Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2019-574-RC2, 2019

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Soil moisture: variable in space but redundant in time" *by* Mirko Mälicke et al.

Anonymous Referee #2

Received and published: 4 December 2019

I am not a hydrologist, so I cannot say anything about the level of novelty of the current work with respect to the published literature of which I am not well aware. On the other hand, the proposed methodology seems very reasonable and effective to me. I liked the paper and the interpretation of the results. On the other hand, I personally found that the text should be improved: some sentences are open to ambiguity or unclear (the meaning or idea to be conveyed is there intuitively, but the structure of the sentence leaves it open to misinterpretations), and there are some repetitions that could be cut out in order to make the paper easier to follow (sometimes is hard). I suppose that the confusion in some sentence stem from the young age of the first Author: put some more efforts in making the text clearer and more specific in order to honor your work.

Comment 1: pp 9, lines 26-27: 'This distance is the foundation of Mean shift on the one

hand and the compression quality calculations presented in section 2.5.2 below on the other hand'. The relevance of the distance in (3) for the mean shift clustering algorithm is somehow hidden during the description of the method in Sec. 2.3.1 (i.e., 'We tested different bandwidth parameters at a few examples and set the bandwidth to the 30% percentile of all pairwise vector distances between the dispersion functions of one year and depth'). If possible, I would like to rephrase this aspect putting more focus on the relevance of (3) within the clustering algorithm.

Comment 2: In the Section 2.5.2 Compression quality, there are some unclear aspects to me. 'To assure comparability we use one binning for all calculations of H (across years and depths). To achieve this, all pairwise distances between all spatial dispersion functions of all four years in all three depths are calculated. The discrete frequency distribution is formed from 0 up to the global maximum distance (between two dispersion functions) calculated using equation (3). The bins are formed equidistant using a width of the maximum function distance that still lies within the error margins calculated using equation (9). Thus, the information content of the spatial heterogeneity is calculated with respect to the expected uncertainties. This way we can be sure to distinguish exclusively those spatial dispersion functions that lie outside the error margins.' It is my understanding that the binning scheme is grounded on the distance function in (3), which make me to think that subsequent calculation (e.g., entropy, KL-divergence) will involve the distance in (3), but then 'To calculate the mean information content of the compressed series each cluster member is substituted by the respective cluster centroid. This substitution is obviously not a compression in a technical sense, but necessary to calculate the Kullback-Leibler divergence. Then a frequency distribution for compressed series X and the uncompressed series Y can be calculated. The Kullback-Leibler divergence DKL of X,Y is given in equation (12)' which compare compressed and uncompressed dispersion functions and does not involve the distance in (3) in any way. It seems that the binning (size of the bin and edges of the diverse bins) entails the diverse distance according to (3) (and the associated uncertainty, according with (9)), but the KL is evaluated for the dispersion function themselves (and not their

distance)? What am I missing? Could you please further clarify how the distance (3) is involved in the evaluation of (12)? I would also briefly describe the meaning of the KL-divergence which is just introduced, but not commented.

Comment 3: pp.12 lines 6-7: 'Dispersion declines with separating distance, as small values correspond to observations which have similar values while large values suggest the opposite' looking at Fig. 4 the dispersion function increases with the spatial lag, such that dispersion increases with separating distance. What am I missing here?

Comment 4: pp.12 lines 18-23: 'As the spatial dispersion functions in the presented example are redundant in time, we compressed the information by replacing the dispersion function within one cluster by the cluster centroid. All four representative functions shown in Figure 4 c) exhibit increasing dispersion with separating distance. For the blue and green cluster this happens step-wise at a characteristic distance of 500 m. That reminds us of a Gaussian variogram, which can also show a step-wise characteristic. The small grey cluster shows an increase at 500 and another one at 1000 m separating distance. In contrast the orange cluster, however, shows a only a gentle increase with distance.' Are there any reasons for these stepwise increases or some physical related explanations for these behaviors? Comment 5: with reference to Fig. 4 and its discussion in Sec. 3.1 vegetation period is mentioned several times, would it be nice to have on Fig. 4.c this period highlighted (also in other figures where vegetation period is of relevance), for example as a light green bar along the x-axe or similarly, in order to understand when this vegetation period is 'on/off'.

Comment 6: pp. 12, lines 24-26: 'In the vegetation period observations are similar even at large separating distances. As the orange cluster (Figure 4 c and d) covers significant parts of the vegetation period, the influence of vegetation on spatial soil water dynamics is considered to be dominant'. The latter sentence is misleading: during the so referred vegetation period which are the concurrent factors along with the vegetation-related influence that could possibly influence the soil moisture? How is it possible to discern the impacts of other factors in order to say that vegetation influence

is the dominant one? Or, if the vegetation-related influence is the only factor, no surprise that it is the dominant one. Please consider revise the sentence or better support it.

Comment 7: pp.13, lines 7-8: 'At the same time the observations get spatially more homogeneous in summer, particularly when the blue cluster emerges, because the dispersion at large lags decreases significantly.' I would substitute 'because' with 'i.e.', the decrease of the dispersion function is a consequence not a cause of the more homogeneous nature of the data during summer.

Comment 8: pp. 13 lines 13-15: 'The green clusters emerge with strong rainfall events after longer previous dry spells (Fig. 5). We would have expected a third occurrence at the beginning of August, but the soil may already be too dry to bear a detectable dependency on separating distance'. It is not clear at which green cluster the Authors are referring to in this sentence, please clarify.

Comment 9: pp. 13 lines 17-18: 'The 50 cm dispersion functions (Fig. 5 f) show a clear spatial dependence and are similar'. I can see the trends of the two dispersion functions with respect to the lag and I can see that these trends are consistent among each other, but I don't see the similarity between the two dispersion functions that are characterized by fairly different values especially at larger lag. Please revise the sentence having care of the specificity of the wording.

Comment 10: pp. 13 lines 18-19: 'Not only the soil moisture observations have become much more homogeneous with depth, also the dispersion functions are more similar in shape.' The first part of the sentence is vague, more homogeneous with respect to what? In time or space? Looking at Fig. 5c, at fixed time (e.g., 2016-02), I can see a great level of heterogeneity across the moisture data, even larger than that recorded in 5.a-b. Furthermore, the dispersion function orange in Fig. 5f reaches values comparable to that of the dispersion functions in Fig. 5d (orange) and Fig. 5e (grey and green clusters).

Comment 11: pp. 14 lines 1-2: ‘We find dispersion functions with characteristic length of 500 m and the blue cluster persists throughout most of the year.’ How is the characteristic length of the dispersion function defined? Please clarify.

Comment 12: pp. 16 lines 4: ‘water water dynamics’, water is repeated.

Comment 13: pp. 18 line 27: ‘In line with H4 spatial patterns of soil moisture were found to be persistent over longer time periods’ Longer than what?

Comment 14: pp. 19 lines 17-18: ‘We thus conclude that there is dependence of the dispersion on the rainfall pattern, which is reflected in their shape and characteristic lengths.’ The sentence, as written, means that the shape and characteristic lengths are referred to that of the rainfall pattern, while I am imagining that are the shape and characteristic lengths of the dispersion function the ones that changes. Please revise the sentence.

Comment 15: pp 20 lines 16-18: ‘This Euclidean distance does, however, not provide information on the underlying cause of dissimilarity and thus a simple shift along the y-axis can result in the same level of dissimilarity as a change in the shape of the dispersion function.’ Despite being clear from an intuitive point of view, this sentence can be sloppy to the most rigorous reader: the y-axe of what? Please revise, like ‘a minor difference in the values of the dispersion functions, even though characterized by the a very similar shape, could results in . . ..’.

Comment 16: pp 22 line 1: ‘A drying and then dry soil exhibits dispersion functions without spatial structure. Interestingly, these functions flatten out by minimizing the dispersion on large distance lags and we can thus see how the soil acts as a low pass filter.’ The first sentence is obscure, especially when linked with the second one. Why the Authors claim that there is no spatial structure during drying and dry periods, when the associated dispersion functions clearly show a flat behavior for the majority of the spatial lags? As far as I have understood, the latter behavior is a sign of homogeneity in the soil moisture across space, which is a clear sign of a structure in space (maybe

not that interesting, though) to me.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-574, 2019.