Hydrology and Earth System Sciences Discussions

Hydrol. Earth Syst. Sci. Discuss., referee comment RC1
https://doi.org/10.5194/hess-2022-122-RC1, 2022

Comment on hess-2022-122

Anonymous Referee #1

Referee comment on "Machine-learning-based downscaling of modelled climate change impacts on groundwater table depth" by Raphael Schneider et al., Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2022-122-RC1, 2022

The manuscript "Machine learning-based downscaling of modelled climate change impacts on groundwater table depth" by Schneider et al. presents a novel downscaling method which uses hydrological model simulation data at a coarse scale (500 meters) together with ancillary data (e.g. topography and hydrogeologic information) to derive indicators for groundwater changes for future climate scenarios at higher spatial resolution (100 meters). Model simulations at a scale of 100 meters for five selected catchments and five input data sets from different regional climate model simulations are used as training data for the downscaling algorithm which is based on the Random Forest method. Estimates of groundwater changes at high resolution are made by using hydrologic simulations at coarse scale (500 meters) with input from 18 regional climate model simulations. The downscaling method is verified with data from a high resolution (100 meters) simulation for one additional catchment.

The topic of the paper is relevant to the hydrologic community as it describes an interesting possibility to provide stakeholders with high resolution information on potential changes in groundwater resources with an affordable computational cost. Generally, the paper is well written but there are a few issues that need to be clarified in my opinion.

General comments:

- The proposed downscaling method can be seen as a data-driven surrogate model for generating high resolution data out of the simulation results of the 500 meter model. This allows to avoid the computationally expensive direct simulation at this higher resolution, but adds some additional uncertainties and errors. In order to judge the quality and usefulness of these high resolution data, the user would still require some information on how the predictions improve when going from 500 to 100 meter resolution. Currently, the manuscript only provides information on how well the downscaling algorithm works but it does not describe the pratical benefits and improvements of the higher resolution. Hence I

would suggest to add a paragraph (e.g. around line 123) that summarizes the main advantages of the high resolution model as inferred from previous comparisons of the low and high resolution model with observation data.

- In section 2.4.3 (lines 248-258) it is mentioned that additional points outside the five 'calibration' catchments were used in the calibration procedure of the algorithm to improve the robustness of the method. Can you explain in more detail what kind of robustness issues you detected? Do you have any explanation why these additional points were necessary although the five chosen 'calibration' catchments closely resempled the statistical properties for whole Denmark (Figure 2)? Which additional information did these 'dummy points' provide?

- Additionally, the selection of additional calibration data through the 'dummy points' is not really in line with the argumentation in the rest of the paper which only refers to a calibration procedure with data from the five subcatchments. I would suggest to clearly state in all relevant parts how the calibration dataset was chosen (i.e. also mentioning the 'dummy points').

- Is it possible to provide guidelines on the size of the training data set? This would be an important information when applying the proposed downscaling method to other regions.

- Some plots are difficult to understand and need to be revised (see specific comments below).

Specific comments:

- Line 150: "...aggregated as described below." Please add the section number you are referring to.

- Line 154: It is not clear how the initial conditions were determined. Did you choose any random simulation time step between 1991 and 2100 as initial conditions or did you e.g. use the mean of this simulation period?

- Equation 1: Please make clear also through the notation that these statistics are calculated individually for each grid cell of the model.

- Line 218: Please provide details on the "...differences between a historic dry and wet period,...".

- Line 222: Why is the 500 meter model output interpolated to 100 meters although this does not provide further information to the downscaling method? Is it a hard requirement of the algorithm to operate on equally sized vectors? Is there any explanation why the algorithm works better with interpolated TBDV data?

- Line 392: Unit missing.

- Figure 4: The scale break in the figure is a bit counterintuitive and misleading. I would suggest to show the different factors on a plot with the same scale (0 to 1) and add an additional plot (either separate or as an inset) with the second scale.

- Figure 5: Legends for the plots in the uppermost row seem to be missing. Generally, it is not readily clear with legend applies to which subplot.

- Figure 6: It is difficult to grasp what part of the verification data is shown in the different subplots (i.e. model input or output of the downscaling method). I would suggest to improve the figure headings and the caption text to guide the reader better through the figure.

- Figure 7: Please clarify the abbreviations in the figure, e.g. nf and ff. This might be guessed from the manuscript text but should also be made clear somewhere in the figure or the figure caption.