Supplement of Hydrol. Earth Syst. Sci., 28, 851–871, 2024
https://doi.org/10.5194/hess-28-851-2024-supplement

Hydrology and Earth System Sciences
Open Access
EGU

Supplement of

Flow intermittence prediction using a hybrid hydrological modelling approach: influence of observed intermittence data on the training of a random forest model

Louise Mimeau et al.

Correspondence to: Louise Mimeau (louise.mimeau@inrae.fr)

Supplement

1 Adequacy of the JAMS-J2000 river networks with the observed river networks

Figure S1, S2 and S3 show the comparison between the synthetic river networks generated for the hydrological modelling with JAMS-J2000 and reference river networks provided by local DRN teams from the DRYvER projetc (Datry et al., 2021). For the Albarine and Lepsämänjoki catchments, the spatial resolution of the generated river networks had to be increased in order to capture all reaches with available observed flow intermittence data.

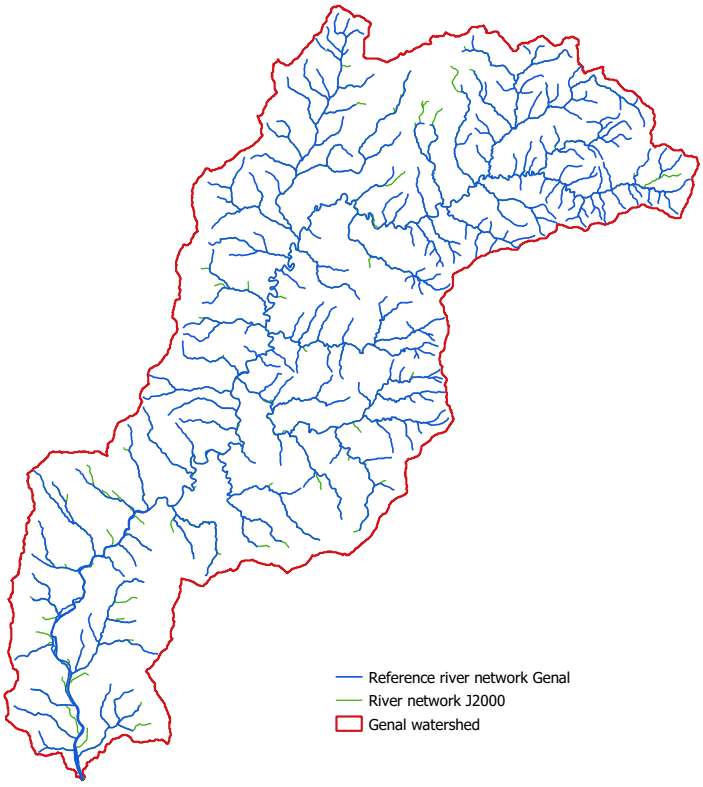

Figure S1. Generated (green) and observed (blue) river networks for the Genal catchment.

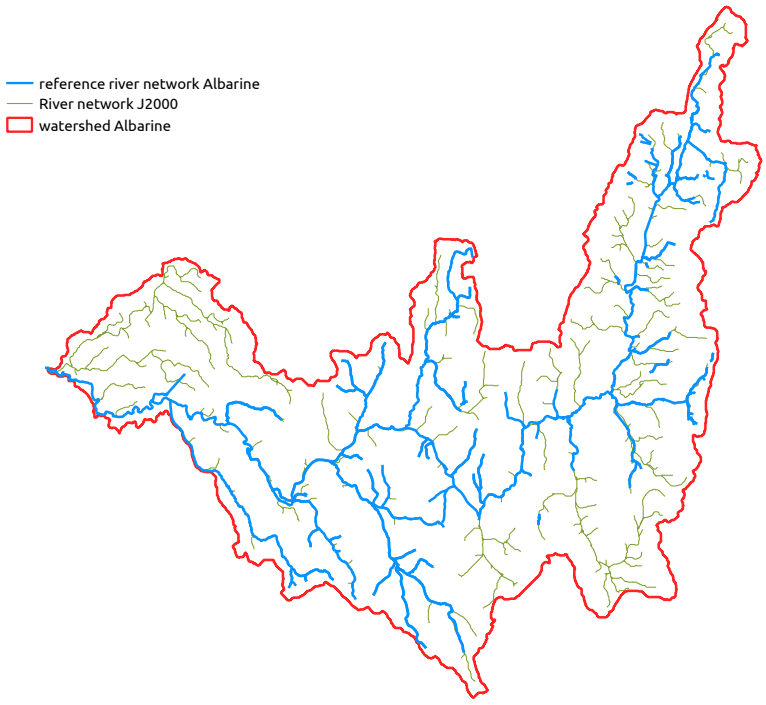

Figure S2. Generated (green) and observed (blue) river networks for the Albarine catchment.

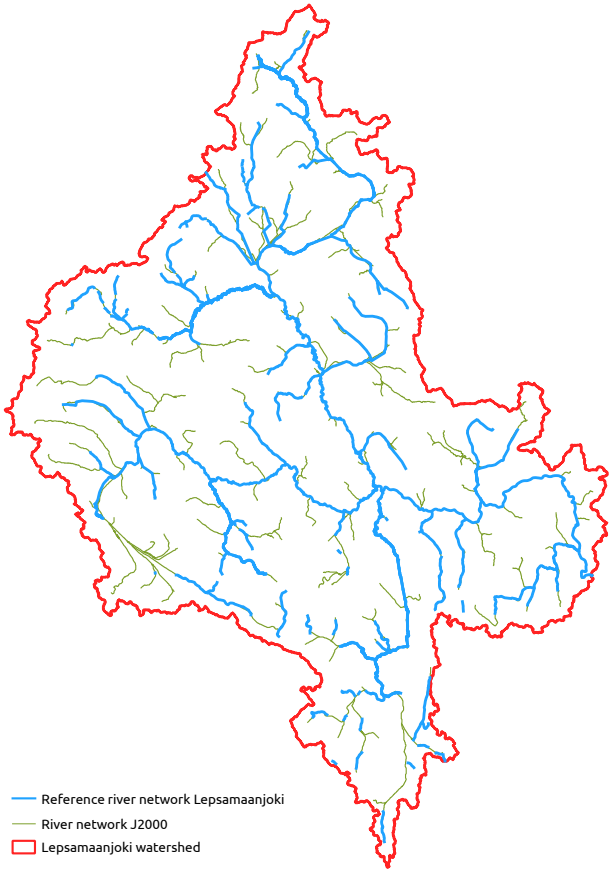

Figure S3. Generated (green) and observed (blue) river networks for the Lepsämänjoki catchment.

2 Calibration of the JAMS-J2000 model

2.1 Catchments for JAMS-J2000 calibration

The calibration of JAMS-J2000 parameters was performed on larger catchments (1500 to 3700 km^2) corresponding to the intermediate-scale basins studied in the DRYvER project (to bridge the gap between the DRN scale and the continental scale).

The larger catchments characteristics are presented in Figure S4.

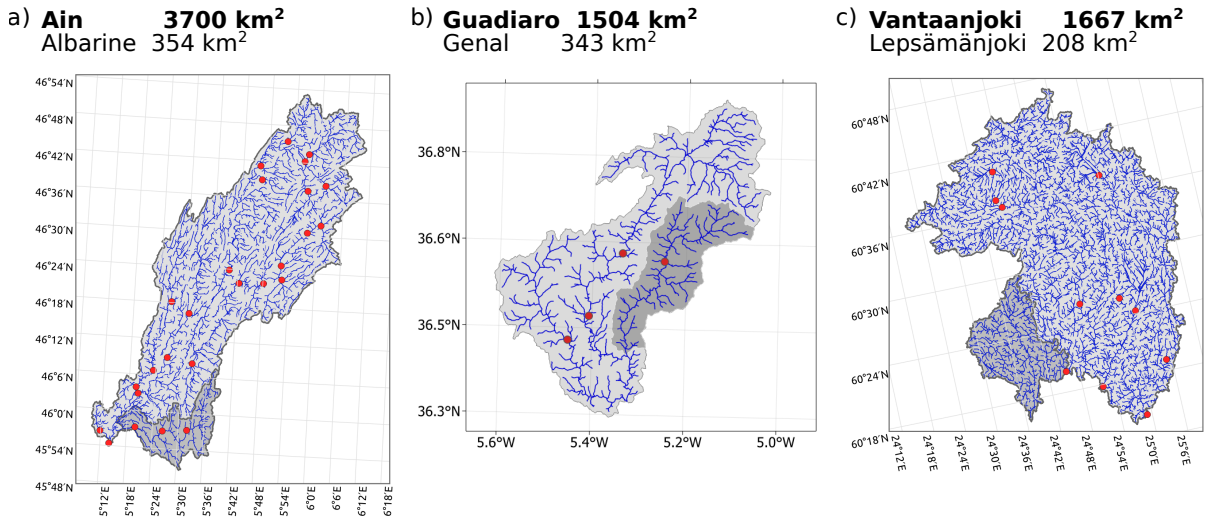

Figure S4. Catchments used for JAMS-J2000 calibration (larger catchments in light grey). Small catchments in dark grey are the catchments used for the flow intermittence modelling. Gauging stations used for the calibration are represented in red. a) French study sites : Ain and Albarine catchments, b) Spanish study sites : Guadiaro and Genal, c) Finnish study sites : Vantaanjoki and Lepsämänjoki.

2.2 Calibration of snow parameters

Snow parameters (Table S1) were calibrated separately by comparing the simulated snow cover with the catchments' fractional snow cover area from MODIS10A2 datasets using the Kling-Gupta Efficiency (KGE) (Gupta et al., 2009). For that purpose, the catchment fractional snow cover area (fSCA) from the MOD10A2 dataset available at a 8-day resolution (Hall et al., 2007) was used as observed data. MOD10A2 fSCA was downloaded for the period 2000-10-15 to 2021-05-25 and aggregated at the catchment scale. The NSGA-II algorithm (Deb et al., 2002) with 1000 iterations was used for the automatic calibration, and the KGE as an objective function. The model time series were split into a period of initialization (1995 to 2000), a period of calibration (even years from 2000 to 2020), and a period of validation (remaining years from 2000 to 2020).

For the Ain catchment, KGE is respectively equal to 0.81 and 0.74 for the calibration and validation period. For the Vantaanjoki catchment, KGE is respectively equal to 0.80 and 0.74 for the calibration and validation period. Figures S5 and S6 show simulated and observed snow cover areas in the Ain and Vantaanjoki catchments. The Guadiaro catchment is not characterized by a regular winter snow cover, and therefore, snow calibration was neglected for this catchment.

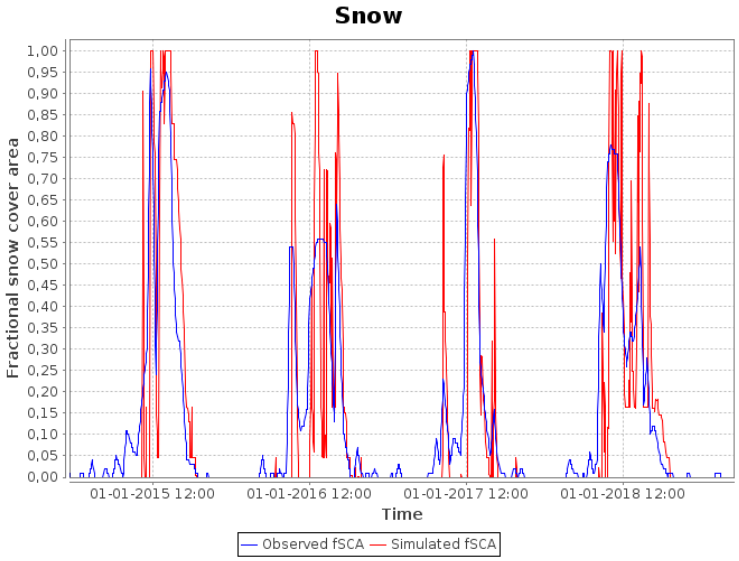

Figure S5. Simulated (red) and observed (blue) snow cover area (fSCA) for the Ain catchment. KGE calibration = 0.81, KGE validation = 0.75

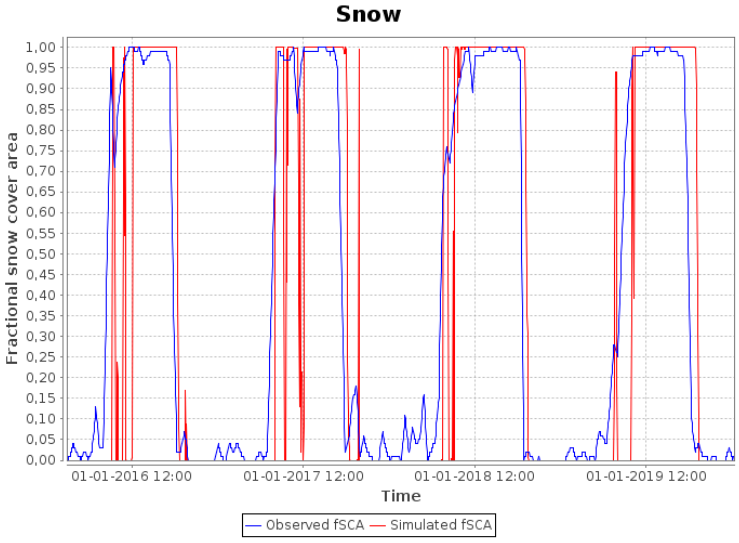

Figure S6. Simulated (red) and observed (blue) snow cover area (fSCA) for the Vantaanjoki catchment. KGE calibration = 0.79, KGE validation = 0.76

2.3 Calibration of discharge

Measured streamflow available at different gauging stations was used for model calibration and validation. For this purpose, the available streamflow data was split into two periods for calibration and validation (Table S3). Due to different catchment char-

acteristics, process dynamics, and data availability, the calibration approaches were slightly different for the three catchments. The calibration procedures are therefore presented separately for the Guadiaro catchment on the one hand and for the Ain and Vantaanjoki catchments on the other hand. The list of the calibrated parameters with their calibration range is presented in Tables S1 and S2 and all the final parameter values can be found in the JAMS-J2000 model datasets available on request.

Abbreviation	Description	Unit	Calibration range	
			Ain and Vantaanjoki	Guadiaro
snow_trans	Half width of the transition zone rainfall-snowfall	K	0 - 3.5	
snow_trs	Threshold temperature for precip phase (the temp. in which 50% of precip fall as snow and 50% as rain)	°C	0 - 3	
t_factor	Temperature factor for snow melt	mm*°C-1	0 - 8	
ccf_factor	Cold content factor	-	0.0001 - 0.01	
CropCoef_aAF	Crop coefficient additive adaptation factor	-	-0.2 - 0.2	
CropCoef_mAF	Crop coefficient multiplicative adaptation factor	-	0.5 - 2	
FCAdaptation	Multiplier for field capacity	-	0.5 - 5	0.5 - 3
ACAdaptation	Multiplier for air capacity	-	0 - 3	0.5 - 3
soilPolRed	Polynomic reduction factor for potential evapotranspiration	-	0 - 10	
soilMaxInfSnow	Maximum infiltration for snow covered areas	mm	5 - 200	
soilMaxInfSummer	Maximum infiltration in summer (Apr - Sep)	mm	5 - 200	
soilMaxInfWinter	Maximum infiltration in winter (Oct - Mar)	mm	5 - 200	
soilMaxPerc	Maximum percolation rate	mm	1 - 20	
soilLatVertLPS	LPS lateral-vertical distribution coefficient	-	0 - 10	
soilOutLPS	LPS outflow coefficient	-	0 - 10	
SoilConcRD1	Recession coefficient for surface runoff	-	1 - 5	1 - 10
SoilConcRD2	Recession coefficient for interflow	-	1 - 10	
gwRG1RG2dist	Distribution factor between shallow and deep groundwater aquifer	-	0 - 10	
flowRouteTA	Flow routing coefficient TA	-	1 - 20	1 -30

Table S1. List of the calibrated global parameters and their calibration range

Abbreviation	Description	Unit	Calibration range	
			Ain and Vantaanjoki	Guadiaro
RG1_max	Maximum storage capacity of the upper ground-water reservoir	mm	10 - 300	not calibrated
RG1_k	Storage coefficient of the upper ground-water reservoir	day	2 - 30	0.3 – 3-times of physically determined parameter
RG2_max	Maximum storage capacity of the lower ground-water reservoir	mm	100 - 1500	not calibrated
RG2_k	Storage coefficient of the lower ground-water reservoir	day	10 - 600	0.3 – 3-times of physically determined parameter

Table S2. List of the calibrated spatially distributed parameters and the calibration range

DRN	Initialization period	Calibration period	Validation period
Albarine	1990-1995	1995-2009	2009-2020
Genal	1998-2001	2001-2004	2012-2018
Lepsämänjoki	2000-2005	2005-2014	2014-2020

Table S3. Calibration and validation periods. Hydrological years start on the 1st of October and end on the 30th of September.

2.3.1 Model calibration and validation for the Guadiaro catchment

The models' performance of simulating the measured streamflow at multiple gauging stations was evaluated using a semi-automatic calibration method, which utilizes automatic and manual calibration techniques. To assess model performance, different performance criteria were used, which focus on different evaluation criteria, such as low-flow, high flows, and bias (Kundzewicz et al., 2018) (Table S4).

Efficiency criteria	Definition and reason for selection
Nash-Sutcliffe Efficiency (NSE)	Multi-objective function, strong focus on simulation of peak flows, widely used
Logarithmic Nash-Sutcliffe Efficiency (NSE_{log})	Like NSE, but logarithm focuses on the representation of simulation of low flows
Relative Volume Error (pBias)	Representing overall under or overestimation
Kling-Gupta-Efficiency (KGE)	multi-objective function, representing bias, correlation, and flow variability

Table S4. Efficiency criteria used for automatic model calibration and performance evaluation (Gupta et al., 2009)

Overall, 15 global model parameters were calibrated, which showed a moderate to high sensitivity on processes related to infiltration, evapotranspiration, percolation, soil, groundwater, and runoff routing (Tables S1 and S2). Besides, hydrogeological parameters influencing the recession of the water from shallow and deep groundwater aquifers were calibrated in a spatially distributed manner. For the automatic optimization, the multi-objective, non-dominated sorting genetic search algorithm NSGA-II was applied (Deb et al., 2002). Here, the three performance criteria NSE, NSElog and pBias at different gauging stations were used in 5000 iterations to optimize the 15 parameters and hence, simulated streamflow in the Genal DRN. Additionally, the process was repeated for different spatially distributed hydrogeology parameter sets addressing the varying groundwater recession from shallow and deep groundwater aquifers. This resulted in several pareto-optimal model solutions, which still inherited strong differences of hydrological process patterns, considering, for example, the overland flow or groundwater contribution to the overall runoff. Even though statistical measures have the advantage to objectively classify model performance and allow comparison across different models, they do not substitute for visual interpretation of simulated and observed hydrographs and interpretation of process dynamics by domain experts (Legates and McCabe Jr, 1999; Moriasi et al., 2007). Therefore, the models were manually calibrated in a second step to fine-tune the results of the automated calibration procedure. The focus here was particularly on the representation of the runoff recession and groundwater contribution. Besides, due to the modeling of flow intermittence, the performance of modeling low flows was weighted higher than the model's ability to simulate high flows accurately. Further, when selecting the final parameter sets for the DRN, sets showing higher performance at the smaller basin were given preference over sets showing higher performance at the larger basin. During calibration, the following hydrological characteristics were taken into account:

1. runoff components (Hortonian and Hewlettian runoff, subsurface flow from soil and upper groundwater zones as well as baseflow), precipitation, actual evapotranspiration, and soil saturation;
2. seasonal and annual water balances;
3. spatially distributed processes at the HRU level: runoff generation, interflow, soil water balancing, evapotranspiration, and groundwater recharge.

Finally, from all pareto-optimal solutions identified through automatic and manual calibration, the most plausible model in terms of process representation according to the observed data and the knowledge about environmental characteristics was selected.

Figures S7 and S8 show the performance of the JAMS-J2000 model in the Guadiaro catchment.

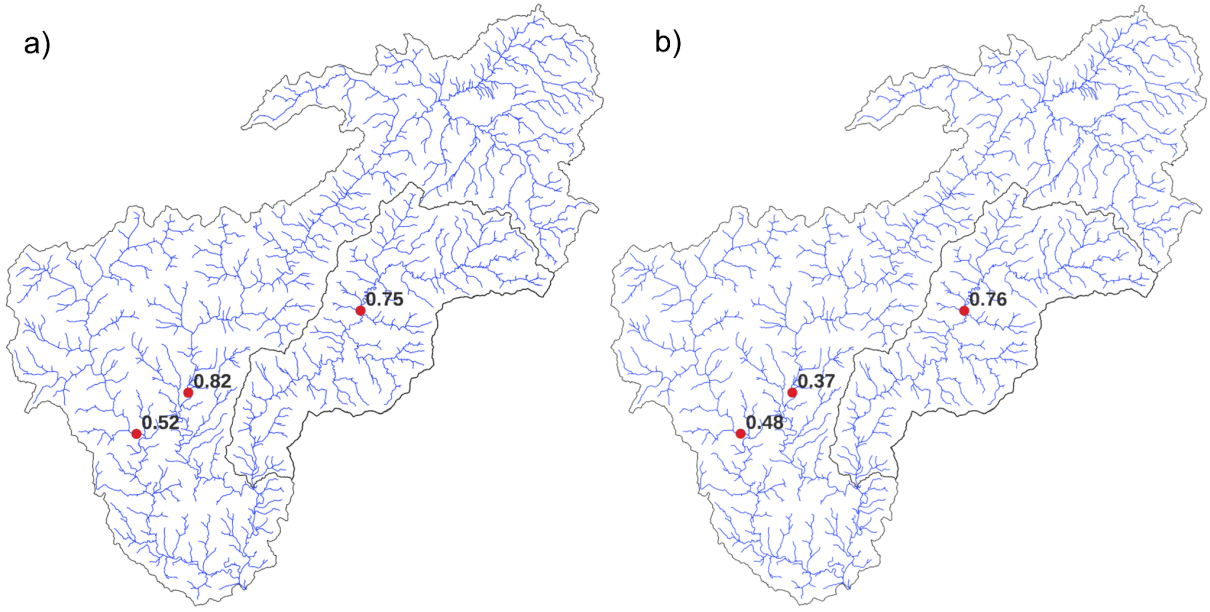

Figure S7. KGE for the a) calibration period (2002 - 2012; Genal: 2001 - 2004) and b) the validation period (2012 - 2018) for the Guadiaro catchment

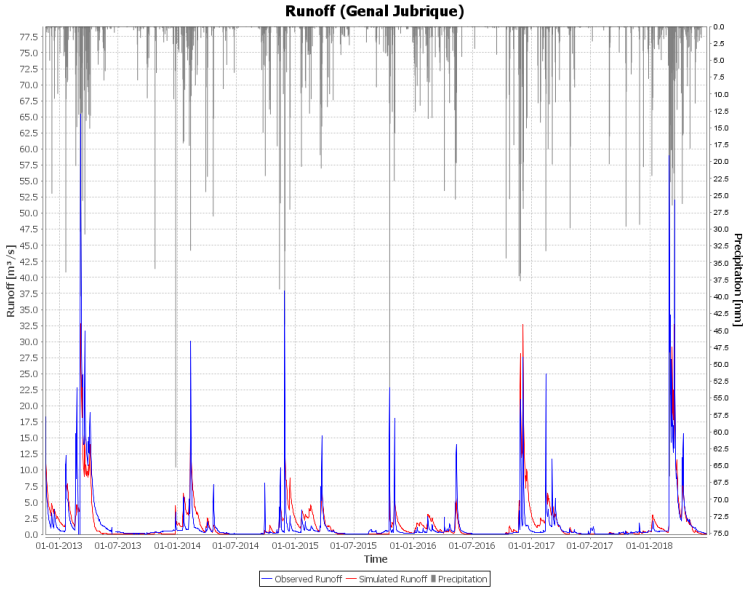

Figure S8. Simulated (red) and observed (blue) discharge at the Jubrique gauging station (in the Genal catchment). KGE calibration = 0.75, KGE validation = 0.76

2.3.2 Model calibration and validation for the Ain and Vantaanjoki catchments

The calibration for the Ain, Fekete, and Vantaanjoki catchments also uses a multi-stations and multi-objectives approach, but using a different method. 15 global parameters related to evapotranspiration, infiltration in the soil layer, and percolation to the groundwater layer, (Table S1) as well as 4 spatially distributed parameters related to the groundwater reservoirs (Table S2), are calibrated. The calibration and validation periods (Table S4) were selected based on the availability of the observed discharge data in the different river basins. In a first step, the Latin Hyper Cube Sampling (LHS) is used to generate 5000 model runs for each pilot river basin (the calibration ranges are provided in Tables S1 and S2). Then the calibrated set of parameters is selected among the 5000 parameters sets so that the model performs best on (i) a multi-objective function (MOF) representing KGE, low-flows (10^{th} percentile Q10), and mean annual outflow (Q_{yr}), (ii) all the stations. The MOF function is computed for all stations and all parameter sets (Eq. 1), then a weighted average over the stations is calculated to prioritize the stations located at the outlets of the large and small river basins (Ref_1) and the other stations located in the small river basin (Ref_2) (Eq. 2) (other stations located in the large catchment are referred as Ref_3). The final calibrated set of parameters is selected among the model runs leading to the lowest $MOF_{\text{all stations}}$.

$$\text{MOF} = 0.6 * (1 - \text{KGE}) + \frac{0.2 * |\text{Q10}_{sim} - \text{Q10}_{obs}|}{\text{Q10}_{obs}} + \frac{0.2 * |\text{Q}_{\text{yr}_{sim}} - \text{Q}_{\text{yr}_{obs}}|}{\text{Q}_{\text{yr}_{obs}}} \quad (1)$$

$$\mathrm{MOF} = \frac{w_{Ref_1} \sum_{i \in Ref_1} \mathrm{MOF}_i + w_{Ref_2} \sum_{j \in Ref_2} \mathrm{MOF}_j + w_{Ref_3} \sum_{k \in Ref_3} \mathrm{MOF}_k}{\sum_{i \in Ref_1} w_{Ref_1} + \sum_{j \in Ref_2} w_{Ref_2} + \sum_{k \in Ref_3} w_{Ref_3}} \quad (2)$$

with w_{Ref_1} = 5, w_{Ref_2} = 2, w_{Ref_3} = 1

The Ain catchment is characterized by karstic areas which have a strong impact on the hydrological response of the sub-catchments. As the JAMS-J2000 model does not represent the karst-related processes, a correction factor k was applied to the observed discharges at the gauges before comparison with simulated discharges to consider water input or water losses in sub-catchments through the karstic network (Eq. 3).

$$k = \frac{\mathrm{P} - \mathrm{ET_{act}} - \mathrm{Q}}{\mathrm{P} - \mathrm{ET_{act}}} \quad (3)$$

with P the observed mean annual precipitation in mm (from the Safran reanalysis; Vidal et al. (2010)), $\mathrm{ET_{act}}$ the mean annual actual evapotranspiration simulated with JAMS-J2000 in mm (forced with Safran reanalysis as climate input data), and Q the observed mean annual outflow in mm.

Figures S9 and S10 show the performance of the JAMS-J2000 model in the Ain catchment.

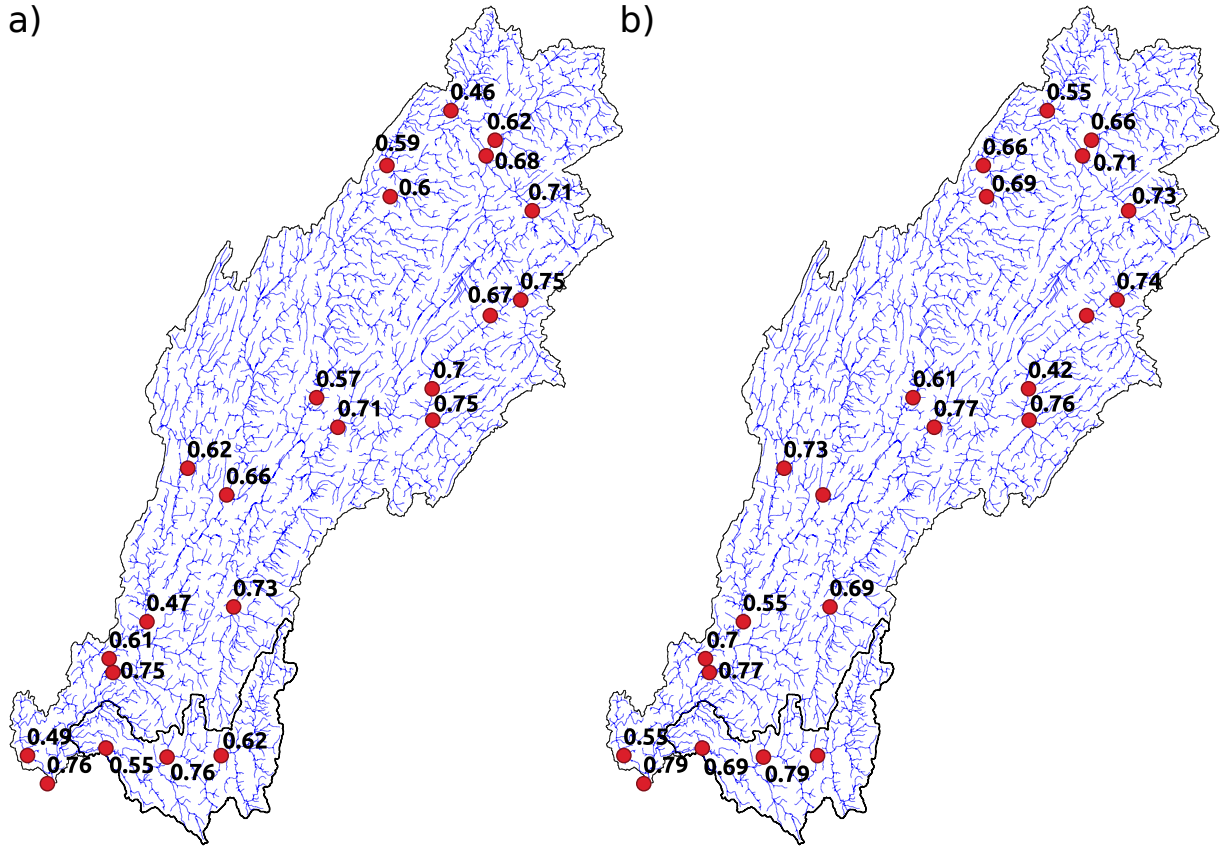

Figure S9. KGE for the a) calibration period (1995-2009) and b) the validation period (2009-2019) for the Ain catchment. Missing values during the validation period are due to the shutdown of gauging stations

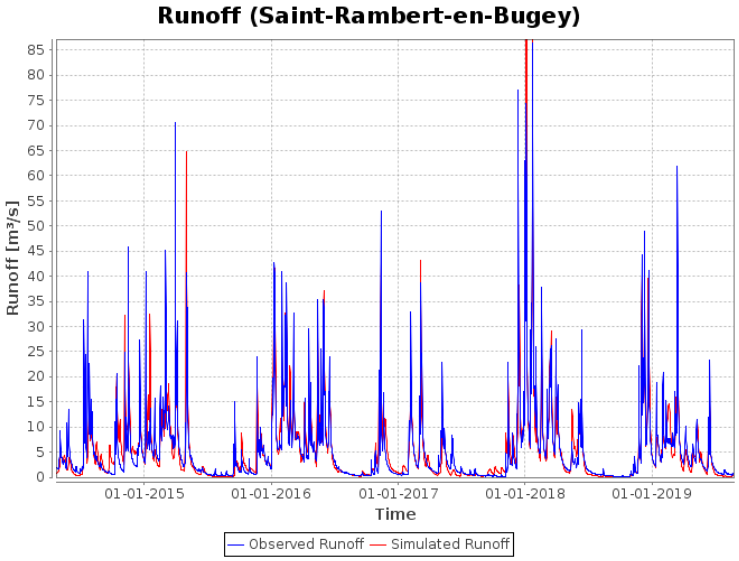

Figure S10. Simulated (red) and observed (blue) discharge at the Saint-Rambert station (in the Albarine catchment) for the validation period. KGE calibration = 0.76, KGE validation = 0.79

Figures S11 and S12 show the performance of the JAMS-J2000 model in the Vantaanjoki catchment.

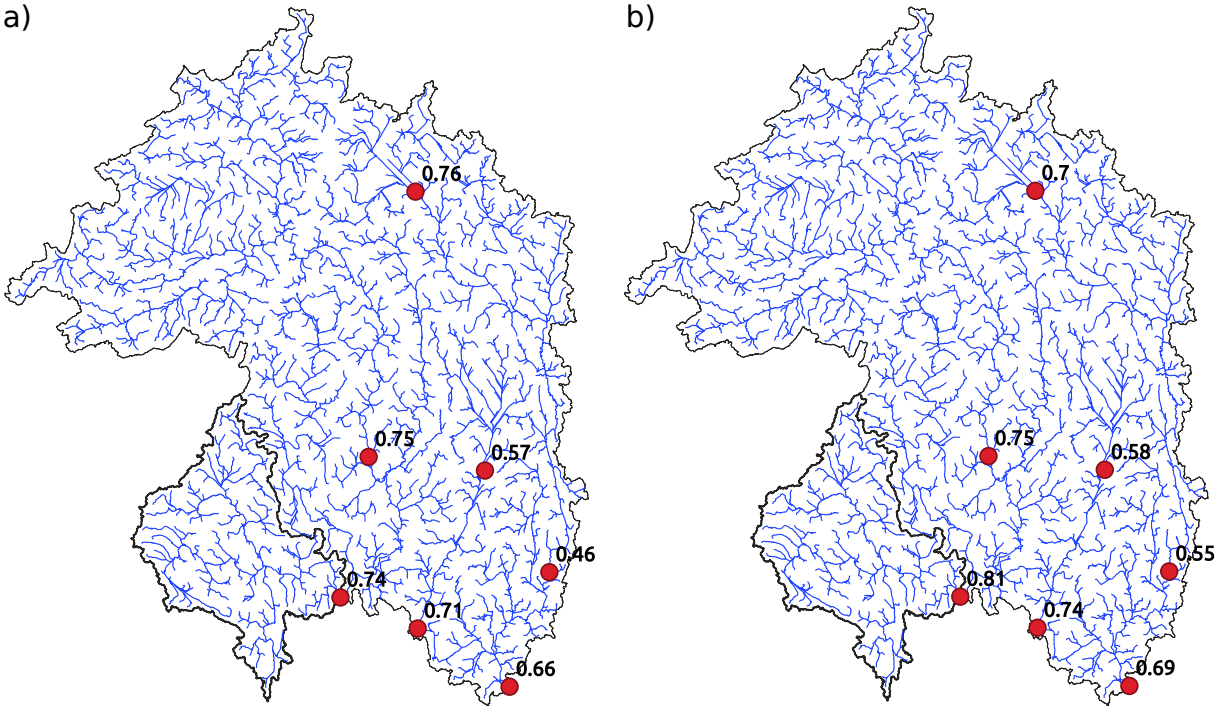

Figure S11. KGE for the a) calibration period (2005-2014) and b) the validation period (2014-2020) for the Vantaanjoki catchment

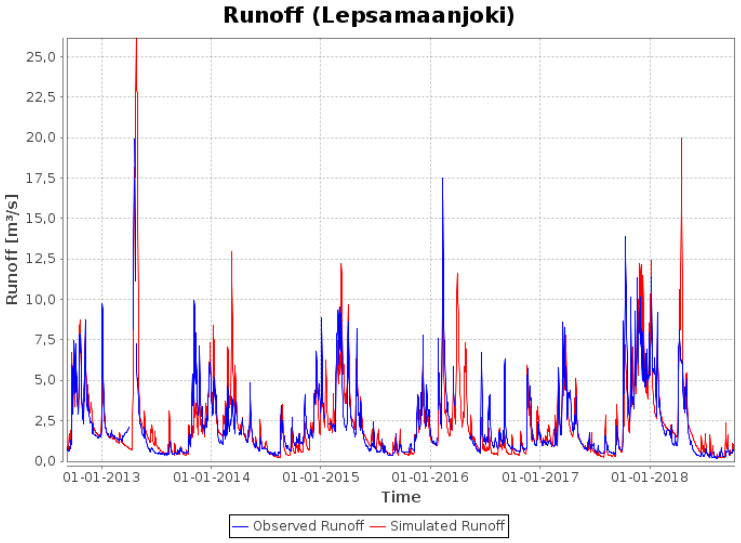

Figure S12. Simulated (red) and observed (blue) discharge at the Lepsämänjoki station for the validation period. KGE calibration = 0.74, KGE validation = 0.81

2.3.3 Validation of the JAMS-J2000 model for low flows

In this section we analyse the performance of the JAMS-J2000 model to simulate discharges below a defined threshold at the gauging stations in the 3 catchments. The threshold is taken as the 10^{th} percentile of observed discharges on the total period (calibration and validation), keeping only the values for hydrological years with less than 10% missing values.

Then the performance metrics were computed as follows:

$$\text{Sensitivity} = \frac{a}{a+c}$$

$$\text{Specificity} = \frac{d}{b+d}$$

$$\text{Accuracy} = \frac{a+d}{a+b+c+d}$$

$$\text{FAR} = \frac{b}{a+b}$$

with a, b, c, and d the number of simulated days for each of the 4 conditions defined in the confusion matrix (Table S5).

	Observed discharges	
Simulated discharges	Q_{obs} < threshold	Q_{obs} > threshold
Q_{sim} < threshold	a	b
Q_{sim} > threshold	c	d

Table S5. Confusion matrix of matches and mismatches of predictions and observations of low flows with JAMS-J2000.

Figure S13 shows the results of the analysis of the performance of JAMS-J2000 to simulate low flows. For the validation period, Sensitivity is respectively equal to 0.84, 0.62 and 0.91 for the St-Rambert (Albarine), Lepsämänjoki, and Jubrique (Genal) gauging stations, which shows that the hydrological model simulates low-flow periods very well (for Saint-Rambert and Jubrique) and fairly well (for Lepsämänjoki). However, for the St-Denis gauging station in the Albarine catchment, Sensitivity is equal to 0. This is due to the fact that the river is intermittent and sometimes completely dries at this station. For the St-Denis station the 10th percentile is equal to 0 m^3/s, and as the JAMS-J2000 is not able to simulate complete drying, the model performance is poor. These results show that the JAMS-J2000 hydrological model can provide correct simulations of the alternation between periods of low flow and medium-high flow to the RF model, but that the J2000-RF coupling is essential for correctly simulating the flow on intermittent river sections.

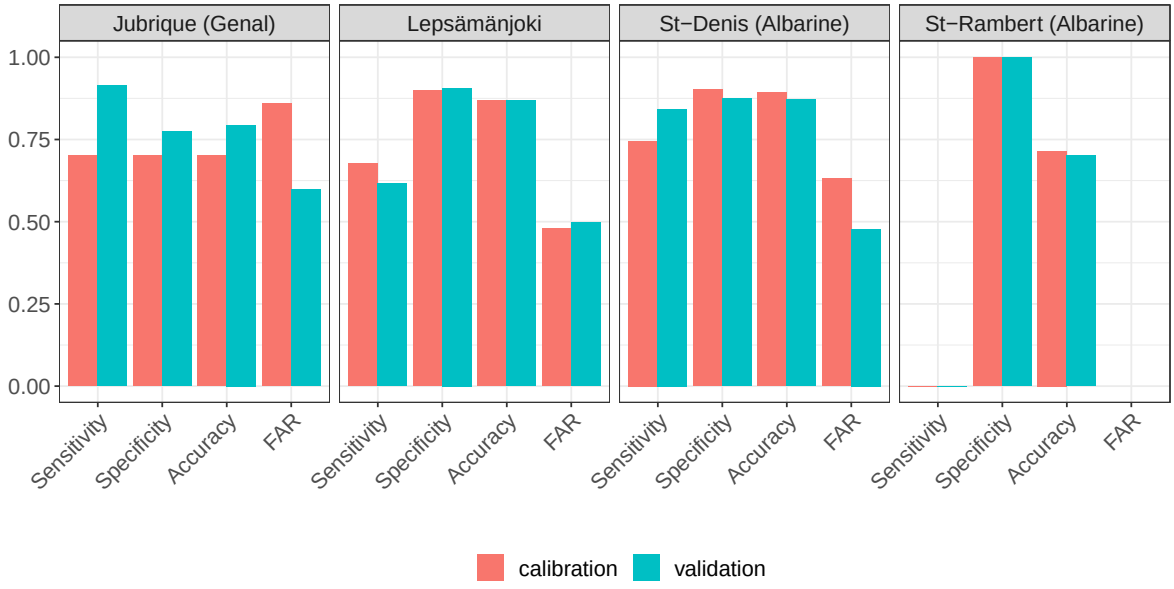

Figure S13. Metrics for the prediction of low flows (< 10^{th} percentile of observed discharges) with JAMS-J2000 model at the gauging stations for the calibration and validation periods.

3 Validation of the simulated spatial flow intermittence pattern in the Albarine and Genal DRNs

Some classification of the reaches flow regime (permanent or intermittent) were provided for the Albarine and Genal DRN by local teams based on their observations. Figures S14 and S15 show the comparison between the simulated spatial pattern of flow intermittence in the Albarine and Genal DRNs with the observed flow regimes in the river networks.

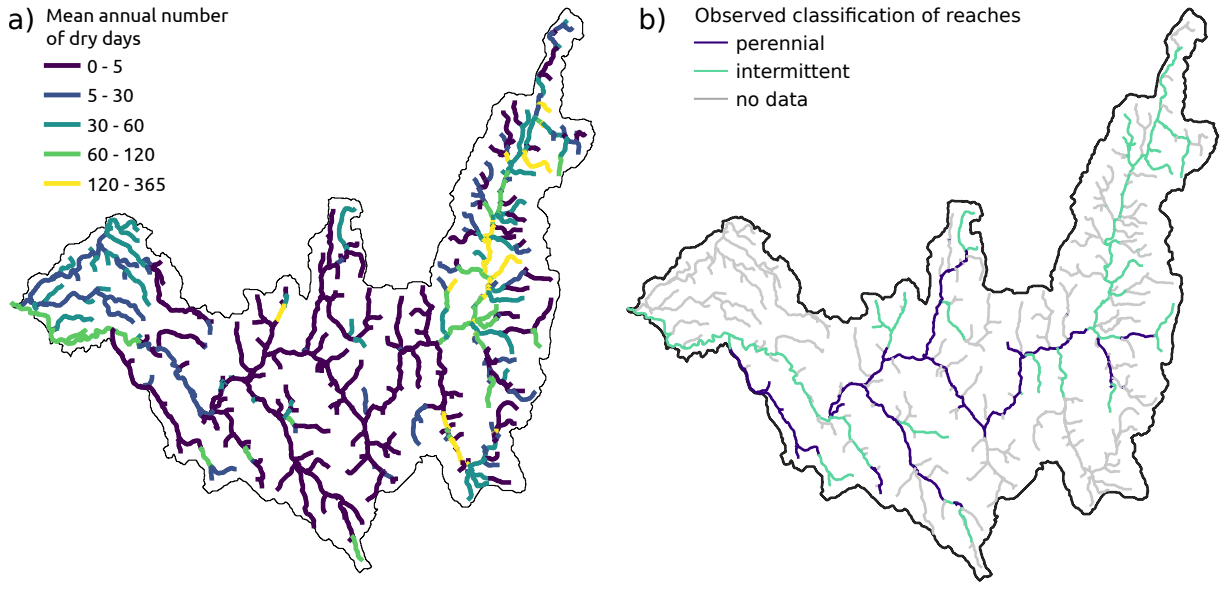

Figure S14. a) Mean annual number of days with a dry condition predicted in the Albarine DRN with the flow intermittence model, b) classification of the reaches (perennial or intermittent) provided by the DRN local experts based on observations.

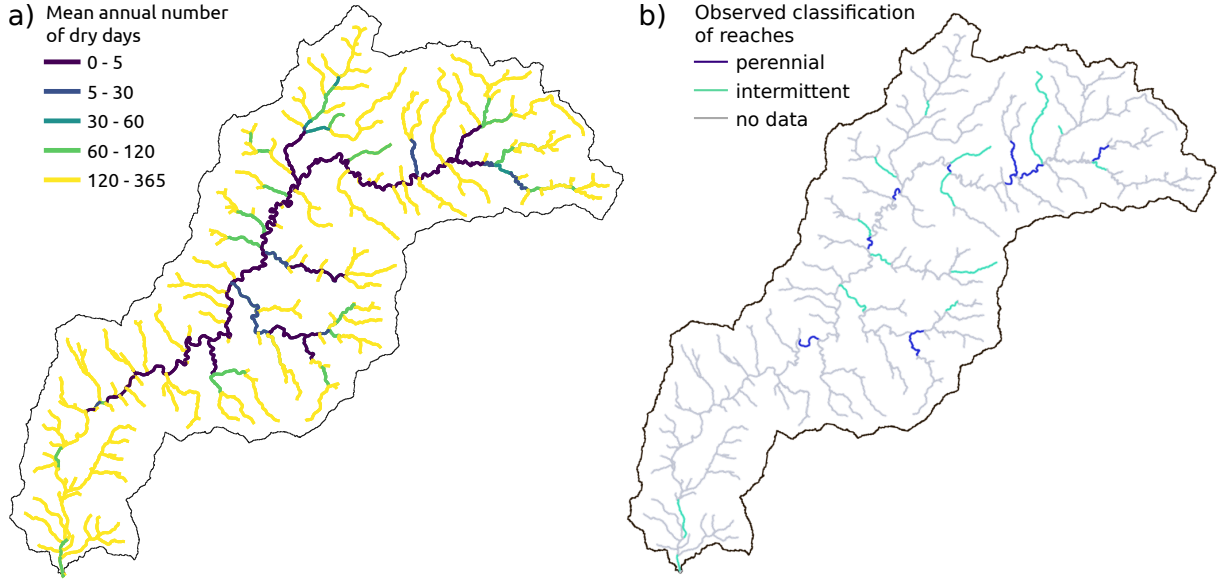

Figure S15. a) Mean annual number of days with a dry condition predicted in the Genal DRN with the flow intermittence model, b) classification of the reaches (perennial or intermittent) provided by the DRN local experts based on observations.

4 Covariates analysis with configuration 1

Figure S16 shows the importance of the covariates with configuration 1 in the 3 catchments. It is very similar to Figure 10 from the manuscript (with configuration 0) for the 5 most important covariates, which shows that for this study the importance of the covariates is not very sensitive to the size of the training sample, but rather to the quality of the covariates (cf Section 3.5.3 of the manuscript).

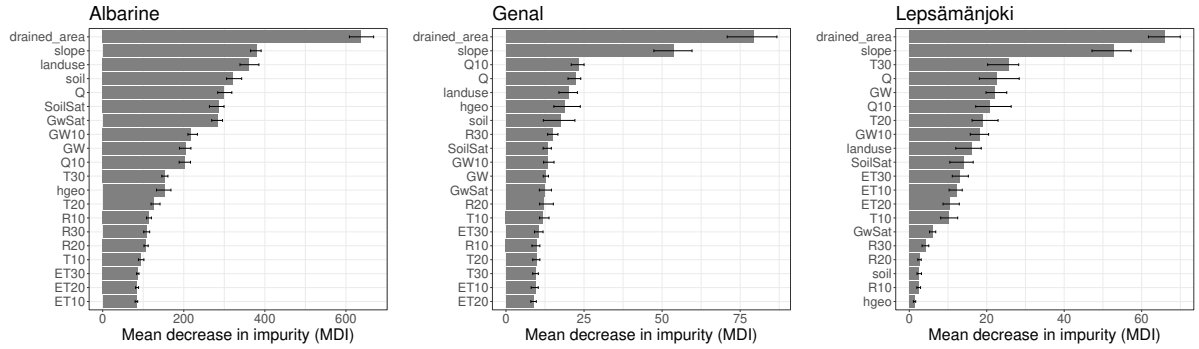

Figure S16. Importance of the covariates in the RF models (mean decrease in impurity (Archer and Kimes, 2008)) for the 3 DRNs. Bars represent the mean MDI and the error bars the minimum and maximum values of MDI for the 20 runs of the RF model with configuration 1.

References

Archer, K. J. and Kimes, R. V.: Empirical characterization of random forest variable importance measures, Computational statistics & data analysis, 52, 2249–2260, https://doi.org/10.1016/j.csda.2007.08.015, 2008.

Datry, T., Allen, D., Argelich, R., Barquin, J., Bonada, N., Boulton, A., Branger, F., Cai, Y., Cañedo-Argüelles, M., Cid, N., Csabai, Z., Dallimer, M., de Araújo, J.-C., Declerck, S., Dekker, T., Döll, P., Encalada, A., Forcellini, M., Foulquier, A., Heino, J., Jabot, F., Keszler, P., Kopperoinen, L., Kralisch, S., Künne, A., Lamouroux, N., Lauvernet, C., Lehtoranta, V., Loskotová, B., Marcé, R., Martin Ortega, J., Matauschek, C., Miliša, M., Mogyorósi, S., Moya, N., Müller, S. H., Munné, A., Munoz, F., Mykrä, H., Pal, I., Paloniemi, R., Pařil, P., Pengal, P., Pernecker, B., Polášek, M., Rezende, C., Sabater, S., Sarremejane, R., Schmidt, G., Senerpont Domis, L., Singer, G., Suárez, E., Talluto, M., Teurlincx, S., Trautmann, T., Truchy, A., Tyllianakis, E., Väisänen, S., Varumo, L., Vidal, J.-P., Vilmi, A., and Vinyoles, D.: Securing Biodiversity, Functional Integrity, and Ecosystem Services in Drying River Networks (DRYvER), Research Ideas and Outcomes, 7, e77 750, https://doi.org/10.3897/rio.7.e77750, 2021.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE transactions on evolutionary computation, 6, 182–197, https://doi.org/10.1109/4235.996017, 2002.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, Journal of hydrology, 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Hall, D., Riggs, G., and Salomonson, V.: MODIS/Terra Snow Cover 8-Day L3 Gobal 500 m Grid V005. National Snow and Ice Data Centre, Digital media, Boulder, 2007.

Kundzewicz, Z., Krysanova, V., Benestad, R., Hov, Ø., Piniewski, M., and Otto, I.: Uncertainty in climate change impacts on water resources, Environmental Science & Policy, 79, 1–8, https://doi.org/10.1016/j.envsci.2017.10.008, 2018.

Legates, D. R. and McCabe Jr, G. J.: Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation, Water resources research, 35, 233–241, https://doi.org/10.1029/1998WR900018, 1999.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., and Veith, T. L.: Model evaluation guidelines for systematic quantification of accuracy in watershed simulations, Transactions of the ASABE, 50, 885–900, https://doi.org/10.13031/2013.23153, 2007.

Vidal, J.-P., Martin, E., Franchistéguy, L., Baillon, M., and Soubeyroux, J.-M.: A 50-year high-resolution atmospheric reanalysis over France with the Safran system, International Journal of Climatology, 30, 1627–1644, https://doi.org/10.1002/joc.2003, 2010.