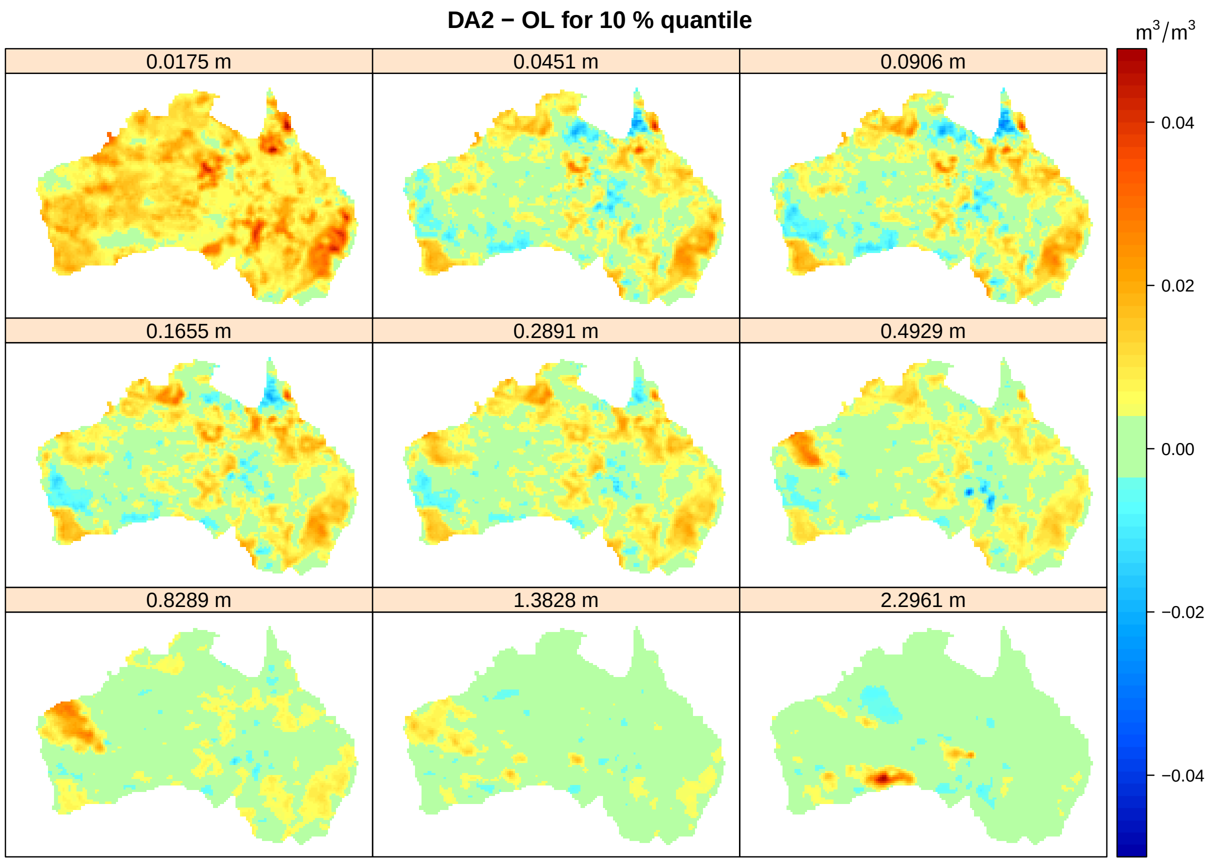

Figure 11. Differences in relative soil moisture ~~%/100~~between ~~open loop~~ open-loop and ~~DA2~~ DA 2 experiment (~~DA2~~ DA 2 - OL) for the 10 % quantile. The individual panels correspond to the top 9 CLM soil layers.

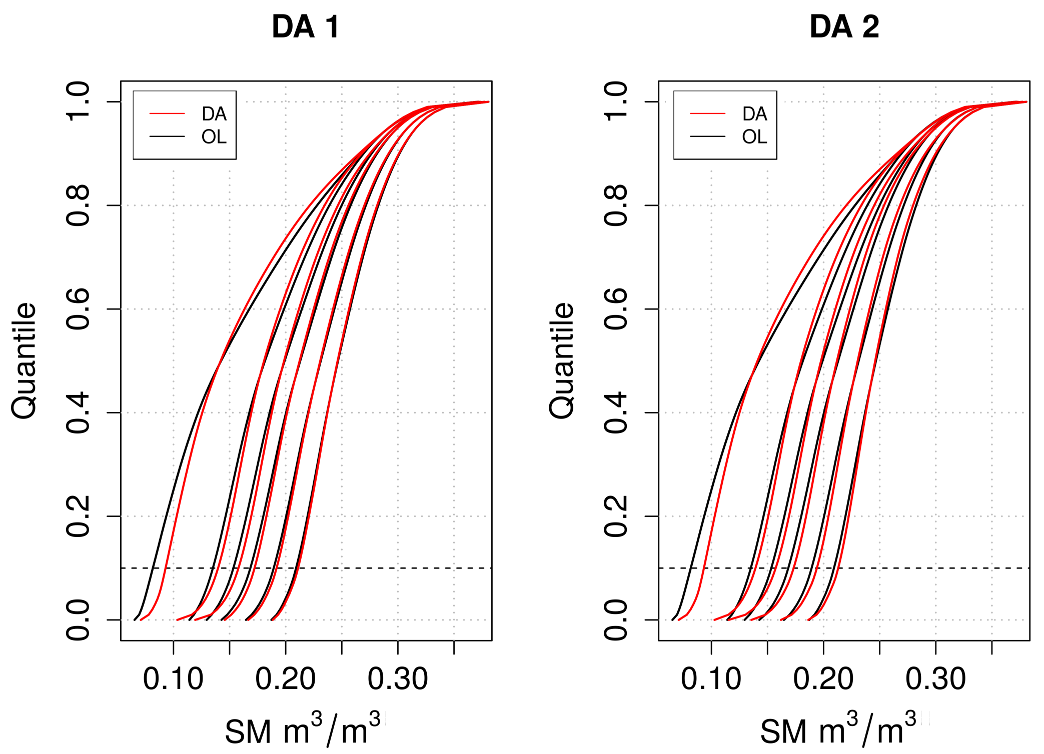

Figure 12. Cumulative distribution functions (CDFs) for the upper 6 CLM soil layers for experiments ~~DA1~~ DA 1 and ~~DA2~~DA 2, based on quantiles computed for all data across the model domain. CDFs for ~~open loop~~ open-loop simulations are shown in black and assimilation results in red. Both panels show changes in CDF behaviour for the layers being updated in the respective experiments, i.e. layers 1-3 for ~~DA1~~ DA 1 and layers 1-6 for ~~DA2.~~ 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 figures 10 and 11.

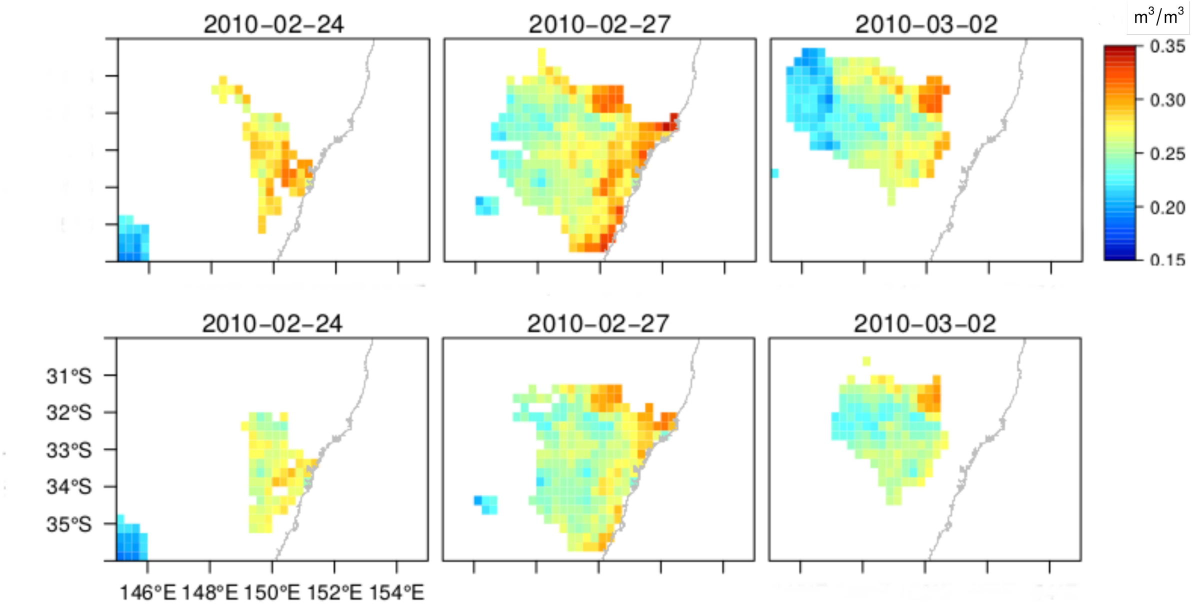

Figure 13. Sample drought event for February 2010, showing only the root zone soil moisture below the 10 % quantile level for the ~~open loop~~ open-loop (above) and experiment ~~DA2~~ DA 2 (below). The different spatial extent and differences in soil moisture itself, depending on the dataset used, at three different days are clearly visible. The figure is centred around the Central coast of New South Wales.