Supplement of:
Teaching hydrological modelling: Illustrating model structure uncertainty with a ready-to-use computational exercise

Wouter J. M. Knoben[1] and Diana Spieler[2]

[1]University of Saskatchewan Coldwater Laboratory, Canmore, Alberta, Canada
[2]Institute of Hydrology and Meteorology, Technische Universität Dresden, Dresden, Germany

Correspondence: diana.spieler@tu-dresden.de

This document contains the following elements:

- S.1 Survey on teaching uncertainty in hydrology
  - S.1.1 Survey outline: a description of the survey circulated among educators in hydrology
  - S.1.2 Survey questions and answers
  - S.1.3 Main survey results
  - S.1.4 Text-based responses to question 15 in the survey
- S.2 Examples hand-outs that guide students through the exercise and towards the learning objectives
- S.3 Course Evaluation
  - S.3.1 Informal Course Evaluation
  - S.3.2 Course preceding Questionnaire
  - S.3.3 Course succeeding Questionnaire
- S.4 Parameter Sampling with full parameter range

S.1 Survey on Teaching Uncertainty

S.1.1 Survey Outline

In order to get a clearer picture on if and how uncertainty is currently taught to students in the field of Earth- and Environmental Sciences a quick survey on "Teaching Uncertainty in Hydrological Modelling" was conducted via the survey software surveymonkey.com. The main questions we wanted to answer were:

1. How commonly is uncertainty in hydrological modelling part of the teaching curriculum in water resources (related) courses?
2. Are data uncertainty, parameter uncertainty and model structure uncertainty equally often covered in the curriculum?
3. Which tools are being used to teach hydrological modelling and, specifically, how often are modelling exercises part of the curriculum?

We wanted to design the survey as short as possible in order to get many participants to take the time to answer it. We hoped that a large number of participants would allow a broad general overview on the topic. The survey was distributed via twitter, different e-mail lists and by explicitly searching for and contacting different hydrology institutes and water research related faculties. Additionally, a flyer to the survey was passed out during AGU 2019. We were able to collect 101 answers in approximately 6 months (the survey was open between 18.09.2019 and 06.03.2020). The survey had between eight and eleven questions depending on the path the participants took through the survey (see Figure S1) and took approximately three minutes to answer.

In the following section the survey questions and answer possibilities are listed with the number of responses in blue. Note that different paths through the survey were possible depending on which answers were given, resulting in consecutive questions occasionally having a different number of respondents. Due to the way the survey tool works, certain questions needed to be duplicated to let them be part of different paths. The different survey paths and their corresponding questions can be seen in Figure S1 displaying a flow diagram of the survey with its introductory text.

S.1.2 Survey questions and answers

Q1 On which level do you teach? - 101 responses

- Undergraduate Level (BSc) - 30
- Graduate Level (MSc) - 28
- Under- and Graduate Level (BSc & MSc) - 40
- Other (please specify) - 3

Specifications of "Other" were: PhD; Public stakeholders in FRM; TX groundwater conservation district.

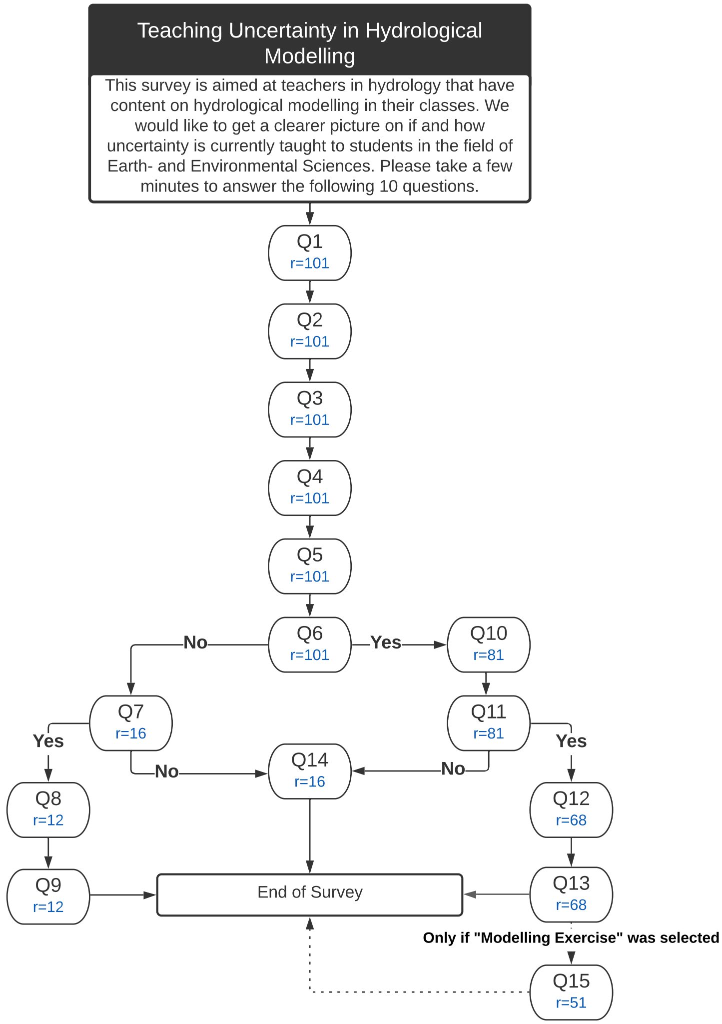

Figure S1. Flow diagram of the survey showing different question paths and the introductory text. The number of responses each question received is given in blue.

Q2 In which field do you teach? - 101 responses

- Hydrology - 50
- Water Resources and Management - 21
- Civil or Environmental Engineering - 21
- Geography - 5
- Other (please specify) - 4

Specifications of "Other" were: Hydrology and Geology; Environmental Modelling; Hydrogeology; Urban Drainage.

Q3 Where do you teach? - 101 responses

- Europe - 52
- North America - 20
- South America - 5
- Africa - 8
- Asia - 8
- Australia - 6
- Other (please specify) - 2

Specifications of "Other" were: Globally; Both - South America and Europe.

Q4 How long have you been teaching a course with hydrological modelling content? - 101 responses

- < 2 years - 35
- < 5 years - 24
- < 10 years - 14
- > 10 years - 28

Q5 What percentage of your total lecture time do you spend on hydrological modelling compared to all other class content? - 101 responses

- state a number between 0 and 100 - The average of all given answers is 44.36 %, but answers vary between 1 and 100 %.

Q6 Does your class include exercises on hydrological modelling? - 101 responses

- Yes - 84
- No - 17
- Other (please specify) - 0

Q7 If there is no exercise on hydrological modelling, is uncertainty in hydrological modelling part of your course? - 16 responses - 1 skipped answer

- Yes - 13
- No - 3

Q8 Please state which sources of uncertainty you are discussing in your hydrological modelling class. Select all options that apply. - 12 responses - 1 skipped answer

- Input Uncertainty - 11
- Model Parameter Uncertainty - 10
- Model Structural Uncertainty - 8
- Uncertainty as Overarching Term - 5
- None - 0
- Other (please specify) - 1

Specifications of "Other" were: initial conditions uncertainty.

Q9 Which tools do you use to communicate uncertainty in hydrological modelling? Select all options that apply. - 12 responses

- Slides & Graphics - 10
- Papers - 5
- Books/Book Chapters - 1
- Other Reading Material - 1
- Case Studies - 6
- Active Discussion - 7
- Other (please specify) - 1

Specifications of "Other" were: serious gaming.

Q10 Which tools do you use to teach hydrological modelling? Please use the "other" option if you wish to specify the tools and/or models you use. Select all options that apply. - 81 responses - 3 skipped answers

- Thought Experiments - 22
- Pen & Paper Exercises - 37
- Labratory Setup - 11
- Computer Exercises - 76
- Other (please specify) - 10

Specifications of "Other" were: Interactive activities such as discussion, presentation of case studies by students, attendance to scientific conferences at the international level; Multimodel exercises, Decision enactments; HEC-HMS; Virtual Reality; Structured group discussion; Literature review assignments on modelling (written and presentation). Actually teach an environmental modelling course with content that is equally applicable to hydrological modelling and other fields.; Lectures; field course - including relationships between field work (to derive parameters) and hydrologic modelling; Hbv- light, building small models with Matlab code; IWG-HW.

Q11 Is uncertainty in hydrological modelling part of your course? - 81 responses

- Yes - 68
- No - 13

Q12 Please state which sources of uncertainty you are discussing in your hydrological modelling class. Select all options that apply. - 68 responses

- Input Uncertainty - 57
- Model Parameter Uncertainty - 60
- Model Structural Uncertainty - 45
- Uncertainty as Overarching Term - 39
- None - 0
- Other (please specify) - 6

Specifications of "Other" were: Formulation of hydrological models in a stochastic physically based framework therefore integrating uncertainty assessment with model identification, estimation and application; Model predictive uncertainty; Scenario uncertainty; Propagation of uncertainty; Forecast uncertainty; Cascade of uncertainties in decision making.

Q13 Which tools do you use to communicate uncertainty in hydrological modelling? Select all options that apply. - 68 responses

- Slides & Graphics - 53
- Papers - 40
- Books/Book Chapters - 21
- Other Reading Material - 10
- Case Studies - 33
- Active Discussion - 44
- Modelling Exercise - 58
- Other (please specify) - 4

Specifications of "Other" were: Open tutorial web pages published in my website. Open videos of all my lectures published on You Tube; Information Theory Approach, Non-linear dynamics, Chaos, Fractals (stochastic), self-similar processes; Blackboard; A model sensitivity of SWMM is part of the homework.

Q14 Why is uncertainty not part of your course? Select all options that apply. - 16 responses

- Covered in another course - 1
- Not enough time to cover it - 13
- No good teaching materials available - 5
- Other (please specify) - 2

Specifications of "Other" were: I wish it was! Just didn't get around to it. Was not in the textbook I used; Bachelor students first need to grasp the concepts of modelling before introducing uncertainty.

Q15 Please give a short description of the modelling exercises you use to teach uncertainty to your students. Are you differentiating different sources of uncertainty? - 51 responses - 7 skipped answers

- Participants had the possibility to give a text based answer.

The text based answers can be found in Section S.1.4.

S.1.3 Main Results

Figures S2 and S3 provide an overview of the main results of each survey question. The text based answers to survey question 15 can be found in Section S.1.4 Note the different possible answer paths and their corresponding questions depicted in Figure S1. Based on these survey results, the answers to the 3 main questions stated in Section S.1 can be summarized as follows. Because the main focus of this survey is generating insight on the use of exercises in teaching modeling uncertainty, answers are split into a group that uses exercises and a group that does not.

General characteristics of respondents

Respondents teach at the undergraduate and graduate level in approximately equal numbers (Figure S2 - Q1). They refer to their field as "Hydrology" in nearly half of cases, with "Water Resources and Management" and "Civil/Environmental Engineering" appearing in 20% of cases each (Figure S2 - Q2). "Geography" and "Other" make up the remainder. Approximately half of respondents are based in Europe, with 20% being based in North America. Between 5 and 10% of respondents are based in each of South America, Africa, Asia and Australia (Figure S2 - Q3). Slightly more than a third of respondents is relatively new to teaching (<2 years experience) and more than a quarter are very experienced teachers (>10 years; Figure S2 - Q4).

How commonly is hydrologic modelling uncertainty taught?

On average the researchers and lecturers that took the survey spend about 44% of their lecture time on teaching hydrological modelling, but this can be as little as 1% or as high as 100%. Figure S2 - Q5 shows a histogram with a bin size of ten percent indicating that 68% of teachers spent less and 32% spent more than half their lecture time on teaching hydrological modelling. Modelling exercises are used by 83% of respondents (Figure S2 - Q6), and in approximately four-fifths of those cases uncertainty is part of the curriculum (Figure S3 - Q11). Those not using any modelling exercises nevertheless teach uncertainty in hydrological modelling in 81% of the cases (Figure S3 - Q7). If modelling uncertainty is not covered in class, the most common reason is lack of time with lack of good materials coming second. Just 6% stated the topic would be covered in another course (Figure S3 - Q14).

Are data, parameter and structure uncertainty equally commonly taught?

The focus on the different sources of uncertainty varies between both groups. Those not using exercises have a slightly stronger focus on teaching input uncertainty (92% compared to 83% parameter uncertainty and 67% structural uncertainty; Figure S3 - Q8) while those using modelling exercises focus sightly more on parameter uncertainty (88% compared to 84% input uncertainty and 66% structural uncertainty; Figure S3 - Q12). Model structural uncertainty is the least taught source of uncertainty, not appearing in the curriculum in a third of all cases.

How is uncertainty in modelling taught and, specifically, how often are exercises used?

The most used tools to teach uncertainty in the groups with and without exercises are slides and graphics as well as active discussion (Figures S3 - Q9 and Q13). The group conducting exercises uses computer exercises in 94% of cases (Figure S3 - Q10) and uses these to communicate modelling uncertainty in 85% of cases. Question 15 reveals a wide variety in the modelling exercises used to teach uncertainty. Notable examples include a combination of fieldwork to obtain observations and modelling based on the fieldwork; exercises to combine calibration, sensitivity analysis and uncertainty analysis; and exercises to trial different model parametrizations.

Figure S2. Results to questions 1 to 6 of the survey. Number of responses to each question is given in brackets. Note that Q5 shows a histogram with a bin size of 10 percent.

Figure S3. Results to questions 7 to 14 of the survey. Number of responses to each question is given in brackets. Q7 to Q9 were answered by respondents that do not conduct exercises on hydrological modelling while Q10 to Q13 describe the responses of the group that does conduct modelling exercises. Q14 was answered by those respondents that stated uncertainty was not part of their course in Q7 and Q11.

S.1.4 Text based responses to question 15

The following is a list of answers to the question "Please give a short description of the modelling exercises you use to teach uncertainty to your students. Are you differentiating different sources of uncertainty?".

- Students are measuring discharge during an excursion and use afterwards in a seminar the hydrological model HYMOD to reproduce the observed discharge, together with given longterm data. They get to know all the measurement instruments (uncertainty discussed), the model structure (uncertainty discussed) and the model parameters (uncertainty discussed). On the gained results, they write a short 2 page thesis.
- The students design their own models for the same catchment. We talk about input uncertainty, but they experience parameter uncertainty during calibration of their own model and structural uncertainty in comparison with the results of their peers.
- The use of Monte Carlo approaches to the assessment of risk
- I teach environmental modeling at the senior level in an undergraduate engineering program. We model pollutants moving across various media including rivers and lakes. We introduce uncertainty through a simple dynamical system that is sensitive to initial conditions. We perform multiple runs with small initial perturbations and study the evolution of the errors as they grow with time until they reach a saturation level. We discuss uncertainty also when performing Mote-Carlo simulations of a recharging aquifer by perturbing the model parameters.
- The class is for Honours years (last year of BSc) and their first exposure to Hydrology, so it is through discussion of results and papers that this is done
- Monte Carlo, Manual calibration by different students, ...
- Rainfall-runoff modelling, groundwater modeling. I am assessing uncertainty in an integrated framework within the stochastic physically based modeling
- Parameter uncertainty of many different models. I consider different sources of uncertainty.
- Assessing different rainfall-runoff model structures (incl. their individual parameter uncertainty) on a range of catchments. Uncertainty quantification and attribution of a hydrologic model in much greater detail (incl. forcing and parameter uncertainty) using Global Sensitivity Analysis methods.
- My approach doesn't fall into your traditional paradigm of uncertainty
- Students represent competing stakeholder groups. Use multimodel analysis to make uncertain forecasts as basis for negotiations. Use models to identify observations. Consider measurement, parameter, and structural uncertainty.

- Students build and implement two models from scratch as part of their learning experience. One is a physics/conceptual watershed-scale state-space rainfall-runoff model. The other is a dynamics systems model for any system of their choosing. Model development includes (1) perceptual-conceptual modeling, (2) mathematical modeling, (3) computational modeling (including numerical implementation), (4) model calibration and evaluation, (5) uncertainty-sensitivity analysis, and (6) model improvement. Students are encouraged to investigate all sources of uncertainty, and to think about models and data from an information theory perspective.

- Monte Carlo simulations of effect of uncertainty in hydrologic properties on the model results; model calibration demonstration that multiple parameter sets can achieve reasonable (or not) matches to uncertain data sets

- For example stochastic rainfall modeling, with uncertain structure and parameters. Also the transfer of such uncertainty to the other components of the hydrologic cycle.

- Primarily looking at how a random hydraulic conductivity field influences model estimates

- We use Rational method and Unit hydro graph method to check input uncertainty on runoff generation

- Calc different process algorithms in excel, play with parameters

- I discuss different uncertainties such as analytical uncertainties of input data (e.g., water table or tracer measurements) and how to propagate these uncertainties for parameters that depend on measured data.

- Creating synthetic rainfall and modifying parameters

- We discuss various sources of uncertainty, but for the computer modeling we mostly focus on parameter uncertainty. We use Excel for undergraduate and MATLAB for postgraduate courses.

- Yes I am: model and input data uncertainty.

- We talk about different sources of uncertainty but in the modelling exercises use the GLUE method to investigate equifinality and parameter uncertainty.

- Comprehensive approach ranging from analysis of how uncertainty has been communicated, application of parameter estimation techniques, to scenario-based exercises. Fundamental idea is to understand the concept that results can change depending on the decisions made in an analysis, and this has a variety of consequences

- use same model with different equation options, e.g. Penman Montieth, Hargreaves and Samani etc. Look at changes in soil parameters and influence on stream flow. impact of rainfall measurements on modelling, nested gauges, radar satellite, diff temps and altitudes. Also look at stochastic flows vs actual, errors in gauges and networks.

- One-factor-at-a-time uncertainty analysis: Varying model parameters by -20%, -10%, 0%, +10%, +20% and then looking at different measures for Goodness of Fit of model results

- not explicitly; students are asked to consider uncertainty in the data, inputs, etc. when calibrating a model
- E.g. using SWAT-CUP mainly for parameter uncertainty and ensemble inputs for input uncertainty.
- Exercise is only on model parameter uncertainty. Students change model parameters, run the model, and look at the results
- yes
- Mainly through consideration of uncertainty in calibrated model parameter values and how these propagate through a model, and by comparing the performance of different model structures.
- Catchment modelling and broader system design problems using optimisation, sensitivity analysis, cross-validation, monte carlo experiments and GLUE and of which are either subject to uncertainty or characterise uncertainty.
- I use GLUE in the computer model simulations we do of hydrological processes
- We set up a hydrological model and discuss sources of uncertainty during each step of the model development.
- hydrological forecasts with ensemble forecasts as input; parameter equifinality in manual and automated parameter identification
- It is about input uncertainty using a normally distributed error model combined with an sensitivity analysis on calibrated parameters.
- GLUE & Top Model, SWAT-CUP, Error propagation, STD in regression parameter
- So far, i did not use modelling exercises
- Evaluate different model structures in different catchments including uncertainty analysis. Tools used are RRMT (Wagener et al., 2011, HESS) and SAFE (Pianosi et al., 2015, EM&S)
- I teach urban hydrology. As a course achievement, the students have to dimension a drainage network using EPA SWMM. Together with a Jupyter notebook, they are asked to perform a parameter sensitivity study acknowledging valid ranges for each parameter in runoff generation, runoff concentration, and in the hydrodynamic part of the model. Structural uncertainty is at least touched slightly, since they analyze differences in the hydrological response for two modeling approaches: the kinematic wave and the full dynamic model. However, this only discussed in brief, since the course on urban hydrology also involves other aspects of the topic such as processes, hydrologic design, blue and green infrastructure, and interdisciplinary approaches.
- ...
- Use different design floods for reservoir dimensioning

– HBV light model is used to simulate streamflow. In a guided exercise, students manually tweak selected parameters to see how the simulations change. We discuss Keith Beven's book Rainfall-Runoff Modelling: The Primer and discuss parameter equifinality. We use Monte-Carlo simulations to create different input parameter sets and test their effect on simulations. We also discuss the use of simulations of different climate models (precipitation and temperature) as input to HBV and the associated uncertainties.

– Uses Swat and Swat cup.

– HBV model combined with automatic calibration, sensitivity analysis, and uncertainty analysis

– Different lectures cover diagnostic methods to approach uncertainty quantification; we cover parameter and input uncertainty in greater depth, structural uncertainty we talk about but no exercises

– Teaching the difference between aleatoric and epistemic uncertainty. Tools to engrave uncertainty in model development and in deciding making

– We discuss epistemic, aleatory Uncertainty and ambiguity. Students design a case specific Uncertainty matrix, addressing uncertainties related to: input, parameters, model structure, context and decision processes. Students analyse input and parameter Uncertainty through model exercises

– I used SWAT for making students understand parameter uncertainty. Input data from different sources was used to assess input data related uncertainty.

– Calibration and dotty plots, Use three differently structured hydrological models

– Download and visualize climate projections or hydrological datasets and examine how they disagree with each other

– comparing own temperature measurements with station data (finding best matching station vs. nearest)

S.2 Suggested exercise handouts

This section contains two example exercise descriptions that can be given to students. Please note that LaTeX source files of these exercise hand-outs, as well as separate PDF files, can be found on GitHub:
https://github.com/wknoben/Dresden-Structure-Uncertainty.

Teaching hydrologic modelling

Showcasing model structure uncertainty with a ready-to-use computational exercise

Exercise 1: MARRMoT basics

1 Introduction

The MARRMoT toolbox comes with 4 prepared workflow examples that form the basis of this introductory exercise. Each workflow script is a stand-alone application of MARRMoT to a specific question. Students are encouraged to run the workflow examples as they are given, and then adjust each workflow example based on the exercise below. By the end of this exercise you are expected to:

- Have basic understanding of MARRMoT functionality;
- Be able to calibrate a hydrologic model and create diagnostic graphics that show the simulation results.

2 Install instructions

The MARRMoT source code can be obtained through GitHub. If you have a GitHub account:

1. Fork the repository https://github.com/wknoben/marrmot to your own GitHub account;
2. Create a local clone of the forked repository.

If you do not have a GitHub account:

1. Go to https://github.com/wknoben/marrmot;
2. Download the repository as a `.zip` file ([Code] > [Download ZIP]);
3. Extract the downloaded files.

Once the files are on your local machine:

1. Remove the folder `Octave` from the downloaded files if you plan on using Matlab;
2. Start Matlab or Octave;
3. Add the downloaded `MARRMoT` folder to your Matlab/Octave path.

3 Assignment description

The workflows that form the basis of the exercises below can be found in the downloaded `MARRMoT` folder, inside the subfolder `User manual`.

3.1 Workflow example 1 – use a model

Workflow summary: use the HyMOD model (MARRMoT ID m29) for streamflow simulation in an example catchment with a pre-defined parameter set and arbitrary initial conditions. Assess the accuracy of the simulation through the Kling-Gupta Efficiency metric.

Exercises:

1. Run the workflow example and compare the resulting streamflow simulations with observations. What would you improve in the model setup?

2. The given parameter set performs poorly in this catchment. Try a set of different parameters. Did the simulation improve?

Notes:

- You can use the function `m_29_hymod_5p_5s_parameter_ranges` in `./MARRMoT/Models/Parameter range files/` to view plausible parameter ranges.

- Adjust the initial storage values if required. It is good practice to not initialize a store above its maximum capacity.

3.2 Workflow example 2 – parameter sampling

Workflow summary: use the HyMOD model for streamflow simulation by randomly sampling parameter values from MARRMoT's provided parameter ranges. Plot the ensemble of simulations.

Exercises:

1. Run the workflow example and compare the resulting streamflow simulations with observations. What would you improve in the model setup?

2. It is possible that the HyMOD model is not suitable for this particular catchment. Choose a different model structure and re-do the sampling. Does performance improve? Why (not)?

Notes:

- The MARRMoT documentation (https://gmd.copernicus.org/articles/12/2463/2019/gmd-12-2463-2019-supplement.pdf) contains an overview of model structures.

– You may need to adjust the initial storage values to fit the new model choice. Ensure that you specify the same number of initial storage values as your chosen model has stores. Set the initial storages to values that fall within the possible range of each store in the new model.

3.3 Workflow example 3 – model sampling

Workflow summary: use three different models with a randomly selected parameter set for streamflow simulation.

Exercises:

1. Run the workflow example and compare the resulting streamflow simulations with observations. What would you improve in the model setup?
2. Plot the evaporation simulations from each model. Do these look better than the streamflow simulations?

Notes:

– All simulation outputs can be found in the cell array "results_sampling". Check the model function call in the workflow script to see which results are stored where (lines 93-111).

3.4 Workflow example 4 – model calibration

Workflow summary: calibrate the HyMOD model for streamflow simulation.

Exercises:

1. Run the workflow example and compare the resulting streamflow simulations with observations. What would you improve in the model setup?
2. Plot timeseries of HyMOD's simulated storage during the evaluation period. Do these look reasonable?
3. Estimating appropriate initial storage values is not attempted in these workflow examples. Is there a way to find better initial guesses of these initial conditions? If not, how can the model simulations be used in a way that reduces the impact of poor initial storage estimates?
4. (Optional) Implement a way to improve the initial storage guesses or a way to reduce the impact of poor guesses.

Notes:

– The workflow code currently doesn't save the simulated storage values. Adjust line 166 so that it does. See the script of workflow example 3 for an example of how to do this.

– The script for workflow example 1 contains code that plots time series of simulated storage values.

Teaching hydrologic modelling

Showcasing model structure uncertainty with a ready-to-use computational exercise

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.

S.3 Course Evaluation

S.3.1 Informal Course Evaluation

The first application of the proposed MARRMoT exercises was a workshop in June 2019 where the attendees were both students (2 PhD & 4 MSc students) and academic or scientific staff (5). The intent of the workshop was to trial prototype exercises which could potentially be included in the curriculum of the first semester "Hydrological Modelling" module of the Hydrology Master Program at TU Dresden. The course was run over the span of two afternoons. During the first afternoon, attendees attended a 1-hour seminar about model structure uncertainty and spent approximately 2.5 hours working on Exercise 1. During the second afternoon, attendees spent approximately 4 hours on Exercise 2. Attendees were divided into groups and each group initially worked with a single combination of one model and one catchment. The second afternoon included a final classroom discussion to tie the insights from individual groups together.

The attendees were asked to fill in a short anonymous feedback form after the course was completed. The form had several questions that had to be answered on a 1 to 5 scale and three open questions. A summary of responses (4 MSc students, 1 PhD student, 3 Postdoctoral Fellows; senior faculty members provided verbal feedback) that could be answered on a 1-5 scale is shown in Figure S4. Attendees unanimously reported that the course was easy to follow and complete, and that the main messages were clear. The number of models and catchments used in the exercises was sufficient and attendees were able to improve their understanding of the implications of model structure and parameter choice. Various attendees noted that the initial setup for sharing modelling results of Exercise 2 between the different groups was somewhat unwieldy. Consequently, the provided example handout for Exercise 2 is set up to work for an individual student and avoids the need to define groups and share results.

S.3.2 Course Preceding Questionnaire

MARRMoT Course Module - Preparation

Thank you for answering this survey.

Over the next few weeks, you will be doing several computer exercises that teach you valuable lessons about hydrological modeling. The goal of this survey (and it's follow up at the end of the semester) is to investigate how well these computer exercises transfer the intended learning objectives. This first survey is intended to measure your existing knowledge of the topic the exercises will address. A follow up survey at the end of the semester will asses the changes in your knowledge. Completing this survey should take you approximately 5 minutes.

Please think about the following questions and select your appropriate answer. Numbers roughly correspond to

1 = strong negative rating

2 = negative rating

3 = undecided/average

4 = positiv rating

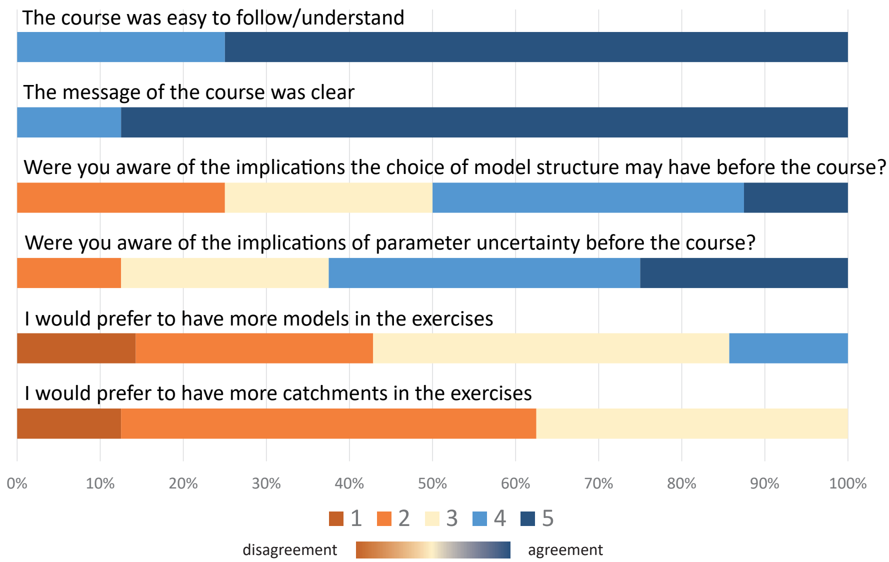

Figure S4. Responses to the feedback form distributed after the course. Only responses to the questions that had to be answered on a one to five scale are shown; responses to open questions are summarized in the main text. One indicates disagreement and five indicates agreement.

5 = strong positive rating

Questions marked with an * are mandatory.

1. Assume you start an internship in an engineering consultancy tomorrow. Your task is to predict the discharge time series for a given watershed by applying a hydrologic model. The expected output will be used for the planning and design of a rainfall retention basin, which the company wants to present to the local water authority in a few weeks. You are told which model is typically used by the consultancy and that all necessary input data is available. However, supervision won't be available because the hydrologist is currently ill. Thinking about the task at hand and your previous education (Bachelor + Master up till today). How prepared do you feel to complete the task? * - 15 responses
   - 1 (not prepared at all) 0 %
   - 2 (not very prepared) 20 %
   - 3 (a little prepared) 40 %
   - 4 (prepared) 26.7 %
   - 5 (very well prepared) 13.3 %
2. Please indicate why you feel you are (not) prepared? [open question] - 12 responses (not shown)
3. How familiar are you with the following sources of uncertainty related to hydrological modelling (theoretical knowledge about uncertainty)? * - 15 responses

Numbers correspond to: 1 (I don't know what this is), 2 (I know this exists but don't know details), 3 (I have some theoretical knowledge), 4 (I know a method to address this), 5 (I can confidently handle this aspect)

	1	2	3	4	5
Data uncertainty		6.7 %	53.3 %	33.3 %	6.7 %
Parameter uncertainty		26.7 %	46.6 %	26.7 %	
Model structural uncertainty	6.7 %	26.7 %	39.9 %	26.7 %	

4. Please provide some details on what you've learned so far regarding the theoretical aspects of uncertainty in hydrological modelling. [open question] - 7 responses (not shown)

5. You work on the previously described task to the best of your capabilities. How likely is it that you consider the following sources of uncertainty in your modelling? (practical knowledge about uncertainty) * - 15 responses
Numbers correspond to: 1 (not likely at all), 2 (unlikely to quantify but I might mention it), 3 (I'd adress this qualitatively), 4 (I'd address this in a basic quantitative way), 5 (I'd address this in a thorough quantitative way)

	1	2	3	4	5
Data uncertainty	6.7%	26.7 %	26.7 %	39.9 %	
Parameter uncertainty		33.3 %	26.7 %	33.3 %	6.7 %
Model structural uncertainty	6.7%	33.3 %	33.3 %	20.0 %	6.7 %

6. Following the previous question, can you briefly describe how you would account for each type of uncertainty you think you would consider? [open question] - 4 responses

   – parameter uncertainty -> sensitivity analysis, parameter variation, maybe monte-carlo approach

   – Data uncertainty is discussed based on the available measurements and their quality. Parameters are judged on the models ability to replicate former events by using the NSE or the coefficient of determination. I don't know how to assess model structural uncertainty.

   – Monte carlo analysis and Bayesian statistics can be used for all three types of uncertainties. However for parameter uncertainties multi-model analysis, inverse modelling can also be used.

   – Data uncertainty with time series analysis (homogeneity, data stationary, seasonality, data correction and gap filling, MLR, trend analysis); Parameter uncertainty with cross validation, LOO analysis, curve fitting and regression etc; Relative analysis of different models structures to get best results eg. NSE, RMSE etc

7. The model you are supposed to use in this project has not been applied to your catchment before, but has been used in a large number of other catchments. How confident are you, that the model will work well in your catchment if it worked well in the others? * - 15 responses

   – 1 (very unconfident) 0 %

– 2 (unconfident) 6.7 %
– 3 (undecided) 60 %
– 4 (confident) 33.3 %
– 5 (very confident) 0 %

8. Please explain why you gave the previous answer. [open question] - 10 responses (not shown)

9. You will need to calibrate this model for use in this particular catchment and measure the models accuracy with some performance metric (e.g. Nash-Sutcliffe efficiency, Kling-Gupta efficiency). Would you agree, that a high value on this score shows that your model is a realistic representation of the physical catchment processes? * - 15 responses

   – 1 (strong disagreement) 0 %
   – 2 (disagreement) 13.3 %
   – 3 (undecided) 60 %
   – 4 (agreement) 20 %
   – 5 (strong agreement) 6.7 %

10. Please explain why you gave the previous answer. [open question] - 9 responses (not shown)

11. You were able to calibrate your model with a satisfying score on your chosen calibration metric. You now want to predict the catchment behaviour for future conditions. Would you agree, that if a model works well for historic data, it will also work well for any future condition? * - 15 responses

   – 1 (strong disagreement) 0 %
   – 2 (disagreement) 6.7 %
   – 3 (undecided) 53.3 %
   – 4 (agreement) 40 %
   – 5 (strong agreement) 0 %

S.3.3 Course succeeding Questionnaire

MARRMoT Course Module - Evaluation

Thank you for taking the time to answer this survey.
Over the last 4 weeks, you solved several computer exercises that were meant to teach you valuable lessons about hydrological modeling. The goal of this survey is to investigate how well these computer exercises transferred the intended learning objectives of the MARRMoT course module. Please think about the following questions and select your appropriate answer.

Numbers roughly correspond to

1 = strong negative rating

2 = negative rating

3 = undecided/average

4 = positiv rating

5 = strong positive rating

Questions marked with an * are mandatory.

1. The MARRMoT course module was easy to follow/understand * - 10 responses
   - 1 (strong disagreement) 0 %
   - 2 (disagreement) 0 %
   - 3 (undecided) 20 %
   - 4 (agreement) 60 %
   - 5 (strong agreement) 20 %
2. The message of the MARRMoT course module was clear. * - 10 responses
   - 1 (strong disagreement) 0 %
   - 2 (disagreement) 0 %
   - 3 (undecided) 10 %
   - 4 (agreement) 50 %
   - 5 (strong agreement) 40 %
3. What were the main lessons you learned from this course module? * [open question] - 10 responses
   - How many different models there are and how important it is to choose the "right" one for the right reasons.
   - Question the model output, as there are too many uncertainties possible (model/input data/etc., indicate results with uncertainties
   - Model structure matters!, parameter calibration is not clear, existing models are used too hastily without checking their usefulness, general modeling process
   - Handling hydrological models in general. What to look for when it comes to choosing models. Modeling with matlab.
   - questioning different models for different catchments
   - Useage of the MARRMoT Toolbox and model seletion. How models from the toolbox behave.How model calibration works.Practicing to work with Matlab

- to find out a suitable model
- What MARRMoT is, what and how you can work with it respectively to the amount of time we had, of course its not possible to go very deep into it...here maybe a single module about MARRMoT or respective Hydrological Modelling Frameworks
- there are plenty options of possible models, learned how to calibrate and vary parameters
- There is a broad variety of conceptual hydrological models. Before applying them it is reasonable to think about what the model is capable of simulating and where it has shortcoming due to its model structure. Good performance in regard to objective funtions is not sufficient to claim to overall model performance as we search for "right results for the right reasons".

4. After attending the MARRMoT course module, how familiar are you with the following sources of uncertainty related to hydrological modelling (theoretical knowledge about uncertainty)? * - 10 responses Numbers correspond to: 1 (I don't know what this is), 2 (I know this exists but don't know details), 3 (I have some theoretical knowledge), 4 (I know a method to address this), 5 (I can confidently handle this aspect)

	1	2	3	4	5
Data uncertainty			50 %	50 %	
Parameter uncertainty		10 %	20 %	60 %	10 %
Model structural uncertainty		10 %	20 %	50 %	20 %

5. Please provide some details on what you've learned since taking the course on the theoretical aspects of uncertainty in hydrological modelling. [open question] - 6 responses

- While the course also discussed data and parameter uncertainty, the focus laid on raising awareness for model structural uncertainty. It reflected on how the model structure is related to the catchments and their climatic/topographic characteristics.
- learned how to vary different models and parameters by MatLab as well as evaluate the results
- nothing new so far but already attended other modelling/simulation courses
- How to asses data uncertainty and how to deal with it.
- The input data already shows various uncertainties, since measurement errors often occur during the collection. The parameters that are spit out by models can be faulty, since the numerics in the model are not always clear and a model can never fully depict reality.
- Model structure-related uncertainties have never really been discussed before, or rather this was my first conscious contact with the topic; Comprehensive calibration needs a certain amount of computing power or time due to the number of necessary calculation runs!; Each catchment area has its own characteristics and is therefore best described by other models; Humid areas can generally be better described by models than arid areas

6. How capable do you feel to consider the following sources of uncertainty in your modelling after attending the MAR-RMoT course module? (practical knowledge about uncertainty) * - 10 responses Numbers correspond to: 1 (not likely at all), 2 (unlikely to quantify but I might mention it), 3 (I'd address this qualitatively), 4 (I'd address this in a basic quantitative way), 5 (I'd address this in a thorough quantitative way)

	1	2	3	4	5
Data uncertainty		20 %	40 %	40 %	
Parameter uncertainty		10 %	40 %	40 %	10 %
Model structural uncertainty		10 %	40 %	20 %	30 %

7. Following the previous question, can you briefly describe how you would account for each type of uncertainty you think you would consider? [open question] - 4 responses

   – Data: check if the data can be corrected, analyse how good the measured data fit to the catchment. Parameter: optimize the parameter with calibration. Model: document the model selection. Find out the model which fits best, via comparing the outputs.

   – Data uncertainty - Bias correction; Parameter uncertainty - ; Model Structure uncertainty - Comparison with original Data, KGE, NSE, RMSE

   – As for the data and parameter uncertainty I would not change a lot, but I will reflect more strongly on the capabilities on the used model and potentially try to quantify the shortcomings through comparison.

   – Data: Description of data-related uncertainties; Parameters: sensitivity analysis, Monte Carlo sampling; Model: Comparison of different model outputs

8. The model you are supposed to use in a given project has not been applied to your catchment before, but has been used in a large number of other catchments. How confident are you, that the model will work well in your catchment if it worked well in the others? * - 10 responses

   – 1 (very unconfident) 10 %

   – 2 (unconfident) 40 %

   – 3 (undecided) 40 %

   – 4 (confident) 10 %

   – 5 (very confident) 0 %

9. Please explain why you gave the previous answer. Did the course contents change or manifest your previous opinion? In what way? [open question] - 9 responses

   – I guess its not advisable to just use what others have used before without understanding how and why it should work. Testing and evaluating beforehand might be the better choice.

- Because models are not universal. Just because they worked for other catchments doesn't mean they can easily be applied to other catchments. A new plausibility check is required every time a "common model" is used on a "new" catchment.
- The course taught me that models have to be handled careful with. I need to analyze and decide which model to use. Not only relying on the fact that it has worked on many other catchments.
- manifest
- It depends on whether the characteristics of the "new" catchment agree with the catchments already described by the model. But even if this is the case, it is not clear to what extent structural uncertainties were taken into account in the previous modelling. The course definitely changed my perspective on modeling practice in this regard, as I had previously not considered structural uncertainties as relevant or existent.
- Some models fit good to many catchments if the available data is good. But sometimes even a good model does not fit to a specific catchment even if it worked fine for many other.
- ATM i don't know if the model that worked for a lot of other catchments meet the requirements of what i want to do, but if the catchments in which the model performed well already are comparable to my catchment, i would be at least confident that the results can be further used
- very unconfident, because we have seen that there are a huge variation of results depending on the choice of model (and there are a lot) and depending on the parametrization (parametrization is also very uncertain cause some parameters are not easy to estimate and different parameter combination can lead to the same results)
- It really depends on the catchments that it was applied to and 1) how similar they are to the new catchment and 2) how good the model performance was in these. If there was only one model used before, I would also consider it beneficial to test alternative models.

10. You will need to calibrate your model for use in the previously mentioned catchment and measure the models accuracy with some performance metric (e.g. Nash-Sutcliffe efficiency, Kling-Gupta efficiency). Would you agree, that a high value on this score shows that your model is a realistic representation of the physical catchment processes? * - 10 responses
    - 1 (strong disagreement) 0 %
    - 2 (disagreement) 40 %
    - 3 (undecided) 10 %
    - 4 (agreement) 20 %
    - 5 (strong agreement) 30 %
11. Please explain why you gave the previous answer. Did the course contents change or manifest your previous opinion? In what way? [open question] - 8 responses

- The course thaugth me that a good KGE does not automatically mean the pysical processes of the chatchment are represented correctly.
- Yes, I was not cofident in using the KGE before.
- As long as the parameters I use are in the parameter range of the model or at least logic/plausible a higher value says that the model performs accordingly to the observed data, but not that it is a projection of reality
- A high score of these performance metrics can also achieved by unlogical parameter kombinations which are characteristically incorrect or do not describe the behaviour in a logical way (for example baseflow is not present in some periods but score is high)
- I still think that a high KGE/NSE is a hint towards the fact that the model is doing something right, however it is not sufficient to consider the score in these metrics.
- A good performance score can be achieved with parameter combinations that do not occur in reality.
- Since I hadn't heard about either of theses efficiency tests the course created my opinion. Either NSE or KGE = 1 means simulation aligns perfectly with observation, >0 is acceptable as it is still better than the mean of the input data. Below 0 means basically simulation is a waste of time since the mean of observation would be better to estimate.
- In some cases, the values can be close to 1 and the parameters are still not realistically calibrated, or the quality criteria do not provide information about the realism of the model representation of the catchment

12. You were able to calibrate your model with a satisfying score on your chosen calibration metric. You now want to predict the catchment behaviour for future conditions. Would you agree, that if a model works well for historic data, it will also work well for any future condition? * - 10 responses
    - 1 (strong disagreement) 10 %
    - 2 (disagreement) 30 %
    - 3 (undecided) 40 %
    - 4 (agreement) 20 %
    - 5 (strong agreement) 0 %
13. Through this course I gained knowledge and confidence in the general area of hydrological modelling. * - 10 responses
    - 1 (strong disagreement) 0 %
    - 2 (disagreement) 0 %
    - 3 (undecided) 10 %
    - 4 (agreement) 40 %

- 5 (strong agreement) 50 %

14. Through this course I gained knowledge and confidence in the area of model structure uncertainty. * - 10 responses

    - 1 (strong disagreement) 0 %
    - 2 (disagreement) 0 %
    - 3 (undecided) 10 %
    - 4 (agreement) 60 %
    - 5 (strong agreement) 30 %

15. Through this course I gained knowledge and confidence in the area of parameter uncertainty. * - 10 responses

    - 1 (strong disagreement) 0 %
    - 2 (disagreement) 0 %
    - 3 (undecided) 20 %
    - 4 (agreement) 60 %
    - 5 (strong agreement) 20 %

16. Through this course I gained knowledge and confidence in the area of data uncertainty. * - 10 responses

    - 1 (strong disagreement) 0 %
    - 2 (disagreement) 30 %
    - 3 (undecided) 20 %
    - 4 (agreement) 50 %
    - 5 (strong agreement) 0 %

17. Through this course my enthusiasm for the subject of hydrological modelling and it's uncertainties grew. * - 10 responses

    - 1 (strong disagreement) 0 %
    - 2 (disagreement) 0 %
    - 3 (undecided) 10 %
    - 4 (agreement) 70 %
    - 5 (strong agreement) 20 %

18. Through this course I gained knowledge about the differences between hydrologic realism and high KGE scores. * - 10 responses

- 1 (strong disagreement) 0 %
- 2 (disagreement) 0 %
- 3 (undecided) 20 %
- 4 (agreement) 30 %
- 5 (strong agreement) 50 %

19. Through this course I gained knowledge on how certain model structures may be more suitable for certain places than others. * - 10 responses
    - 1 (strong disagreement) 0 %
    - 2 (disagreement) 10 %
    - 3 (undecided) 0 %
    - 4 (agreement) 70 %
    - 5 (strong agreement) 20 %
20. Through this course I gained knowledge on how model structures can represent very different processes. * - 10 responses
    - 1 (strong disagreement) 0 %
    - 2 (disagreement) 0 %
    - 3 (undecided) 40 %
    - 4 (agreement) 50 %
    - 5 (strong agreement) 10 %
21. Through this course I am now more confident to apply methods to consider uncertainty in hydrological modelling. * - 10 responses
    - 1 (strong disagreement) 0 %
    - 2 (disagreement) 0 %
    - 3 (undecided) 20 %
    - 4 (agreement) 80 %
    - 5 (strong agreement) 0 %
22. Through this course I gained insights into the importance of communicating to clients the uncertainties of hydrological model simulations. * - 10 responses
    - 1 (strong disagreement) 0 %

- 2 (disagreement) 0 %
- 3 (undecided) 30 %
- 4 (agreement) 40 %
- 5 (strong agreement) 30 %

23. What I liked about the MARRMoT course module: [open question] - 10 responses

- It was really informative and gave an insight on how professionals work, especially with matlab, and modelling in general.
- The course gave new insights into hydrological modeling in a very clear way and a lot of new knowledge was conveyed. The very helpful and student-oriented teaching made it pleasant and all questions asked were answered comprehensively. The teaching material was very well prepared and presented clearly. Different media were used.
- Basically everything but in particular the hybrid offer and entering your own results in the Excel sheet (animated to participate)
- Very well prepared! Also nice interaction tools were used. The MARRMoT course was more helpful and interesting than all the other exercises of the module. The use of actual hydrological models makes way more sense to me than watching somebody typing in a MATLAB Code.
- the real life issues and practical work
- very practical. Interesting presentation by Wouter Knoben. Put a lot of input in it.
- it was a new tool for me to learn. It was quite new for me to use code for hydrological model run.
- Very flexible hydroligical modelling framework, with additional features as it is base on MatLab e.g.: implementation of printing, output assessment, multicore usage, data manipulation, ...
- I really like the practical work, not only talk about modeling/calibration -> we did modelling and evaluated results.
- practical modelling application, new and well communicated information, guest lecture

24. What could be improved in the MARRMoT course module: [open question] - 9 responses

- Offer it as an extra course in the middle of the semester.
- more time (more MARRMoT in general maybe)
- Maybe first a practical lesson before the lecture to get into the topic. Would have made it maybe easier for me to follow the presentation.
- Sometimes I didn't quite know what to do, since it wasn't clear what exactly we had to do. I would consider myself rather firm with matlab syntax and knew what the scripts were doing (programmatic at least). I could imagine that people with less knowledge of programming in general and especially with matlab have a hard time even trying to

understand what it means to change certain parameters. Even though the scripts are self explaining they are still the old and basic looking scripts. For non matlab-savy people it might be worth looking to use the newer live-scripts or even transforming the MARRMoT-Suite into an matlab-app.

- short intros to necessary termini such as benchmarks and scores before each exercise might be helpful to get into the topic
- the hybrid course curriculum was confusing for me. I preferred to have this course fully online. also it could be better to provide additional videos/lecture videos of this tool. Also, before running the model, I prefer to have a description/ theories about the task.
- More time (so expanding the course), maybe also a lil' term work to have more time to experiment with the program
- the aim could be referenced more throughout the course
- Sometimes it was not quite clear what to do (whats the task), so I had to ask some times (especially when you not follow the speaker for some seconds). There were a lot of tasks on the sheet and it would be nice if you can just say which task (for example "now work on number 4 on sheet 3") is the current task. Another thing was the folder structure of the MARRMoT. I was always searching the files/sheets in the folders and were confused sometimes (it took me some time).

25. Additional comments: [open question] - 4 responses

- I wish all my "exercises" were like this! :)
- All in all a very very good teaching "block", very useful and show insights of practical hydrological modelling. Those lessons bring typical modelling issues closer to the people/students and they were more aware of this during "normal" lessons. Should be more of this! Unfortunately these lessons were to close to the exams/presentations so you couldn't really focus and spent a lot of time on this. This should be expanded.
- it was great to learn a new tool
- Due to my limited Matlab knowledge, it was sometimes difficult to keep up with the pace and to properly think everything through.

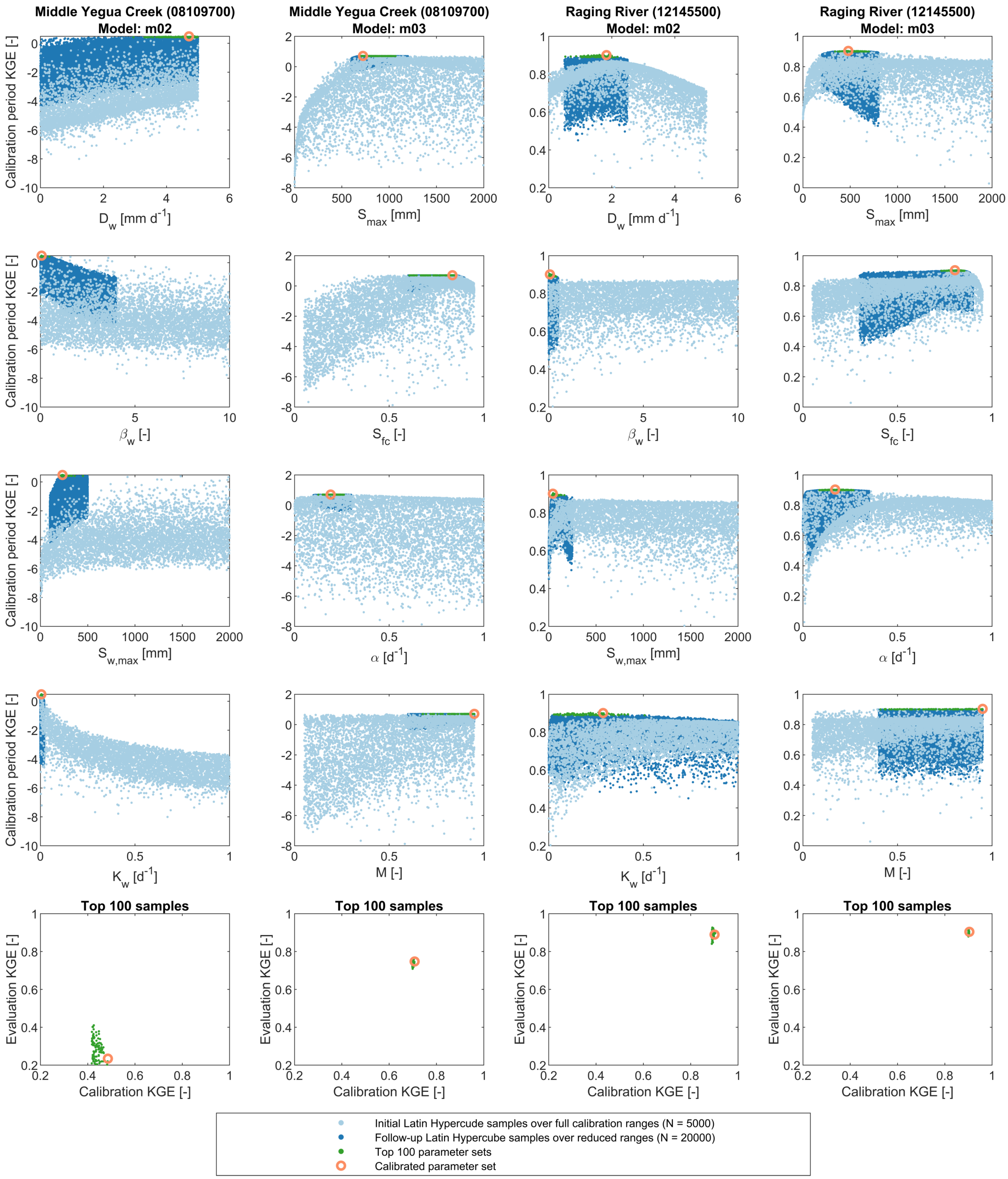

Figure S5. Top four rows: Latin Hypercube parameter samples for both models and catchments, and asscoiated KGE scores for the calibration period. Initial sampling (light blue) covered the full parameter ranges as used during model calibration to identify regions of interest for more thorough sampling (dark blue). Bottom row: comparison of calibration and evaluation KGE scores of the top 100 samples (defined as highest KGE scores for the calibration period) and calibrated parameter sets.

References

Addor, N., Newman, A. J., Mizukami, N., and Clark, M. P.: The CAMELS data set: catchment attributes and meteorology for large-sample studies, Hydrology and Earth System Sciences, 21, 5293–5313, https://doi.org/10.5194/hess-2017-169, 2017.

Gupta, H. V., Wagener, T., and Liu, Y.: Reconciling theory with observations : elements of a diagnostic approach to model evaluation, Hydrological Processes, 3813, 3802–3813, https://doi.org/https://doi.org/10.1002/hyp.6989, 2008.

Knoben, W. J., Freer, J. E., Fowler, K. J., Peel, M. C., and Woods, R. A.: 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, 2463–2480, https://doi.org/10.5194/gmd-12-2463-2019, 2019.