Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-617-RC2, 2017

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Interactive comment on "Upscaling of soil moisture content from surface to profile: multi-station testing of observation operators" *by* Xiaodong Gao et al.

W. Wagner (Referee)

wolfgang.wagner@geo.tuwien.ac.at

Received and published: 3 January 2017

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 R2 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

Printer-friendly version

Discussion paper

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. 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.

SOME FURTHER COMMENTS

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

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

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

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

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

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.

Page 3, line 17: Computational efficiency is not a problem for most analytical methods. Also in data assimilation one can fine 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.

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

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.

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

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.

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

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

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

Page 8, lines 6-8: This is obvious and should not be necessary to state in this context.

Page 9, line 1: Why do you write “cross correlation analysis”? “Correlation analysis” should suffice.

Figures 2 and 3: What is the purpose of these two figures?

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.

Figure 5: The mismatch between calibration and validation period is much too large for RMSE and NS. How can it be that R2 changes only modestly in comparison?

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-617, 2016.

Printer-friendly version

Discussion paper