Dear anonymous referee #1:

We thank you for the comments that were very insightful to improve our manuscript. We highlighted our replies in blue after each original comment in below. The sentences or paragraphs that were added to the revised manuscript are in red.

General comment

This paper presents a performance analysis using information theory to better understand the characterization potential of the data (TB) and the performance of the inversion algorithms (MCDA). To my knowledge it is a very original approach in the field of application which is targeted and the approach seems to be very relevant. As a naïve reader with regard to the analysis method used, I had a little difficulty to follow the details of the calculation (some quantities would gain to be defined), but the essence of the method is well restored and allows a non-specialist reader to understand the approach. My main criticism lies in the scope of the data used to make the analysis Indeed, 58 data sets corresponding to 58 stations located in the USA are treated independently. It seems useful to me to recall that the MCDA method aims at exploiting the H and V polarizations in order to separate the reflectivity of the soil from the scattering phenomena linked to vegetation and roughness, the latter being represented by the difference between Tbh and Tbv. By working locally, the variability (humidity, vegetation) is only partially taken into account, taking into account only the annual variations which at the scale of a SMAP pixel present small variations. In fact, by limiting ourselves to a stationary analysis, we underestimate the interest of the MCDA algorithm which is applicable everywhere and allows an estimation of humidity whatever the vegetation cover. This leads to find that the quality of the estimates (here seen by the correlation coefficient between the moisture retrieved and the observed moisture) is all the better as the redundancy term is high, a criterion which is proposed for the following analysis of the quality of the algorithm. The interpretation of R could be better described in the material and method and in particular it is important to specify if a good model is characterized by large value of R, meaning that the model outputs and its input data are well interdependent. The largest R is probably found in low vegetation situations where the ranges of moisture and Tb are greatest. This is a known feature and it seems to me that the quality of the MCDA model is more in its ability to represent the diversity of ecosystems and the associated plant formations. Would it be possible to process a data set of all the stations?

Response: We thank the reviewers for these overall constructive comments concerning about this work. Following the reviewer's suggestion, we have partitioned our study sites into different landcovers. The results after partitioned our study sites into different landcovers are shown Table 1 and Table 2 below. We found that Additionally, we switched the 9km SMAP datasets to 36km SMAP datasets to address the comments from another reviewer who would like to know how the choice of different resolution of SMAP products may affect the overall analysis. Therefore, we obtained the 36km SMAP product and we found that the newly obtained 36km SMAP product no longer provides the MDCA soil moisture and is replaced by the Dual Channel Algorithm (DCA) soil moisture with some data updates. Thus, we decided to switch to the newest 9km and 36km SMAP data products. We also included the soil effective temperature (T_{eff}) in the uncertainty decomposition analysis because it constitutes a non-trivial information component of the model. We found that there's no pronounced difference between 9km product and 36km product as shown in Figure 1 below ($p > 0.05$, based on two sample t-test). While there is no set in-stone interpretation of the redundant components, we have expanded our description of this aspect of our study. Generally, it should be interpreted with respect to a specific system. For the SMAP DCA, we found that higher R is an indication of better model performances (better Pearson correlation between *in situ* and DCA soil moisture). Finally, more equations were provided in the revised manuscript regarding how we computed each of the quantity involved into this analysis.

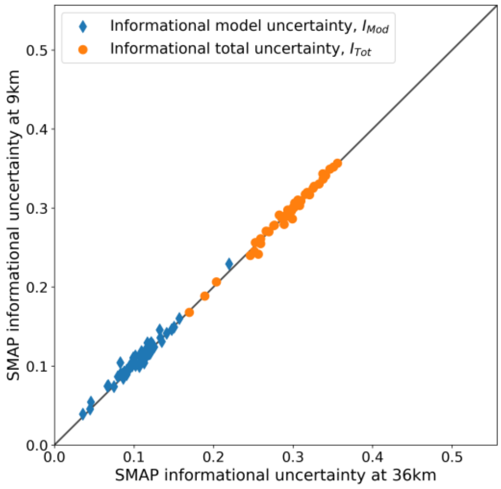

Figure 1 Figure 1. Informational uncertainty comparisons between 36km and 9km SMAP DCA products

Detailed Comments L85 I think that part of uncertainty is due to the scale of the pixel with mixed surface and in situ moisture that is sparsely sampled (here I think it is local measurement) while the moisture is strongly variable within the pixel.

Response: We agree and admit that part of the uncertainty is due to scale mismatch between point measurements of *in situ* and SMAP data product. We have added the following "It is important to acknowledge that we used the point based *in situ* soil moisture as the ground truth in this analysis. Due to course spatial resolution of SMAP products, we acknowledge that *in situ* soil moisture may not be able to represent the spatial averaged soil moisture well. Although the nominal sensing depth of L-band SMAP soil moisture is 5 cm, the penetration depth was found to be even shallower in wetter regions (Shellito et al., 2016). In fact, the L-band sensing depth was found to as little as ~1cm (Jackson et al., 2012) and can be more sensitive to surface meteorological conditions and more random than the actual *in situ* soil moisture. The heterogeneity in each pixel relative to the *in situ* observations together with the sensing depth disparity may negatively influence the results of this study result in an overestimate the actual informational uncertainties. We also acknowledge the existence of upscaling methods for matching the *in situ* soil moisture to satellite footprint (Crow et al., 2012). However, most of upscaling methods are achieved under the assistance of additional reference soil moisture datasets. This process introduces additional pieces of information in the DCA system making the separation of the uncertainty induced by the upscaling algorithm or additional dataset from other informational uncertainties much harder. Additionally, we used the hourly *in situ* data to best match the SMAP DCA soil moisture retrievals in time (within an hour). Therefore, it is hard to find such reference dataset at such a high frequency time domain. Here we consider the informational uncertainty caused by the spatial mismatch and sensing depth mismatch between *in situ* and DCA soil moisture as part of the informational random uncertainty (I_{Rnd}). Because the DCA essential is a mathematical function and does not inherently requires the inputs of a specific resolution. The spatial resolution is often the inherent attribute of the data. The sensing depth is more of imperfection L-band sensor. The reader should also keep these in mind while interpreting and adopting the results in this study." to the methodology to address this aspect."

Eq 5, I suggest to the equation I(TB H or V; Yobs) which is used In equations 7 and 8. It would help me to follow the text

Response: We thank the reviewer for this comment. We did not change the representation $I(T_{Bh}, T_{Bv}; Y_{obs})$ as suggested in the revised manuscript because we wish to follow standard mathematical notation for this quantities. The reason is that the I(A, B; C) represents the information of random variable A and B together (as a set of random variables {A, B}) shared with the random variable C. The notation in the manuscript follows the notation in other information studies and it also follow the convention (Cover and Thomas, 2005) and other information studies in earth sciences (Goodwell and Kumar, 2017a, 2017b). The notation proposed by the reviewer may interpreted differently since the "or" means the information specifically from T_{Bh} or T_{Bv} but it this information should be jointly in the notation of the manuscript.

Eq 9: RMMI is not defined

Response: We thank the reviewer for pointing out this. We now have added definition of R_{MMI} as

$$R_{MMI} = \min(I(T_{Bh}; DCA), I(T_{Bv}; DCA))$$

L209: an explanation how to interpret The quantity in the context of the study. A good model should lead to high or low values of U, R and S. At least for S which is the most commented quantity;

Response: We thank the reviewer for this comment. In the context of this study, we found that R is has the largest values and mostly closely related to the performance of SMAP DCA. Therefore, we conclude that a good DCA model/performances should corresponded to higher values of R. Therefore, it is expected that higher R should also correspondent to smaller values of S, U_h and U_v. We have added "Good DCA model performance (as measured by Pearson correlation between *in situ* and DCA soil moisture) is more likely to be found in locations where the redundant information of brightness temperatures shared with DCA soil moisture is high and is more dominant relative to other components." to the revised manuscript.

L245: I(h, v ; in situ)? rather than I(MCDA, insitu)

Response: We thank the reviewer for this comment. In the original manuscript. Line 245 "The information gap between $H_{CN}(in\ situ)$ and $I(\mathrm{MDCA};\ in\ situ)$ is the overall SMAP uncertainty in which 88% is contributed by the random uncertainty in the systems explanatory variables (Fig. 3)" The abbreviations in the original sentence is correct since the overall SMAP uncertainty is defined as $H_{CN}(in\ situ) - I(\mathrm{MDCA};\ in\ situ)$. In order to avoid this type confusion, the following equations and paragraphs were added to the manuscript

"

$$I_{Rnd} = H_{CN}(in\ situ) - I(T_{Bh}, T_{Bv}, T_{eff};\ in\ situ), \tag{1}$$

$$I_{Mod} = I(T_{Bh}, T_{Bv}, T_{eff};\ in\ situ) - I(\mathrm{DCA};\ in\ situ), \tag{2}$$

and

$$I_{Tot} = H_{CN}(in\ situ) - I(\mathrm{DCA};\ in\ situ) = I_{Rnd} + I_{Mod}. \tag{3}$$

where I_{Rnd} is the informational random uncertainty, I_{Mod} is the informational model uncertainty, I_{Tot} is the informational total uncertainty, $H_{CN}(in\ situ)$ is the entropy of *in situ* soil moisture, $I(T_{Bh}, T_{Bv}, T_{eff};\ in\ situ)$ is the mutual information between horizontally (T_{Bh})- and vertically-polarized brightness temperature (T_{Bv}), $I(\mathrm{DCA};\ in\ situ)$ is the mutual information between DCA soil moisture and *in situ* soil moisture."

L245 and 247: honestly i don't see where 0.88 and 0.12% come from. Not evident to see such values in Fig3

Response: We thank the reviewer for this comment. We have replaced the Figure 3 of the original manuscript with the figure below. A new table that contains these summary statistics is provided (Table 1 below).

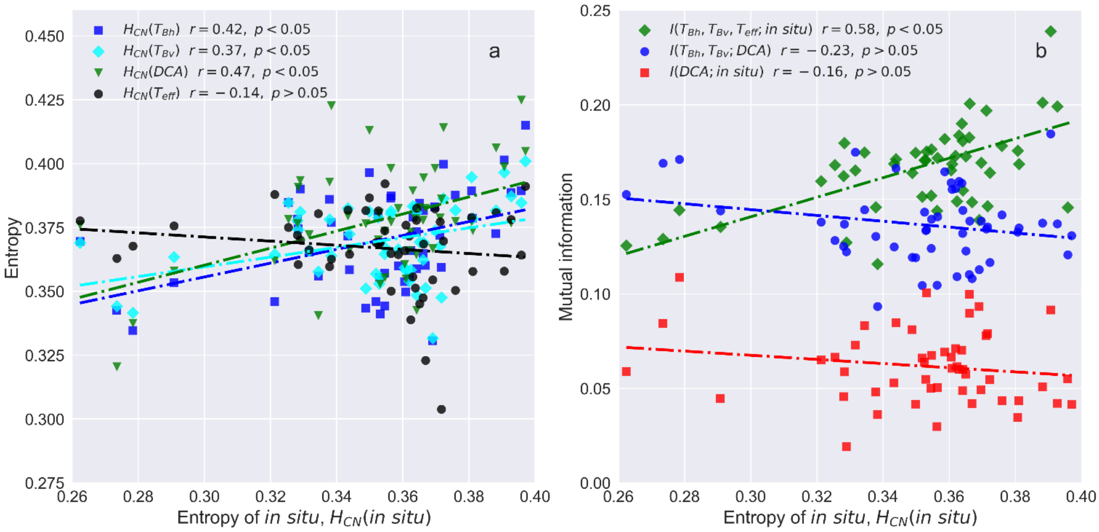

Figure 1. Entropy of *in situ* soil moisture against the entropies of DCA soil moisture, horizontally polarized brightness temperature (T_{Bh}), vertically polarized brightness temperature (T_{Bv}) and soil effective temperature (T_{eff}) (a) and mutual information quantities (b)

Landcover	Informational random uncertainty, I_{Rnd} (and its % of I_{Tot})	Informational model Uncertainty, I_{Mod} (and its % of I_{Tot})	Informational total uncertainty, I_{Tot} (and its % of H_{CN}(*in situ*))
Shrublands	0.22 (69%)	0.10 (31%)	0.32 (88%)
Grasslands	0.20 (62%)	0.09 (38%)	0.29 (83%)
Croplands	0.18 (65%)	0.10 (35%)	0.28 (79%)
Mixed	0.20 (68%)	0.09 (32%)	0.29 (81%)
Overall	0.18 (64%)	0.11 (36%)	0.29 (82%)

Table 1 The amount of informational uncertainties in percentage. The values in the table are the average of each landcover. The values in "Overall" is the average of all the sites.

L251: what are the fraction of model uncertainty

Response: we thank the reviewer for this comment. In the original manuscript, we mean the proportion of model uncertainty to the overall uncertainty. We have corrected this in the revised manuscript.

L261;263 : how I cand tale 0.55 of I

Response: We thank the reviewer for this comment. We have provided a table for these summary statistics in the revised manuscript (Table 2 below)

Landcover	Unique information of T_{Bh} (U_h) (and its % $I(T_{Bh}, T_{Bv}; DCA)$)	Unique information of T_{Bv} (U_v) (and its % $I(T_{Bh}, T_{Bv}; DCA)$)	Synergistic information of T_{Bh} and T_{Bv} (S) (and its % $I(T_{Bh}, T_{Bv}; DCA)$)	Redundant information of T_{Bh} and T_{Bv} (R) (and its % $I(T_{Bh}, T_{Bv}; DCA)$)	Mutual information ($I(T_{Bh}, T_{Bv}; DCA)$)
Shrublands	0.03 (27%)	0.017(15%)	0.03 (26%)	0.036 (32%)	0.113
Grasslands	0.029 (21%)	0.014 (10%)	0.02 (14%)	0.077 (55%)	0.14
Croplands	0.017 (12%)	0.013 (9%)	0.016 (12%)	0.095 (67%)	0.141
Mixed	0.014 (12%)	0.007 (6%)	0.01 (8%)	0.091(74%)	0.122
Overall	0.026 (19%)	0.013 (10%)	0.019 (14%)	0.08 (57%)	0.137

Table 2 The partial information decomposition components. The values in the table are the average of each landcover. The values in "Overall" is the average of all the sites.

L264 Uv likely takes greater value if data from different sites are merged
Response: We thank the reviewer for this comment. The statistics of U_v are shown in Table 2. We found that U_v is consistently the smallest while compared with other components

L266:268: yes but at local scale only. Independence of H and V will be much stronger when different location with different ecosystem are taken into account
Response: We thank the reviewer for this comment. We have extended this analysis to some contrasted landcovers in the revised manuscript.

L303:304 : not only : see comment on L85
Response: We thank the reviewer for this comment. We have added a new paragraph in methodology to address this issue.

L312:315: What are the parameter considered (tau is derived from H an V) here
Response: We thank the reviewer for this comment. We consider the parameter such as vegetation single scattering albedo (ω), surface height standard deviation s etc. We have specified these parameters in the revised manuscript.

L315:317: speculative ? references
Response: We thank the reviewer for this comment. The study from (Konings et al., 2017) has been added as the reference.

L332:332: I am not you can say date. The correlation between H and V is well known, the expected ortogonality is more on V-H and H, that is expressed using various ecosystems. Here we are lacking interpretation key. But correlation between inputs does not means that inputs and output are redundant, which my understanding of R.
Response: We thank the reviewer for this comment. We have removed such statements in the revised manuscript.

L355: making the analysis on individual station is a strong limitation, as MSDA capacity were not fully analysed
Response: We thank the reviewer for the comment. The analysis has been extended to different landcovers in the revised manuscript.

L358:361: speculative (reference – difficult to understands without additional information)
Response: We thank the reviewer for this comment. We found that it may confuse the reader without providing specific information. Therefore, we decided to remove such statements.

L370: I don't what is the HESS policy. It would be better to have codes in open repository
Response: We thank the reviewer for the comment. The python codes and datasets used in this study has been upload to https://github.com/libonancaesar/HESS_Information_Uncertainty.

Figure 2 : remove MI in Y legend
Response: Removed as suggested.

Figure 7d : the y axis of the embedded graph is not described. The interest of th H V correlation is really limited (see comment above). I suggest to remove it.
Response: We thank the reviewer for this comment. The embedded graph has been removed as suggested.

In conclusion beside the minor improvement suggested in my comment I expect the authors: 1) better defining the interpretation scheme of the R S and U quantity 2) extending the analyse to merged data set, or at least a subset gathering sites having contrasted ecosystems. This will give stronger overview of the MCDA models and its interes. This might have an impact on the discussion and conclusion.

Response: We thank the reviewer for this comment. (1) the definition of these components has been defined in the methodology of the revised manuscript (2) Additional analysis regarding different landcovers has been added

References

Cover, T. M. and Thomas, J. A.: Elements of Information Theory, Wiley., 2005.

Crow, W. T., Berg, A. A., Cosh, M. H., Loew, A., Mohanty, B. P., Panciera, R., de Rosnay, P., Ryu, D. and Walker, J. P.: Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products, Rev. Geophys., 50(2), 634, doi:10.1029/2011RG000372, 2012.

Goodwell, A. E. and Kumar, P.: Temporal information partitioning: Characterizing synergy, uniqueness, and redundancy in interacting environmental variables, Water Resour. Res., 53(7), 5920–5942, doi:10.1002/2016WR020216, 2017a.

Goodwell, A. E. and Kumar, P.: Temporal information partitioning: Characterizing synergy, uniqueness, and redundancy in interacting environmental variables, Water Resour. Res., doi:10.1002/2016WR020216, 2017b.

Jackson, T. J., Bindlish, R., Cosh, M. H., Zhao, T., Starks, P. J., Bosch, D. D., Seyfried, M., Moran, M. S., Goodrich, D. C., Kerr, Y. H. and Leroux, D.: Validation of Soil Moisture and Ocean Salinity (SMOS) Soil Moisture Over Watershed Networks in the U.S., IEEE Trans. Geosci. Remote Sens., 50(5), 1530–1543, doi:10.1109/TGRS.2011.2168533, 2012.

Konings, A. G., Piles, M., Das, N. and Entekhabi, D.: L-band vegetation optical depth and effective scattering albedo estimation from SMAP, Remote Sens. Environ., 198, 460–470, doi:10.1016/j.rse.2017.06.037, 2017.

Shellito, P. J., Small, E. E., Colliander, A., Bindlish, R., Cosh, M. H., Berg, A. A., Bosch, D. D., Caldwell, T. G., Goodrich, D. C., McNairn, H., Prueger, J. H., Starks, P. J., van der Velde, R. and Walker, J. P.: SMAP soil moisture drying more rapid than observed in situ following rainfall events, Geophys. Res. Lett., 43(15), 8068–8075, doi:10.1002/2016GL069946, 2016.