Hydrology and Earth System Sciences Discussions

Hydrol. Earth Syst. Sci. Discuss., referee comment RC2
https://doi.org/10.5194/hess-2022-122-RC2, 2022

Comment on hess-2022-122

Anonymous Referee #2

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-RC2, 2022

Schneider et al proposed a RF-based downscaling method to downscale changes in the simulated water table depth over Denmark from 500 m resolution to 100 m resolution under different future climate scenarios. The method was trained on data from five submodels that cover a wide range of geologic, topographic, and hydrologic variability occurring across Denmark, and validated on data from another submodel (VI). The results obtained by the proposed method outperformed 500m-resolution water table depth and its bilinear interpolation in showing the climate change-induced changes to the shallow groundwater table. The paper would be of interest to the hydrological community. Overall, it is well-written and the related questions are discussed thoroughly. However, I have the following concerns regarding the paper.

General comments:

1. Traditional downscaling techniques downscale a product at a coarse resolution to the same product at a finer resolution. Here the authors used different statistics calculated from the coarse-resolution product (TBDV). Why didn't the authors directly use the 500m water table depth as a covariate here?

2. Which criteria did the authors use to select their validation submodel (VI. Aarhus Å/Aarhus)? From Fig.1, the submodel has a very shallow mean water table depth (0.5-2.5m). There are many areas in Demark having water table depth > 10 m, like the submodel V. I wonder if the selection of VI would give a biased conclusion for the RF validation.

3. I really like the idea to study the importance of each covariate (feature) used in RF. I also think that determining the feature importance based on ML model performance is a

feasible method. However, the authors may need to check the independence of their covariates before implementing such an approach. If two or more covariates are strongly correlated, perturbing one of them may not impact the ML performance, which leads to wrong results. I would like to know how the authors dealt with this issue.

4. Please improve the quality of the figures.

Specific comments:

1. Line 102, Page 4: "referred to the provided literature". Which literature? (Abbott et al., 1986; DHI, 2020)? Please specify there.

2. Line 112-114, Page 4: I am not an expert in hydrological model simulation, and I am a bit confused here. The authors mentioned that precipitation, temperature, and potential ET used for historic climate forcing to the DK-model HIP have various resolutions, 10 km or 20 km. However, in Line 105, they mentioned that all input data have a spatial resolution of 100 m. Therefore, did they downscale historical climate forcing data to 100 km or use them directly?

3. Climate models, Page 5: Can the authors clarify which 17 RCMs they chose and which 5 RCMs are used as a subset?

4. Line 173, Page 6: Why did the authors use changes to the 1m exceedance probability? Can the authors explain the practical meaning of this statistic?

5. Line 235, Page 8: "RF is a supervised ML learning method; that means it requires training data". This statement is wrong in my opinion. Unsupervised ML methods also require training data. I think here the authors meant supplementary teacher signals that are used to guide the training process. In addition, ML is the abbreviation for machine learning. Please delete the extra "learning" here.

6. Please mark the locations of dummy points used in RF training in Fig.1 if possible.

7. Line 276, Page 9: I believe there should be Table 3.

8. Line 335, Page 11: which statistic does "the climate change-induced changes to the

shallow groundwater table” indicate?

9. Fig.7: Please explain the legends (e.g., 500m HM intp) in the caption.