

Responses to the comments of Dr. Thierry Pellarin

The manuscript presents a statistic approach to derive Root-zone soil moisture (0-100 cm) from near-surface soil moisture measurements (5cm). This is an important and current topic since remote sensing techniques devoted to measure soil moisture are only sensitive to the top 5 cm soil moisture and most of applications (agriculture, meteorology, hydrology) are interested in deeper soil moisture estimates. The proposed methodology consists in using a CDF-matching procedure applied to the 5 cm soil moisture data to predict 0-100 soil moisture dynamic. To my knowledge, this has not been tested in previous studies. Authors address the problem of three issues: the calibration period-length, the temporal sample, and the climatic effect on the surface/deep soil moisture link. Dataset comes from the US SCAN network. Although the subject is interesting and deserves publication in HESS, the paper presents some deficiencies. For this reason, I would recommend major revisions prior to potential publication in HESS.

>> Thank you very much for your constructive comments on our paper. Here are our point-to-point responses to all of the comments.

General comments:

The main problem of the proposed methodology rely on the fact that it can't be used in ungauged regions (as state by authors in conclusions). In my opinion, this is not totally true and the use of the SCAN database is probably underutilized in this study. There is no interest to use a method which can be used only in regions where we have the truth. However, the paper can be much more improved if a last paragraph is added to discuss about the way to derive the 5 required fifth-order polynomial coefficients from climatic/soil/vegetation characteristics related to other 190 SCAN sites. Depending on the results, this publication can be useful to explain that this solution is probably (or not) the right direction to search.

>> Thanks for this valuable suggestion and we agree. It is interesting to link the polynomial coefficients with the climatic/soil/vegetation properties of the SCAN stations. In this manner, we can obtain the observation operators by using these properties. Here is our way. First, we choose the SCAN stations having at least two-year continuous soil water observations, and then establish observation operators by using the fifth-order polynomial coefficient for each selected station. Second, the

basic and easily measured environmental variables are collected for all of the selected SCAN stations. These variables include, for examples, mean annual precipitation in climate, soil texture fractions (sand, silt and clay) in soil, and aboveground net primary production in vegetation. Finally, the multivariate linear regression (MLR) model is employed to associate these environmental variables (inputs) with each respective polynomial coefficient. And then the five MLR models can be used to estimate the polynomial coefficients. The text has been edited.

The effect of soil freezing is also not addressed in the paper while there is a clear signal in humid continental pixels (Fig.6). Sudden decreases of the surface soil moisture (5cm) in winter periods (around 1/1 and 3/1 in Centralia Lake and Molly Caren for instance) are probably occurring only at the surface and not in deeper layers. This water doesn't disappear contrary to sudden decrease during summer periods. Consequently, the CDF model generate too strong decrease in winter. This point has to be discussed in the paper.

>> Thanks for this valuable suggestion and we agree. Yes, we notice that surface soil moisture and the CDF-matching predictions suddenly decreased in freezing periods under the humid continental climate. This could suggest that soil freezing influences the surface-profile soil moisture relations and the prediction accuracy of the CDF-matching model. Therefore, we suggest that the soil moisture observations during soil freezing period in cold climates should be excluded because they do not represent the true soil moisture content. The text has been edited.

It is necessary to use a coherent formulation for the different soil moisture profile: in-situ soil moisture profile (sometime called "Profile", or "ThetaP"), the CDF profile ("ThetaP" or "Profile (CDF)"), as well as the exponential filter ("wp" or "profile(SWI)", I would suggest Profile(EF)).

>> We agree. They have been edited in the manuscript.

Specific comments

1) US-SCAN network consists of over 200 soil moisture station across US but only 12 of them were used without any explanation of that choice (p.3 line 29). Authors should clarify this point.

>> We agree. There are two primary grounds for our selection. First, it is because a lot

of stations contained considerable missing data at one or several depths especially in humid continental and humid subtropical climates. Second, three stations were used as three replicates for each climate and also for optimizing data resolution and length because three replicates generally are enough in statistics.

2) “The outliers were then excluded from the analyses” (p.4 line 5). Does this mean that the station is excluded from the analysis or only the period concerned with outliers?

>> It means that only the period concerned with outliers were excluded.

3) P.6 line 26. The term “reference” is somewhat confusing. Actually, the exponential filter is not used as a reference method to judge the performance of observation operators but as an alternative method. The reference dataset is in-situ soil moisture. Same remark p.9 line 2 and p. 12, line 2.

>> We agree. The term “reference” in these lines has been changed into “independent”.

4) P.7 lines 15-20. I do agree that exponential filter (EF) provides index values (between 0 and 1) and need to be rescaled using w_{2max} and w_{2min} values. But this solution only required 2 parameters whereas the CDF matching procedure required a large amount of 5 cm soil moisture measurements to derive the 5 coefficient of the fifth-order polynomial procedure. The sentence seems to indicate that these 2 coefficients are a strong limitation of the EF method. Authors should reformulate this paragraph.

>> First of all, the exponential filter method also requires preceding soil moisture data in surface and subsurface layers to obtain the optimal T parameter. The predicted second-layer soil moisture (SWI_2) by the exponential filter method was rescaled by using w_{2max} and w_{2min} just in order to compare with the CDF-matching method which used original moisture values. In most cases (e.g., Wagner et al., 1999; Albergel et al., 2008; Brocca et al., 2011), however, the scaled soil moisture (0-1) is usually used for the exponential filter method and is not needed to do rescaling. Therefore, the two parameters of w_{2max} and w_{2min} would be not a limitation for the EF method in these cases. The text has been edited for clarity.

Albergel, C., et al.: 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, Hydrol. Earth Syst. Sci., 12, 1323–1337, 2008.

Brocca, L., et al.: Soil moisture estimation through ASCAT and AMSR-E sensors: an intercomparison and validation study across Europe, Remote Sens. Environ., 115, 3390–3408, 2011.

Wagner, W., Lemoine, G., and Rott, H.: A method for estimating soil moisture from ERS scatterometer and soil data, Remote Sens. Environ., 70, 191–207, 1999.

5) I don't think Eq. 12 is required as Eq. 11 is supposed to provide a 0-100cm soil moisture content. I expect results of each equation to be quite similar ($L2(=100) \gg L1(=5)$).

>> In the exponential filter method, the second layer refers to the soil depth in the 5-100 cm, but in the CDF-matching method the profile refers to the top 100 cm. Therefore, in our opinion, it is needed to use the Eq. 12 to get the profile soil moisture for the exponential filter although the results are very similar.

6) Section 3.2.1. It is not clear how the three correlation coefficients (0.92, 0.75 and 0.59) can be observed in Fig. 4. Please clarify

>> The correlation coefficient equals to the maximum cross correlation value. The three coefficients of 0.92, 0.75 and 0.59 are the arithmetical means of the maximum cross correlation coefficients in the humid subtropical, humid continental and semiarid climates, respectively. It has been clarified in the text.

7) P.10 lines 18-19 and Figure 9. I can understand why the bias related to the exponential filter method is not equal to 0 since the 0-1 time-series is scaled using the min and max value. However, looking at Fig. 8 (top-middle graph for instance), the green curve does not seem to be scaled to the black one. Therefore, I expect the bias (Fig. 9) to be slightly overestimated.

>> We are not sure whether the non-zero MBE values are necessarily related to the rescaling of the predictions (SWI_2) by using the max and min values for the exponential filter. The difference between the predicted (green line) and observed (black line) profile moisture content for the EF method might be resulted from the assumption that the water flux between surface and subsurface layers is proportional

to the difference in soil moisture between these two layers. In fact, this relationship should be nonlinear with soil moisture content. Furthermore, this assumption may result in relative low Nash-Sutcliffe coefficient (< 0.7 for the majority of stations and even negative) even at the optimal T (T_{opt}) in calibration period (Figure 5). This means that there is a mismatch between the predicted and observed profile moisture content (Figure 6-8).

8) P.11 line 10. I do not agree with the sentence “that only surface (5 cm) soil moisture is needed as the input”. Actually, the method also requires deeper layer soil moisture measurements (up to 100 cm) to derive fifth-order polynomial coefficient of the CDF procedure. Please reconsider this sentence.

>> Here we mean that as soon as the observation operators are built only surface soil moisture is needed as the input to predict profile soil moisture. We have rewritten this sentence for clarity.

Technical corrections

1) Table 2: Authors should indicate in the legend which are the time-serie compared in this Table. Idem in Fig.3

>> We agree. Soil moisture observations in the years of 2014 and 2015 were used for the Shagbark Hills and Sevilleta stations and that in the years of 2013 and 2014 for the Perdido Riv Farms station. It has been edited.

2) Fig. 4, indicate that the 3 top graphs correspond to “humid continental” pixels, and idem for middle and bottom graphs.

>> We agree. It has been edited.

3) Idem Fig.4 remark

>> We agree. It has been edited.

4) Fig.9: Authors should better explain the meaning of “a”, “b” and “ab”

>> We agree. Here the Least Significant Difference (LSD) method was used to do Post Hoc multiple comparisons for the one-way ANOVA. First, one statistical metric (RMSE, R^2 or MBE) was ordered from the highest to the lowest by value among the three climates. Next, the climate having the largest value (C1) was labeled with “a”.

Then C1 was compared with the climate having the medium value (C2). If they were significantly different ($P < 0.05$), the C2 was labeled with “b”, and then C2 was compared with the climate with the lowest value (C3). Similarly, if they were significantly different ($P < 0.05$), the C3 was labeled with “c”. However, if C1 and C2 were not statistically ($P < 0.05$) different, C2 was labeled with “a” as well and C1 would further compare with C3. If C3 and C1 were significantly ($P < 0.05$) different, C3 was labeled with “b”. In this case, C3 should be also compared with C2. If they were significantly ($P < 0.05$) different, the comparison ended; or else, C2 was further labeled with “b” and then the label of C2 became “ab”.

In fact, we had given a short explanation in the title of Figure 9. Because the above explanations can be accessible in statistical textbooks, they would not be indicated in the revisions.

5) P.5 line 17: a word is missing after “then”. Probably “used” ?

>> We agree. It has been edited.

6) P10. Line 30. A word is missing after “clearly”. Maybe “shows”

>> We agree. It has been edited.