Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-24-RC3, 2020

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Soil Moisture Sensor Network Design for Hydrological Applications" *by* Lu Zhuo et al.

Anonymous Referee #3

Received and published: 10 March 2020

General comments to the authors The proposed approach to identify the optimal number and location of a limited number of soil moisture sensors is sound and leverages well-known statistical techniques such as Principal Component Analysis and K-means clustering analysis. Clearly the research community is active in this topic and approaches of this type will be valuable to the scientific community, state and federal agencies, and watershed managers. While the approach is sound, it has two important cons in my opinion: 1) The approach requires a substantial amount of available information about the variable in question to decide the architecture of the network. This somewhat conflicts with the idea of deploying a new network, which is trying to resolve the problem of no having soil moisture data available. The second drawback is that watershed managers will likely not spend substantial amount of time running complex

simulations to generate a dataset to inform the proposed approach unless they are assisted by a group of scientists. I think the manuscript deserves publication. Below are some suggestions and questions to allow the authors improve the manuscript.

Specific comments to the authors Line 49: Remove comma between "probe" and "and Time.." Line 50: Compared to other technologies such as neutron scattering and the gravimetric method, sensors relying on electromagnetic principles are probably not that old, particularly from the point of view of automated systems. They are about 30 years old, but not sure whether it classifies as one of the oldest. Line 51: When calibrated, these sensors can be accurate. Otherwise measurements can have substantial bias. Line 54: Economic considerations about what aspect? Please clarify. Line 95: I suggest being more explicit and write "Po river plain and surrounding hilly areas" Line 97-98: Please define acronym "ESA CCI" Line 152: Please, consider a modification along these lines "5-minute angle soil database" or "5-minutes geographic resolution soil database" Line 157: Please use square brackets when nesting parentheses. Line 197: Remove "the" from "since the eigenvectors of the X" What is the depth of the soil moisture sensing by the current stations? How will this approach handle multiple sensing depths? It seems that the final network configuration may be different for different soil layers. In Figure 9 it is unclear what is the number of stations considered and the method (CA-max, CA-median) employed to create the "designed network". Please provide more details in the figure caption. Do the current soil moisture monitoring stations only measure soil moisture? Do they observe other hydrological or meteorological variables that need to be taken into account at the time of designing a new network? To account for the network designed in probabilistic terms, have the authors considered using a probabilistic clustering method such as the Fuzzy C-means approach? Does the proposed approach account for the layout of existing monitoring stations? By looking at Figure 10, it seems that a complete re-arrangement of stations is required, which may conflict with existing agency resources and manpower. Judging from Figure 11 it seems that the whole effort of re-designing the network will remove some existing bias, but may not lead to a substantial improvement. Can the authors provide more details

on the logic behind the location of existing stations? It seems that the CA-max proposed clustering technique tends to initially place stations at or near the edges of the watershed. While this is the result of the proposed objective method, there are reasons why this might not be ideal. The CA-median seems to be less sensitive to this. How is the PCA superior to defining the number of clusters using something like the Silhouette method? How is the k-means part of this method different than variance minimization methods based on correlograms or semi-variograms? How does this approach include or handle the changing land use/land cover? Is it How can this method be applied to smaller catchment areas where a single pixel from remote sensing sources or distributed models is equal or even larger than the entire watershed?

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-24, 2020.