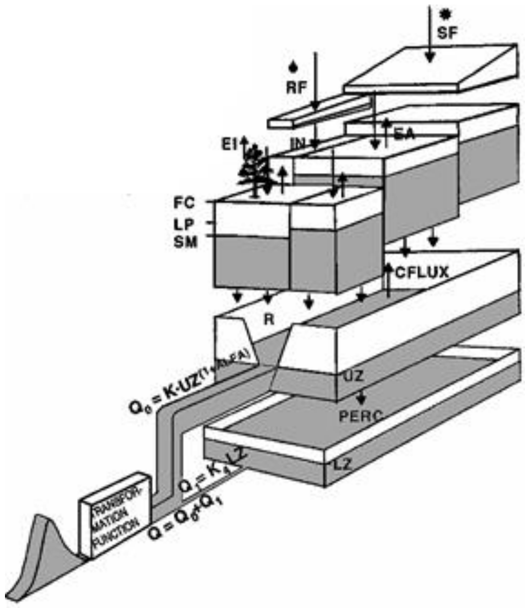
BAYESIAN CALIBRATION FOR MODEL PARAMETER UNCERTAINTY (TUWmodel)



In the previous labs we focused on propagation of input uncertainty through a model. Input uncertainty, however, is not the only source of uncertainty. Also, uncertainty about the values of model parameters and structural error contribute to output uncertainty. In this lab we demonstrate the use of Markov Chain Monte Carlo (MCMC) for handling parameter, model and observational error simultaneously in uncertainty analysis. Owing to time constraints we use a relatively simple, lumped (non-spatial) model to demonstrate the procedure. However, the approach is equally applicable to spatially distributed models. We will use a rainfall-runoff model (TUWmodel) for which an R package and example data are available (https://cran.r-project.org/web/packages/TUWmodel/index.html).

The TUWmodel, developed at the Vienna University of Technology, is a lumped conceptual rainfall-runoff model, following the structure of the HBV model (see e.g.: https://en.wikipedia.org/wiki/HBV_hydrology_model) as illustrated in the figure above. It runs on a daily or shorter time step and consists of a snow routine, a soil moisture routine and a runoff routing routine. Details are provided in Parajka et al. (2007) [1], Seibert and Vis (2012) [2] and Nijzink et al. (2018)[3]. In this lab we will focus on the runoff routing on hill slopes, which is governed by a set of parameters (briefly explained below) which we calibrate and for which we assess uncertainty.

The model represents runoff routing on hill slopes by an upper and a lower soil reservoir. Excess precipitation enters the upper reservoir and leaves it through three paths: (1) outflow from the reservoir based on a fast storage coefficient k_1 (expressing the proportionality of outflow to the water storage (S_{xxx}) in the reservoir); (2) percolation to the lower zone with a constant percolation rate *cperc*; and (3), if a threshold of the storage state *lsuz* is exceeded, through an additional outlet based on a very fast storage coefficient k_0 (see Eq. 1). The lower zone releases water based on a slow storage coefficient k_2 (Eq. 2). Note that the way the storage coefficients are expressed (and hence the units) may vary between different implementations of the HBV model.

$$Q_{upp} = \begin{cases} S_{upp} \,/\, k_1 + S_{upp} \,/\, k_0 & \text{if } S_{upp} > lsuz \\ S_{upp}/k_1 & \text{otherwise} \end{cases} \qquad [1]$$

$$Q_{low} = S_{low}/k_2 \qquad [2]$$

[1] Parajka, J., Merz, R., Bloeschl, G., 2007. Uncertainty and multiple objective calibration in regional water balance modelling: case study in 320 Austrian catchments, *Hydrological Processes*, 21, 435-446.

[2] Seibert, J., Vis, M.J.P., 2012. Teaching hydrological modelling with a user-friendly catchment-runoff-model software package, *Hydrology and Earth System Sciences*, 16, 3315-3325. https://doi.org/10.5194/hess-16-3315-2012.

[3] Nijzink, R.C., Almeida, S., Pechlivanidis, I.G., Capell, R., Gustafssons, D., Arheimer, B., et al. (2018). Constraining conceptual hydrological models with multiple information sources. *Water Resources Research*, 54, 8332–8362. https://doi.org/10.1029/2017WR021895.

The outflow from both reservoirs is modified by a triangular weighting function with base B_Q (i.e. the time length over which outflow is dispersed) using two parameters: a maximum base at low flow (bmax) and a scaling parameter (croute) (Parajka et al., 2007) (Eq. 3). This weighting function smooths outflow over time to represent delays caused by flow routing over streams within the catchment before reaching the outlet where discharge is measured.

$$Q_{tmp} = Q_{upp} + Q_{low}$$
$$B_Q = \max(bmax - croute \cdot Q_{tmp},\ 1) \qquad [3]$$

In this lab you will:
1. run the TUWmodel on a provided data set using a given set of plug-in values for its parameters;
2. use prior distributions to express *a priori* belief about possible values for selected model parameters (k_1, *lsuz*, *cperc* and *croute*), two parameters of a first order autoregressive (AR) structural error model (β_1 and σ_δ^2) and
3. update these prior distributions by MCMC using a time series of measured outflows; thus obtained posterior distributions and residual variance are measures of parameter uncertainty and structural model uncertainty, respectively;
4. do an uncertainty analysis by simultaneously considering parameter, model- and observational error in the discharge;
5. assess whether parameter distributions obtained from Bayesian calibration improved model fit;

Input data

Similar to Wadoux et al (2020)[4], which is provided in the literature folder, we study data from the Thur basin (~1700 km^2) located in the North-East of Switzerland. However, unlike that work we use a daily time interval over the period 01/01/2004 – 31/12/2011. The following data are provided (`Thurdata.zip`; unzip the data in a sub-folder named `Wednesday`):

`agERA5prec.txt`	Spatially aggregated daily precipitation over the study area retrieved from the agERA 5 dataset (DOI: 10.24381/cds.6c68c9bb), which has been found useful for hydrological modelling[5].
`evap.txt`	Spatially aggregated daily evaporation over the study area computed from wheather station data provided by Swiss Federal Office for Meteorology and Climatology (MeteoSwiss)
`temp.txt`	Spatially aggregated daily mean temperature over the study area computed from wheather station data provided by MeteoSwiss
`runoff.txt`	Daily cumulative discharge data for the period 2004-2011 from the Swiss Federal Office for the Environment (FOEN). The discharge measuring station is located near Andelfingen.
`dates.txt`	date strings for the period 01/01/2004 – 31/12/2011

[4] Wadoux, A.M.J.C., Heuvelink, G.B.M., Uijlenhoet, R., de Bruin, S., 2020. Optimization of rain gauge sampling density for river discharge prediction using Bayesian calibration. *PeerJ* 8:e9558 http://doi.org/10.7717/peerj.9558

[5] Tarek, M., Brissette, F.P., Arsenault, R., 2020. Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. *Hydrol. Earth Syst. Sci.*, 24, 2527–2544. https://doi.org/10.5194/hess-24-2527-2020

The river Thur is a tributary of the Rhine river and it is the largest non-regulated river in Switzerland. The elevation within the basin ranges from 356 to 2437 m above sea level (with an average height of 765 m. The climate is submontane, relatively cool and dominated by high precipitation, most of which falls during the summer months (June-August) in the form of rain, whereas the winter months (December-February) are characterized by snow.

Script

The R script for this practical is provided in MS Teams (`BayesTUWmodel2024.R`).

Default run

Lines 1-25 of the script `BayesTUWmodel2024.R` verify whether the package TUWmodel has been installed, load and attach the package, and load the data for the Thur basin. The TUWmodel is run with default plug-in parameter values and results are plotted in lines 27-39. Lines 42-55 prepare a calibration and validation dataset and lines 57-62 compute statistics of the relative deviations (%) between model predictions and discharge measurements over the validation period.

Q1 *Run lines 1-62 of the script. You may have to change the working directory with* `setwd()`. *Are the discharges predicted by the model biased with respect to the measured values? What could be causing this?*

Bayesian calibration

We will now use Bayesian calibration for computing a sampling distribution of k_1, *lsuz*, *cperc*, *croute* and parameters of multiplicative discharge measurement error and model structural error, given the time series of daily discharges. That is, the joint probability distribution of the parameters is to be derived from the data by comparing model outputs with independent observations, while taking uncertainty owing to observation errors (in the discharge) into account.

Using Bayes theorem, we aim to update prior probability distributions (shorthand: prior) of the model parameters and residual error using evidence from our data to express uncertainty about the parameter values. Unfortunately, we can only compute quantities that are proportional to the probability densities of the parameters given the data. To overcome this problem, we use the Metropolis–Hastings algorithm for implementing the Markov Chain Monte Carlo (MCMC) approach. The procedure has been more extensively explained in the lecture. In summary, based on Bayes' theorem the likelihood of observed data given an arbitrary starting sample of the model parameters is multiplied with the prior. This produces a quantity, say *denominator*, that is proportional to the desired probability density of the parameters given the data. The procedure is repeated for a new proposed sample (often referred to as proposal) which is obtained by applying a jumping distribution. Let the second quantity be denoted *numerator*. If the ratio *numerator* / *denominator* > 1, the proposed sample is accepted; otherwise it is accepted with a probability *numerator* / *denominator*. If rejected, the previous sample is kept. By repeating those steps many times, MCMC produces a chain of samples that form a sample distribution. After many iterations and *thinning*, this sample distribution converges to the true distribution of the parameters given the data and the prior.

Since the time series used in this lab is fairly long —potentially causing numerical issues— all calculations in the script are done on the log-scale. This implies that multiplication is replaced by summation while division is replaced by subtraction. The underlying principle of the Metropolis–Hastings algorithm remains the same, however.

It is unrealistic to assume that the multiplicative model structural errors are temporally independent. Therefore, this error component is represented by a first order autoregressive (AR) model having three parameters, two of which (β_1 and σ_δ^2) are free and need to be calibrated. Including the variance of the measurement error (σ_η^2) there are thus seven parameters to be calibrated (four model parameters, two parameters of the autoregressive model structural component and one parameter for the discharge measurement error).

Lines 122-267 of the script BayesTUWmodel2024.R implement the MCMC procedure. Its main component, the function metropolis, is called in line 284-285. In turn, metropolis calls the function loglik that returns the log-likelihood of the data given the sampled model parameters. The priors are defined and plotted in lines 64-119. The initial proposal is given in line 275, while the parameters of the jumping distribution are defined in line 278. It has been established that the latter parameters should be set in such a way that the acceptance rate (i.e. accepted proposals / evaluated proposals) roughly equals 0.25 [6].

Q2 *Study and run the code where the priors are set (lines 64-119). Which priors seem to be most informative?*

Q3 *First run lines 122-278 to load the required functions and set the initial proposal and step parameters. Next —if needed— repeatedly run lines 278-285 and adjust the step parameters in line 278 until a suitable acceptance rate is obtained. Examine the trace plots of the sampled model parameters (lines 287-297). Do you observe any striking behaviour? When satisfied, replace the 2000 in line 281 by N and generate a full MCMC sample. This will take a few minutes.*

Q4 *Run lines 300-341. In lines 300-302, the MCMC sample is subsampled. Why? Which priors were most influenced by Bayesian calibration? Do you see any unexpected results? Are there any priors that seem to have been overly restrictive?*

Q5 *In lines 343-345 scatter plots between the sampled parameters are drawn and correlations between the parameters are computed. What do you observe? Does that make sense?*

Uncertainty analysis

Next, we use the thinned sample of parameter values along with the parameters of the multiplicative error components concerning model structure and discharge measurement error for assessing uncertainty about discharge at the Andelfingen measurement station. For this purpose, we zoom in on the period 2009-2011. Different sets of computations and plotting are done to assess the uncertainty contributions of (1) all considered uncertainty components, (2) only model parameter uncertainty and (3) both model parameter uncertainty and discharge measurement error.

Study and run lines 348-472.

Q6 *Did Bayesian calibration improve the model predictions? How do you know?*

Q7 *Which uncertainty component contributed most to uncertainty about discharge at the Andelfingen measurement station? Which the least? Do the 90% prediction intervals cover most of the observed discharges?*

Q8 *Did we forget some contribution to total uncertainty or is it somehow covered? How can we improve upon the current approach?*

[6] For example: Roberts, G.O., Gelman, A. and Gilks, W.R., 1997. Weak convergence and optimal scaling of random walk Metropolis algorithms. *The Annals of Applied Probability* 7(1): 110-120. https://projecteuclid.org/euclid.aoap/1034625254