To:

Harrie-Jan Hendricks Franssen,

Editor, HESS

Re: Manuscript ID hess-2015-181

Dear Dr. Hendricks Franssen,

Thank you very much for overseeing the review process for the article and for your guidance regarding the revisions. Please find below our detailed responses to the questions and concerns raised by the two anonymous reviewers. In each response, we have indicated the modifications (including their position in the updated manuscript) that were made in the article. In the PDF of the updated manuscript, we have highlighted the changes to the manuscript in yellow (in response to reviewer #1's comments) and in blue (in response to reviewer #2's comments).

Overall, we feel that the comments and concerns brought forward by the two anonymous reviewers call for what many would probably consider fairly "minor" revisions. In any case, we have addressed all of the reviewers' comments thoroughly, and we are confident that the reviewers will be satisfied with our modifications to the paper.

Sincerely,

Sujay V. Kumar

REVIEWER #1

In this paper, the authors evaluate a number of a number of soil moisture products to compensate for unmodelled processes. In essence the idea of the paper is good, but I do have a number of comments.

- Page 5972: it is stated that is is common practice "to remove the bias between the observations and the model, and use a bias-blind assimilation approach...by rescaling the observations prior to assimilation". I am not sure I agree with this. Over the last decade or so, quite a bit of literature has been developed on the online estimation of biases, as the authors explain a bit earlier. None of this literature is mentioned. This is really not placing the research in the correct frame, and a discussion on this literature should at least be included.

Thank you for the comment. We have modified this statement as follows, so that the efforts to employ bias-aware methods are also captured in the description. These updated statements appear on page 5, lines 1-4.

"Though there have been a number of studies that rely on online estimation of biases (De Lannoy et al. (2007); Reichle et al. (2010)), the common practice in land data assimilation studies, is to remove the bias *between* the observations and the model, and use a bias-blind assimilation approach to correct only short-lived model errors". Note also that at the beginning of the paragraph, we make references to the two types of data assimilation systems in terms of their approach to handling biases.

- Page 5983: is 0.02 volumetric soil moisture a realistic observation error? Please justify.

The expected error of some of the newer soil moisture sensors (SMAP/SMOS) are in the range of 0.02 – 0.05 m3/m3. The value of 0.02 used in the simulations is an optimistic estimate of the expected error. The following text has been included on page 15, second paragraph.

"Random gaussian noise with an error standard deviation of 0.02 $m^3\ m^{-3}$ is added to the truth soil moisture values to mimic measurement uncertainties, which is an optimistic estimate of the error levels in the current space-borne L-band radiometers (SMOS and SMAP)."

- Page 5983: an ensemble size of 12, is that not a bit on the small side? Is there a particular reason why a larger ensemble size has not been chosen?

Larger ensemble sizes would be helpful to improve the sampling density, but here we chose a size of 12 based on the 1-d EnKF employed in this work. The 1-d EnKF updates each grid cell independently of all other grid cells, and in each grid cell there are only 4 layers of soil moisture (i.e., the state vector has only 4 elements). Thus, the small size of the state vector justifies the use of a small ensemble size. In addition, prior studies (see Kumar et al. 2008) have characterized the tradeoff in accuracy and computational cost as a function of the ensemble size and we chose 12 in this experiment based on the prior works (Reichle et al. 2007, Kumar et al. (2008, 2009, 2012)).

We have modified the text on page 16, last paragraph to say "An ensemble size of 12 is used in the simulations with perturbations applied to both meteorological fields and model prognostic fields to simulate uncertainty in the model estimates. The determination of 12 as the ensemble size was based on the prior works (Reichle et al. (2007); Kumar et al. (2008, 2009, 2012)) and because the size of the model state vector is small (4 Noah soil moisture state variables)."

- Discussion on page 5987: In this context, there is a paper by Dara Entekhabi in which he presents a number of metrics to evaluate soil moisture products. Perhaps it is not a bad idea to discuss this paper in this context.

Thank you for the suggestion. We have added the following statement to the discussion on page 20, first paragraph:

"For real data assimilation systems, metrics recommended by Entekhabi et al. (2010) that compute estimates of soil moisture accuracy while accounting for biases, may be more appropriate."

Also, please see below for our response to the next question.

- As a general comment on this discussion, would everything not depend on the way soil moisture is defined? If soil moisture is defined as what we can measure in the ground, then the argument could be raised that a higher RMSE does mean a worse product. To me, this means that soil moisture in a model is not really soil moisture, but a variable that is used to calculate ET and runoff etc. Since we are having a philosphical discussion here, I would add this kind of discussion as well.

The reviewer is right that the soil moisture in the model is not really soil moisture, but rather an index of wetness. This is a long-established fact and cited in many papers and explained in detail in Koster et al. (2009) article (Koster, R. D., Z. Guo, R. Yang, et al. 2009. "On the nature of soil moisture in land surface models." *J Climate*, 22: 4322-4335). In the synthetic example used in the paper, we know that the sources of differences are from irrigation alone and therefore, the use of RMSE is appropriate. Based on the comment, we have modified the discussion on pages 19 and 20 as follows:

“An important philosophical point, however, is warranted here. Implicit in the above discussion of Fig. 11 is the assumption that a higher RMSE reflects a poorer performance. Depending on application, this may not be true at all. It is a well-established fact that the soil moisture estimate from the model is essentially an index of wetness and a highly model- dependent quantity (Koster et al. (2009)). As a result, care must be exercised when comparing model soil moisture directly to in-situ or satellite measurements. The whole point of the scaling exercise is to convert a satellite-based soil moisture value, prior to its assimilation, to a value consistent with that of the LSM used. This allows the further use of the assimilated soil moisture value in that LSM, e.g., to initialize a forecast. If, once the data assimilation process is finished, a soil moisture value is needed that reflects a more “correct” climatology (e.g. with an irrigation-influenced seasonal cycle, as in the Control simulation), the data assimilation product can easily be scaled back to that climatology using the reverse of the original scaling approach. Viewed in this light, the data assimilation approach, with scaling, is essentially designed to capture the year-to-year or short-term variations in soil moisture anomalies rather than the structure of the seasonal cycle. Also note that though the seasonal cycle of RMSE is lowest in the DA-NOBC integration, this configuration is not really viable in real data assimilation systems where biases are unavoidable. The DA-NOBC integration is included in the suite of experiments, as we have the knowledge of the exact sources and magnitudes that contribute to the biases in this synthetic configuration. For real data assimilation systems, metrics recommended by Entekhabi et al. (2010) that compute estimates of soil moisture accuracy while accounting for biases, may be more appropriate.”

- A general comment is also that the results of the study do make sense. Given this, in the section with the Summary the limitations of cdf-matching are discussed. I would also add that cdf-matching will not help your model much if the objective of the model is to model ET or runoff or any other soil-moisture related variable. If you do cdf-matching you will lose a lot of the important information in your data (the way I understand it).

We agree that the quantile mapping methods make no distinction of the source of the biases (unmodeled or from other sources) and therefore will lead to loss of signals related to unmodeld artifacts. As the reviewer correctly notes, this limitation is not just limited to soil moisture. We have emphasized this point in the Summary section, on page 23, first paragraph:

"As the a priori bias correction approaches make no distinction of the source of the biases (un- modeled or from other sources) they treat all systematic differences between the model and observations as biases. As a result, all a priori bias correction strategies considered above cause the signal from seasonal irrigation (or other unmodeled processes) to be excluded in the DA results, though the analysis of the DA internal diagnostics indicate near optimal performance for such configurations. "

Overall I think that with these improvements the paper can be published.

Anonymous Referee #2

Received and published: 21 July 2015

This paper addresses two topics. Firstly it investigates the potential of several remotely sensed soil moisture products to detect irrigated areas. Secondly it investigates soil moisture data assimilation in irrigated areas when the irrigation process is not accounted for in the model. This study shows limitations in the data assimilation system that prevent from making an optimal use of the observations in these conditions. The results and discussion focus on comparing different bias correction approaches or no bias correction, concluding that none of the approach is fully satisfactory.

The paper is well written and results are clearly exposed. I suggest it is published after the comments below are accounted for.

Page 5970, lines 20-22: "Therefore, in this article we focus on irrigation as an analog of a human engineered process that is typically not represented in land surface models." This paper uses irrigation to illustrate a process typically not represented in land surface models (LSMs). This gives the wrong impression that irrigation is not at all represented by LSMs, which is not true. The authors should acknowledge that several land surface models account for irrigation, such as for example ORCHIDEE (de Rosnay et al. GRL 2003), CLM4 (Leng et al. JGR 2011), WRF (Lawston et al. JHM 2015) and many others including Noah as discussed later in the paper. The point is that the bias correction problems addressed in this paper only concern the specific cases/studies/applications where irrigation is not represented in the LSMs. A few lines in the introduction to clarify the context would be useful. In the conclusion, recommendations are discussed to investigate alternative bias correction approaches by grouping model and observations depending on vegetation type for example. This discussion is interesting, however in irrigated areas irrigation is a major process that drastically affects water reservoir and fluxes. The reader wonder if it is the purpose of data assimilation to correct for such a major process when it is not represented. So, one of the main recommendation should also be to account for irrigation in LSMs whatever the application is.

As noted in the lines 20-22, the focus on irrigation in this paper is mainly to use it as an analog of a human engineered process that is often not included in the models. This is certainly not meant to disregard some of the efforts reported in the literature towards developing conceptual formulations of irrigation in the models. We have added the suggested references and modified the statement as follows on page 3, first paragraph:

"Though recent studies have reported the development of conceptual representations of irrigation in land surface models (de Rosnay et al., 2003; Ozdogan et al., 2010; Leng et al., 2013; Lawston et al., 2015), capturing and representing the subjective nature of irrigation practices remains a hard problem. Therefore, in this article we focus on irrigation as an analog of a human engineered process that is often not represented in land surface models."

As noted here, the conceptual representations of irrigation are inherently limiting, as it is very difficult to capture and represent the onset, duration and the timing of irrigation practices in physical models. In that respect, we disagree that representing these processes within the model is the solution to the bias issues, because arguably the conceptual irrigation formulations will still be limited in capturing the complexity of human practices. Observations (if reliable) are probably the only practical way to capture and represent such features. The DA (or other techniques) must be modified to preserve such features to include them in the model simulations. The bias issues resulting from unmodeled processes are obviously not limited to irrigation alone, but are also applicable to other unmodeled processes that may be hard to represent in physical models.

Page 5981, lines 1-6: It is interesting to notice that SMOS and AMSR2 do not capture the irrigation signal whereas ASCAT does. The resolution of the raw data is a possible explanation as discussed in the text. An other explanation could be related to the effect of intercepted water on the signal which, when it is underestimated (or not accounted for) in the retrieval algorithm, leads to opposite effect on retrieved soil moisture from active and passive sensors. So, the fact that ASCAT captures the irrigation signal may be an artefact due to the intercepted water contribution to the signal.

Thank you for the comment. We have added the following statement to the text on page 13, second paragraph:

"Thus, because Fig. 6 focuses on 0.125 grid cells with at least 30 % irrigation, the SMOS and AMSR2 data (interpolated to that resolution) will necessarily include some soil moisture information from areas outside those defined by the 30 % threshold areas that are, almost by definition, drier. ASCAT, with a raw resolution of ~25 km does not seem as affected by this, perhaps in part due to its finer base resolution. Another possible reason may be related to the influence of intercepted water, which has opposite effects on the active and passive sensors. More analysis is needed, however, to understand the different behaviors of the sensors."

Page 5994, line 9: Draper et al., 2014 should be 2015. Also update the text when the reference is cited.

Both the reference and the text where it is cited have been updated. Thanks for pointing this out.