

HESS review responses

Author response to R1

We thank the reviewer for their kind comments which have helped to strengthen this manuscript. Below are our responses:

1) The reviewer noted that the DA framework in this study is very similar to previous studies.

We agree that the method is similar to the studies noted, we have updated the text to clarify this. P7L7:

“This is a similar framework to that introduced in Yang et al. (2007, 2009) for data assimilation with the Simplified Biosphere model 2 (SiB2) (Sellers et al., 1996), ...”

2) The reviewer asked if we had optimized soil porosity.

Soil porosity is updated implicitly via the pedo-transfer functions we use for the JULES model. We have updated the text to make this clear. P7L8:

“...with the exception that the soil porosity parameter of JULES is updated implicitly within the pedo-transfer functions rather than explicitly included in the optimisation.”

3) The reviewer asked for more explanation of the reference to *Maidment et al. 2017* P6L28-29.

We have updated the text with more explanation. P8L30:

“This finding is consistent with the comparisons of precipitation between v3.0 and v2.0 presented by Maidment et al. (2017), where TAMSATv3.0 was shown to reduce a dry bias present in TAMSATv2.0 when compared to ground station data.”

4) The reviewer disagreed that we were overfitting to the data at some grid cells P8L3-4.

We agree that “overfitting” is not the right terminology here. We believe that missing processes within the model and deficiencies of our pedo-transfer functions may cause us to retrieve unrealistic soil texture values for some grid cells. We have updated the text to remove the previous explanation. We have also included a comparison to in-situ soil texture observations from the African Soil Profile Database. P11L10:

“The inability of our data assimilation to improve soil texture estimates at certain points is most likely due to issues of spatial representativity between the modelled soil map and the in-situ data. It is also possibly impacted by errors in our pedo-transfer functions, which may perform better if specifically calibrated for Ghanaian soils (Patil and Singh, 2016).”

5) The reviewer didn't agree with our discussion on P8L10-13.

We agree with the reviewer that discussion about the biases was not meaningful and have removed this discussion from the text.

6) The reviewer commented that the CCI soil moisture observations are believed to be representative of a depth of less than 10cm P12L23-25.

Reviewer 3 also thought that 10cm was too deep a layer for comparison with the CCI soil moisture. We have therefore updated the JULES model to run with a top layer of 5cm depth and re-run our experiments. P3L15:

“JULES is typically run with 4 soil layers, with the top layer being 10 cm deep. In this paper we have updated JULES to run with a top layer of 5 cm to be more representative of the ESA CCI soil moisture observations. Another option to deal with the issue of representativity would be an exponential filter (Albergel et al., 2008) which has been used in sequential data assimilation studies previously (Massari et al., 2015; Alvarez-Garreton et al., 2016).”

Author response to R2

We thank the reviewer for their comments which have certainly helped to strengthen the manuscript. Please find our responses below:

1) The reviewer suggested that we change the title of the paper as the previous title was not well aligned with the paper contents.

We agree with the reviewer that a more appropriate title is necessary and have updated the title accordingly:

“Impact of remotely sensed soil moisture and precipitation on soil moisture prediction for a data assimilation system with the JULES land surface model”

We feel that this more accurately brings out the novel aspects of the paper, i.e. and exploration of the useful information two different EO data sources – one driving the model (precipitation) and the other being used as a target variable (soil moisture).

2) The reviewer asked us to either remove discussion of agricultural drought from the abstract or complement the Results, Discussion and Conclusion with how the work here would be useful for better decision making in crop cultivation.

We agree that agricultural drought is not discussed enough throughout the paper to warrant its inclusion in the abstract. We have removed the relevant text.

3) The reviewer asked that we include the reduction in root-mean squared differences for both the North and South within the abstract P1L5-6.

We have included the statistics for the two regions instead of the whole of Ghana. We have also updated the performance statistic used at the request of other reviewers to unbiased root-mean squared error. P1L5:

“The use of an improved remotely-sensed rainfall dataset contributes to 6% of this reduction in error in northern Ghana and 10% in southern Ghana.”

3) The reviewer asked us to remove mention of forecasting from the abstract P1L10

We have removed this from the abstract

4) The reviewer asked us to comment on the novelty of this research, compared to *Bolton et al. 2010* P2L13-14.

Bolten et al. 2010 conduct state estimation using a sequential data assimilation technique (Ensemble Kalman Filter), meaning that whenever an observation is available the model state (in this case soil moisture) will be updated, but if no data is available the model estimate to soil moisture will not be updated. In this paper we are conducting parameter estimation using a variational technique. Once we have ingested data over some time window the model parameters governing soil moisture will be updated and so will future and past estimates to soil moisture. We have clarified this in the text. P2L16:

“It has been shown by *Bolten et al. (2010)* that assimilation of remotely sensed surface soil moisture can significantly improve the prediction of root-zone soil moisture and drought modelling, where a sequential DA technique is used for soil moisture state estimation.”

5) The reviewer asked us to remove the term “observations” from P3L4-5.

We have removed this.

6) The author asked us to highlight that we update only two parameters of soil texture, with the third being updated implicitly P4L31.

We have amended the text to show this. P7L5:

“we updated the percentage of sand and silt in the soil at each minimisation step (with clay being updated implicitly)”

7) The reviewer asked us to unify our usage of terminology “grid cell” and grid box” P6L9 and P6L17 with Figure caption 2.

We have updated the text accordingly.

8) The reviewer commented that the amplitude of soil moisture is decreasing for the model (2010-2014) and wondered if the precipitation data showed the same pattern P6L13-14.

We agree there is a noticeable trend in the amplitude of soil moisture and this is also seen in the precipitation data for both TAMSAT products. We have added a comment on this in the paper. P8L15:

“In Figure 4 we can also see the amplitude of this seasonal cycle slightly decreasing, this is a pattern also seen in both TAMSAT products which exhibit a drying over the period 2010-2014 for this grid cell”

9) The reviewer suggested it would be good to show a skill of JULES over daily averages of ESA CCI data for 2009 for Figure 2 and 3.

We have included root-mean squared differences between JULES and CCI data in the text. In relation to Figure 2, P8L13:

“The model skill for predicting this seasonal cycle is markedly improved after data assimilation, with a root-mean squared difference (RMSD) of 0.035 after data assimilation compared to a RMSD of 0.094 before for 2009”

In relation to Figure 3, P8L20:

“(RMSD of 0.059 after data assimilation compared to RMSD of 0.102 before for 2009)”

10) P6L20-21: “Although we do improve the fit to the observations after data assimilation in Figure 3...” – Reviewer asked us to add a quantitative metric.

We have included RMSD reduction as a quantitative metric, please see previous comment (no. 9).

11) The reviewer asked us to add a plot of the difference in precipitation between TAMSAT versions P6L30-31.

We have added the suggested figure (See new Figure 1) and also included a figure of cumulative rainfall over the period of the experiments (2009 – 2014) to show TAMSAT v2.0 is the drier product (see new Figure 2).

We have added text to section 2.2 discussing these figures, P4L15:

“In Figure 1 we show yearly cumulative rainfall averaged over 2009 - 2014 for TAMSAT v2.0 and v3.0, we can see the different spatial distributions of rain with v3.0 being wetter in the south and v2.0 wetter in the east. To show the difference in the amount of rainfall for the two products we also show cumulative rainfall for the period 2009 - 2014 averaged spatially over Ghana in Figure 2, this shows TAMSAT v2.0 to be the drier of the two products overall, as expected.”

12) The reviewer asked us to update Figure 5 to have monthly or weekly RMSEs (rather than yearly) P6L32

We have updated this figure accordingly to display monthly rather than yearly averages and now display two new Figures (see Figure 7 and 8 in attached manuscript of proposed changes), as the reviewer has suggested this helps to understand dynamic patterns better. P8L35:

“Figure 7 and 8 show experiment monthly root-mean-square differences (RMSD) for north and south Ghana respectively. For Figure 7 this shows that the most accurate model run overall is experiment 4 (TAMSAT v3.0 with DA). We see in the majority of years that towards the start of the season as soils are wetting up it is experiment 3 and 4 (TAMSAT v3.0 no DA and with DA respectively) that have the lowest RMSD, suggesting that it is precipitation, as

opposed to the assimilation of soil moisture, that is most important for improving soil moisture estimates during this period. This relationship changes towards the end of the rainy season with experiment 2 and 4 being the most accurate (TAMSAT v2.0 with DA and TAMSAT v3.0 with DA respectively) suggesting that assimilation of soil moisture estimates is most important in this period..."

13) The reviewer asked if there was another database of soil properties for comparison with HWSD and JULES-DA results P7L12-15.

We have used in-situ observations from The African Soil Profile Database compiled by ISRIC to compare with our results. We find an improvement in soil texture in the north of Ghana compared to the in-situ measurements. This is not the case for the south, where we slightly degrade our estimate to soil texture. Relevant text has been added to P10L13:

"Comparing estimates of soil texture derived from CCI soil moisture to in-situ observations is inevitably problematic due to spatial scales of representativity. However, independent sources of verification are difficult to find over Ghana. We therefore compare our soil maps to in-situ observations from The African Soil Profile Database (Leenaars et al., 2014). This database is compiled by the International Soil Reference and Information Centre (ISRIC) with the quality of the data being rated from 1 (highest quality) to 4 (lowest quality), here we compare only to observations with a quality flag of 1 or 2. In table 1 we show the root-mean-squared error (RMSE) for our soil maps when compared to 21 in-situ observations of soil texture in the north of Ghana and 36 in-situ observations in the south (locations shown as red dots in Figure 9). For the north of Ghana where we have most confidence in our results we find a reduction in RMSE for both sand and clay (almost halving the RMSE in clay %). However, we do increase our RMSE for the silt %. In the south of Ghana, we do not manage to recover a better estimate to soil texture after data assimilation, with an increase in RMSE for silt and clay but a decrease in RMSE for sand. The inability of our data assimilation to improve soil texture estimates at certain points is most likely due to issues of spatial representativity between the modelled soil map and the in-situ data. It is also possibly impacted by errors in our pedo-transfer functions, which may perform better if specifically calibrated for Ghanaian soils (Patil and Singh, 2016)."

14) The reviewer asked us to calculate summary statistics for the north and south of Ghana rather than the country as a whole P8L5-12

We have updated the summary statistics in Table 2 accordingly. At the request of other reviewers, we have also changed the statistics used in this table from RMSE and bias to unbiased RMSD and correlation.

15) P12L5- 8: The reviewer asked us to describe in details why "updating of soil parameters" (= updating all of the parametrization of water cycle processes in JULES) sometimes works worse than precipitation forcing updating.

We have added more explanation here. P15L10:

"This can be seen in Figure 7 where TAMSAT v3.0 without DA outperforms TAMSAT v2.0 with DA at certain points in the season. This is because at these points the data assimilation system is not able to overcome the errors in the precipitation forcing data to improve the estimates further. If there is too little rainfall, there is a point where the DA system cannot make the soil any wetter (because we are not changing the model soil moisture state - only the soil texture)."

16) The reviewer commented that it is not possible to improve predictions with no access to forcing data P12L12-14.

We have removed this text.

17) The reviewer commented that discussion of the data assimilation system was too technical for the discussion P12L15-22.

We have moved this text to the methods section 2.4

18) The reviewer commented that the conclusions were too technical and miss the HESS aims and scope P13L9-19.

We have updated conclusions to try and address this point

Typos:

P12L11: "after after".

amended.

Author response to R3

We thank the reviewer for their comments which have helped to strengthen our manuscript. Please find the responses below:

1) The reviewer suggested that we change the notation from RMSE and BIAS to RMSD and mean relative error as the prior notation makes the assumption that the ESA CCI estimates are close to the truth. The reviewer also commented that we should make an effort to include some independent observations.

We agree that the notation suggested by the reviewer is more appropriate and we have updated the manuscript accordingly. Another reviewer has also raised the concern over the quality of the observations with which we are comparing, especially as they are CDF matched to another land surface model. We have taken their advice and also included temporal correlation and the unbiased RMSD (ubRMSD) as metrics to show that the model still improves after removing any competing bias. We have added text commenting on the issues for the ESA CCI combined soil moisture product. We have also included independent in-situ observations of soil texture from the African Soil Profiles Database to judge our models soil maps and added text discussing the performance of CCI soil moisture over West Africa in section 2.3, P4L9:

“Dorigo et al. (2015) also show that the ESA CCI product performs well over Western Africa when judged against in-situ soil moisture observations from the AMMA network (Cappelaere et al., 2009) with stations in Benin, Mali and Niger. When judged against the AMMA network CCI soil moisture was shown to have a high correlation (~ 0.7) and one of the lowest unbiased root-mean squared differences (~ 0.04) of the 28 worldwide networks used in the study. This bodes well for our comparison over Ghana, which has a similar climate regime in the north to the sites in the AMMA network.”

2) The reviewer suggested we included different performance scores as from the two included it appeared that the DA mainly acted to reduce the BIAS (mean relative error).

We agree that an additional performance score would be beneficial. As another comment from the author of a short comment on the paper has also suggested. We have included their suggestion of temporal correlation and ubRMSD to show that the DA technique also reduces random errors within the model, rather than just acting to rescale the model predictions. P11L14:

“Satellite soil moisture products can be subject to larger errors and biases associated with data processing. This is particularly true for the CCI level 3 combined active and passive product used in this paper, as in order to merge information from 11 different sensors data is CDF matched to the GLDAS-Noah v1 model (Rodell et al., 2004). Therefore, any bias within the GLDAS-Noah model will be included in the level 3 soil moisture product used here. To

make sure we are not just correcting the bias of the JULES model to that of GLDAS-Noah we include summary statistics of unbiased root-mean squared difference (ubRMSD) and temporal correlation in table 2. In every case we find that after data assimilation we improve both ubRMSD and correlation and in the majority of cases find the best results for experiment 4 (TAMSAT v3.0 with DA). For the north of Ghana, we reduce the ubRMSD by 18% from experiment 3 (0.0622 m³ m⁻³) to experiment 4 (0.0508 m³ m⁻³). From experiment 2 to 4 we can see that, after data assimilation, using TAMSAT v3.0 rainfall over v2.0 has contributed to a 6% reduction in ubRMSD when calculating statistics over the whole period. In the south of Ghana we reduce the ubRMSD by 21% from experiment 3 (0.0590 m³ m⁻³) to experiment 4 (0.0467 m³ m⁻³), here improved rainfall data has contributed to 10% of this reduction.”

3) The reviewer commented that the assumption made that CCI soil moisture observations are representative of the JULES top 10cm soil layer is unrealistic. They suggested we reformulate this assumption or use an exponential filter to address this.

We agree that the assumption made originally is perhaps not realistic. We have therefore updated the JULES model to run with a 5cm soil top layer and have re-run our experiments. We have updated the text in section 2.1 to reflect this and have also included the references mentioned by the reviewer commenting that this is another option. We find similar results as when running with a 10cm soil depth except a larger dry mean relative error in the north after data assimilation (see Figure 6). This is understandable as a shallower soil layer will dry more quickly. P3L16:

“In this paper we have updated JULES to run with a top layer of 5 cm to be more representative of the ESA CCI soil moisture observations. Another option to deal with the issue of representativity would be an exponential filter (Albergel et al., 2008) which has been used in sequential data assimilation studies previously (Massari et al., 2015; Alvarez-Garreton et al., 2016).”

4) The reviewer asks us to demonstrate the soil texture after DA more closely reflects reality than the HWSD.

We have included comparisons of our retrieved soil maps to in-situ observations of soil texture from the African soil profile database. We find that we can improve soil texture estimates in the north but are unable to do so in the south. We have added discussion about this in the text. P10L13:

“Comparing estimates of soil texture derived from CCI soil moisture to in-situ observations is inevitably problematic due to spatial scales of representativity. However, independent sources of verification are difficult to find over Ghana. We therefore compare our soil maps to in-situ observations from The African Soil Profile Database (Leenaars et al., 2014). This database is compiled by the International Soil Reference and Information Centre (ISRIC) with the quality of the data being rated from 1 (highest quality) to 4 (lowest quality), here we compare only to observations with a quality flag of 1 or 2. In table 1 we show the root-mean-squared error (RMSE) for our soil maps when compared to 21 in-situ observations of soil texture in the north of Ghana and 36 in-situ observations in the south (locations shown as red dots in Figure 9). For the north of Ghana where we have most confidence in our results we find a reduction in RMSE for both sand and clay (almost halving the RMSE in clay %). However, we do increase our RMSE for the silt %. In the south of Ghana, we do not manage to recover a better estimate to soil texture after data assimilation, with an increase

in RMSE for silt and clay but a decrease in RMSE for sand. The inability of our data assimilation to improve soil texture estimates at certain points is most likely due to issues of spatial representativity between the modelled soil map and the in-situ data. It is also possibly impacted by errors in our pedo-transfer functions, which may perform better if specifically calibrated for Ghanaian soils (Patil and Singh, 2016).”

5) Section 2.2 Define here which are the differences between TAMSAT 2.0 and 3.0

We have added text explaining the difference between the 2 products and added two figures showing how the products differ over Ghana. P4L10:

“TAMSAT v3.0 differs from TAMSAT v2.0 in that it uses an updated calibration against in-situ data that is more representative of local scales. [...] For more information on the differences between the two TAMSAT products see Maidment et al. (2017).”

See new Figures 1 and 2 in attached manuscript of proposed changes.

6) Equation 1. Define N. Also xi should not be x0?

The reviewer is correct xi should be x0, N is the number of observations within the chosen time window. We have updated the text to clarify this.

7) Section 2.4. The reviewer asked us to describe how the matrices B and R are estimated.

We have described how B and R are specified in this section. P6L13:

“As we do not have a good estimate to the error in the prior estimates of model parameters we chose a conservative 5% standard deviation for the prior error covariance matrix B. This ensures we do not retrieve unrealistic estimates of soil texture after data assimilation. For the observational error covariance matrix R we have a diagonal matrix with variances estimated from the standard deviations included in the ESA CCI soil moisture product.”

8) P12L15 The reviewer asked us to remove the phrase “retrieve hydraulic parameters”.

Removed

Author response to R4

We thank the reviewer for their comments which have helped to strengthen this manuscript. Please find our responses below:

1) The reviewer shared the concerns of others with respect to the use of the CCI data and performance metrics used within the paper.

We have included extra text explaining the limitations of the used observations and have also included unbiased RMSD and correlation statistics for our experiments to show that the DA is not just acting to match the climatology of GLDAS1-Noah soil moisture. This is the same text from our response to reviewer 3’s second comment, P11L14:

“Satellite soil moisture products can be subject to larger errors and biases associated with data processing. This is particularly true for the CCI level 3 combined active and passive product used in this paper, as in order to merge information from 11 different sensors data is CDF matched to the GLDAS-Noah v1 model (Rodell et al., 2004). Therefore, any bias within the GLDAS-Noah model will be included in the level 3 soil moisture product used here. To make sure we are not just correcting the bias of the JULES model to that of GLDAS-Noah we include summary statistics of unbiased root-mean squared difference (ubRMSD) and temporal correlation in table 2. In every case we find that after data assimilation we improve both ubRMSD and correlation and in the majority of cases find the best results for experiment 4 (TAMSAT v3.0 with DA). For the north of Ghana, we reduce the ubRMSD by 18% from experiment 3 (0.0622 m³ m⁻³) to experiment 4 (0.0508 m³ m⁻³). From experiment 2 to 4 we can see that, after data assimilation, using TAMSAT v3.0 rainfall over v2.0 has contributed to a 6% reduction in ubRMSD when calculating statistics over the

whole period. In the south of Ghana, we reduce the ubRMSD by 21% from experiment 3 (0.0590 m³ m⁻³) to experiment 4 (0.0467 m³ m⁻³), here improved rainfall data has contributed to 10% of this reduction. We find the highest correlations in the north of Ghana for the whole period (2010 - 2014), this is mainly due to the seasonal cycle being much more pronounced in this region.”

2) The reviewer asked us to explain the larger improvement and lower overall errors shown for the South in Figure 5, given that we mention the products have lower skill in the south. In Figure 5 we are comparing our model estimates to the ESA-CCI observations which we believe will also have higher errors in the south given the denser vegetation cover and higher cloud cover. Therefore, lower errors in Figure 5 do not necessarily mean that our model estimate is closer to the truth in the south if, as expected, the CCI observations are also of poorer quality in this region. There is also a much more pronounced seasonal cycle in the north, so that the model mistiming this cycle slightly can lead to large peaks in error. We have since re-run our experiments with a 5 cm soil depth and find similar results. The new figures relating to this point are 7 and 8. We have updated the text to clarify this point.

P9L3:

“Experiments 2 and 4 have a lower RMSD in the south (Figure 8) compared to the north (Figure 7), this seems surprising given that we consider the quality of the data to be poorer in the south. However, this is in part due to the much more pronounced seasonal cycle in the north leading to peaks in RMSD when the seasonal cycle is even slightly mistimed by the model. We also have less confidence in the CCI soil moisture observations in the south so a lower RMSD in comparison to this product over this region is perhaps not indicative of a better soil moisture estimate overall.”

3) The reviewer asked us to indicate the version of the ESA CCI soil moisture we are using. We are using ESA CCI level 3 v03.2. We have included this in the manuscript in section 2.3.

Author response to SC1

We thank the author for their short comment which has helped to strengthen our manuscript. Please find our response below:

1) The author asked if the CCI product used is CDF matched to the GLDAS- NOAH model and if so requested that we include different performance metrics.

We are using the level 3 combined product which is CDF matched to the GLDAS- NOAH model, we agree that including different performance metrics is best to show the reduction in random errors. We have included ubRMSD and temporal correlation as suggested.

2) The author comments that it would be very desirable to repeat this experiment over a data rich area in order to have much better ground based verification.

We agree this would be very beneficial to look at and we would be excited to talk to the author more about possibilities for the future. The work in this paper formed part of the ERADACS project which was based in Ghana, it was therefore necessary that the focus of the paper was over this region. Please do contact me for further discussion:

e.pinnington@reading.ac.uk