

Referee comment: Technical note: HydroModPy - a Python toolbox for deploying catchment-scale shallow groundwater models

Gauvain et al. present the framework *HydroModPy* that constitutes a method for script-based construction of groundwater models. The framework uses existing functions from well established packages like *WhiteboxTools*. The framework allows to extract watershed delineation information, setup the model with the required parameter data, calibrate the model against observation data, and export model results for further analysis or directly visualise them.

I appreciate the author's work to develop a framework that simplifies the creation of groundwater models for catchments. While referee 1 already pointed out quite some important points, I feel that currently the paper is missing a recognition of how uncertainty and equifinality might influence the model calibration framework. Coming from a hydrological modelling perspective, uncertainty of input data, model structure, model parameters and even scenario uncertainty are relevant because those influence the model results. While I don't think a technical implementation has to be included at this point for this paper, the relevance and its implications for the framework application should be discussed. This especially applies to potential equifinality of model parameters. Here, a recognition of this aspect should be included and, in for future version, technical approaches to evaluate uncertainty and equifinality in model calibration could be discussed.

Major comments:

Lines 41 to 47 and lines 455 to 461: In the introduction, you mention the disadvantage of GUI-based model development environments in their limitation to explore parameter sensitivity and uncertainty of model parameters. However, *HydroModPy* currently does not allow for sensitivity analysis or estimation of parameter uncertainty as well. Then, later in lines 455 to 461 you mention that *HydroModPy* is further developed to incorporate a graphical interface. This seems to clash with what you wrote in the introduction.

Line 114, Figure 1: Why is model calibration not a step in this workflow overview if it is quite central for model development?

Line 127: That sounds like the buffer value is chosen quite arbitrarily? Wouldn't it make more sense to have a workflow where this model domain extension is chosen based on the discrepancy between discharge-effective precipitation and measured discharge? In *KarstMod*, for example, the parameter RA (Recharge Area) scales the effective area of the catchment to close the water balance when precipitation and ETa are well constrained.

Line 144: If the model takes groundwater recharge as input rather than precipitation, isn't that limiting the applicability? I would guess that groundwater recharge time series data would have to be derived from hydrological models.

Line 321: Does that mean the whole 30 m aquifer is simulated as a single layer? I feel like that could be a calibration rather than parameterisation parameter.

Lines 339 to 340: Does that mean that the evaluation was done over 10 values? I think the sentence can easily be misunderstood.

Minor comments:

Line 3: development

Line 157: coefficient

Line 158: Point within sentence

Line 159: either implemented or specified

Lines 159 to 161: Both sentences say the same

Lines 164 to 165: Why is Figure 2d linked here?

Line 359: perennial

Line 361: Citation incorrectly embedded in text