Hydrology and Earth System Sciences Discussions

Hydrol. Earth Syst. Sci. Discuss., referee comment RC9
https://doi.org/10.5194/hess-2022-96-RC9, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on hess-2022-96

Anonymous Referee #9

Referee comment on "FarmCan: a physical, statistical, and machine learning model to forecast crop water deficit for farms" by Sara Sadri et al., Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2022-96-RC9, 2022

Review of „FarmCan: A Physical, Statistical, and Machine Learning Model to Forecast CropWater Deficit at Farm Scales " by Sadri et al.

This paper describes a climate-informed machine learning (ML) framework to predict crop water demand at the farm scale that was tested at four farms in the Canadian Prairies Ecozone.

The topic is of general interest due to the increasing drought and overexploitation of water resources in many regions of the world due to global climate change and fits well within the scope of the journal. The manuscript is written in a comprehensible way for the most part, but it needs to be revised as it often contains incorrect wording and phrasing.

Unfortunately, in my opinion, this work is not yet mature enough to be recommended for publication because of numerous methodological shortcomings (see general and specific comments). Furthermore, the presentation of the results and the discussion are not adequate and need to be revised.

General comments:

There are already numerous tools available to predict water demand for crop management. The novelty of this study is for the user to select a specific farm location, which alone is not sufficient for publication. Therefore, the novelty of this study needs to be better explained.

It is problematic that the ML method was only tested at sites with relatively high precipitation and thus no need for irrigation. In particular, the method should be tested in climate zones where irrigation is actually needed to be carried out due to the climate conditions.

The modelling procedure shows some weaknesses:

- The model uses P to predict ET and PET, although both variables were found not to correlate well with P in the datasets used.
- The model predicts 8 day-sums of NI, which are then disaggregated to daily values based on daily P values. This seems to be an unnecessary oversimplification. A more accurate way would be to directly forecast daily NI values.
- The model uses PET to calculate the NI, which can lead to an overestimation of the NI, as PET represents the maximum ET at optimal soil water supply. In this way, NI will also be greater than 0 when the crops are not under water stress. Water stress only sets in when soil moisture falls below a certain values called MAD (maximal allowable demand), which is crop and soil specific (see e.g. Taghvaeian et al., 2020)
- Switching from SM to RZSM after the second harvest stage is questionable, as the plants reach a greater rooting depth than 5 cm much earlier.

The analysis in Chapter 3.2 shows mainly weak correlations and incomprehensible argumentations. In addition, it is detached from the modelling part. Therefore, I suggest removing it.

A discussion of the results in the light of existing literature and potential limitations is missing.

Specific comments:

L14: “four” instead of “4”

L16-18: This statement already indicates that the ML method was not sufficiently tested.

L32-35: It is unclear why and to what extent irrigation demand forecasts are important to rainfed farmers. It is very unlikely that they would change their farm management just because irrigation demand forecasts are available. This statement is also not included in the citations.

L60: Use consistent capitalization

L64: I wonder why the length of the growing season cannot be determined by the inputs, i.e. coordinates and crop type.

L66: What do you mean by “subfield”? Please consider that the SMAP data has only 36 km resolution.

L67: I don't understand why this method is tailored to this area, as the method seems to be generic.

L110: This equation does not correspond to the one shown in the cited reference, e.g. effective precipitation is used. In addition, using actual ET instead of potential ET provides more realistic estimates of irrigation demand and should therefore be used when available, as in this study.

L118: “highly accurate” seems to be exaggerated. Please use quantitative information on accuracy from SMAP validation studies, e.g. Montzka et al., 2017.

L148-150: Please remove the 0.1 factor discussion, which is not important.

L167: There seems to be a citation missing.

L173: What do you mean be “relevant depth of soil moisture”?

L174: The total number of expected growing days seems to be very uncertain. How will this uncertainty influence the forecast results?

L181-185: This procedure is not clear to me. How do you calculate the radii? How do account for the different spatial resolution of the different data sets? Are you averaging over the SMAP grid? How representative is the SMAP soil moisture for a specific field?

L187-189: It is unclear how the extension of the SMAP data to 2010 was done in detail. Furthermore, it is likely that a machine learning method will lead to very uncertain

estimates of soil moisture, and I therefore do not see its benefit for the predictive modelling. Either explain in more detail how and why the extension was done, or better leave it out.

L190-192: From a viewpoint of a soil hydrologist, it is very strange and arbitrary to first predict RZSM and than use this for the prediction of SM, as SM should be much stronger controlled by P than RZSM due to infiltration processes. Please explain in more detail reasoning behind this. In addition, explain why you predicting SM at all?

L194: Again, why are you not using ET instead of PET, see comment L110?

L195-202: Eq. 2 is not correct. To obtain the correct weights, the result from Eq. 2 must be divided by 800. Furthermore, this procedure is a strong simplification, as it does not distinguish on which day P fell within the 8 days, which makes a big difference in reality. Let us assume, for example, that 100 mm P fell on the first day of the week. In this case, the irrigation requirement for the crops for the following days would be much lower compared to if 100 mm P would fell on the last day.

L229: No precipitation predictions are used in this study.

L240-242: Given the data shown in this paper, only the first explanation seems plausible. This, however, questions the motivation of this study, i.e. that the CPE region is susceptible to droughts and that predictions of NI are important for agricultural management.

L259: This statement is an exaggeration, because this is clearly not always the case. Otherwise, farmers would use irrigation regularly. The available soil water at the beginning of the growing season also contributes to meeting the water needs of the plants because the soil type in this region is predominantly chernozemic clay, which has an extremely high water storage capacity.

L260-261: Just because irrigation is not used is not an argument for drought vulnerability in agriculture. See also comment above.

L262-263: One should be careful with these claims, because if this were indeed the case, farmers in this region would have already started irrigating to achieve high crop yields. Conversely, this may also indicate large uncertainties in the used P and PET products.

L281-282: This argumentation is not plausable.

L284-285: But high PET values can be found also for positive delta values of RZSM and SM for all regions.

L292-294: This statement contradicts the time series of soil moisture shown in Fig. 6. If it were true, one would expect soil moisture to decrease continuously during the growing season. However, the delta values of SM and RZSM both show fluctuations around zero, suggesting that P is sufficient to compensate for ET losses. This discrepancy may be due to uncertainties in the PET and ET data.

L309: In Fig. 8, M1 and M2 show strong P events of about 30 mm on 1 July. This results in an increase of SM to about 0.5 m^3/m^3 indicating soil saturation. Nevertheless, the model predicts a decrease in ET which is not plausible.

Table 1: The source for the crop water needs is missing and not all crops are covered. The values for P are much too low (in the text values between 400-1100 during the vegetation period are given). The values of P/PET seems to be too low as well.

Figure 4: I don’t see the benefit for showing the spatial distribution of P, ET, PET, SM and RZSM as this study only concerns time series analysis. Instead, monthly climate diagrams would be very useful to better understand the climate and soil hydrological situation in the four farms used in this study to develop and test the model.

Figure 6: Colours for ET and SM not distinguishable.

Figure 7: The two subplots are redundant and should be merged.

Figure 8: There should be no subtitles under the subplots. The plots are difficult to understand because of the large number of symbols. The meaning of the growth stage line is unclear.

Figure 9: Font size is too small. NI is not an observed value.

Literature

Taghvaeian, S., Andales, A. A., Allen, L. N., Kisekka, I., O'Shaughnessy, S. A., Porter, D. O., ... & Aguilar, J. (2020). Irrigation scheduling for agriculture in the United States: The progress made and the path forward. *Transactions of the ASABE*, *63*(5), 1603-1618.