Hydrology and Earth System Sciences Discussions

Hydrol. Earth Syst. Sci. Discuss., referee comment RC2
https://doi.org/10.5194/hess-2022-63-RC2, 2022

Comment on hess-2022-63

Anonymous Referee #2

Referee comment on "Integrating process-related information into an artificial neural network for root-zone soil moisture prediction" by Roiya Souissi et al., Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2022-63-RC2, 2022

This paper discusses the relationship between surface moisture and moisture at depth in the soil layer occupied by roots. This is an important issue for analyzing the time series of surface moisture observed by satellite by providing much more relevant information to characterize the functioning of ecosystems. Indeed, deep moisture has a very important impact on water fluxes by controlling both transpiration and deep drainage. The proposed approach is to use neural networks trained on a large dataset as offered by the ISMN. The main innovation is to introduce variables into the network training that can account for factors impacting the relationship between surface moisture and humidity in the RZSM. I find that the introduction does not insist enough on the processes that govern this relationship. Indeed, it is strongly linked to 1) root uptake, which depends on the canopy and the root profile in interaction with the climate and 2) capillary rise, which means that, depending on the properties of the soil, the water that evaporates from the surface layer is more or less compensated. This very important process, especially for lightly covered soils, is never mentioned. This knowledge of physical processes could have been put forward to justify the process based variables.

I remain a little dubious about the choice of process based variables.

NDVI: for me there is no doubt that this variable must be taken into account. On the other hand, the use of ndvi modis variable does not seem to me to be adapted to the sites used. Indeed, many measuring stations are placed on sites where the vegetation is not representative of the nearby environment. SMOSMANIA is placed on a meteorological station with a non-irrigated fallow land placed in the middle of an agricultural zone (probably dominant at the scale of the modis pixel). The stations of the plain of Kairouan are on bare soil (probably to simplify the management of sensors) while the plain is an agricultural area. So I think there may be a big difference between the modis ndvi and the ndvi on the representative area of the measurement. This is illustrated in table 4 where the model including the ndvi led to degraded results in comparison to ANN-SSM.

The evaporative fraction as calculated is directly related to the surface moisture. There is therefore no introduction of information except via LEP which acts as a second order factor on the evaporative fraction.

The recursive exponetial filter completes the filtering by averaging over 10 30 and 90 days. It would have been interesting to compare them in order to identify to what extent these filters are complementary

The surface temperature could have been an indicator of the evaporative intensity. However, below the vegetation cover, the interpretation is far from obvious and requires knowledge of the air temperature in the canopy. Here it is not clear at what depth it is measured (probably at the depth of the first sensor). In this case the only interest seems to me to be to be able to flag the periods of freezing to eliminate the data which do not have a physical meaning. I would make this cleaning before training the neural networks.

A more thorough discussion of the process based variables would be necessary, showing in particular on examples how they allow a better understanding of the relationship between surface and depth.

I would now focus on the results. Are the results presented in table 3 qualitatively good? For example, for the RMSE, the introduction of co-variables has a positive impact in only 57% of cases at best. This also means that in 43% of the cases the results are worse. I think that a more rigorous statistical analysis would be necessary to decide whether or not the gain is significant.

Looking at Figure 4, I am impressed with the quality of the results. Using complex process models and measuring all the soil properties, I have never been able to simulate the water dynamics in different layers with such realism. I am impressed that a neural network trained with data from all over the world is able to reproduce with such fidelity the moisture levels between layers and the temporal variations in deep layers. No spurious variations are observed while the surface signal is particularly impacted by many rain events. If we can highlight the association of variables that allows such quality results, we have a major result for the understanding of water dynamics. This point must absolutely be highlighted.

Finally, on the form, the article does not seem to me well written. Not being an English speaker, I find the English not very good and the text not always clear. An subtancial editing work is for me essential.

In conclusion, I find the submitted draft article has not reached a stage of maturity allowing a publication. Some of the results are potentially extremely important and I therefore invite further analysis.