Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2019-121-AC1, 2019

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Characterization of Hillslope Hydrologic Events Using Machine Learning Algorithms" *by* Eunhyung Lee and Sanghyun Kim

Eunhyung Lee and Sanghyun Kim

kimsangh@pusan.ac.kr

Received and published: 30 May 2019

Thanks for your comments. Please find our responses for your comments.

Specific comments: 1. Stability indices: The study performs an analysis of temporal soil moisture variability. Thus the "Index of Temporal Stability" (ITS) is introduced. Unfortunately, it is defined in a really counterintuitive and misleading way: It measures exactly the opposite of what the term suggests (l. 155-158). Please rename to "Instability index" or the like.

- The term "Index of Temporal Stability (ITS)" had been widely used by many re-

searchers who had worked on spatial distribution of soil moisture measurements. Following papers used ITS for soil moisture variability.

Zhao, Y., S. Peth, X. Y. Wang, H. Lin, and R. Horn (2010), Controls of surface soil moisture spatial patterns and their temporal stability in a semi-arid steppe, Hydrol. Process, 24, 2507-2519. Zhao, W., Z. Cui, J. Zhang, and J. Jin, 2017, Temporal stability and variability of soil-water content in a gravel-mulched field in northwestern China, J. Hydrol., 552, 249-257. Zhang, P., and M. Shao (2013), Temporal stability of surface soil moisture in a desert area of northwestern China, J.Hydrol., 505, 91-101. Penna, D., L. Brocca, M. Borga, and G. D. Fontana (2013), Soil moisture temporal stability at different depths on two alpine hillslopes during wet and dry periods, J. Hydrol. 477, 55-71. Biswas, A., and B. C. Si (2011), Scales and locations of time stability of soil water storage in a hummocky landscape, J. Hydrol., 408, (1-2), 100-112. Gao, L., and M. Shao (2012a) Temporal stability of soil water storage in diverse soil layers, Catena, 95, 24-31. Gao, L., and M. Shao (2012b) Temporal stability of shallow soil water content for three adjacent transects on a hillslope, Agri. Water Manag., 110, 41-54. Gao, L., M. Shao, X. Peng, and D. She (2015) Spatio-temporal variability and temporal stability of water contents distributed within soil profiles at a hillslope scale, Catena, 132, 29-36. Gomez-Plaza, A., J. Alvarez-Rogel, J. Albaladejo, and V. Castillo (2000), Spatial patterns and temporal stability of soil moisture across a range of scales in a semiarid environment, Hydrol. Process., 14, 1261-1277. Li, D. F., and M. Shao (2014), Temporal stability analysis for estimating spatial mean soil water storage and deep percolation in irrigated maize crops, Agric. Water Manage., 144, 140-149. Junquira Junior, J. A., C. R. Mello, P. R. Owens, J. M. Mello, N. Curi, and G. J. Alves (2017), Time-stability of soil water content in an atlantic forest-latosol site, Geoderma, 288, 64-78. Lee, E., and S. Kim (2017), Pattern similarity based soil moisture analysis for three seasons on a steep hillslope, J. Hydrol., 551, 484-494. Vachaud, G., P.P. de Silans, P. Balabanis, M. Vauclin, 1985, Temporal stability of spatially measured soil water probability density function, Soil Sci. Soc. Am. J., 49, 822-828. Brocca, L., F. Melone, T. Moramarco, and R. Morbidelli, 2009, Soil moisture temporal stability over experimental areas in Central

Italy, Geoderma, 148(3-4), 364-374. Xiaoxu, J., M. Shao, X. Wei, and Y. Wang, 2013, Hillslope scale temporal stability of soil water storage in diverse soil layers, J. Hydrol., 498, 254-264. Hu, W., L.K. Tallon, and B.C. Si, 2012, Evaluation of time stability indices for soil water storage upscaling, J. Hydrol, 475, 229-241. Z. Liu, Y. Wang, P. Yu, A. Tian, Y. Wang, W. Xiong, and L. Xu, 2018, Spatial pattern and temporal stability of root-zone soil moisture during growing season on a larch plantation hillslope in Northwest China, Forests, 9, 68.

- Therefore, changing ITS as "instability index" seems not appropriate simply because majority of hydrology community used ITS for many years.

Secondly, the mathematical definition is unnecessarily complex and hardly comprehensible (equation 2). Why do you add squared mean and variance of normalized differences (δ)? What information does that index provide that would not already given, e.g., by the mean of normalized differences δ? In addition, both ITS and the normalized difference δ (equation 1) can easily be mixed up with the "soil moisture difference index" ?θ (equation 4). The latter seems to have been used only for the SOM classification. Why do you need that many different indices for soil moisture variation? This is very confusing and not comprehensible for the reader.

- We referred widely used indices, such as ITS and SDRD, for soil moisture representation. Of course, we used equation (4) for SOM. We presented Its or other terms because we will compare the performance of SOM in delineating the representative point and those of other existing approaches. We think presenting new development should be done based on the context of existing studies. Through this way, we may highlight what is new and how our approach can contribute in analyzing soil moisture data.

2. Classification of hydrological events: Assignment of hydrological events is a crucial point in this study, as clusters were assigned to processes. Unfortunately, the clustering is a very weak point of the study. First of all, I do not understand why an SOM has been

performed prior the cluster analysis. I acknowledge that SOM is a nice tool to visualize multivariate data sets (cf., Fig. 5). However, assigning data to the nodes of an SOM is not without loss of information, as even slightly different instances of the data set are assigned to the same node. Thus I would highly recommend performing the cluster analysis directly on the data rather than on their aggregated representation in an SOM.

- Considering complex soil moisture dynamics for numerous rainfall events in hillslope area, two primary objectives in applying self-organizing map to rainfall and soil moisture dataset are as follows; one is the characterization of non-linear relationships among hydrologic variables and clustering of rainfall events based on trained self-organizing map. - Firstly, the visualization capability of self-organizing map, which you mentioned above, is much better than other conventional approaches such as PCA (Principle component analysis) and MDS (Multi-dimensional scaling) and other alternatives, such as linear tools which previously frequently used (Reusch et al., 2005; Liu et al., 2006). Especially, previous researches reported that 2-dimensional space of self-organizing map provides more reliable data representation compared to the other linear dimension reduction approach (Brosse et al., 2001; Chon et al., 1996; Hilton and Salakhutdinov, 2006; Pinto et al., 2008; Reusch et al., 2007). - Secondly, the clustering of rainfall events based on trained self-organizing map is the other main objective. The representative direct clustering method consists of two parts like Hierarchical clustering (Single linkage, Average linkage, Centroid method, Ward's method, Kth Neighbor) and Non-hierarchical clustering(K-means algorithm, K-median algorithm). - Comparison of performance between self-organizing map and direct clustering methods have frequently been reported in previous researches. The self-organizing network is superior to hierarchical clustering in both sensitiveness for noisy dataset and accuracy aspects (Mangiameli et al., 1996; Vesanto and Alhoniemi, 2000; Samsonova et al., 2005; Abbas, 2008; Sivasankari et al., 2014). Using hierarchical clustering approach with dendrogram, it was very difficult to characterize the cluster characteristics as well as to select optimized cluster number for researcher's objective (Samsonova et al., 2005). - The figure (below, attched) is the agglomerative hierarchical clustering results (ward method) dendrogram which was suggested by reviewer. Extracting or finding any useful explanation from this figure is almost impossible. The cluster number obtained from the coefficient delta (see Table) was two, which is almost useless in terms of hydrological interpretation. Therefore, the self-organizing map is comparably feasible to recognize the relationship among rainfall events. This was pointed out by many researchers (Mangiameli et al., 1996; Vesanto and Alhoniemi, 2000; Samsonova et al., 2005; Abbas, 2008; Sivasankari et al., 2014).

Figure for the hierarchical clustering of hydrological dataset (from this study) using ward method.

Table (another figure)

- Furthermore, many previous researches reported that even non-hierarchical approaches showed worse performances in clustering accuracy compared to self-organizing map (Jassar and Dhindsa, 2015; Dhingra et al., 2013; Toor and Singh, 2013; Abbas, 2008; Kiang and Kumar, 2001). The performance of non-hierarchical method is very sensitive to the heuristic selected initial setting (Jassar and Dhindsa, 2015), but self-organizing map is less prone to the change in initial network setting (Kiang and Kumar, 2001). According to characteristics of dataset, the comparison of performance among clustering method can be changed, but many relevant scientists have proved both applicability and higher performance of self-organizing in many disciplines. This means that the clustering with the SOM instead of directly clustering the data is not only computationally effective approach but also have better or comparable performance with significantly reduced uncertainty and much better robustness compared to direct clustering method (Vesanto and Alhoniemi, 2000).

Brosse, S., J.L. Giraudel, and S. Lek, 2001, Utilisation of non-supervised neural networks and principal component analysis to study fish assemblages, Ecological Modelling, 146, 159-166. Reusch, D.B., R.B. Alley, and B.C. Hewitson, 2005, Relative performance of self-organizing maps and principal component analysis in pattern extrac-

tion from synthetic climatological data, Polar Geogr, 29(3), 227-251. Hinton, G.E., and R.R. Salakhutdinov, 2006, Reducing the dimensionality of data with neural networks, Science, 313(5786), 504-507. Liu, Y., R.H. Weisberg, and C.N.K. Mooers, 2006, Performance evaluation of the self-organizing map for feature extraction, J. Geophys. Res., 111, C05018. Pinto, N., D.D. Cox, and J.J. DiCarlo, 2008, Why is real world visual object recognition hard?, PLoS Computational Biology, 4(1), 0151-0156. Chon, T.S., Y.S. Park, K.H. Moon, E. Cha, and Y. Pa, 1996, Patternizing communities by using an artificial neural network, Ecological Modelling, 90, 1403-1409. Reusch, D.B., R.B. Alley, and B.C. Hewitson, 2007, North Atlantic climate variability from a self-organizing map perspective, Journal of geophysical research, 112, D02104, doi:10.1029/2006JD007460. Abbas, O.A., 2008, Comparison between data clustering algorithms, the international arab journal of information technology, 5(3), 320-325. Mangiameli, P., S.K. Chen, and D. West, 1996, A comparison of SOM neural network and hierarchical clustering methods, European Journal of Operational Research, 93, 402-417. Toor, A.K. and A. Singh, 2013, Analysis of clustering algorithms based on number of clusters, error rate, computation time and map topology on large dataset, International Journal of Emerging Trends & Technology in Computer Science, 2(6), 94-98. Samsonova, E., T. Back, J.N. Kok, and A.P. IJzerman, 2005, Reliable hierarchical clustering with the self-organizing map, in Proc. 6th International Symposium on Intelligent Data Analysis, 397-413. Dhingra, S., R. Gilhotra, and Ranishanker, 2013, Comparative analysis of kohonen-SOM and k-means data mining algorithms based on academic activities, International Journal of Computers & Technology, 6(1), 237-241. Vesanto, J., and E. Alhoniemi, 2000, Clustering of the self-organizing map, IEEE TRANSACTIONS ON NEURAL NETWORKS, 11(3), 586-600. Kiang, M.Y., and A. Kumar, 2001, An evaluation of self-organizing map networks as a robust alternative to factor analysis in data mining applications, Information systems research, 12(2), 177-194. Sivasankari, A., S. Sudarvizhi, and S.R.A. Bai, 2014, comparative study of different clustering and decision tree for data mining algorithm, International Journal of Computer Science and Information Technology Research, 2(3), 221-232. Jassar, K.K., and K.S. Dhindsa,

2015, Comparative study and performance analysis of clustering algorithms, International Journal of Computer Applications, 0975-8887, 1-6.

Secondly, the number of clusters was selected in “a heuristic approach aiming to achieve a hydrologically meaningful classification of events and parsimonious clustering” (l. 289-290) without any clearly defined and understandable criterion. Likewise, no criterion is given for deciding on the superiority of one cluster approach compared to another (l. 211-213). Thus arbitrary decisions seem to have a major effect on the assignment to clusters, and subsequently to hydrological processes.

- Thanks for your comment. One important criterion was the parsimony in clustering with hydrologically meaningful event classification. We need to make 7 clusters to identify difference between cluster 5 and 6. As explained in the context, the identification of cluster 5 provide completely distinct interpretation (preferential flow) in hillslope hydrological process. We can divide SOM maps into 3 parts or 7 parts depending on the distinction of vertical flow and lateral flow in upslope and downslope (see Tables in paper). But classifying SOM maps as 7 numbers was found to extracting a meaningful hydrologic information about hydrological process characteristics in this hillslope. Further classification looks redundancy in terms of physically meaningful interpretation. The criterion cannot be given as an explicit formula (because natural process can be formulated) but existences of distinct hydrological processes (vertical flow and lateral flow) either upslope or downslope as criterion and interpretation for as illustrated in Table 2. The distribution of soil moisture indices and soil water storage (Table 1 and Table 2) and its statistical (volumetric soil moisture) results in Figure 8 were made from soil moisture measurements. These are not postulation but evidence. These results are not arbitrary decision but the summary of comprehensive analysis of soil moisture and rainfall for 396 events. - Of course, this is heuristic approach, simply because all clustering algorithms are heuristic approach. As we mentioned, both hierarchical and non-hierarchical method also need heuristic approach for selecting cluster number in dendrogram and initial setting.

Last but not least, according to Fig. 5, and except for antecedent soil moisture and rainfall (?) duration. there is a rather smooth and continuous increase of the soil moisture difference indices. The only exception concerns cluster 5 and 6. In cluster 5, the soil moisture difference index at the upslope sites below 10 cm depth is clearly less compared to the downslope sites, whereas the soil moisture difference index is high both at upslope and downslope sites in cluster 6. Except for these two clusters the cluster analysis obviously subdivides the data set more or less arbitrarily along a single continuous gradient rather than identifying clearly distinct groups (cf., Fig. 8). Thanks for comment. We partially agreed to your comment for continuous increase of soil moisture difference index. What you pointed out is somewhat natural in SOM classification simply because the distance between the input vector and weighting vector is mainly used for SOM classification (Actually this trend is common for all SOM application).

- The response of soil moisture on rainfall events consists of rainfall characteristics, antecedent soil water condition, and soil moisture difference index. Although components in same category have inter-connection, every component individually represents each characteristic which we would like to show multiple relationships among components. The dendrogram with dissimilarity matrix among components vectors in SOM map shows distinctively the available number of clusters. Among every cases of clusters, our choice considering hydrological processes was also heuristic, but the optimized number of cluster was 7 for identifying the difference between cluster 5 and cluster 6.

Identification of representative soil monitoring points: A major topic of the study is identification of “representative soil moisture monitoring points” (e.g., l. 30-31). It remains unclear in what regard these sites should be representative. According to l. 50-51 “high stability is an important criterion for determining the best location for the monitoring spatially averaged soil moisture of a given area”. The study focuses on the temporal dynamics. But here the term “representativeness” cannot refer to the temporal dynamics because “high stability points” would systematically underestimate that dynamics. On the other hand, “high stability” points could not be representative for the spatial mean either. Highest temporal stability likely occurs at sites where the soil is close to saturation all the time, that is, at sites with the highest soil moisture values. Or do you mean “representative” in regard to ascription to hydrological event clusters (L. 114-116, l. 485 et seq., l. 511 et seq.)? But then temporal stability of soil moisture would not be a relevant criterion.

- Following quotations are from one of best papers in this topic (Penna et al. 2013) “A positive value of MRD for a certain sampling location indicates that the location is wetter than the hillslope average whereas a negative value of MRD indicates that the location is drier than the hillslope average. The SDRD specifies how variable such as estimate is. The sampling location with the highest temporal stability shows the lowest value of ITS and is selected as representative point. Whereas points with high values of ITS are identified as the wettest and driest field location (Penna et al., 2013).” Penna, D., L. Brocca, M. Borga, and G. D. Fontana (2013), Soil moisture temporal stability at different depths on two alpine hillslopes during wet and dry periods, J. Hydrol. 477, 55-71.

- Several studies also reported that the most temporally stable points were significantly related to the soil texture (Jacobs et al., 2004; Zhang and Shao, 2013; Xiaoxu et al., 2013) or topographic indexes (Brocca et al., 2009; Lee and Kim, 2017), and the other studies also noted that the time-stable points were poorly correlated to topography and soil properties (Tallon and Si, 2004). These various findings indicate that the controlling factors of the time-stable points (the representative points) are very complicated and dependent on the characteristics of the study area (e.g., climate, soil, and vegetation). An increase in temporal stability with depth was expected due to the reduced dependence on the climatic, biological, and hydrological factors that determined the SWC dynamics (Martinez-Fernandez and Ceballos, 2003; Hu et al., 2009), which was also observed by Cassel et al. (2000) for cropland and by Lin (2006) for forest watersheds.

Zhang, P., and M. Shao (2013), Temporal stability of surface soil moisture in a desert

area of northwestern China, J.Hydrol., 505, 91-101. Xiaoxu, J., M. Shao, X. Wei, and Y. Wang, 2013, Hillslope scale temporal stability of soil water storage in diverse soil layers, J. Hydrol., 498, 254-264. Brocca, L., F. Melone, T. Moramarco, and R. Morbidelli, 2009, Soil moisture temporal stability over experimental areas in Central Italy, Geoderma, 148(3-4), 364-374. Tallon, L.K., and B.C. Si, Representative soil water benchmarking for environmental monitoring, J. Environ. Inf, 4, 31-39. Jacobs, J.M., B.P. Mohanty, E.C. Hsu, and M. Choi, 2010, Time stability and variability of electronically scanned thinned array radiometer soil moisture during Southern Great Plains hydrology experiments, Hydrol. Process, 24, 2807-2819.

- In this study, we expand the representative point from existing definition “the point showing least variation between points” to the point which address efficient characterization of hydrologic event as well as traditional temporal stability. This is why we presented ITS, SDRD as so on.

Assigning clusters to processes: Results from the upslope and downslope sites, respectively, are aggregated (per depth), and differences between these groups are interpreted in terms of systematic differences between these groups. However, it has not even been tested whether these groups differ significantly at all. Data of all single sites are presented in Fig. 2 only, revealing substantial heterogeneity even within the upslope or downslope sites, respectively, and suggesting more small scale variability (consistent with my own experience) rather than systematic differences.

- The main objective of this study is the analysis of data through SOM under the context of hillslope hydrological processes. Therefore, the difference of hydrological processes between upslope and downslope was the main topic and the small scale variability is associated with the local heterogeneity, which is not the scope of this study. Besides, it is almost impossible to characterize different non-linear patterns through SOM for every single points and interpret complexity and heterogeneity depending on measuring locations and measuring depths. Characterization of soil moisture for two parts (upslope and downslope) is consistent to several previous researches of the study area and many other hillslope studies (Haga et al., 2005; Kim, 2009, 2016; Lee and Kim. 2017, 2019; Tromp and McDonnell, 2004; Uchida et al., 2004). Thanks for your understanding.

Haga, H., Y. Matsumoto, J. Matutani, M. Fujita, K. Nishida, and Y. Sakamoto, 2005: Flow paths, rainfall properties, and antecedent soil moisture controlling lags to peak discharge in a granite unchanneled catchment. Water Resour. Res., 41, W12410, doi:10.1029/2005WR004236. Kim, S.: Characterization of soil moisture responses on a hillslope to sequential rainfall events during late autumn and spring, Water Resour. Res. 45, W09425, https://doi.org/ 10.1029/2008WR007239, 2009. Kim, S.: Time series modeling of soil moisture dynamics on a steep mountainous hillside, J. Hydrol., 536, 37-49, https://doi.org/ 10.1016/j.jhydrol.2016.02.027, 2016. Lee, E., and Kim, S.: Pattern Similarity Based Soil Moisture Analysis for Three Seasons on a Steep Hillslope, J. Hydrol., 551, 484-494, https://doi.org/ 10.1016/jhydrol.2017.06.028, 2017 Lee, E. and Kim, S.: Wavelet analysis of soil moisture measurement for hillslope hydrological processes, J. Hydrol, https://doi.org/10.1016/j.jhydrol.2019.05.023 Tromp van Meerveld, I., and McDonnell, J.J.: Comment to "Spatial correlation of soil moisture in small catchments and its relationship to dominant spatial hydrological processes, J.Hydrol., 286, 113-134", J.Hydrol., 303, 307-312, https://doi.org/10.1016/j.jhydrol.2004.09.002, 2005. Uchida, T., Y. Asano, T. Mizuyama, and J. J. McDonnell, 2004: Role of upslope soil pore pressure on lateral subsurface storm flow dynamics. Water Resour. Res., 40, W12401.

At least there seems to be substantial overlap between upslope and downslope sites. The study aims at assigning single hydrological events, characterized by meteorological and soil moisture data based indices, to clusters, which in turn are interpreted in terms of hydrological processes (l. 90, l. 111-113). The latter step is of fundamental importance for the study. Unfortunately, that step remains completely obscure to me even after having studied the manuscript again and again. There is no clear and comprehensible reasoning at all. How do the indices relate to the respective processes?

Relationships are postulated but without sound justification (e.g., l. 317-319, l. 322-324, l. 334-l.338, l. 402-462). Assigning differences in soil moisture at a scale of a few hours to lateral subsurface flow over a distance of roughly 50 m (Fig. 1) would require fairly high lateral flow velocities. Is there any additional evidence (e.g., tracer experiments) for that? Did you account for the spatial heterogeneity of throughfall and stemflow? What about surface runoff that might have re-infiltrated on its way downslope?

- This is about interpretation of hydrological processes using results of analysis. Table 1, Table 2 and Figure 8 are not postulation but the evidences obtained from field measurement. The soil moisture difference index is made to present soil moisture variation impact in terms of normalized impact for antecedent soil moisture. This is an effective way to express the impact of rainfall to soil moisture for big data (30 points in 396 events). l 317-319: The soil moisture difference index was made from soil moisture measurements. Table 1 showed the vertical distribution of soil moisture difference indices are different between upslope and downslope. In case of upslope, UP10 > UP30 > UP60 for clusters 1 to 6 indicating dominance in vertical infiltration, but DO10 > DO60 > DO30 for clusters 3, 4 and 7 and DO60 > DO10 > DO30 for clusters 5 and 6 in downslope. This indicated that the generation of lateral subsurface flow in downslope. Table 1 is not a postulation but the evidence of different hydrological processes between upslope and downslope. l 322-324: The soil moisture difference index (192.6%) was obtained from saturation data of soil moisture (48% - 50 %) in surface of downslope during rainfall event. This is the evidence of surface saturation for extreme events (cluster 7) which most likely introduce overland flow in downslope. This analysis was done based on field measurement of 64 events. l 334-339: The analysis appeared in Table 2 was based on field measurements. The water storage analysis for clusters 5, 6 and 7 showed that the storage changes (47,40, 116mm) in deeper depth(60cm) were greater than those (19.6, 11.6, 49.4 mm) for intermediate depth (30cm) at downslope area. Considering scales of throughflow (20, 35, 85.5), the storage changes in deeper depth (60cm) can be only explained by subsurface lateral flow over bedrock from upslope. Again this is not postulation but results from field measurements. We did not measured the stem flow for table 2 but there is no systematic difference in vegetation between upslope and downslope and throughfall can be assumed to almost uniform. Furthermore, the impact of canopy interception for significant events (clusters 5,6 and 7) can be negligible. l 402-462: The generation of lateral flow in deep soil layers was supported by results from Table 1 and Table 2, which were obtained from soil moisture measurements. Supplementary explanations of component planes (Figure 5) also support spatial distributions of hydrological processes between upslope and downslope. Discussion with Figure 8(made from measurements) and other references (Kim, 2009; Lai et al., 2016; Uchida et al., 2001; Weinjofer and Zehe 2013; Haga et al., 2001; Feng and Liu, 2015) also support hydrological process interpretation for significant events (cluster 5, 6, and 7). Reviewer asked further field measurements (tracer experiment and stem flow). The evidence of soil moisture analysis (Table 1, Table 2 and Figure 8) indicated distinct hydrological processes between upslope and downslope. Unfortunately, we did not collect tracer data and stem flow for 396 events and presentation or analysis of these additional data seems not the scope of this paper. Besides data like stem flow and tracer are not the scope this study. The scope of paper is the application of SOM into big dataset of rainfall and soil moistures and interpretations of hillslope hydrological processes about classification. Thanks for your understanding.

The manuscript requires substantial language editing. Technical corrections: - we corrected English from native speaker. Thanks,

L. 23: Please be more precise. According to Fig. 2 soil moisture was measured at 10 sites but at three different depths each. - We corrected it into “30 points (three depths in 10 locations)”

L. 142: Were the trenches re-filled after installation of the soil moisture probes? If not, how to deal with resulting artefacts? - Of course, we refilled all locations. Otherwise, how can 10 years monitoring can be sustainable?

L. 186-188: Please be more precise. Logarithm transformation is one out of a set of Box-Cox transformations. Did you apply other Box-Cox transformations as well? If so, for which observables, and how? Besides, neither the logarithm nor other Box-Cox transformations centralize the means of the variables to zero. Instead, these transformations are usually applied when Gaussian distribution is required. However, that is not the case for any of the applied approaches.

- Of course, we explored Box-Cox transformation and we adapted logarithmic transformation simply because it provides the best result in terms of centralization (such as smallest skewness). This is widely used transformation for centralization (see Salas et al., 1988). Salas, J. D., Delleur, J. W., Yevjeuch, V., Lane, W. L., 1988. Applied Modeling of Hydrologic Time Series. Water Resource Publication. Chelsea, Michigan.

L. 279-281: I do not understand why you select soil moisture measured at one single point “as the representative soil moisture before the event for the SOM analysis”. That introduces an unnecessary bias. Why not taking the mean of the values measured at the different sites? - The selection of representative points based on temporal stability method had been widely used and applied in the relationship among soil moisture points in many paper. The measurement data for 10 years cannot be 100% perfect, there can be some missing data and the perfect mean of every soil moisture points is not available for all rainfall events. Therefore, the soil moisture of representative point was replaced for calculating antecedent soil water condition. Again the temporal stability analysis is widely accepted approach for evaluation of mean soil moisture variation.

L. 574: “https://www.re3data.org/” is not a repository but gives an overview over numerous repositories. Please be more precise: Where will the data be published? - Of course, if the paper is accepted, the data will be uploaded in open repository.

References: Some references are out of alphabetical order (Minet et al. 2013, Montero and Vilar 2014, Zhu et al. 2014). - We corrected it. Thanks.

Fig. 5: Figure caption: Missing explanation of the lower panels (cf., caption of Tab. 1). - We corrected as follows. Thanks. Figures 5. (a)-(j) Component planes of variable weightings for the rainfall amount (a), antecedent soil moisture (b), soil moisture difference indices for the upslope at depths of 10(c), 30(d), and 60 cm(e), rainfall duration(f), rainfall intensity(g), and soil moisture difference indices for the downslope at depths of 10(h), 30(i), and 60 cm(j).

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-121, 2019.

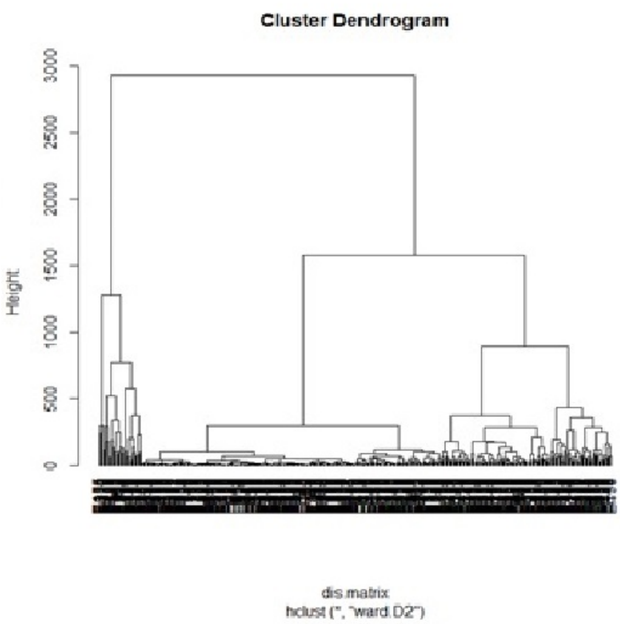

Fig. 1.

# of clusters	Sum of square	Coefficient delta
1	8767491	
2	4455593	4311897.58
3	3207137	1248456.58
4	2395312	811824.39
5	1988096	407216.49
6	1688206	299889.83
7	1521255	166950.81
8	1383726	137529.06
9	1285079	98647.41
10	1212425	72653.42

Fig. 2.