Answers to Reviewer #1 comments

We thank the reviewer for the constructive comments on the manuscript. We will detail in our response below how we plan to address the reviewer comments.

General comments

1) The paper is quite long and it contains many figures (17+3) and analyses. The intercomparison between the soil moisture and rainfall datasets is carried out in space and time, considering absolute and anomaly values. The reader can be lost to exactly catch the paper message and the main results. I suggest making an attempt to summarize the analyses made in the paper, For instance, a table listing them might be beneficial. Moreover, a clear distinction between the spatial and temporal analyse should be made.

Reviewer #2 has raised a similar point, asking for a more condensed version of the manuscript. We agree that the manuscript might contain too many figures and that this might distract the reader from the main messages. We think that some of the figures could be removed and will move others to the Annex. In detail, we plan to do the following:

- Remove Fig. #1, Fig.#6
- Combine Fig. #2,#3
- Move Figures #6, #9, #11 to the Annex
- Combine Fig #4 and Fig#5 into a single figure
- Shorten the text where still possible, in particular section 4.3

The reviewer is also asking for an overview Table of the main results. In fact, the manuscript already contains such a table (Table 1) which is summarizing the main potential application areas and caveats of the new soil moisture dataset. In addition we plan to include a table in the discussion section of a revised manuscript that summarize the analysis performed and results obtained.

2) In the paper, the datasets are compared in space and time by using, mainly, a correlation and partial correlation analysis. While I believe that the temporal analysis is appropriate and give insights on the behaviour of the ECVSM dataset, I have some doubts about the spatial analysis (shown in Figures 6, 9 and 10). As underlined by the authors (page 3545, lines 8-10 and page 3563, lines 4-13), the current version of the ECVSM dataset (likely to be changed in the near future) uses the GLDAS as a common scaling reference. Therefore, the global maps of the mean, median or percentiles values simply reflect the corresponding values of the GLDAS. Therefore, their intercomparison with the reanalysis and modelled data might be not correct as it does not reflect the spatial behaviour of the ECVSM dataset. I suggest clarifying this aspect and, possibly, to reduce the space given to this analysis.

Alternatively, it could be interesting to investigate the evolution in time of the spatial correlation between the different datasets, e.g. at a global scale and subdivided by continents or climatic regions. In my opinion, this analysis could give the correct representation of the capability of the ECVSM dataset to reproduce the spatial variability of soil moisture at large scale.

The reviewer is addressing here two different points. The first is the question, how the ECVSM spatial patterns correspond to the spatial patterns of other soil moisture datasets used in the study. One of the conclusions of the paper is that the soil moisture statistics (e.g. mean field) is largely dominated by the GLDAS soil moisture distribution due to the rescaling procedure which is discussed in the paper. The reviewer states that "*it does not reflect the spatial behavior of the ECVSM dataset".* In fact, it does

reflect the ECVSM data product spatial behavior, but this variability is dominated by the GLDAS field. We agree that the results might differ, if the original input datasets (passive/active product) would be used, but this was not the objective of this study as we here analyze the ECVSM data product (which is the official product provided to the scientific community) and not its input datasets. The motivation for the spatial comparison of the percentiles is rather simple. From a climate modeling perspective, it is of largest interest, if the general dynamic is captured well by the model, both in space and time. This can be analyzed using the approach chosen. While we have already emphasized the problem with the GLDAS matching, we will emphasize this also in the discussion of the percentile distribution results.

The reviewer then suggests to perform an alternative analysis, focusing on the temporal evolution of the spatial correlation in different regions. This is a very good idea and we have explored this idea. Figure 1 shows the temporal evolution of the correlation between ECVSM and JSBACH soil moisture for different regions, like used by Brovkin et al., 2013. Except for the fact that the correlation is seasonally dependent in some regions, one can only observed major differences between different regions on the globe. This pattern in the correlation coefficient is presented in more detail in the Figures which are already included in the manuscript. We have therefore decided to not include these additional results as we think that they don't provide additional insight and would result in a more comprehensive, rather than more condensed manuscript.

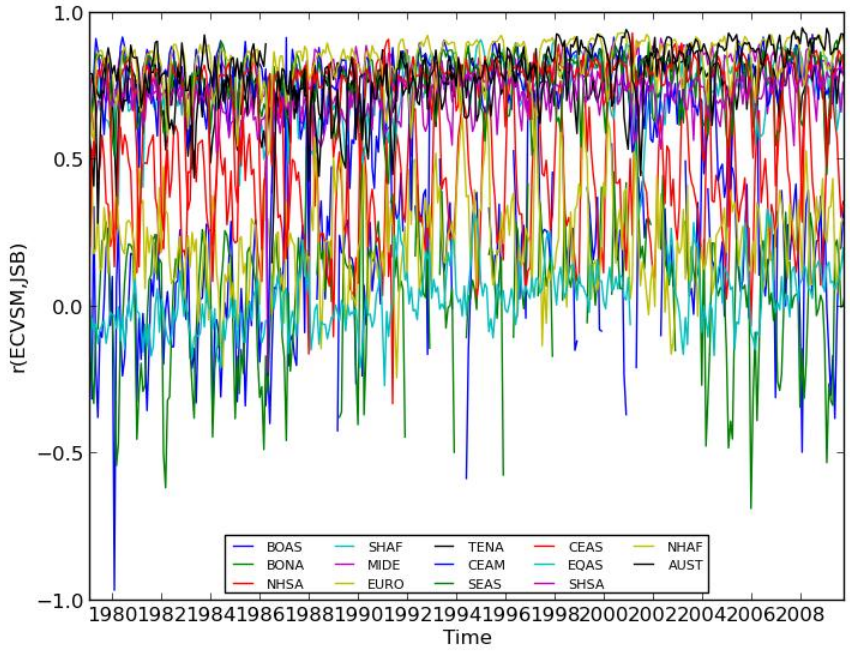

Figure 1: Temporal evolution of spatial correlation pattern between ECVSM and JSBACH soil moisture for different regions like defined in Brovkin et al., 2013

Reference:

- Brovkin, V. et al., 2013. Evaluation of vegetation cover and land-surface albedo in MPI-ESM CMIP5 simulations. Journal of Advances in Modeling Earth Systems, p.n/a–n/a. Available at: http://www.agu.org/pubs/crossref/pip/2012MS000173.shtml

3) One of the main outcomes of the paper is that the agreement between the ECVSM dataset and the JSBACH modelled data is better than the one with the ERA-Interim dataset. This is likely due to some inconsistencies in the ERA-Interim soil moisture data that sometimes show different range of variability in the different years. For this issue, the ECMWF has delivered a new soil moisture dataset, called

ERA-Land, to be employed for long-term analysis (as the one in this paper) that avoids the inconsistencies found in the ERA-Interim dataset. I suggest mentioning this aspect as (possible) additional explanation of the obtained results.

We thank the reviewer for this comment and will discuss this point in a revised version of the manuscript and will give reference to the new ERA-Land dataset (Balsamo et al., 2012, Albergel et al., 2013)

Reference:

- Balsamo, G., C. Albergel, A. Beljaars, S. Boussetta, E. Brun, H. Cloke, D. Dee, E. Dutra, F. Pappenberger, P. de Rosnay, J. Muñoz Sabater, T. Stockdale, F. Vitart (2012): ERA-Interim/Land: A global land-surface reanalysis based on ERA-Interim meteorological forcing. ERA Report Series (12). http://www.ecmwf.int/publications/library/do/references/show?id=90553
- Albergel, C. et al., 2013. Skill and global trend analysis of soil moisture from reanalyses and microwave remote sensing. J. Hydrometeorol., doi: 10.1175/JHM-D-12-0161.1

Specific Comments (P: page, L: line or lines)

P3542, L7-8: Actually, in the paper a detailed comparison of the ECVSM dataset with modelled data from the ERA-Interim re-analysis dataset and the JSBACH land-surface model is shown. The potentials and limitations of the ECVSM dataset for climate modelling applications is not the main content of the paper. Please rephrase the abstract, even simply changing the order of the sentences.

The paper analyzes the ECVSM dataset together with ERA-interim and a land surface scheme used in a state-of-the-art Earth System Model. The perspective how we analyze the data (monthly,global,anomalies) is mainly driven by the needs to evaluate a land surface scheme on average: Is the seasonal cycle well captured? Is the general response to precipitation well captured? What is a characteristic spatial and temporal dynamic?. Are pronounced climate anomalies captured?

We agree with the reviewer that a general application for climate modeling might be broader than in the present study (e.g. model parameterization, model initialization). The paper is however addressing major aspects of climate model evaluation using observations. We will clarify this difference in the abstract of a revised version of the manuscript.

P3542, L9: Only the land-surface model JSBACH is used in this study, not the climate model. I suggest correcting, also later in the text, to avoid misunderstanding.

The reviewer is right that the full coupled (land/atmosphere/ocean) model is not used in the present study. This is however clearly detailed in section 2.1.2 where we describe the general setup and say that the model is forced with the WATCH forcing data set only. We will emphasize in a revised version that the study is **not** dealing with coupled model simulations (no feedback to the atmosphere).

P3546, L12: (Hagemann and Stacke, 2013) is missing in the references list.

Thanks for this comment. We will consider it in a revised manuscript.

P3546, L14: I guess that the first soil layer is used in the subsequent analyses but it should be clearly specified. The same should be done for the ERA-Interim soil moisturedata.

The reviewer is right that the first soil layer is used for comparison. We will clarify this for JSBACH and ERA-interim in a revised manuscript.

P3550, L14-16: The rationale for removing the trend of the timeseries (explained later in the text) should be specified also here.

We thank for this suggestion and will change the manuscript accordingly.

P3551, L17: In the analysis of percentiles, only the spatial analysis is carried out. Please remove "temporal".

We agree with the reviewer that this formulation might be misleading, as the main focus of the analysis is on the similarity of the spatial pattern. What we mean by “temporal” patterns is, that the percentile distribution for each grid box is compiled based on all times. The shape of the resulting pdf therefore provides information on the statistics of the temporal dynamics at each gridpoint.

We will clarify this point in the revised version of the manuscript by rephrasing the sentence.

P3557, L10-11: The employed precipitation datasets represent rainfall plus snowfall or only rainfall? Please specify.

It is precipitation + snowfall. This will be clarified.

P3562, L23-24: As specified in the General Comments, I am not sure that the spatial analysis with the percentiles represents a useful approach to evaluate the ECVSM dataset.

We have responded already on this point above (#2).

P3563, L1-3: The results of the correlation and partial correlation analyses between the ECVSM, JSBACH and ERA-Interim soil moisture anomalies show very similar results. The mean correlation values are equal to 0.36 (0.41) and 0.15 (0.23) for the correlation and partial correlation analyses, respectively, in the comparison of ECVSM and ERA- Interim (JSBACH) datasets. Please rephrase this part.

We thank the reviewer for this comment and will rephrase this part in the paper as suggested by the reviewer.

P3575, Fig.1: Please change "WFD" with "WFDEI" as in the text WFD refers to ERA40 reanalysis data.

Thanks, we will do so.