Dear Prof. Wolfgang Wagner,

We thank you very much for the constructive comments and suggestions on our manuscript (hess-2016-617). We have responded to all the questions raised. In the following pages are our point-by-point responses. We hope that our responses are clear enough to all of your questions. Thanks for your consideration, and we are

Sincerely yours

Xiaodong Gao, Xining Zhao, Luca Brocca, Gaopeng Huo, Ting Lv, Pute Wu

----

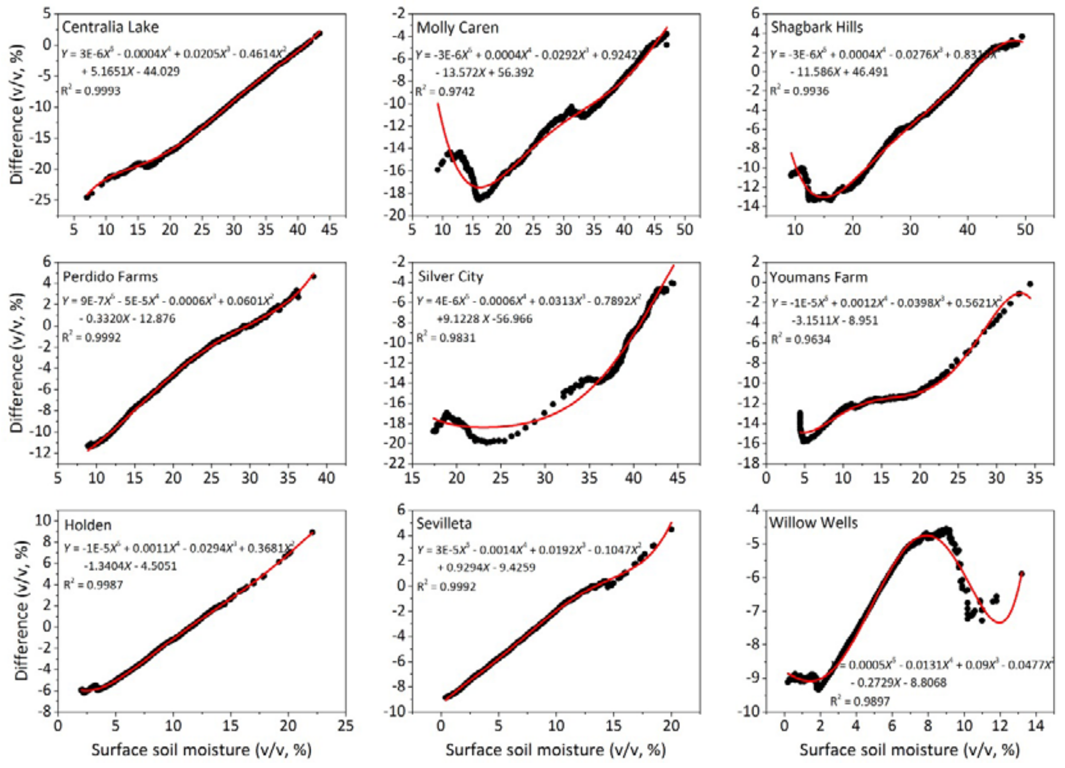
Responses to comments

In this study the authors investigated the potential of the Cumulative Distribution Frequency (CDF) matching method to predict the profile soil moisture (PSM) content from surface soil moisture (SSM) data. They used in situ soil moisture data collected at different depth at several SCAN stations (apparently only from 12 stations rather than from 31 stations as described in the manuscript). While I have no doubts that CDF matching may give good results - and under certain circumstances even very good results - achieving R^2 values that are consistently larger than 0.9 is in my view unrealistic. Looking at Figures 7 to 9, I also do not see how this could work. Consider, for example, station Molly Caren as shown in Figure 7. The CDF matching function of this station should be more or less monotonic (judging from the left CDF plots) which means that for any given SSM value there should only be one corresponding PSM value. However, as shown in the middle plot, for a SSM value of 16 m3/m3 there are multiple PSM values anywhere in the range between about 19 m3/m3 (close to 11/1 of the second year) and 45 m3/m3 (close to 1/1 of the second year). Overall, as long as the CDF matching function is near-monotonic (even though highly non-linear) one should be able to visually match the timing and relative magnitude of fluctuations in the SSM and PSM time series. This is however not possible in many instances in Figures 7 to 9 (neither in the calibration- nor the validation period). Maybe I missed an important point in the description of the methodology.

>> Sorry for confusion. For clarity, we give more details of the methods in the corresponding responses of the questions. Here are our explanations to the raised issue in this comment.

First, the cumulative distribution frequency (CDF) is generally monotonic for one dataset according to its definition. In statistics, it is needed first to count the frequency of each score or score interval, and then calculate the cumulative frequency. Particularly, soil moisture is a state vector. Even several soil moisture contents have the same values they correspond to different dates and also may have different directions (decreasing or increasing) in time series. Therefore, in the CDF matching method, all of the soil moisture values are ranked and same values correspond to different cumulative frequency (Please see the attached excel files; Attachment 1). This can be also found in other studies with respect to the CDF matching method, for example, Drush et al. (2005), Han et al. (2012) and Gao et al. (2013).

Second, the key of this approach is using the observation operators (polynomial) built by the CDF matching method to adjust the difference between soil moisture in surface layer and profile. The observation operators used here are five-order polynomial (for reasons please see our responses in pages 7 and 8) and thus are highly nonlinear (Please see the figure below). Therefore, the monotonicity of CDFs is not directly related to how well the deviation can be eliminated between surface and profile soil moisture content, and the point lies in the robustness of observation operators.



Finally, we are sorry because there was something wrong in the original method of CDF matching which resulted in the unrealistic good agreement between measured and predicted profile soil moisture. We have corrected the method and presented the details in our responses in page 7. Furthermore, we also give the files which contain the detailed processes of profile soil moisture prediction for all of the nine stations under the three different climates. Please find the attachment (Attachment 1).

References

Drusch, M., Wood, E.F., and Gao, H.: Observation operators for the direct assimilation of TRMM microwave imager retrieved soil moisture. Geophys. Res. Lett. 32 (15), 2005.

Han, E., Heathman, G.C., Merwade, V., and Cosh, M.H.: Application of observation operators for field scale soil moisture averages and variances in agricultural landscapes, J. Hydrol., 444-445, 34-50, 2012.

Gao, X., Wu, P., Zhao, X., Zhou, X., Zhang, B., Shi, Y., and Wang, J.: Estimating soil moisture in gullies from adjacent upland measurements through different observation operators, J. Hydrol., 486, 420-429, 2013.

Nonetheless, considering that there are many more problems and open questions with this paper (as identified by the first reviewer, Na Li, and below), I do not see this study fit for publishing.

>> We have responded to all the questions raised by Prof. Na Li. We hope these responses are clear and reasonable. Please see the attached response letter.

SOME FURTHER COMMENTS

The term “upscaling” is usually used in a different context. Please avoid it.

>> We agree. It has been changed into “depth scaling” in both title and text.

Page 2, line 6: Confine to “microwave remote sensing”

>> We agree. It has been edited in the text.

Page 2, lines 12-13: What is he difference between “statistical” and “computational statistical”?

>> We learn this word from Wikipedia. According to Wikipedia, *computational*

statistics, or statistical computing, is the interface between statistics and computer science. Computational statistics may also be used to refer to computationally intensive statistical methods including resampling methods, Markov chain Monte Carlo methods, local regression, kernel density estimation, artificial neural networks and generalized additive models. https://en.wikipedia.org/wiki/Computational_statistics. For clarity, only the term "statistical" were retained.

Page 2, line 18: How to you define "robust estimates" here in this context?

>> In this context, the robust estimates mean that these data assimilation methods apparently improve prediction accuracy of profile soil moisture compared to open loop modeling. But the sentence including this term has been deleted in order to compress this part according to the suggestion of the other reviewer.

Page 2, line 20: Explain in which sense data assimilation is the "most promising approach".

>> We agree. Data assimilation is the most promising approach in predicting profile soil moisture by combining satellite-based soil moisture with those from land surface models. But the sentence including this term has been deleted in order to compress this part according to the suggestion of the other reviewer.

Page 3, line 10: Rather than saying that "the time stable depth is not necessarily the surface layer" one should not that the stability/persistence of soil moisture increases with the layer depth.

>> We agree that soil moisture time stability increases with soil depth which has been reported by several literatures. The time stability method in Hu et al. (2014), however, means to characterize the statistical relationship between soil moisture at different depth intervals and profile soil moisture, and then identify the most time stable depth at which soil moisture can be used to represent profile soil moisture. This relationship can be expressed by using the equations as follows:

$$\delta_{ij} = \frac{\theta_{ij} - \bar{\theta}_j}{\bar{\theta}_j} \quad (1)$$

$$\bar{\delta}_i = \frac{1}{N}\sum_{j=1}^{N}\delta_{ij} \quad (2)$$

$$\sigma(\delta)_i^2 = \frac{1}{N-1}\sum_{j=1}^{N}(\delta_{ij} - \overline{\delta}_i)^2 \quad (3)$$

where θ_{ij} and $\overline{\theta}_j$ represent soil moisture at the layer i and profile-mean soil moisture, respectively, at time j; δ_{ij} and $\overline{\delta}_i$ represent relative difference and mean relative difference, respectively; $\sigma(\delta)_i$ represents the standard deviation of relative difference; and N is the number of sampling occasions. A soil layer is identified as the most temporally stable layer when it has the lowest value of $\sigma(\delta)_i$. However, the value of $\sigma(\delta)_i$ does not necessarily decrease with soil depth.

Hu, W., and Si, B.C.: Can soil water measurements at a certain depth be used to estimate mean soil water content of a soil profile at a point or at a hillslope scale, J. Hydrol., 516, 67–75, 2014.

Page 3, line 17: Computational efficiency is not a problem for most analytical methods. Also in data assimilation one can find efficient workarounds if necessary. Hence, this is not an argument in favor for statistical methods. The same applies for the second argument ("wide range of environments") as statistical methods are at least as difficult to transfer to other environmental conditions as more physical approaches.

>> We agree. We have changed the words with respect to the computational efficiency of analytical and data assimilation methods and the application range of statistical methods.

"The analytical methods require fewer input parameters and are computationally more efficient than data assimilation methods. …. In addition to the two above methods, statistical models are also introduced to do depth scaling of soil moisture due to its simplicity and are completely data driven. …. However, the existing statistical methods usually defined surface soil deeper than 20 cm even down to 40 cm which is far beyond the scope of satellite sensors. This restricts the application of statistical methods to profile soil moisture estimation because in many cases only surface measurements (≤ 5 cm) are available. Despite the existing deficiency, robust statistical methods are still appealing in predicting profile soil moisture because of their simplicity and applicability to a wide range of environments."

Furthermore, “applicability to a wide range of environments” here means that statistical models can be built easily under various environments, such as the observation operators here, although its parameters needs calibration in other environments as is true for both analytical and physical methods.

Page 4, line 12: For the purpose of this study, 12 (31) stations are by far not.

>> Sorry for this confusion. We have noticed this problem as the discussion paper was online. Thereafter, we have uploaded an erratum to correct this mistake in December 12, 2016.

Page 4, line 16-18: Please describe the methods for outlier removal in more detail. The two methods you mention (check for rainfall events and fluctuations in adjacent layers) may have a large impact on the results.

>> We agree. In normal conditions, the fluctuation of soil moisture is a result of precipitation, evapotranspiration and/or groundwater. Furthermore, soil moisture in one layer is usually highly correlated with that in adjacent layers. Soil moisture values can be identified as outliers for the two instances. On the one hand, if soil moisture in one layer clearly increases during some period but no rainfall events occurs before and the soil moisture in adjacent layers do not show clear increase, then the soil moisture values in this layer during this period can be identified as outliers. On the other hand, if soil moisture in one layer clearly decreases whereas soil moisture in adjacent layers do not clear decrease, then these soil moisture values are also identified as outliers. The outliers were then excluded from the analyses.

In fact, we almost detected no outlier in the stations used in this study except for the Little Red Fox (in the resubmitted paper it has be replaced by another station) according to the method in the paper because data managers have carefully check the quality of the soil moisture observations.

Page 5: Methods must be described in much more detail. Show, e.g., CDF matching function and discuss their properties.

>> We agree. More details have been given of building observation operators by using the CDF matching method. Please see the methods in the revised manuscript.

The technical procedure of this method progressed as follows:

(I) The *in situ* measured surface (θ_s) and profile (θ_p) soil moisture datasets were ranked.

(II) Next the differences (Δ) in soil moisture between corresponding elements in the surface and profile datasets were calculated as:

$$\Delta = \theta_s - \theta_p \tag{4}$$

(III) A polynomial fit was then used to quantify the relationship between θ_s and Δ. This study employed pre-experiments to identify the optimal order (details are shown in our response to the next comment), and a fifth-order polynomial was finally used when considering the accuracy of fitting and the principle of parsimony.

$$\hat{\Delta} = k_0 + k_1 \cdot \theta_s + k_2 \cdot \theta_s^{\ 2} + k_3 \cdot \theta_s^{\ 3} + k_4 \cdot \theta_s^{\ 4} + k_5 \cdot \theta_s^{\ 5} \tag{5}$$

where $\hat{\Delta}$ is the predicted difference of Δ, and k_0, k_1, k_2, k_3, k_4 and k_5 are parameters. The polynomial Eq. (5) serves the observation operators here to eliminate the systematic difference between θ_s and θ_p.

(IV) Profile soil moisture could then be estimated by using the observation operators to rescale surface measurements.

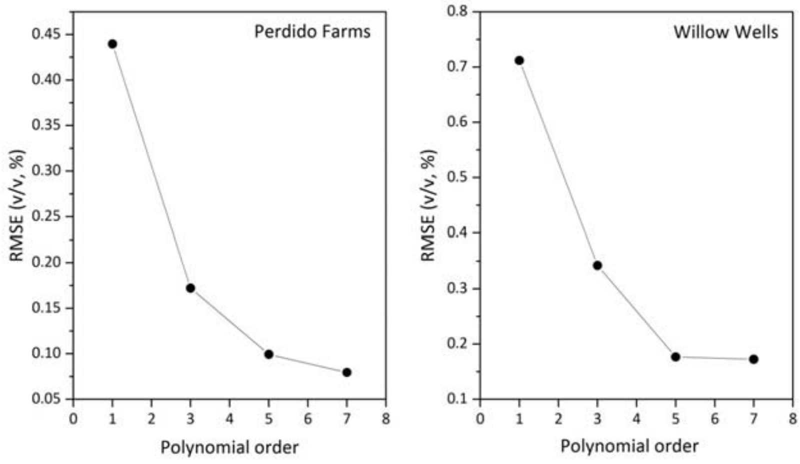
$$\hat{\theta}_p = \theta_s - \hat{\Delta} \tag{6}$$

where $\hat{\theta}_p$ is the predicted profile soil moisture.

Page 5, line 4: CDF matching with a fifth-order polynomial is prone to many problems (overfitting, non-monotonicity, extreme non-linearity). Please justify your choice based on a solid analysis.

>> Here we calculated the root mean square error (RMSE) and R^2 under different polynomial orders to determine the optimal order. According to the figure in page 3, Perdido Riv Farm station shows near-linear relationship between surface soil moisture (θ_s) and the difference ($\Delta\theta$) between surface and profile soil moisture. Willow Wells station, however, shows significantly non-linear relationship between θ_s and $\Delta\theta$. Hence we take these two stations as examples to analyze the effects of polynomial order on the prediction accuracy of profile soil moisture. Here we only show the changes of RMSE with polynomial order because the R^2 had weak changes as order

increases. The figure below shows that the RMSE values decrease clearly as the polynomial order increases from one to five but change slightly at higher orders. Therefore, we used the five-order polynomial fitting in this study.



Page 5, line 14: What do you mean by "was then incorporated"?

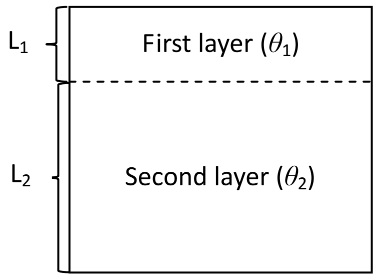
>> Here we first used soil moisture at hourly resolution to build observation operators and these observation operators were applied at daily and weekly resolutions. Therefore, it means that daily/weekly surface soil moisture was used as input to predict profile soil moisture by using the observation operators built by hourly data.

Page 7; equation 5: The first layer is not a "SWI".

>> We agree. It has been corrected in the text.

Page 7, equation 8: This is not the original SWI method. Has it been published by other authors before?

>> This equation just shows how to use soil moisture at the first and second layers to obtain the profile soil moisture. The figure below illustrates it:



$$\theta_p = \frac{\theta_1 \times L_1 + \theta_2 \times L_2}{L_1 + L_2} \tag{7}$$

where θ_p refers to the profile soil moisture content (m^3 m^{-3}); θ_1 and θ_2 refer to soil moisture content at the first and second layer, respectively (m^3 m^{-3}); and L_1 and L_2 refer to the depth of the first and second soil layers (m). It has been edited in the text.

Page 8, lines 6-8: This is obvious and should not be necessary to state in this context.
>> We agree. It has been deleted in the text.

Page 9, line 1: Why do you write "cross correlation analysis"? "Correlation analysis" should suffice.
>> Cross correlation analysis measures the similarity of two time series at different time lags, and this term can be found universally in literatures (e.g., Mahmood et al., 2012; Ford et al., 2014; Guber et al., 2016). But correlation analysis is a broader concept and can include Pearson correlation, cross correlation and autocorrelation among others. Furthermore, we added autocorrelation analysis in the text according to the suggestion below. In order to avoid confusion, the term of "cross correlation analysis" continues here.

References

Mahmood, R., littell, A., Hubbard, K.G., 2012. Observed data-based assessment of relationships among soil moisture at various depths, precipitation, and temperature. Applied Geography 34, 255-264.

Ford, TW, Harris, E., Quiring, S.M., 2014. Estimating root zone soil moisture using near-surface observations from SMOS. Hydrology and Earth System Sciences 18, 139-154.

Guber, A., Su, C.H., Crow, W.T., Zwieback, S., Dorigo, W.A., Wagner, W., 2016. Estimating error cross-correlations in soil moisture data sets using extended collocation analysis. Journal of Geophysical Research-Atmospheres 121, 1208-1219.

Figures 2 and 3: What is the purpose of these two figures?
>> These two figures were used just to show the surface and profile soil moisture time series among the stations. They have been deleted in the revised paper.

Figure 4: Improve figure caption. In addition to showing the correlation between SSM

and PSM for different time lags, you may also have a look at the auto-correlation for both SSM and PSM to better interpret the results.

>> We agree. But we moved the correlation analyses to the part of applications in climate regions where cross correlations of SSM and PSM for different time lags were computed as well as the autocorrelations for each of SSM and PSM for the nine stations under three climates. And these can be used to interpret the results of differences in predictions of PSM under various climates.

Figure 5: The mismatch between calibration and validation period is much too large for RMSE and NS. How can it be that R^2 changes only modestly in comparison?

>> We have recalculated these values by using the corrected procedure of CDF matching method as is shown in the figure below. Here, we used McCracken Mesa station in Utah instead of the Little Red Fox station because significant outliers were identified in the latter station. Furthermore, we employed the mean bias error (MBE) as another metric to judge the difference of observation operators and dropped the Nash-Sutcliffe coefficient because we found that the NSC was highly linearly correlated with R^2 for the updated calculations. As shown in this figure, the R^2 values decreased clearly in validation period.