Comment on hess-2021-590
Anonymous Referee #3

This paper discusses a way of modeling pan evaporation using 3 methods: the Penman-Monteith equation, multiple step-wise regression, and the Kohonen self-organising map. The novel element to this work appears to be the fact that the pan evaporation was measured with 3 class A pans but two contained sediment and one contained submerged macrophytes on top of sediment. To me, this is the part that sets the paper apart and is important, because the authors are trying to create more realistic conditions for the observational experiment before using the data for a modelling study.

I liked the part of the discussion lines 248-250 where the authors discuss the low winds in this study and in earlier studies. I think that the fact that they honestly state a kind of negative result is helpful and proper.

1. I was disappointed that the methods were not discussed in more detail, in particular, the Kohonen self-organising map method. I can see that proper citations are given, and that this method has been applied to evaporation modelling before. However, I think that the paper would benefit greatly from an introduction to the Kohonen method. Then all 3 methods could be explained to the reader alongside the relevance to the physical process being studied. Why were these methods chosen in the first place? What are the pros and cons to these methods?

2. Figure 5 should be explained. I can see that this is related to figure 3 but the links are not made clear. The x and y axes are not labelled in either figure. How do the hexagons map to the inputs or outputs? I can't figure this out from reading the manuscript. Also, I found no definition for "importance" in the caption. This is part of the lack of time spent discussing and explaining the relevance of the methods used in the study.
3. Tables 1 and 2 are very large and comprehensive. While I think that the information is critical to explaining the conclusions of the paper, I do not think that the information in them is easy to make sense of. Is there any way that the tables could be re-organised or even some of the information could be turned into figures to display the information in a clearer way? One suggestion for table 2 is to keep information for the correlation values only in the tables, but to put all the statistics (max, min, mean, std. dev.) in a figure with sub panels. Are all the statistics relevant? Perhaps the authors could be a bit more selective? I agree that information on the full, training and testing data should be presented.

4. Figure 1 was not displayed with very high resolution. Could the authors provide a higher quality figure?

5. Figure 4 box plot whiskers and circles are not clearly defined. Software packages which compute these types of diagrams are not all the same. Please could the range (and meaning) of the box length, whiskers and circles be stated explicitly.

6. Table 1 has three lines at the bottom saying "Based on observed means" but I don't know what these lines are referring to. Is it the full, training and testing data sets?

Minor corrections

In figures with multiple panels, the sub-panels should be labeled with letters (a, b, c,...) in order to make the discussion of the results clearer in the body of the paper and make it easier to clarify the definitions in the captions.

L167 "displays a regular pattern" does not make sense to me. Can the authors make clearer what the Relative humidity and global radiation statistics are exhibiting and why it is important to notice this?

L178 Where is F defined? I could not find it.

L241 Please change "few" to "little"

L247 Please change "researches" to studies"

L273 Please change neural network to "a neural network approach".
L290 the range or tuple presented is unclear. Is "(-0.42-0.44)" really (-0.42 to -0.44) or (-0.42, -0.44)? Also, it seems strange to put the larger number -0.42 in front of -0.44 if indeed they both are negative.

L296 Please change "has high priority" to "is superior to". I'm not sure about the phrase "prediction precision". Precision sounds like a computational error here. Are you talking about skill? High correlation or low RMSE? Is there a better way to phrase this?