Answers to Reviewer #2 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.

Major Comments

It is not very clear to me why the partial correlation analysis is performed against precipitation or how to interpret the partial correlation values. Precipitation is the major driver for soil moisture dynamics and if the information from precipitation is completely removed, there won't be much left. I have some doubts on how effectively the removal is (see specific comments). Even the precipitation factor can be well removed in partial correlation, neither ERA-interim nor JSBACH uses GPCP precipitation – so what is exactly removed from them? Or the authors are looking for the "added" value of ECVSM on top of precipitation products?

The motivation of the partial correlation analysis is to see whether the ECVSM dataset shows skill in representing soil moisture dynamics that goes beyond the major driver, which is precipitation like the reviewer correctly stated.

If one compares some signal with a seasonal component, one already gets good "model" skills by using simple zero-order approaches (e.g. a sine curve). Thus the overall idea is to remove the major driver (precipitation) and then look, if the remaining data still shows some correlation. In this case we assume that ECVSM provides additional skill in monitoring soil moisture dynamics. The control variable, which has been removed, is the precipitation dataset that was used for forcing the models. This was clearly stated in the manuscript as:

"The simulated soil moisture fields were correlated against ECVSM using partial correlations where the effect of the precipitation forcing was removed (control variable). For ERA-interim, the ERA-interim precipitation was removed while WFDEI precipitation was removed for JSBACH simulations. As the true precipitation is unknown, the GPCP dataset is assumed to be the precipitation dataset which best captures the temporal and spatial precipitation dynamics. It is therefore used as a control variable on the ECVSM soil moisture fields for the partial correlation analysis."

In a revised version we will more clearly describe the motivation for the partial correlation analysis.

I think the authors should perform time breakdowns for some of the correlation and trend analyses to show how different satellites, though somehow homogenized, agree (or disagree) with each other and how those differences affect (or contribute) to the overall conclusions. For example, break it down to the periods where SMMR, SSM/I and AMSR-E tend to be the major contributor at the time. Multiple satellites may create artifacts in the long-term analysis so this is especially important for studying the potential for "climate model evaluations".

We agree in general with the reviewers comment that a temporal breakdown would be an interesting exercise. Dorigo et al. (2013) have compared the ECVSM product performance compared to in situ measurements and showed comparable performance of the harmonized product compared to the individual data products from different time periods. However, the accuracy varies over time and the actual reasons for the varying product accuracy is currently under investigation (Dorigo, pers. comm.).

The major objective of the present study was to assess the general applicability of the final ECVSM data product and not its individual components. We think that this is interesting, but beyond the scope of this paper, as it would add much additional content to the paper. Like the reviewer is saying in his next comment, the present version of the manuscript is already lengthy and contains a lot of figures and results. We would like to further condense the manuscript instead of widening the scope of the

paper. We have therefore decided not to add the additional analysis suggested by the reviewer to the present paper, but plan to conduct the analysis with temporal breakdowns and publish these results in addition with other ongoing analyses in a separate publication. Besides, a manuscript is currently in review for publication in Remote Sensing of Environment (Albergel et al., 2013) where the product performance of ECVSM is monitored over time (also focusing on the individual blending periods) using ERA-Land soil moisture estimates.

References

- Albergel, C., Dorigo, W., Balsamo, G., Muñoz-Sabater, J., de Rosnay, P., Isaksen, L., Brocca, L., de Jeu, R., & Wagner, W. (in review). Monitoring multi-decadal satellite earth observation of soil moisture products through land surface reanalyses. Remote Sensing of Environment, In review

There are too many figures and a lot of them contain multiple sub-figures. Please consider condensing the results (try to stress on those leading to major conclusions of the paper) and delivering the messages in a more powerful way.

Reviewer #1 had a similar comment. We agree with the general comment and will revise the manuscript accordingly. Please see response to Reviewer#1 question 1) for details.

Specific Comments

P3543, L27: spell out the acronym ECVSM if possible.

ECVSM stands for "Essential Climate Variable Soil Moisture". We will clarify this on its first occurrence in the revised manuscript.

P3545, L8-9: ". . . statistics . . . resembles . . . dynamics . . ." consider rephrasing it.

We will modify the sentence as:

… statistics… product is generated by the soil moisture dynamics …

P3545, L10: what is "percentile distribution"? I thought the percentile distribution is always uniform. Can you clarify?

What is meant are the percentiles of the probability density function. For each grid cell, the frequency distribution is calculated from the entire time series and the percentiles are estimated. We will clarify this in the revised manuscript.

P3545, L17: is RMSE 0.05 [m3m− 3] calculated at daily or monthly scale? Besides RMSE, any correlation number reported against ground observations?

The study of Dorigo et al. (2013) estimated the RMSE based on daily values. Mean correlations are in the order of 0.5 but show a strong variation among sites. For further details, the reader is referred to the paper (in revision). We will add the above details to the manuscript.

We will also replace RMSE by RMSD, as this is more appropriate as the true value is not known.

P3545, L6: spell out JSBACH if possible. There are similar cases elsewhere.

As the original acronym is not longer valid, it was decided to use JSBACH as a name of the MPI-ESM land surface scheme. Consequently, no acronym can be provided.

We agree with the reviewer that soil moisture signal is predominantly a damped precipitation signal. There have been a multitude of approaches to simulated soil moisture evolution as function of precipitation data only, like the API model or the exponential filter (Wagner et al., 1999; Albergel et al., 2008). We therefore agree that at short timescales a linear correlation is not likely and linear correlation measures might be not so suitable. However on monthly timescales, the seasonal cycle is predominant in both variables and the relationship between soil moisture and precipitation can be approximated by a linear relationship, as shown in the following figure:
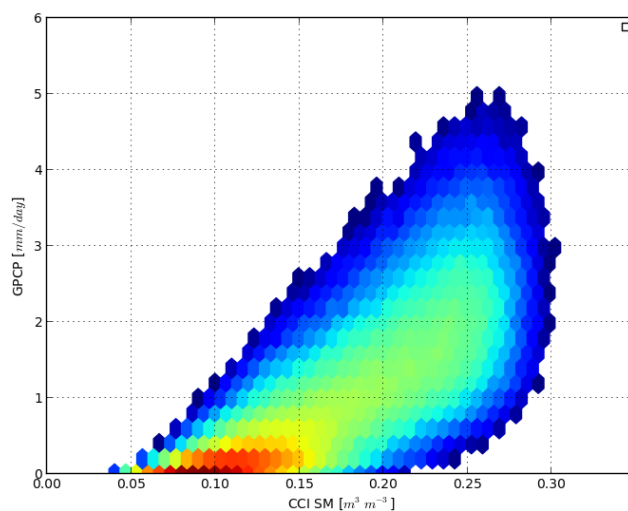

Figure 1: Joint frequency distribution of monthly CCI soil moisture and GPCP precipitation. Note in general linear relation between the two variables on monthly timescales. Highest frequencies are found for very low precipitation P<<0.2; Note that very small precipitation rates are likely to be very uncertain.

The idea of the reviewer to transform the precipitation signal into a soil moisture signal by means of the API is very much appreciated. It would allow for an easier comparison of precipitation and soil moisture data at <u>short</u> (e.g. daily) timescales. We see however two major problems in following this approach for the present study: A) the API requires input data at high temporal resolutions to result in meaningful results. Input data on monthly timescales, like those used in the present study, are not useful. B) API requires some parameterization for the loss of water due to percolation and evaporation. Simple approaches have been developed for that purpose (e.g. Crow et al., 2007) which are however not taking into account regional differences in the loss of water in the soil due to evaporation processes.

We therefore think that for the present study an application of the API model would not make sense (global analysis, monthly mean data), but will use that approach in future daily data analysis. At a first approach the linear correlation measures are considered to be appropriate due to its linearity on monthly timescales as shown above.

- Albergel, C. et al., 2008. From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in-situ observations and model simulations. Hydrology and Earth System Sciences Discussions, 5(3), pp.1603–1640. Available at: http://www.hydrol-earth-syst-sci-discuss.net/5/1603/2008/ [Accessed January 21, 2013].
- Wagner, W., Lemoine, G., Rott, H. (1999). A method for estimating soil moisture from ERS scatterometer and soil data. Remote Sens. Environ., 70, 191-207
- Crow, W.T. & Zhan, X., 2007. Continental-Scale Evaluation of Remotely Sensed Soil Moisture Products. IEEE Geoscience and Remote Sensing Letters, 4(3), pp.451–455.

P3552, L1: "percentile correlation" is usually referred as "Spearman's rank correlation" or "Spearman's ρ" in other literatures.

The Spearman correlation coefficient is defined as the Pearson correlation coefficient between the ranked variables. The percentile correlation referred here is the Pearson correlation of the spatial field of a certain soil moisture percentile (see maps on Fig. 9). This is different from the definition of the Spearman rank correlation coefficient.

P3553, L10: 1980ies -> 1980's

Thanks, we will revise this.

Fig 1, 6, and 9: maps too small and hard to read. Suggestions – the text font can be a little smaller, white gaps in between maps can also be reduced, no need to repeat and same color bar 3 times and reduce the number of color bars from 6 to 2, part of the map captions can be labeled within the map over the ocean part, ...

Thanks for these suggestions. We will optimize the representation of all figures in a revised version of the manuscript.

Fig 8: text is very hard to read.

This is actually a problem with the HESSD page size. On a normal letter sized page, the figure size is o.k., but here it is squeezed too much. We will modify the figure to be more suitable for the journal's format.

Fig 8, 11, and A1: colors reverted? According to the color bar, correlation values on the maps are mostly in the negative range?

We thank the reviewer for this comment, as this is actually an error in the edited version of the manuscript. The colorbar in these figures is represented by an own subfigure. Here we used the wrong order of the colorbar. We will correct the colorbars in the figures accordingly.

Fig 9: should "Soil moisture percentile [m3m− 3]" be "Volumetric soil moisture [m3m− 3]"? In all related figures, ECVSM is labeled as ECV_SM. Please make it consistent throughout the paper. In all related figures and discussion, the letter "p" is used for both "soil moisture percentile" and "p-value in significance test". I suggest the authors spell out "p-value" in full.

- We will rephrase the caption to "Volumetric soil moisture [m**3/m**3] percentiles for …"
- We will ensure to make the acronym ECVSM consistent throughout the paper
- We will ensure consistency of "p-value" in the text. The only case where the percentiles are used is Fig.9, where we use the uppercase 'P', but we agree that this might lead to confusion and will clarify this in a revised version of the manuscript.