

Exercise 2: model structure uncertainty

1 Introduction

In this exercise, you will calibrate two different MARRMoT models for two different catchments and compare model performance after calibration. By the end of this exercise you are expected to:

- Be able to navigate model documentation and the inner workings of hydrologic model code;
- Critically think about the relationship between model structure, catchment structure and model calibration and evaluation procedures.

The model setup approach used in this exercise (i.e. splitting data into calibration and evaluation periods and using an efficiency score to find the mathematically optimal parameter set) is common in hydrology but leaves ample room for improvement. The intent of this exercise is to highlight certain weaknesses of this approach and to encourage critical reflection on model calibration, evaluation and interpretation of model results.

2 Assignment description

You have been provided with data for the catchments Middle Yegua Creek, Texas, and Raging River, Washington, extracted from the CAMELS data set (Addor et al., 2017). You will calibrate MARRMoT models m02 and m03 (Knoben et al., 2019) for these catchments.

2.1 Exercise 1 – Get to know the catchments and models

First, do a brief investigation into the two catchments and models, to better understand what you are working with. Time series of meteorological forcing and streamflow observations are provided in the file `Part 2 - catchment data.mat`. Catchment attributes are provided in the file `Part 2 - catchment attributes.xlsx`. Descriptions of the models can be found in the Supplement of Knoben et al. (2019): <https://gmd.copernicus.org/articles/12/2463/2019/gmd-12-2463-2019-supplement.pdf>.

Exercises:

1. Load the time series data and create figures to show the meteorological forcing and streamflow observations for both catchments. Are these catchments water-limited or energy-limited?

2. Use the catchment attributes file to get an idea of the hydrologic conditions in these catchments. For both catchments, what is the average (1) aridity, (2) snow fraction, (3) runoff ratio, (4) streamflow, (5) fraction of the catchment covered by forest? Based on the information available, which hydrologic processes do you think are important in each catchment?
3. Briefly investigate model structures m02 and m03. For both models, what (1) is the number of parameters, (2) what is the number of stores, and (3) which hydrologic processes are considered?
4. Run each model for each catchment. You can use random parameters and arbitrary initial storage values. What are the Kling-Gupta efficiency scores the models obtain with these random settings?
5. Conduct a minor sensitivity analysis by trialing different parameter sets. Are there differences in parameter sensitivity? What is the best KGE score you were able to obtain?

Notes:

- For loading and plotting data, the following MATLAB commands may be useful: `load`, `figure`, `plot`.
- MARRMoT workflow example 1 may be adapted to complete the modelling part of this exercise.

2.2 Exercise 2 – Model calibration

Now that you have a feeling for the catchments and models, estimate parameter values by calibrating both models for each catchment. Test the calibrated parameter performance on unseen evaluation data.

Exercises:

1. Set up scripts to calibrate each model against streamflow observations for each catchment. Use data for the period 1989-01-01 to 1998-12-31 for calibration and the Kling-Gupta efficiency (KGE, [Gupta et al., 2008](#)) as an objective function.
2. Evaluate the KGE performance of the calibrated parameter sets on unseen data, i.e. data from the period 1999-01-01 to 2009-12-31.

Notes:

- For efficient use of time, it is recommended to set up your calibration and evaluation code using only a small subset of the available data. Only run the full 10-year calibration when you are satisfied that the code works is expected.
- MARRMoT workflow example 4 may be adapted to complete this exercise.

2.3 Exercise 3 – Comparative assessment of model performance

Having calibrated two models for two catchments, critically assess difference and similarities in model performance.

Exercises:

1. Compare calibration and evaluation scores of both models for catchment 08109700 (Middle Yegua Creek). Based on what you know of this catchment’s streamflow regime and the two model structures, which difference between the model structures do you think causes the difference in performance?
2. Compare calibration and evaluation scores of both models for catchment 12145500 (Raging River). Based on the objective function scores, is it clear which model is the most appropriate choice for this catchment?
3. Compare calibration and evaluation performance for both catchments for model **m02**. Based on these scores, do you think you can safely apply model **m02** in any catchment, regardless of the catchment’s conditions?
4. Compare calibration and evaluation performance for both catchments for model **m03**. Based on what you know of both catchments’ streamflow regime and the structure of model **m03**, is it possible that this model truthfully reflects the dominant hydrological process in both catchments? If so, does that mean your interpretation of the model’s behaviour and fluxes is different for both catchments? If not, do you think the KGE scores alone are sufficient to determine in which catchment the model produces “the right results for the right reasons?”
5. Given the comparisons above, formulate at least three take home messages about model structure uncertainty.

References

- Addor, N., Newman, A. J., Mizukami, N., and Clark, M. P. (2017). The CAMELS data set: catchment attributes and meteorology for large-sample studies. *Hydrology and Earth System Sciences*, 21:5293–5313.
- Gupta, H. V., Wagener, T., and Liu, Y. (2008). Reconciling theory with observations : elements of a diagnostic approach to model evaluation. *Hydrological Processes*, 3813(22):3802–3813.
- Knoben, W. J., Freer, J. E., Fowler, K. J., Peel, M. C., and Woods, R. A. (2019). Modular Assessment of Rainfall-Runoff Models Toolbox (MARRMoT) v1.2: An open-source, extendable framework providing implementations of 46 conceptual hydrologic models as continuous state-space formulations. *Geoscientific Model Development*, 12(6):2463–2480.