Answer to interactive comment by Referee #2 Dr. Teuling

The authors thank Dr. Teuling for providing comments and suggestions, which will help us to improve our manuscript. An initial reply to the reviewer's comments is provided below following the original review comments shown in Italic font.

1. General comments

COMMENT:
The manuscript by Iwema et al. addresses the use of new Cosmic-Ray Neutron Sensor (CRNS) data in land surface model calibration. By using CRNS and in situ sensor data from 12 sites across the U.S., the authors systematically investigate to which degree calibration against this novel data leads to an improvement in the simulation of observed surface fluxes. Unfortunately the results do not show a clear improvement of using CRNS data over in situ data, but this does not in any way affect the quality of the research. Overall, I have a positive impression of the manuscript which I believe would make a good contribution to HESS, but it will need improvements on several aspects. In particular, some of the main conclusions do not seem to follow from the results (a problem also identified by the other referee), the authors should make clear that the in situ data is not used to its full potential so the study is not a clean comparison between datasets but rather a comparison of scale, and the authors seem to have been somewhat selective in the selection of references to identify the knowledge gap. These issues are discussed in more detail below. However I believe these comments can be addressed by minor (mostly textual) revisions.

ANSWER:
We thank the Referee for his effort and for his positive comments regarding the quality and relevance of our study to the HESS community. We will revise the conclusions, so they do follow from the results, as also provided in the answer to Referee #1 and Referee #3. We will also appropriately discuss the relevant points based by the referee in the revised manuscript and include additional supporting literature as mentioned.

COMMENT:
A first remark concerns the title, which does not seem to reflect the contents of the manuscript. The study really looks at calibration using soil moisture observations at different scales, and it does not look at measurement scales as such. I suggest a title along the lines of "Land surface model performance using cosmic-ray and in situ soil moisture data for calibration". Also, the use of "reducing" is seemingly at odds with the results, which does not show an improvement when using observations at the "model" scale.

ANSWER:
We thank the Referee for making this valuable point. We will change the title appropriately to reflect the reviewer's suggestion.

COMMENT:
As was also pointed out by the other referee, the conclusion that JULES has a weak coupling between soil moisture and ET does not follow from the results. In fact, JULES has a strong coupling by definition, and the coupling in reality can only be less strong. The fact that ET estimates do not improve with improvements in soil moisture is likely because ET is not so sensitive to changes in soil parameters, even for different climate conditions. This is for instance shown in a paper I wrote in 2009 (Teuling et al., 2009), in which I investigated effects of soil parameters on soil moisture and ET. In short, effects of soil parameters on soil moisture are generally large, but effects on ET are small. This is primarily caused by the main effect of soil parameters which is a shift in the mean, rather than

the dynamics (this is also consistent with the strong contribution of bias to MSD as reported by the authors). This should be discussed better.

ANSWER:
We thank the Referee for providing this very important comment. The revised manuscript will focus mainly by addressing more specifically the fact that changes in soil parameters limited the improvement to soil moisture causing little effect on on evapotranspiration.

COMMENT:
Related to this is the question why rooting depth was not optimised along with the soil parameters. If the model rooting depth does not reflect the actual root profile, optimization of soil parameters along will not lead to a better estimation of available soil water. These choices should be better motivated.

ANSWER:
We used site specific data in order to prescribe rooting depth for the analysed sites (as mentioned in original manuscript Page 7, line 32 – Page 8, line 2). The decision was made to avoid prescribing more than one rooting depth parameter for individual sites given that rooting depth in JULES is defined per Plant Functional Type. This is indeed a limitation in our analysis and will be mentioned in our revised manuscript.

COMMENT:
In making a case for the validity of their research question, the authors miss out on an important body of literature on spatio-temporal characteristics of soil moisture fields. It has been shown by numerous studies that while soil moisture generally shows a large spatial variability, individual points maintain their rank while the mean changes. This insight started with the "classical" Vachaud et al. study, but numerous other studies (for instance Teuling et al., 2006, Mittelbach and Seneviratne, 2012, among others) have reported similar behaviour. This behaviour implies a relatively small spatial variability of fluxes, which was explained from a theoretical perspective by Albertson and Montaldo (2003) and explored Teuling and Troch (2005), among others. In effect, based on these studies, the hypothesis could also be formulated more neutral in the form of a null-hypothesis: "A reduced scale mismatch does not lead to LSM flux estimates closer to eddy covariance observations..". The results could subsequently be interpreted as insufficient evidence to reject this hypothesis. In any case the introduction should be changed to include a discussion on the relation between point-scale and large-scale soil moisture. This can then also be used to explain the reference to Franz et al. (2012), which now conflicts with the hypothesis (if CRNS and in-situ soil moisture have already been reported to be similar, how can the authors still expect to find a difference in fluxes?). These comments do not make the research any less relevant, but the phrasing of the hypothesis should be in line with the "state-of-the-art", and not a convenient selection thereof.

ANSWER:
We thank the Referee for this comment. We will expand the discussion and include additional literature as suggested by the reviewer. Based on the reviewer's suggestion, we will also revise our hypothesis accordingly. In addition, please notice that Franz et al. (2012) compared a network of multiple point-scale sensors within a single CRNS. This network consists of multiple profiles of soil moisture estimated by point-scale sensors within the same CRNS footprint.

COMMENT:
In applying CRNS and PS soil moisture products, the authors only consider the shallower PS soil moisture observation in order to comply with the CRNS observation depth. While this makes sense given the goal of the study (scale mismatch and not a comparison of observation techniques), it is not

sufficiently recognized that this introduces an unfair disadvantage to the PS data. There might be important information in deeper PS observations, in particular during soil moisture-limiting conditions, that is not being considered in this study. It is thus crucial to make a clear distinction between the scale aspect and the observation technique, and acknowledge in the discussion in the discussion that PS soil moisture might give better results when all available observations are used. A better alternative would be to redo the analysis using all observation depths in addition to using only the shallow observations, but this would likely require a substantial amount of work.

ANSWER:
We thank the reviewer for this comment. We agree that not using the deeper soil moisture observation can introduce disadvantages to the point-scale measurement in the comparison. Notice, however, that most Ameriflux sites provided only data from shallow sensors with exception of WR (down to 50 cm) and MO (down to 100 cm), respectively. Understanding the impact of deeper soil moisture dynamics is indeed very interesting to be investigated but it is beyond the original scope of our study which focus on understanding more directly the impact of horizontal footprint. We will include a statement to better clarify this in our revised manuscript.

REFERENCES
Franz et al, 2012. Field validation of a cosmic-ray neutron sensor using a distributed sensor network. Vadose Zone Journal 11 (4), doi:10.2136/vzj2012.0046.

Teuling et al, 2009. Parameter sensitivity in LSMs: An analysis using stochastic soil moisture models and ELDAS soil parameters. J. Hydrometeorol., 10(3), 751-765.