Review of High-resolution drought simulations and comparison to soil moisture observations in Germany by Boeing et al.

Manuscript Number hess-2021-402

A recommendation: moderate revision

Using observed SM data collected from 40 locations and in 4 different measurement methods in Germany, the authors evaluate the performance of second generation operational German Drought Monitor in simulating soil moisture (SM). Two major research questions within this paper have been adequately addressed and can be summarized as follows: 1. how well the GDM capture the SM dynamics; 2. will GDM with higher spatial resolution produce SM estimates with higher quality compared with the GDM of former edition. Through the research, it was found that 1. SM dynamics simulations could be moderately improved; 2. higher resolution drought information at the one-kilometer scale can be met.

This research is a report of the improvement in the model performance which is evaluated in the perspective of comparisons between the model simulations and the observations. The article conforms to the journal-specific instructions and is relevant to HESS. The work is appropriate to be published in this journal after some revisions. In the following part, I will state the major arguments in detail.

The comparison between the observations and the model simulations are mainly indicated in the form of Spearman rank correlation coefficient. Nevertheless, the results of the comparison are in fact not so ideal in the perspective of solely the Spearman correlation coefficient, not to mention that significance values have been neglected when the observations and model results are compared. For example, when the observations from all the sites are included, the coefficients are normally lower. In addition, p-value has only been mentioned when both versions of the mHM and observations are compared (table 2). No mentioning of significance values also makes the comparison results not so validated. The significance value should be mentioned in the research to make results more reliable.

In addition, the analyses are focused on the explanations for the correlation coefficients, few further knowledge regarding the inner uncertainty within the models such as misrepresentations or no involvement of certain natural factors are included to be accounted for the discrepancy between the observations and model results. In this way, the detailed explanations in the conclusion and discussion part of this paper are needed to clarify the differences between observations and model results.

The four kinds of observations seem not to be standardized. In this way, some values of correlation coefficient which concern a specific observation technique, are not guaranteed to be valid in indicating the performance of GDM in simulating SM dynamics. A standardization method should be carried out to ensure a consistent comparison between observations and model simulations.

In order to facilitate a better standardization, not only the differences in the methods of the

observations, the observation sites' conditions should be considered in explaining the research results, such as the landuse, elevation, precipitation of different sites. Additionally, the change of the landuse dataset and geology dataset in the two versions of the models has not been verified to have a positive effect on the simulation of soil moisture as the improvement in the second version of the model is not so evident (especially in spring and summer the change of Spearman rank correlation coefficient is negative). Further clarification for the effects of the model setting is needed in this research or in the future studies to be conducted by the authors' team.

Moreover, the limited length of observed soil moisture data (< 10 years for most locations) causes some uncertainty in the comparison between observations and model results in the whole. More observations are needed to facilitate a more reliable model evaluation using the observation datasets as the existing observations within this paper are only validated in representing some regions of the Germany.

The authors can try making better simulations of the water cycle including soil moisture, ground water, and precipitation, while a higher-quality observational soil moisture dataset is applied in the future study. Other methods to indicate the correlations between the observations and model results can be used and also other indices apart from SMI can be utilized in indicating the severity of soil moisture drought.

Last but not least, it would be better to have some discussions on the applicability of the model to other regions other than Germany. Are there some future plans on extending the regional applicability of the model? Besides, some discussions on what can we learn from a small-scale modelling to improve large-scale modelling can be stated.

Some specific revisions regarding some parts of the manuscript are listed as follows:

1. The abstract is complete and correctly summarize the content of the paper, but it may need to be reduced a little to be more concise.
2. The process of parameter calibration and optimization needs to be in more detail to facilitate a reproducibility in the future study.
3. Some of the conclusions are overstated. The explanation should be in more detail regarding these issues.

   (1) Figure 3 shows the time series of both the observations and model simulations. It seems that the coefficient is much higher than those when all the sites were selected. How the sites are selected may need to be mentioned if there are other sites that contain both Cosmic Ray Neutron Sensing (CRNS) and Spatially Distributed Measurements (SDM).

   (2) There is lower agreement between observations and simulations in winter.

   (3) There is improvement of second version of the model in representing the upper soil but stagnation in representing the whole soil.

(4) The values of the Spearman rank correlation coefficient are not high enough to conclude a definite improvement of the first version of the mHM (Table 2).

4. In general, the authors have given proper credit to related work and clearly indicate their own original contribution. The references are appropriate to the research, but it would be better if some more papers are referenced especially those in which multibasin model calibrations and the SMI were applied.

The suggestions regarding some minor flaws and typos are described as follows:

Page 5, line 125: delete the ".However".

Page 5, line 131: delete the "," between "1.23" and "km".

Page 6, line 138: move "that were used in the analysis" before "are located".

Page 9, line 212, 213: remove "as"

Page 9, line 212, 213: add "," before "including", "the estimating", and "is hampered".

Page 11, line 270: remove "," after the "both".

Figure 3: add (a) to (l) for each sub panel to facilitate a better reference to the figures in the text when making the explanations.

Page 22: change the subtitle to "Conclusions and Discussions".

Page 23, line 436: remove "that".

Page 23, line 438: change "constitute" to "conclude".

The research is sound and fundamental. Some language edits could be good for improving the paper's quality.