European Geosciences Union
General Assembly 2021
Gather Online | 19–30 April 2021

Session HS2.3.7: Multi-dataset, multi-variable, and multi-objective techniques to improve prediction of hydrological and water quality models and their Bayesian applications

Large-scale calibration of conceptual rainfall-runoff models for two-stage probabilistic hydrological post-processing

Georgia Papacharalampous[1,*,#], Hristos Tyralis[1,2], Demetris Koutsoyiannis[1], Alberto Montanari[3]

[1] Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, Iroon Polytechniou 5, 15780 Zographou, Greece

[2] Hellenic Air Force General Staff, Hellenic Air Force, Mesogion Avenue 227–231, 15561 Cholargos, Greece

[3] Department of Civil, Environmental and Materials Engineering (DICAM), University of Bologna, via del Risorgimento 2, 40136 Bologna, Italy

[*] Correspondence: papacharalampous.georgia@gmail.com, gpapacharalampous@hydro.ntua.gr

[#] Currently also at the Department of Engineering, Roma Tre University, Via V. Volterra 62, 00146 Rome, Italy

H.F.R.I. Hellenic Foundation for Research & Innovation

The research work of Georgia Papacharalampous was supported by the Hellenic Foundation for Research and Innovation (HFRI) under the HFRI PhD Fellowship grant (Fellowship Number: 1388).

1. One-slide summary

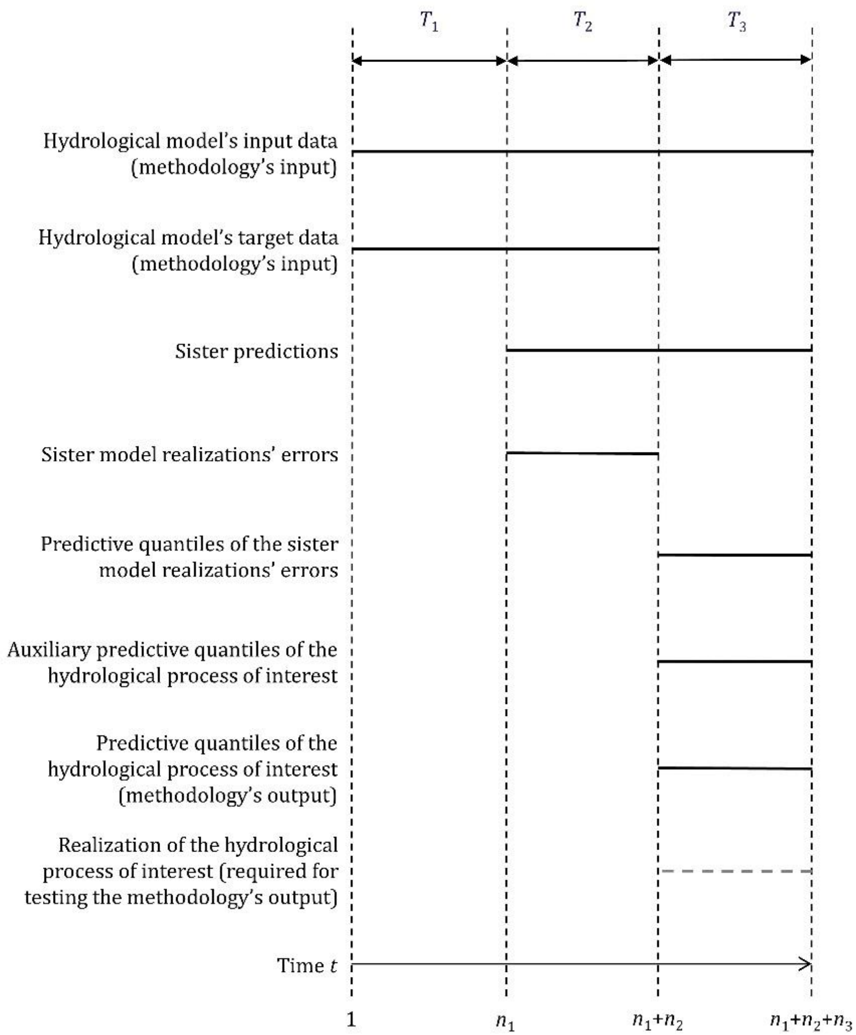
- **Background:** Probabilistic hydrological modelling methodologies often rely on statistical post-processing schemes for benefitting from the "hydrological experience" encompassed in conceptual and physically-based rainfall-runoff models (see e.g., Montanari and Brath 2004; Montanari and Grossi 2008; Tyralis et al. 2019). These schemes might require issuing an ensemble of rainfall-runoff model simulations by using different input data and/or different parameters (see the blueprint by Montanari and Koutsoyiannis 2012 and its summary in **Slide 2**). For obtaining a large number of rainfall-runoff model parameters in this regard, Bayesian large-scale calibration schemes have been adopted in the literature (e.g., in Montanari and Koutsoyiannis 2012; Sikorska et al. 2015).
- **Objective:** Here, we investigate a possible replacement of Bayesian with non-Bayesian schemes for large-scale rainfall-runoff model calibration within probabilistic hydrological modelling methodologies of the above-defined family.
- **Motivation:** Bayesian rainfall-runoff model calibration schemes are accompanied by computational limitations, which are well-recognized in the literature and may prohibit applications "at scale".
- **Methodology:** Starting from a Bayesian rainfall-runoff model calibration scheme, we define a computationally convenient calibration scheme (see **Slide 4**). We then apply both these schemes as parts of six diverse variants (hereafter, referred to as "ensemble schemes") of the probabilistic hydrological modelling methodology by Papacharalampous et al. (2020a) and Papacharalampous et al. (2020b). This latter methodology (see its summary in **Slide 3**) retains some robust features from its mother blueprint-method (Montanari and Koutsoyiannis 2012) and simultaneously allows for benefitting from machine learning quantile regression algorithms (see e.g., the references in Papacharalampous et al. 2019, Section 2.3). In this specific context, the two calibration schemes are compared using proper scores and large-scale benchmarking (see **Slides 5** and **6**; see also **Slide 11**).
- **Main finding:** Overall, our results (see **Slides 7**–**9**) suggest that the two rainfall-runoff model calibration approaches can lead to mostly comparable probabilistic predictions (see also **Slide 10**).
- **Further reading:** For further information on the experiments summarized herein, the reader is referred to Papacharalampous et al. (2020b, Appendix E).

2. Probabilistic hydrological modelling blueprint

- Our methods and experiments build on top of the blueprint by Montanari and Koutsoyiannis (2012), a theoretically consistent and flexible scheme for predictive uncertainty quantification in hydrological modelling.
- In its basic implementation, this scheme uses (a) one rainfall-runoff model to issue a large number of point predictions-simulations (within an ensemble simulation framework), and (b) one error model to statistically post-process each of these point predictions.
- Its output is an ensemble of statistically post-processed point predictions-simulations, together constituting the probabilistic prediction.
- Its hydrological and error models are estimated/trained in two subsequent stages.
- This two-stage character naturally allows the accommodation of regression-based predictive modelling solutions to the scheme (see e.g., the variants of this blueprint by Sikorska et al 2015, Quilty et al. 2019, Papacharalampous et al. 2020a, Papacharalampous et al. 2020b, Quilty and Adamowski 2020).
- Quantile regression algorithms differ from the typical regression algorithms, and can be incorporated into the blueprint by Montanari and Koutsoyiannis (2012), as detailed in Papacharalampous et al. (2020a) and Papacharalampous et al. (2020b); see also **Slide 3**.
- Basic two-stage probabilistic hydrological post-processing (see e.g., Montanari and Brath 2004; Montanari and Grossi 2008; López López et al. 2014; Dogulu et al. 2015; Papacharalampous et al. 2019; Koutsoyiannis and Montanari 2020) can also be viewed as a subcase of this blueprint.

3. Probabilistic hydrological modelling methodology

For details, see Papacharalampous et al. (2020a) and Papacharalampous et al. (2020b).



Step 1: Obtain m sister model realizations by calibrating the hydrological model with input and target data for the period T_1

Step 2: Obtain m sister predictions by running each of the sister model realizations with input data for the period $\{T_2, T_3\}$

Step 3: Compute the sister model realizations' errors (or a single sister model realization's errors) in the period T_2 by subtracting the target data from its corresponding sister predictions, available for the same period

Step 4: Train the error model in the period T_2 by regressing the sister model realizations' error at time t on selected predictor variables (e.g., the corresponding sister prediction at time t)

Step 5: Obtain the predictive quantiles of the sister model realizations' errors in the period T_3 by using the information about these errors obtained at step 4

Step 6: Transform the predictive quantiles of the sister model realizations' errors to auxiliary predictive quantiles of the hydrological process of interest by subtracting them from their corresponding sister predictions

Step 7: Obtain the predictive quantiles of the hydrological process of interest by grouping the auxiliary predictive quantiles of this process based on their probabilities and by averaging them over each group

In this presentation, we compare two modelling schemes for taking this step. These two schemes are hereafter referred to as "calibration schemes" and are described in **Slide 4**.

4. Rainfall-runoff model and its calibration

Rainfall-runoff model

- We use the GR2M conceptual rainfall-runoff model by Mouelhi et al. (2006).

Calibration scheme 1: A Bayesian calibration scheme

- We simulate the posterior distribution of the parameters of the rainfall-runoff model conditional on the observations of the period T_1 (see **Slide 3**) within a Bayesian MCMC framework.
- We use flat priors for both the parameters θ_1 and θ_2.
- The likelihood error function is defined by the following equation, where y_t is the monthly streamflow discharge observations at time t, $u_t(\theta_1, \theta_2)$ is the prediction at time t and $|T_1|$ is the number of target data points included in the period T_1:

$$L(\theta_1, \theta_2) \propto \left(\sum_t \left(y_t - u_t(\theta_1, \theta_2)\right)^2\right)^{-|T_1|/2}$$

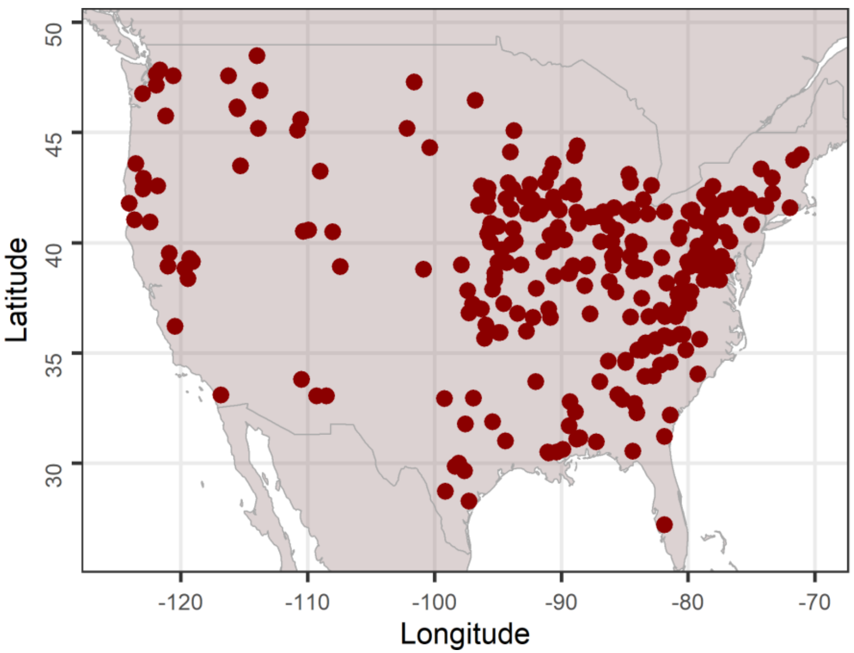
- We run three parallel Markov chains with different initial values, each comprising 2 000 iterations.
- The iterative simulation is performed by using the DRAM algorithm by Haario et al. (2006).
- We assess the approximate convergence of these chains by implementing the algorithm by Brooks and Gelman (1998).
- The simulation process is repeated until approximate convergence is achieved.
- Once the simulation is over, we retain the last 200 values of each chain, i.e., 600 values in total for each catchment (see e.g., the related example in **Slide 6**).

Calibration scheme 2: A computationally convenient calibration scheme

- Instead of retaining the last 200 parameter values of each simulated chain (see above), we retain the first 200 parameter values that have not converged to the posterior distribution of the parameters.

5. Experimental data and large-scale benchmarking

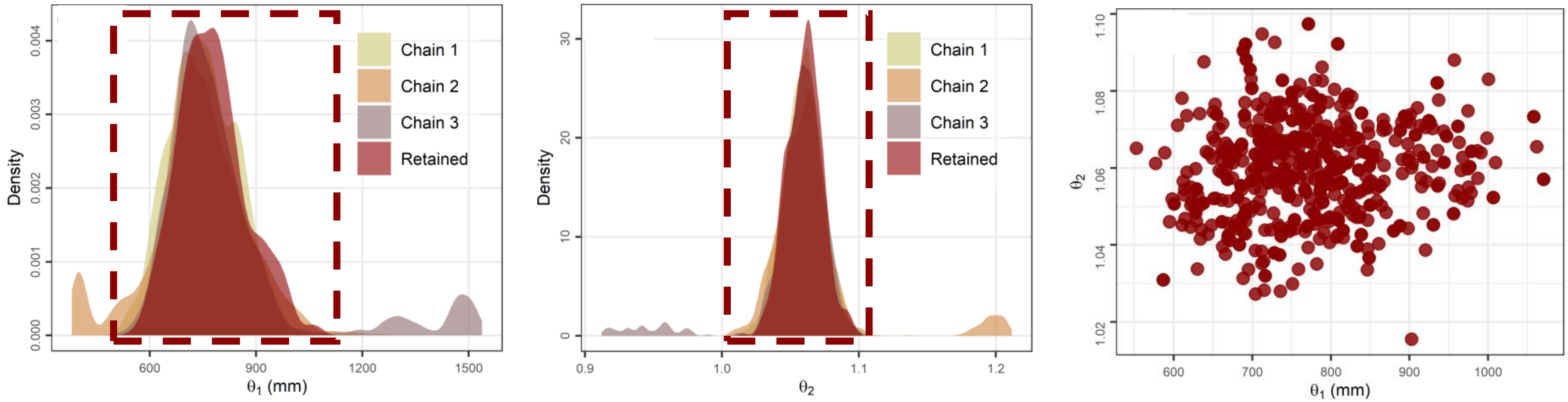
- For our experiments, we use 50-year long monthly precipitation, potential evapotranspiration and streamflow time series originating from 270 MOPEX catchments (Schaake et al. 2006).
- The geographical locations of the streamflow stations are presented on the bottom of this slide.
- Following the notations provided in **Slide 3**, we define the periods T_1 = {13, …, 156}, T_2 = {157, …, 300} and T_3 = {301, …, 600} (respectively corresponding to years 1951−1962, 1963−1974 and 1975−1999). We also define period T_0 = {1, …, 12} (corresponding to year 1950 in our dataset). This period is used for warming up the rainfall-runoff model (see **Slide 4**).
- Additionally, we define six ensemble schemes starting from the selected probabilistic hydrological modelling methodology (see **Slide 3**).
- These ensemble schemes differ with each other in terms of some of their components, as detailed in Papacharalampous et al. (2020b).



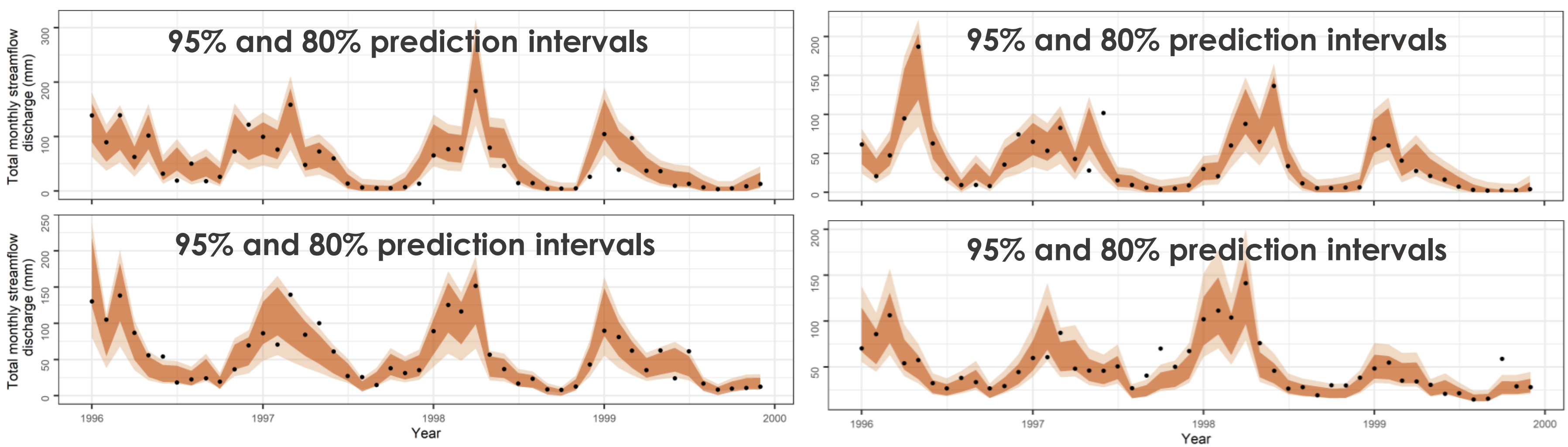
- We apply the ensemble schemes twice, each time using a different calibration scheme (see **Slide 4**).
- Specifically, we predict the quantiles of monthly streamflow at levels 0.005, 0.0125, 0.025, 0.05, 0.10, 0.90, 0.95, 0.975, 0.9875 and 0.995 for the period T_3. These quantiles are then used to form the 80%, 90%, 95%, 97.5% and 99% prediction intervals for the same period.
- Lastly, we assess the quality of our predictions by computing (a) their average interval scores (see e.g., Gneiting and Raftery 2007), and (b) the relative improvements provided by the Bayesian calibration scheme over the computationally convenient scheme (see **Slides 7** and **8**). We further summarize these relative improvements in terms of their means and medians (see **Slide 9**).

6. Application examples within the study's framework

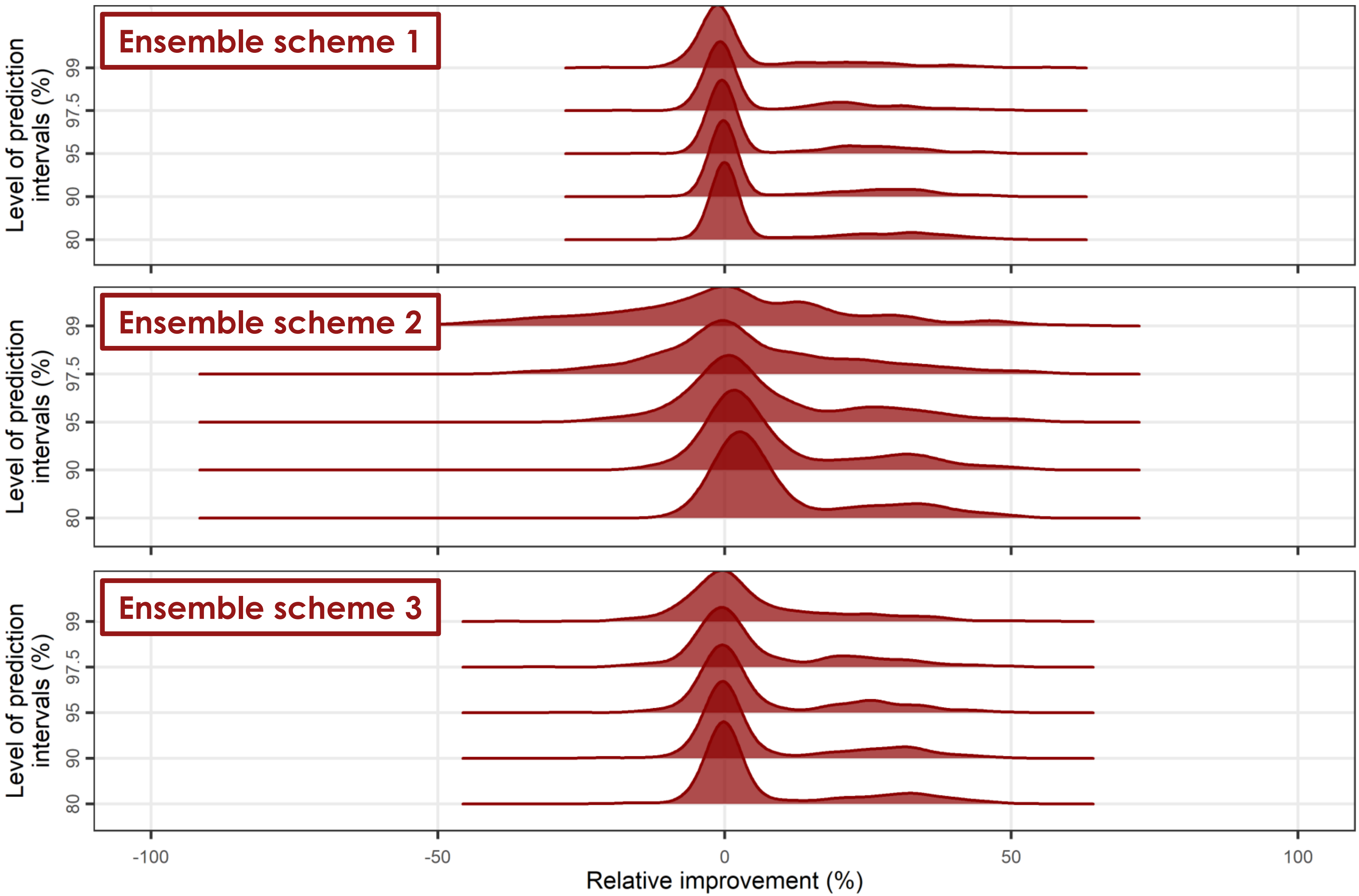
Examples of simulated chains using the Bayesian calibration scheme



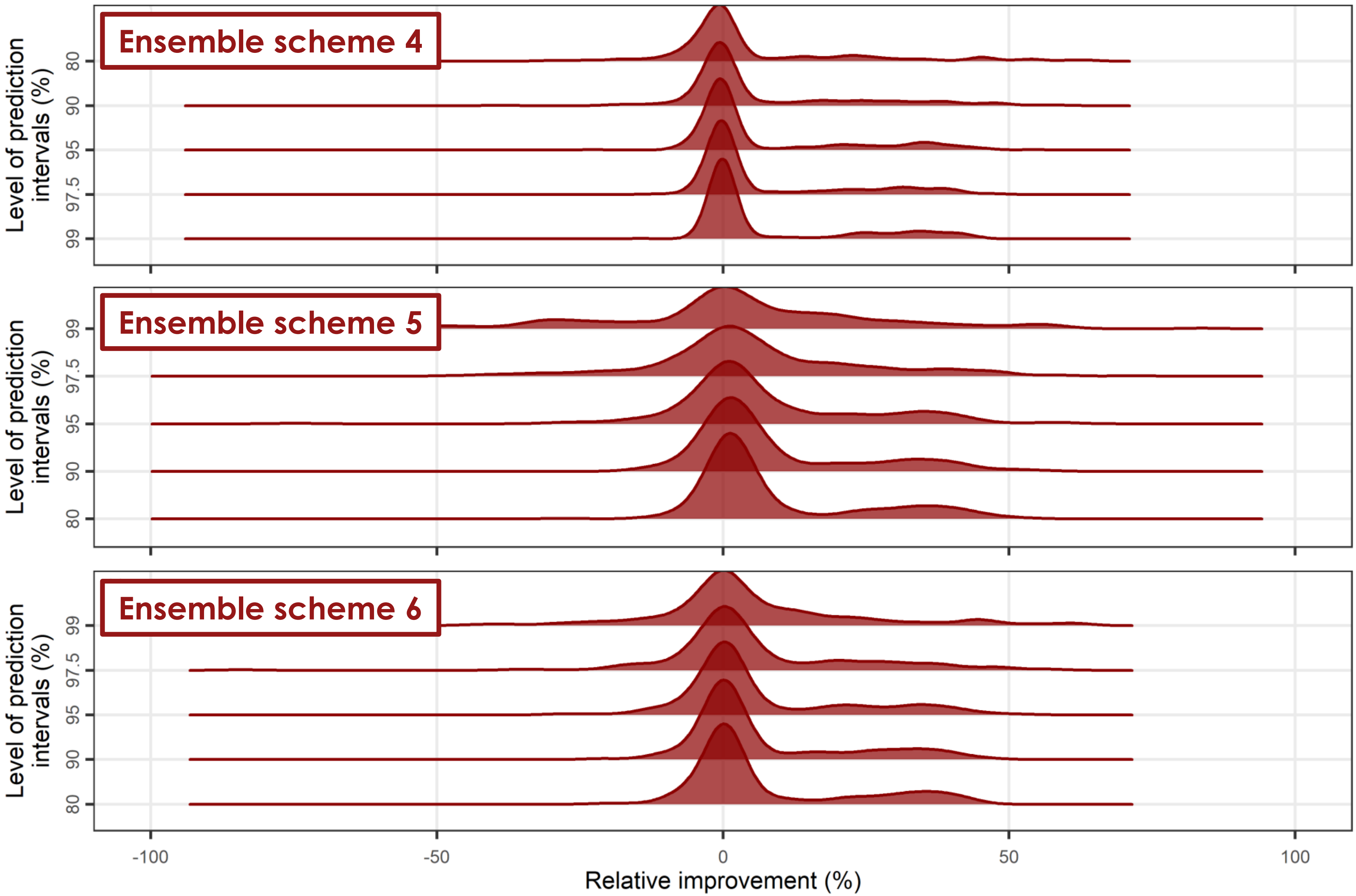
Examples of probabilistic predictions



7. Relative improvements in terms of average interval score

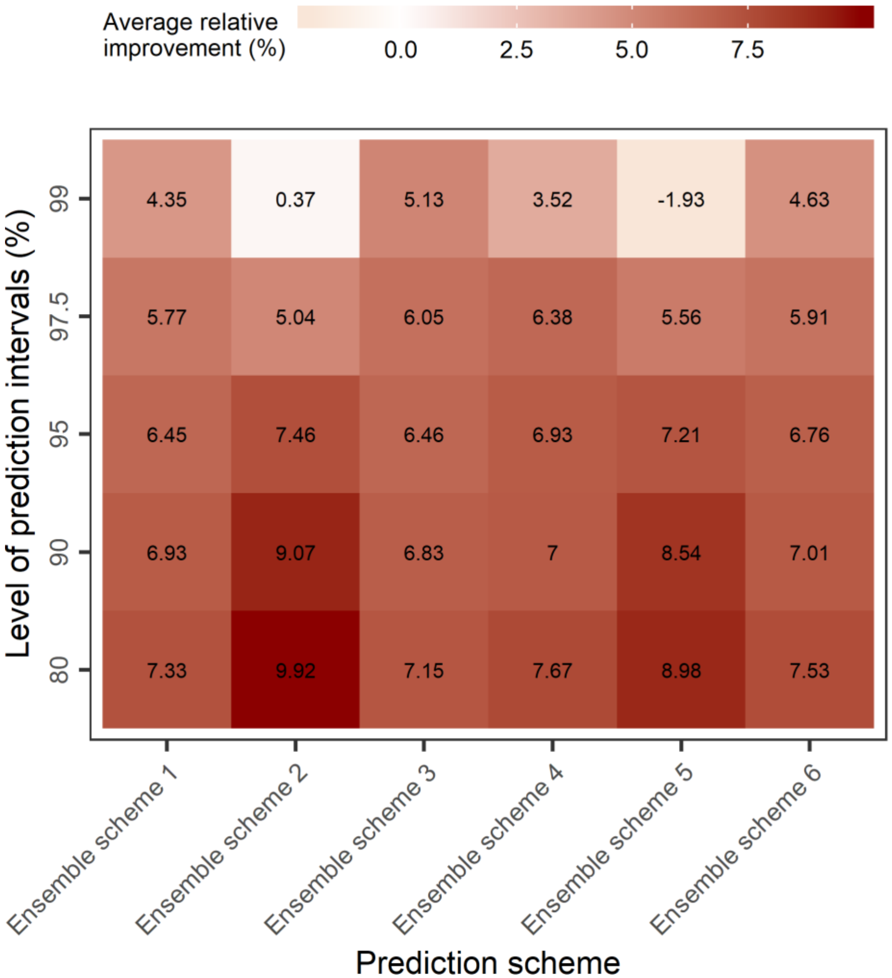


8. Relative improvements in terms of average interval score

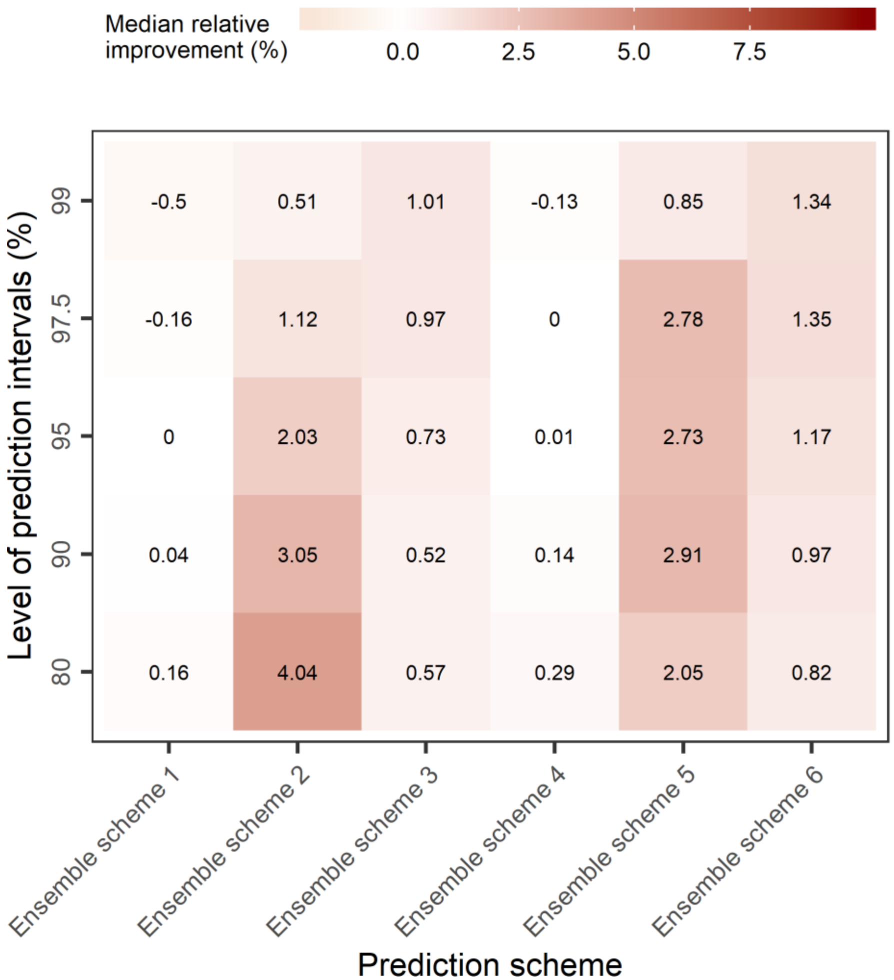


9. Relative improvements in terms of average interval score

Mean relative improvements



Median relative improvements



10. Key findings and conclusions

- Overall, the Bayesian calibration scheme and the computationally convenient calibration scheme can lead to mostly comparable probabilistic predictions in the long run.
- The relative improvements provided by the Bayesian calibration scheme over the computational convenient calibration scheme (see **Slides 7** and **8**) have been found to be either positive or negative for all the ensemble schemes.
- They have also been found to considerably depend both on the examined catchment and the examined prediction interval.
- On average, they favour the Bayesian calibration scheme to a small extent, mostly due to outliers, while their median values are closer to zero (see **Slide 9**).
- We feel that outliers could be fewer if longer time series had been examined.
- We hope that our preliminary experiments will trigger further investigations on the subject.

11. Statistical software information

The analyses and visualizations have been performed in `R` Programming Language (R Core Team 2019). We have used the following contributed `R` packages: `airGR` (Coron et al. 2017, 2019), `bestNormalize` (Peterson 2017, 2019), `coda` (Plummer et al. 2006, 2019), `data.table` (Dowle and Srinivasan 2019), `devtools` (Wickham et al. 2019c), `dplyr` (Wickham et al. 2019b), `FME` (Soetaert and Petzoldt 2010, 2016), `gdata` (Warnes et al. 2017), `ggplot2` (Wickham 2016a; Wickham et al. 2019a), `ggridges` (Wilke 2018), `hddtools` (Vitolo 2017, 2018), knitr (Xie 2014, 2015, 2019), maps (Brownrigg et al. 2018), `matrixStats` (Bengtsson 2018), `plyr` (Wickham 2011, 2016b), `quantreg` (Koenker 2019), `readr` (Wickham et al. 2018), `reshape` (Wickham 2007, 2018), `rmarkdown` (Allaire et al. 2019), `tidyr` (Wickham and Henry 2019) and `zoo` (Zeileis and Grothendieck 2005; Zeileis et al. 2019). We have also followed procedures described in the contributed vignettes of the `airGR` `R` package.

References

Allaire JJ, Xie Y, McPherson J, Luraschi J, Ushey K, Atkins A, Wickham H, Cheng J, Chang W, Iannone R (2019) rmarkdown: Dynamic Documents for R. R package version 1.14. https://CRAN.R-project.org/package=rmarkdown

Bengtsson H (2018) matrixStats: Functions that Apply to Rows and Columns of Matrices (and to Vectors). R package version 0.54.0. https://CRAN.R-project.org/package=matrixStats

Brooks SP, Gelman A (1998) General methods for monitoring convergence of iterative simulations. Journal of Computational and Graphical Statistics 7(4):434–455

Brownrigg R, Minka TP, Deckmyn A (2018) maps: Draw Geographical Maps. R package version 3.3.0. https://CRAN.Rproject.org/package=maps

Coron L, Thirel G, Delaigue O, Perrin C, Andréassian V (2017) The suite of lumped GR hydrological models in an R package. Environmental Modelling and Software 94:166–171. doi:10.1016/j.envsoft.2017.05.002

Coron L, Delaigue O, Thirel G, Perrin C, Michel C (2019) airGR: Suite of GR Hydrological Models for Precipitation-Runoff Modelling. R package version 1.3.2.23. https://CRAN.R-project.org/package=airGR

Dogulu N, López López P, Solomatine DP, Weerts AH, Shrestha DL (2015) Estimation of predictive hydrologic uncertainty using the quantile regression and UNEEC methods and their comparison on contrasting catchments. Hydrology and Earth System Sciences 19:3181–3201. doi:10.5194/hess-19-3181-2015

Dowle M, Srinivasan A (2019) data.table: Extension of data.frame. R package version 1.12.2. https://CRAN.Rproject.org/package=data.table

Gneiting T, Raftery AE (2007) Strictly proper scoring rules, prediction, and estimation. Journal of the American Statistical Association 102(477):359–378. doi:10.1198/016214506000001437

Haario H, Laine M, Mira A, Saksman E (2006) DRAM: Efficient adaptive MCMC. Statistics and Computing 16(4):339–354. doi:10.1007/s11222-006-9438-0

Koenker RW (2019) quantreg: Quantile Regression. R package version 5.51. https://CRAN.R-project.org/package=quantreg

Koutsoyiannis D, Montanari A (2020) A brisk local uncertainty estimator for hydrologic simulations and predictions (Blue Cat), European Geosciences Union General Assembly 2020, Geophysical Research Abstracts, Vol. 22, Vienna. doi:10.5194/egusphereegu2020-10125

López López P, Verkade JS, Weerts AH, Solomatine DP (2014) Alternative configurations of quantile regression for estimating predictive uncertainty in water level forecasts for the upper Severn River: A comparison. Hydrology and Earth System Sciences 18:3411–3428. doi:10.5194/hess-18-3411-2014

Montanari A, Brath A (2004) A stochastic approach for assessing the uncertainty of rainfall-runoff simulations. Water Resources Research 40(1):W01106. doi:10.1029/2003WR002540

Montanari A, Grossi G (2008) Estimating the uncertainty of hydrological forecasts: A statistical approach. Water Resources Research 44(12):W00B08. doi:10.1029/2008WR006897

References

Montanari A, Koutsoyiannis D (2012) A blueprint for process-based modeling of uncertain hydrological systems. Water Resources Research 48(9):W09555. doi:10.1029/2011WR011412

Mouelhi S, Michel C, Perrin C, Andréassian V (2006) Linking stream flow to rainfall at the annual time step: The Manabe bucket model revisited. Journal of Hydrology 328(1–2):283–296. doi:10.1016/j.jhydrol.2005.12.022

Papacharalampous GA, Tyralis H, Langousis A, Jayawardena AW, Sivakumar B, Mamassis N, Montanari A, Koutsoyiannis D (2019) Probabilistic hydrological post-processing at scale: Why and how to apply machine learning quantile regression algorithms. Water 11(10):2126. doi:10.3390/w11102126

Papacharalampous GA, Koutsoyiannis D, Montanari A (2020a) Quantification of predictive uncertainty in hydrological modelling by harnessing the wisdom of the crowd: Methodology development and investigation using toy models. Advances in Water Resources 136:103471. doi:10.1016/j.advwatres.2019.103471

Papacharalampous GA, Tyralis H, Koutsoyiannis D, Montanari A (2020b) Quantification of predictive uncertainty in hydrological modelling by harnessing the wisdom of the crowd: A large-sample experiment at monthly timescale. Advances in Water Resources 136:103470. doi:10.1016/j.advwatres.2019.103470

Peterson RA (2017) Estimating normalization transformations with bestNormalize. https://github.com/petersonR/bestNormalize

Peterson RA (2019) bestNormalize: Normalizing Transformation Functions. R package version 1.4.0. https://CRAN.Rproject.org/package=bestNormalize

Plummer M, Best N, Cowles K, Vines K (2006) CODA: convergence diagnosis and output analysis for MCMC. R news 6(1):7–11

Plummer M, Best N, Cowles K, Vines K, Sarkar D, Bates D, Almond R, Magnusson A (2019) coda: Output Analysis and Diagnostics for MCMC. R package version 0.19-3. https://CRAN.R-project.org/package=coda

Quilty J, Adamowski J (2020) A stochastic wavelet-based data-driven framework for forecasting uncertain multiscale hydrological and water resources processes. Environmental Modelling and Software 104718. doi:10.1016/j.envsoft.2020.104718

Quilty J, Adamowski J, Boucher MA (2019) A stochastic data-driven ensemble forecasting framework for water resources: A case study using ensemble members derived from a database of deterministic wavelet-based models. Water Resources Research 55(1):175–202. doi:10.1029/2018WR023205

R Core Team (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org

Schaake J, Cong S, Duan Q (2006) US MOPEX data set. IAHS Publication 307:9–28

References

Sikorska AE, Montanari A, D Koutsoyiannis (2015) Estimating the uncertainty of hydrological predictions through data-driven resampling techniques. Journal of Hydrologic Engineering 20(1):A4014009. doi:10.1061/(ASCE)HE.1943-5584.0000926

Soetaert K, Petzoldt T (2010) Inverse modelling, sensitivity and Monte Carlo analysis in R using package FME. Journal of Statistical Software 33(3):1–28. doi:10.18637/jss.v033.i03

Soetaert K, Petzoldt T (2016) FME: A Flexible Modelling Environment for Inverse Modelling, Sensitivity, Identifiability and Monte Carlo Analysis. R package version 1.3.5. https://CRAN.R-project.org/package=FME

Tyralis H, Papacharalampous GA, Burnetas A, Langousis A (2019) Hydrological post-processing using stacked generalization of quantile regression algorithms: Large-scale application over CONUS. Journal of Hydrology 577:123957. doi:10.1016/j.jhydrol.2019.123957

Vitolo C (2017) hddtools: Hydrological data discovery tools. The Journal of Open Source Software 2(9). doi:10.21105/joss.00056

Vitolo C (2018) hddtools: Hydrological Data Discovery Tools. R package version 0.8.2. https://CRAN.Rproject.org/package=hddtools

Warnes GR, Bolker B, Gorjanc G, Grothendieck G, Korosec A, Lumley T, MacQueen D, Magnusson A, Rogers J (2017) gdata: Various R Programming Tools for Data Manipulation. R package version 2.18.0. https://CRAN.R-project.org/package=gdata

Wickham H (2007) Reshaping data with the reshape package. Journal of Statistical Software 21(12):1–20

Wickham H (2011) The split-apply-combine strategy for data analysis. Journal of Statistical Software 40(1):1–29

Wickham H (2016a) ggplot2. Springer International Publishing. doi:10.1007/978-3-319-24277-4

Wickham H (2016b) plyr: Tools for Splitting, Applying and Combining Data. R package version 1.8.4. https://CRAN.Rproject.org/package=plyr

Wickham H (2018) reshape: Flexibly Reshape Data. R package version 0.8.8. https://CRAN.R-project.org/package=reshape

Wickham H, Henry L (2019) tidyr: Easily Tidy Data with 'spread()' and 'gather()' Functions. R package version 0.8.3. https://CRAN.R-project.org/package=tidyr

Wickham H, Hester J, Francois R (2018) readr: Read Rectangular Text Data. R package version 1.3.1. https://CRAN.Rproject.org/package=readr

Wickham H, Chang W, Henry L, Pedersen TL, Takahashi K, Wilke C, Woo K, Yutani H (2019a) ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. R package version 3.2.1. https://CRAN.R-project.org/package=ggplot2

Wickham H, François R, Henry L, Müller K (2019b) dplyr: A Grammar of Data Manipulation. R package version 0.8.3. https://CRAN.R-project.org/package=dplyr