Hydrology and
Earth System
Sciences

Discussions

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

N. Li (Referee)

lina@cugb.edu.cn

General comments:

The manuscript by Gao et al. applied a statistic approach (using observation operators built by Cumulative Distribution Frequency matching method) to multi-station in situ soil moisture observations, aiming to predict profile soil moisture from surface soil moisture. They first investigated the effects of temporal resolution (hourly, daily and weekly) and data length (half year in non-growing season, half year in growing season, one year, two years and four years) on the performance of observation operators. Based on the investigation, daily soil moisture data with two-year duration was then used to test the robustness of observation operators, illustrated in three primary climates (humid con-

tinental, humid subtropical and semiarid) of the continental USA. They also compared estimation results with those obtained by the exponential filter method to present the effectiveness of their approach.

Evaluation: The article addresses an important topic in vadose zone research and agricultural management, since such predictions are increasingly being done as the recent advances in soil moisture measurement technology, such as a range of ground-based sensors and remote sensing, providing unprecedented opportunities for mapping moisture dynamics on the soil surface. The Results & Discussion section was described in a clear and concise manner, and the results were presented and illustrated equally good. However, there are some major deficiencies which have to be seriously considered. I will report them in the major comments section. In my opinion, the article deserves publication in HESS however after substantial revision and for this reason I propose major revision.

Major comments

First of all, the structure of the introduction section was in my opinion not good. They have to be accurate and should be focused on their main point.

1) The paper applied a statistic approach to predict profile soil moisture from surface soil moisture. A variety of approaches for the prediction exist, as reviewed by the authors. The authors spent quite a bit of paper (two paragraphs) on the other two kinds of methods - data assimilation methods and analytical methods, which do not directly relevant to this study (except the exponential filter). It is good that the authors cite other's work, however, this could be compressed. Still maintain the citations, but compress the explanations for example.

2) Further, the authors spent one paragraph to list several statistical approaches such as linear and multivariate regression methods and pointed drawbacks of these methods. Two points should be explained: Is these methods directly relevant to the DCF matching method employed in this paper (if not, this paragraph also needs to be com-

pressed); Does the CDF matching method employed by the authors overcome those drawbacks? In any case, I suggest the authors pay more attention to the development and application of the CDF methods.

3) With regard to the CDF match method, I didn't really understand what other studies exist which deal with the same or similar topic (application of CDF matching method to predict profile soil moisture from surface moisture) and in what way this study is different and/or better. This is very crucial for the impact of the article. Furthermore, if similar studies exist I believe that the authors should point out what are the benefits of this method in comparison with existed studies.

4) Further, what is the scientific merit and what is the main contribution of the this study?

5) The authors stated that one of the advantages of the statistical methods is their computational efficiency compared to other two kinds of methods (line 15 in page 3). But this is confused and seems not consistent with previous statement that the analytical methods require fewer input parameters and are computationally efficient. Did the authors compare the computational express between the exponential filter and the CDF methods? Additionally, the first sentence (lines 15-18 in page 3) should be rewritten and more accurate. Actually, in my opinion, it is better to move this sentence to other place, e.g., prior to the introduction of the statistical methods.

6) Lines 28-30 in page 3, to my understanding, this study applied the CDF method to 12 stations respectively. Although they chose these stations on the basis of some differences between them, cross-relations between these stations are not considered. Thus, the sentence (lines 28-30 in page 3) should be rephrased to avoid confusion.

My second concern has to do with the methodology which they used.

1) The CDF matching method developed and applied in this study was not well explained. The authors just described the technical procedure (lines 25-30 in page 4)

with a concept map in figure 1, it is better to demonstrate the method in detail. I noted that the authors presented the detailed formula of the exponential filter method instead.

2) Further, as stated by the authors, the CDF method was first calibrated and then validated in different time period, but which parameters and functions are calibrated? This could be explained in section 2 (Methodology) and the calibration results could be presented in section 3.

3) With regard to the outliers excluded (Lines 13-18 in page 4), some values of the moisture in one given layer were identified as outliers when their variations were inconsistent with values at adjacent depths and rainfall events. Generally, lagged relations exist between the rainfall events and the variation of the soil moisture, especially the subsurface soil moisture.

4) Is it a novelty to use fifth-order polynomial instead of third-order. And is this the only difference compared with previous CDF matching method (Lines 2-5 in page 5)?

5) Section 2.2.3, did the four replicates be conducted in all three stations?

Lastly, simulations conducted in the manuscript are not explained explicitly.

1) It was not explained explicitly soil moisture in which layers are considered in the calibration and validation. Did the soil moisture at all depths of 5, 10, 20, 50, and 100 cm be used to calculate the observation operators? Was the near-surface (5 cm) soil moisture data regarded as the input (surface soil moisture)? Which layer was the "profile" referred to in this study?

2) To this reviewer's understanding, the authors used soil moisture with daily resolution to calculate the observation operations (line 23 in page 8) and then used these observation operations to predict daily soil moisture in profile. Is it true?

3) The authors listed the soil texture in different stations (Table 1). Does the texture impact the prediction results, besides the type of climate.

4) I cannot understand why does figure 5 illustrate the effects of data length on performance of observation operators. Which level of data length is indicated in figure 5?

5) I guess the authors omitted one figure, maybe figure 11 (mentioned in line 9 in page 11).

6) Did the same input (surface and profile soil moisture data) be used when the CDF and the exponential filter methods were employed respectively to do calibration and validation?

7) Uncertainty is one of the main issues when statistical methods are applied. Could the authors explore possible sources of uncertainties for the prediction.

8) Lines 13-16 in page 11. I am not sure whether it is reasonable and substantiated enough to conclude that the CDF matching method is more robust than the exponential filter method based on the application of this study. It is better to constrain such conclusion in specific conditions considered in this study.

Minor comments

1) Title, why use "upscaling" in the title? The main point of the method is to use surface soil moisture to predict profile soil moisture based on a prior calibration using available surface and profile soil moisture data in other time period. I think this procedure is not relevant to upscaling.

2) Line 6 in page 3, "found that multivariate regression and artificial neural network was able to produce reliable profile soil moisture estimations, but required ...", "was" should be "are", "required" should be "require".

3) Line 1 in page 5, "profile" instead of "Profile".

4) Line 5 in page 5, replace "when considering the accuracy of fitting and the principle of parsimony" by "when the accuracy of fitting and the principle of parsimony are

considered".

5) Line 18 in page 9, "soil moisture time series data length" needs to be rephrased.

6) What do the symbols "a", "b", "ab" mean in figure 10? The authors should present the explanation in the caption of this figure.

————————————————