Geosci. Model Dev., 18, 8663–8678, 2025
https://doi.org/10.5194/gmd-18-8663-2025

Geoscientific Model Development
Open Access

Data-driven estimation of the hydrologic response using generalized additive models

Quentin Duchemin[1,★], Maria Grazia Zanoni[2,★], Marius G. Floriancic[2,3], Hansjörg Seybold[2], Guillaume Obozinski[1], James W. Kirchner[3,4,5], and Paolo Benettin[6]

[1]Swiss Data Science Center, ETH Zürich & EPFL, Zürich & Lausanne, Switzerland
[2]Department of Environmental Systems Science, ETH Zürich, Zürich, Switzerland
[3]Department of Civil, Environmental and Geomatic Engineering, ETH Zürich, Zürich, Switzerland
[4]Swiss Federal Research Institute WSL, Birmensdorf, Switzerland
[5]Department of Earth and Planetary Science, University of California, Berkeley, California, USA
[6]Institute of Earth Surface Dynamics (IDYST), Faculty of Geosciences and Environment, Université de Lausanne, Lausanne, Switzerland
★These authors contributed equally to this work.

Correspondence: Paolo Benettin (paolo.benettin@unil.ch)

Received: 4 April 2025 – Discussion started: 25 April 2025
Revised: 30 September 2025 – Accepted: 11 October 2025 – Published: 17 November 2025

Abstract. Estimating the hydrologic response of watersheds to precipitation events is key to understanding streamflow generation processes. Impulse Response Functions, commonly referred to as the Instantaneous Unit Hydrograph (IUH) in hydrology, have been used for over 50 years to predict stormflow and compare catchment behaviors. These response functions are often strongly affected by modelers' choices of parameters and data preprocessing procedures, limiting their diagnostic utility and generalizability across different sites and time periods. With the increasing availability of compiled rainfall-runoff series, there is now a growing opportunity to develop new approaches that fully exploit such datasets. This paper introduces GAMCR, a novel data-driven approach that employs Generalized Additive Models (GAM) to estimate time-dependent Catchment Responses (CR). GAMCR is designed to capture the complex, nonlinear relationships between precipitation and runoff, offering a flexible and interpretable framework for the systematic analysis of hydrological responses. The model is successfully validated on synthetic data, where the true response functions are known. Additionally, we demonstrate the model's potential using observed data from six Swiss basins with distinct hydrological behaviors. Results are fully consistent with those obtained from ERRA, another recent data-driven approach with a very different architecture, as well as with the climate and physical properties of the sites. Overall, GAMCR is a modern and effective tool for leveraging rainfall-runoff datasets to investigate the controls on hydrologic responses in small to midsize basins under conditions similar to those studied here, and it provides a framework that can be further explored in other climatic and physiographic settings in future research.

1 Introduction

Precipitation is generally the main water input to a landscape and the fundamental driver of streamflow generation. Quantifying how much streamflow is produced after a rain event is essential for water resources management and flood prevention, and is also useful to characterize watershed behavior. The hydrologic response (or runoff response) is usually defined as the change in streamflow induced by a given input of precipitation (Ponce, 1995; Kirchner, 2022; Kirchner et al., 2023). Years of tracer studies have clarified that, apart from rare exceptions, such a response does not primarily consist of water that fell as precipitation during the same event, but rather by water already existing in the landscape (in the form of soil water and groundwater) that is quickly mobilized during the storm (Kirchner et al., 2000; McGuire and McDon-

Published by Copernicus Publications on behalf of the European Geosciences Union.

nell, 2006; Botter et al., 2010; van der Velde et al., 2012; Kirchner, 2003; Knapp et al., 2025). The hydrologic response can be interpreted as reflecting the celerity with which increases in hydraulic potentials, induced by the new precipitation, propagate through the subsurface (McDonnell and Beven, 2014). Thus, stream water is generally much "older" than the most recent rainfall (McDonnell et al., 2010), although it may respond within minutes after the onset of precipitation.

The hydrologic response is a fundamental catchment signature, but its estimation is not straightforward, because catchment behavior is often nonlinear and nonstationary, meaning that the effects of precipitation inputs are not simply additive, and the same rain can generate different hydrological responses, depending on when it falls (Kirchner, 2024; Beven, 2001). The first approaches to characterize the hydrologic response came from the need to make streamflow predictions for engineering design. These approaches were based on instantaneous unit hydrographs (IUH, Sherman, 1932), analogous to the concepts of impulse response functions or transfer functions in signal processing, which are probability density functions describing how impulses of precipitation are transformed into runoff. The IUH has been typically modeled as a parametric curve like a Gamma or Weibull distribution. To cope with the complexities of runoff generation processes, the classic IUH approaches rely heavily on the concept of effective rainfall (or rainfall excess, J_e), which is the fraction of rainfall that effectively mobilizes runoff. The effective rainfall is typically modeled as a (nonlinear) function of antecedent wetness (e.g. through the popular SCS Curve Number approach, Soil Conservation Service, 1985) and acts as a filter that separates the rainfall volumes that effectively produce runoff from those that evaporate or that recharge subsurface storage. The IUH is then assumed to be linear and time-invariant, enabling the use of standard convolution approaches to compute streamflow Q from an effective precipitation time series. The IUH theory, pioneered by the work of Sherman (1932) and further developed by Snyder (1955) and by Bruen and Dooge (1992), provided an effective and structured way to represent the relationship between (effective) rainfall and runoff. Several advances to IUH theory have been made over the years, including linking the IUH shape with basins' geomorphological properties (see Rigon et al., 2016). The IUH approach is also popular for teaching the rainfall-runoff transformation in many engineering programs (Mays, 2019).

Although IUH approaches are often successful at reproducing stormflow hydrographs, they typically require pre-processing steps to estimate effective precipitation, and to separate the hydrograph into stormflow vs. baseflow. These pre-processing steps limit the diagnostic capability of the IUH and its generality for comparing different sites and time periods. Rainfall-runoff data from hundreds of watersheds worldwide is increasingly available in harmonized databases that facilitate modeling and cross-site comparisons (e.g. Kratzert et al., 2023; do Nascimento et al., 2024). These emerging datasets create the possibility to characterize hydrological responses from many diverse watersheds, and thus to better understand their controlling factors. To characterize hydrological response without the constraints inherent in the IUH approach, Kirchner (2022) proposed a data-driven approach for estimating impulse response functions that account for nonlinear, nonstationary and heterogeneous system behavior. This approach was further developed for rainfall-runoff data and termed ensemble rainfall-runoff analysis, or ERRA (Kirchner, 2024). Although the ERRA approach shows considerable promise (e.g. Gao et al. (2025)), it is worth considering whether other approaches can be developed to exploit the power of machine learning for innovative explorations of hydrological response.

Building on these advancements and on the widespread availability of rainfall-runoff data, here we introduce GAMCR, a data-driven approach that employs Generalized Additive Models (GAM) to estimate time-dependent Catchment Responses (CR). We present the general model architecture and provide a series of synthetic and observed data examples to: (1) validate GAMCR and compare its performance with the ERRA approach, and (2) showcase the model's potential to estimate hydrological response at diverse watersheds, characterized by diverse properties and behaviors. Differently from ERRA, GAMCR aims to estimate the hydrologic response to each individual precipitation event using combinations of spline basis functions, with coefficients determined through machine learning techniques. This approach, though requiring to fit Generalized Additive Models, allows for greater flexibility since additional information (e.g., temperature, dam operations, or site-specific characteristics) can be incorporated into the model. The goal of GAMCR is to facilitate systematic comparisons of hydrological responses across sites where precipitation-runoff time series are available.

2 Model development

2.1 General convolution model

According to the classic convolution integral, streamflow Q is computed as the convolution of precipitation J with the stationary hydrologic response IUH, which in continuous time is expressed as:

$$Q(t) = \int_0^\infty J(t-\tau)\,\mathrm{IUH}(\tau)\,\mathrm{d}\tau \tag{1}$$

Here we use a discrete-time approximation to Eq. (1), generalized to allow the IUH to vary with time:

$$y_t = \sum_{T=0}^{T_{\max}} x_{t-T}\, h_{t-T}(T)\,\Delta T \tag{2}$$

where y is the output flux (i.e., streamflow) at time t, x is the input flux (i.e., precipitation) T time steps earlier (i.e., at time $t-T$), and $h_{t-T}(T)$ is a time-variable and non-unitary response function that reflects the streamflow response to precipitation falling at time $t-T$, as a function of lag time T. The dependence of h on the precipitation time $t-T$ incorporates any dependence on internal and external forcings, such as precipitation intensity and wetness conditions at the time that rain falls.

At this stage we make very few assumptions about the shape that h can take. It is not a probability density function, meaning that its area can be smaller or larger than one. While in principle h can take negative values (if this is what the system under consideration does, and is reflected in its data), we will assume that h is always non-negative (see Sect. 2.2). By design, h refers to the response to precipitation falling over a specific time step $t-T$. Any two time steps are generally expected to initiate different responses, but Eq. (2) is obviously ill-posed because the array $h_{t-T}(T)$ contains many more unknowns than can be constrained by the vectors y_t and x_{t-T}. Thus it is necessary to evaluate $h_{t-T}(T)$ as an average over one or more ensembles of time steps (for example during which the precipitation intensity or antecedent wetness is within a given range). In particular, the ensemble responses introduced by Kirchner (2024) can be readily obtained in a post-processing step. Given an ensemble of time points $\mathcal{E}$, the Runoff Response Distribution (RRD, units of $1/T$) is the average response weighted by precipitation intensity h over the selected time points $\mathcal{E}$:

$$\mathrm{RRD}_{\mathcal{E}}(T) := \frac{\sum_{t'\in\mathcal{E}} x_{t'} h_{t'}(T)}{\sum_{t'\in\mathcal{E}} x_{t'}}, \tag{3}$$

(where $t' = t-T$ represents the time that precipitation falls), while the Nonlinear Response Function (NRF, units $\mathrm{L/T^2}$) is the average response multiplied by the corresponding precipitation intensity:

$$\mathrm{NRF}_{\mathcal{E}}(T) := \frac{1}{|\mathcal{E}|}\sum_{t'\in\mathcal{E}} x_{t'} h_{t'}(T). \tag{4}$$

2.2 GAMCR model

GAMCR is a machine learning model that estimates transfer functions from flux data. GAMCR models the catchment's response to any single precipitation event as a weighted sum of spline basis functions. The time-varying coefficients of these basis functions are estimated using machine learning techniques, specifically Generalized Additive Models. As a result, we use more technical language in this section and the next, drawing terminology from the data science literature.

The problem of learning time-dependent transfer functions from rainfall-runoff data is ill-posed, meaning that considering a too large model class might result in zero training loss but with poor test error. In the machine learning community, the standard approach to cope with such badly conditioned inverse problems is to exploit prior knowledge on the structure of the studied system to either shrink the class of target functions or to regularize the optimization problem (Arridge et al., 2019). Following this approach, GAMCR is built on three core principles.

1. First, GAMCR specifies a set of features that are assumed to be the main drivers of the catchment response to a given precipitation event. These features can be modified by the user if needed and should typically include information characterizing the catchment condition and the precipitation event considered.

2. Second, we assume that the catchment response to a precipitation event will vary smoothly as a function of this feature vector, a structural assumption similar to the one implicitly used in the approach by Kirchner (2022). This means we expect similar feature vectors to produce similar hydrologic responses.

3. Third, we expect the transfer functions, $T \longmapsto h_{t'}(T)$, to exhibit sharp peaks for short time lags, that progressively smooth out as the lag time T increases.

With these guiding principles, we model the transfer functions as follows:

$$h_{t'}(T) = \sum_{\ell=1}^{L} g_\ell(z_{t'}) b_\ell(T), \tag{5}$$

where $(b_\ell)_{\ell\in[L]}$ are B-splines constructed by considering an irregular spacing of knots, z_t' is a feature vector describing both the catchment conditions and the precipitation event at time t' and g_ℓ is a GAM. The basis functions $(b_\ell)_{\ell\in[L]}$ are illustrated in Fig. 1, highlighting that the knot density is much higher for shorter lags, while the knots become more spaced out for longer lags. This design enables the model to capture the large variability of the transfer functions at short lags, while still accounting for potentially long recessions. The feature vectors $z_{t'}$ used in GAMCR are the intensity of the precipitation event at time t', the weighted averages of both the past precipitation and the past evapotranspiration over different time windows, and the sine and cosine of the fractional year.

Since we model the functions $(g_\ell)_{\ell\in[L]}$ using GAMs, one can write

$$g_\ell(z_{t'}) = \boldsymbol{\xi}_{t'}^\top \boldsymbol{\gamma}_\ell, \tag{6}$$

where $(\boldsymbol{\xi}_{t'}^\top)_{t'}$ is the design matrix of the GAM. Each entry of $\boldsymbol{\xi}_{t'}$ corresponds to one of the spline basis functions evaluated at a given feature (i.e. a specific entry of $z_{t'}$). We have:

$$\begin{aligned} y_t &= \sum_{\ell=1}^{L}\left(\sum_{T=0}^{T_{\max}} x_{t'} b_\ell(T) \boldsymbol{\xi}_{t'}\right)^\top \\ \boldsymbol{\gamma}_\ell &= \sum_{\ell=1}^{L} \mathbf{W}_{t,\ell,:}\boldsymbol{\gamma}_\ell = \vec{\boldsymbol{w}}_t^\top \vec{\gamma}, \end{aligned} \tag{7}$$

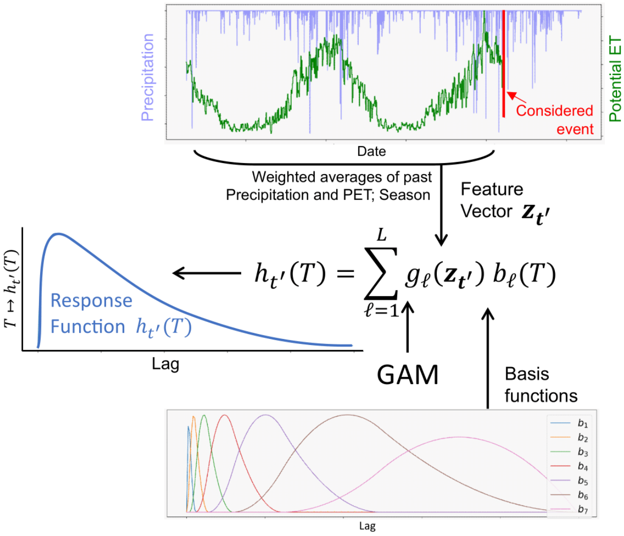

Figure 1. Overview of GAMCR. Given some precipitation event of interest occurring at time t', GAMCR computes a feature vector $z_{t'}$ including information on the system up to time t'. The response function is expressed as a weighted sum of spline basis functions, $(b_\ell)_{\ell\in[L]}$, where the weights are derived from $z_{t'}$ through L distinct Generalized Additive Models $(g_\ell(\cdot))_{\ell\in[L]}$.

where $\vec{\boldsymbol{\gamma}} := \mathrm{vec}(\boldsymbol{\gamma}_\ell,\ \ell\in[L])$ and $\vec{\boldsymbol{w}}_t := \mathrm{vec}(\mathbf{W}_{t,\ell,:},\ \ell\in[L])$ given by $\mathbf{W}\in\mathbb{R}^{n\times L\times d_s}$ such that:

$$\mathbf{W}_{t,\ell,:} = \sum_{T=0}^{T_{\max}} x_{t'} b_\ell(T)\boldsymbol{\xi}_{t'}. \tag{8}$$

Here, d_s is the number of features resulting from the GAM formulation.

2.3 Model training

With GAMs we can encode prior knowledge and control overfitting by using penalties and constraints during training. In our case, we consider two smoothness-inducing penalties. The first one promotes the smoothness of the functions g_ℓ by penalizing the second order derivative, as commonly done in the GAM literature (see Hastie et al., 2017). This penalty ensures that the coefficients of the transfer functions in the basis $(b_\ell)_\ell$ smoothly evolve with respect to the catchment features $z_{t'}$. This penalty acts on the L time-dependent coefficients of the transfer functions in the basis $(b_\ell)_\ell$ independently. The second regularization term promotes the smoothness of the transfer functions globally by adding a similar penalty on the model coefficients.

The final optimization problem considered is:

$$\min_{(\boldsymbol{\gamma}_\ell)_{\ell\in[L]}\geq 0} \frac{1}{n}\sum_{t=1}^{n}\Big(y_t - \sum_{\ell=1}^{L}\mathbf{W}_{t,\ell,:}\boldsymbol{\gamma}_\ell\Big)^2 + \lambda_1\sum_{\ell=1}^{L}\boldsymbol{\gamma}_\ell^\top \boldsymbol{P}_1\boldsymbol{\gamma}_\ell + \lambda_2\sum_{1\leq\ell,k\leq L}\frac{1}{n}\sum_{t'=1}^{n} (\boldsymbol{\xi}_{t'}^\top\boldsymbol{\gamma}_\ell)[\boldsymbol{P}_2]_{\ell,k}(\boldsymbol{\xi}_{t'}^\top\boldsymbol{\gamma}_k), \tag{9}$$

which can be equivalently written using a vectorized formulation as:

$$\min_{\vec{\boldsymbol{\gamma}}\geq 0}\frac{1}{n}\sum_{t=1}^{n}\big(y_t - \vec{\boldsymbol{w}}_t^{\,\top}\vec{\boldsymbol{\gamma}}\big)^2 + \vec{\boldsymbol{\gamma}}^{\,\top}\big[\lambda_1\boldsymbol{P}_1' + \lambda_2\boldsymbol{P}_2'\big]\vec{\boldsymbol{\gamma}}, \tag{10}$$

where, denoting by $\otimes$ the Kronecker product between two matrices we have defined

$$\boldsymbol{P}_1' := \mathrm{Id}_L\otimes\boldsymbol{P}_1,\ \text{and}\ \boldsymbol{P}_2' := \boldsymbol{P}_2\otimes\left(\frac{1}{n}\sum_{t'=1}^{n}\boldsymbol{\xi}_{t'}\boldsymbol{\xi}_{t'}^\top\right) \tag{11}$$

Provided that the hyperparameters λ_1 and λ_2 are not both zero, the optimization problem Eq. (10) has a strongly convex objective function with convex constraints. As a result, it admits a unique optimal solution, and the projected gradient descent algorithm is guaranteed to converge to this solution provided that the learning rate is set small enough (cf. Boyd, 2004). For all experiments, we used $\lambda_1 = 10^{-3}$ and $\lambda_2 = 1$ (we refer to Sect. 3.3 for further details). In practice, the parameters $\vec{\boldsymbol{\gamma}}$ are initialized by solving the unconstrained version of the problem, which involves computing the minimum L^2-norm solution via the pseudoinverse of a matrix. This initial solution is then projected onto the positive orthant, after which the projected gradient descent algorithm is applied. The learning rate starts at a large value (namely 10^{-1}) and is gradually and adaptively reduced throughout the iterations to ensure a strict decrease in training loss at each step.

The matrix $\mathbf{W}$ is precomputed offline prior to running the projected gradient descent algorithm, and parallel computation can be employed to obtain $\mathbf{W}$ quickly. This precomputation significantly accelerates the training process by eliminating redundant calculations.

2.4 Software GAMCR v1.0 description

The model developed in Sect. 2.2 has been implemented in the Python language as the software GAMCR v1.0. The code is publicly available on Zenodo (Duchemin, 2025), and a detailed online tutorial https://quentin-duchemin.github.io/GAMCR/tutorials/ (last access: 30 September 2025) is provided to guide users through usage and reproducibility steps. To use GAMCR, the user must provide time series of precipitation, streamflow, potential evapotranspiration and the corresponding dates and times at equally spaced time intervals. The software operates in a series of steps to ensure accurate and efficient analysis. First, users can use a predefined notebook to ensure that their data has the proper format (e.g. column names that conform to the software's requirements). Next, a predefined script is run to perform key precomputations, including the calculation of the matrix $\mathbf{W}$, which significantly enhance the efficiency of the model training process. These precomputations are completed within a few seconds to a few minutes on a standard laptop for a decade of hourly data. Once these precomputations are completed, users can proceed to train the GAMCR model on their dataset. Let us stress once again that the number of basis functions L used by the model is automatically computed based on the maximum lag $T_{\max}$. With $T_{\max} = 5$, 10 or 15 d the model uses 6, 7 or 8 basis functions, respectively. After the model has been trained, users can launch another predefined script to compute key statistics of interest, such as the NRFs over predefined ensembles (such as different precipitation quantiles) and the RRD. These results are automatically saved for further analysis. A detailed tutorial is provided in the online documentation of the GAMCR package, where users can reproduce the results of this paper for the Euthal catchment. The tutorial offers a step-by-step explanation of each stage, equipping users with the necessary tools to apply GAMCR effectively to their own datasets. Overall, GAMCR can be efficiently used on personal laptops, with model training on 20 years of hourly data typically taking around 30 min for $T_{\max} = 10$ d.

3 Model testing

Developing strategies to rigorously quantify the performance of trained machine learning models is essential. In the case of the hydrologic response, the evaluation step is particularly important because impulse response functions cannot be measured directly, and the model is trained on streamflow data only (Gupta et al., 2008; McDonnell and Beven, 2014; Kirchner, 2022).

Below, we describe two datasets that serve two different purposes. A synthetic dataset (Sect. 3.1) is used to validate the model, because the estimated response can be compared against the benchmark "ground truth" response, which is exactly known (unlike in real-world catchments). An observed dataset (Sect. 3.2), which includes measurements from six diverse catchments across Switzerland, is used to showcase how the model can be used to estimate the hydrologic response at different locations.

3.1 Synthetic data and model validation strategy

3.1.1 Generation of Synthetic Data

The synthetic dataset was generated using precipitation and air temperature measurements available from the Federal Office of Meteorology and Climatology (MeteoSwiss) for the station of Lugano, along with streamflow data from the nearby gauging station Chiasso, Ponte di Polenta, on the Breggia River. These measurements were used to calibrate a lumped nonlinear and nonstationary conceptual model (Sect. S2), allowing us to create a synthetic streamflow time series (40 years at hourly resolution) that realistically reflects observed dynamics (case A) without aiming for exact replication. To explore different hydrological responses, we adjusted the model parameters to represent both a more damped (case B) and a more flashy (case C) hydrologic system. By working with these synthetic yet realistic datasets, we can rigorously assess the model's performance, because the underlying mechanisms are exactly known and the data are free from disturbances such as dams, hydropeaking, or leakages. Details of the approach employed in the model and the parameters used are provided in Sect. S2.

The generated synthetic time series are shown in Fig. 2 over an example 4-month period. The figure shows clearly that, compared to the reference streamflow series (case A), the damped series (case B) has lower peaks and longer recessions, while the flashy series (case C) has higher peaks and similar recessions. The data also clearly show the nonlinearity and nonstationarity of hydrologic systems, as some precipitation events cause almost no streamflow response (e.g. in June 2010) while others may cause a sharp response (e.g. in late August 2010).

To compute the response functions for the synthetic data (ground-truth response), we simply ran the lumped hydrological model as many time as there were time steps with nonzero precipitation. In every simulation, we masked a different time step by setting its precipitation to zero. The hydrologic response to precipitation occurring on a specific time step was then computed as the difference between the modeled series with and without precipitation over that time step. This approach provides responses for each event individually, which can be aggregated to compute ensemble responses over e.g. particular periods, precipitation events or antecedent conditions.

3.1.2 Model validation strategy

For each of the three synthetic study cases, GAMCR was trained, and hydrologic responses were computed from the trained model as NRFs aggregated over six precipitation-intensity quantiles. Validation was performed by comparing the GAMCR predictions the corresponding aggregated ground-truth NRFs from the synthetic datasets. We also compared our estimates with those derived from ERRA. Detailed results of the validation are presented in Sect. 4.1.

3.2 Real-world data

We compiled a 15-year record (2005–2019) of measured, hourly precipitation-runoff data from six Swiss watersheds (Fig. 3). The data from the first 13 years were used for model training, while the final two years of data (2018

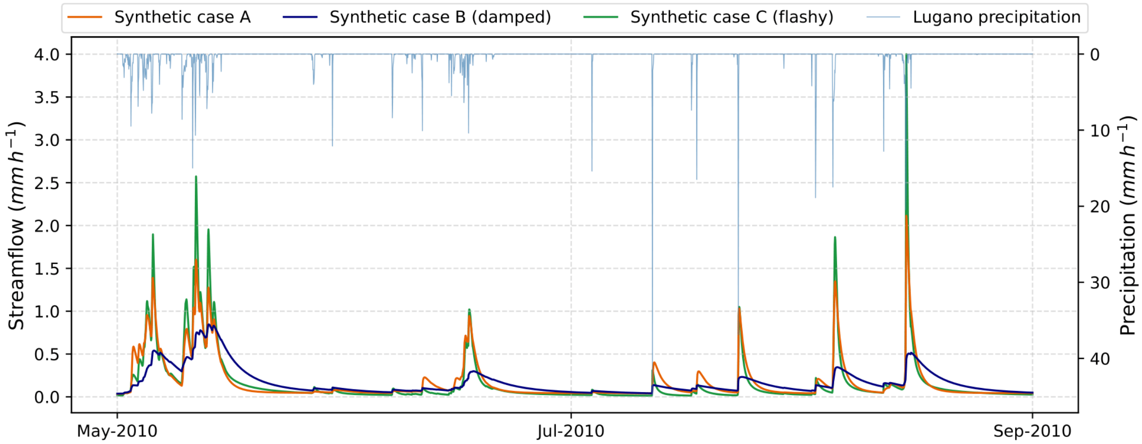

Figure 2. Example of the synthetic streamflow time series (for snow-free months in 2010): case A is the reference (orange curve); case B is a more damped response (purple curve), with lower peaks and longer recessions; case C is a flashier response with higher peaks (green curve). Additionally, the Lugano precipitation time series is shown (light blue curve) with an inverted y axis for comparison.

and 2019) served as the test period. Streamflow time series were provided by the Federal Office for the Environment (FOEN). Precipitation data were sourced from the 'CombiPrecip' product, developed by MeteoSwiss (MeteoSwiss CombiPrecip, 2005). Potential evapotranspiration time series were computed based on air temperatures provided by MeteoSwiss, through the Hargreaves method from Hargreaves and Samani (1985) (implemented through the Python Pyeto package https://github.com/woodcrafty/PyETo, last access: 30 September 2025) and then uniformly distributed across each day at hourly intervals. At each site we also extracted key catchment attributes and hydrological statistics (Tables 1 and 2) and computed the mean monthly precipitation, streamflow, potential evapotranspiration over the study period (Fig. 4), and the flow duration curve for the snow-free period (Fig. 5). We selected these sites because of their diversity in hydrological regimes, elevation and soil depths, which we expect will be reflected in substantially different hydrologic responses. Additional criteria included minimal glacier influence, natural flow regimes (no dams or major abstractions), and complete, reliable data records. The sites have comparable size (between 34–185 km^2), which classifies them as small to small/medium basins. Further catchment characteristic analyses appear in Sect. S1 in the Supplement.

As snow introduces complexities in catchment response, such as delayed runoff generation and temperature-driven melt rates, we focused our analysis on snow-free periods only. We considered as snow-free periods the months from May to October, inclusive, except for two basins at the highest altitudes (Euthal and Lavertezzo, with maximum and mean elevations above 2200 and 1300 m a.s.l., respectively), for which we assumed that first snow-free month is June (Fig. 4).

Figure 5 shows the Flow Duration Curves (FDCs) of streamflow during the snow-free months. The curves highlight clear contrasts among catchments: Chiasso and Salmsach are dominated by low flows typical of pluvial regimes, while Lavertezzo and Euthal display broader distributions with higher discharges, reflecting steep topography and rapid runoff. Magliaso and Sonceboz fall in between, balancing frequent low flows with moderate events. These patterns confirm that the selected basins capture a meaningful gradient of hydrological responses, even when snow-driven processes are excluded.

Key hydrological statistics (Table 2) further quantify these differences. Q10–Q90 ranges emphasize the variability between rainfall-dominated and more responsive basins, with Chiasso and Salmsach showing low median flows and narrow ranges, and Lavertezzo and Euthal displaying much higher values. Runoff ratios vary widely (0.30–0.82), while the center of timing (CT) spans from day 184 to 221, marking earlier peaks in pluvial catchments and later ones in higher-elevation sites. Together, these indicators confirm the diversity of regimes and support the suitability of these basins for testing the proposed methodology.

3.3 Implementation details

While our model is designed to estimate the hydrologic response to each precipitation event, we are primarily interested in the model's ability to reproduce the ensemble responses (RRD or NRF) over given conditions of precipitation intensity or antecedent wetness. Therefore, the model will be tested over ensemble responses. This also offers the opportunity to estimate the hydrologic response – and its main statistics – with ERRA and assess the consistency between GAMCR and ERRA.

We tested the need for optimization of the hyperparameters λ_1 and λ_2 through initial (and computationally expen-

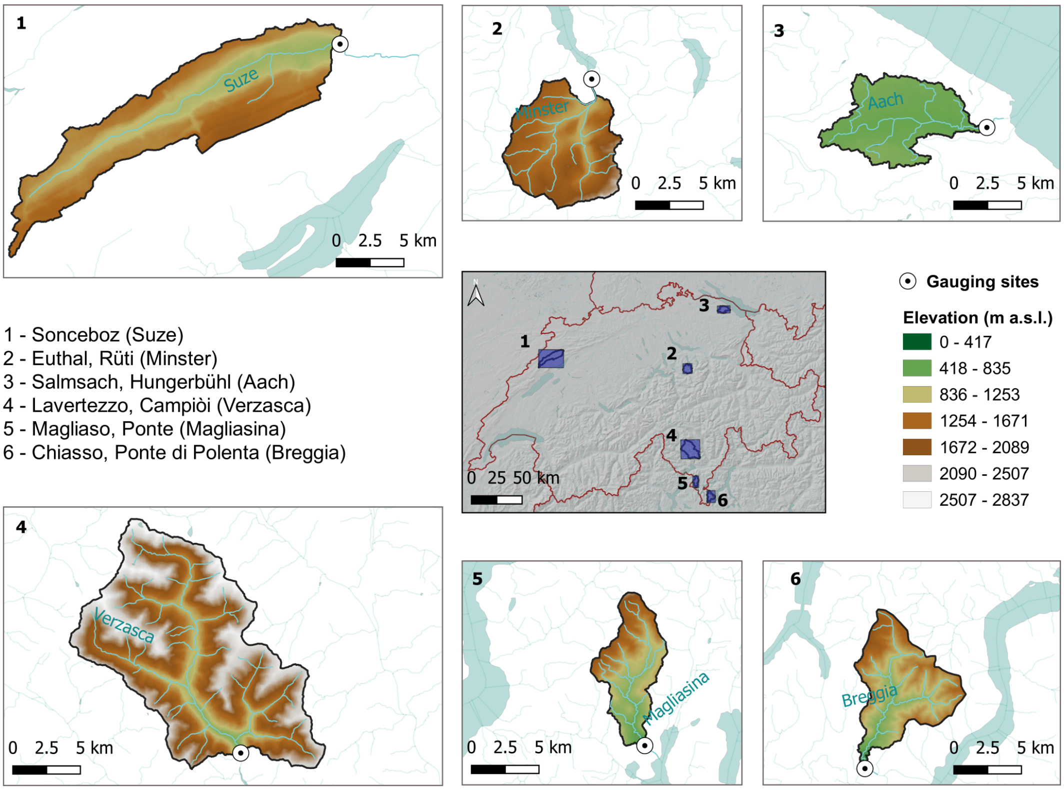

Figure 3. Map of Switzerland showing the six catchments analyzed, along with their corresponding gauging stations (listed on the left with the river names in brackets). Each catchment is displayed in a separate plot for a detailed view of its dimensions and elevation ranges. Numbers mark the catchments' locations within Switzerland and can be seen on the map in the center. The sixth gauging station (Chiasso, Ponte di Polenta) also provided the streamflow time series used to create the synthetic dataset with precipitation data from Lugano.

Table 1. Overview of the gauging stations and their catchment features: associated river; mean elevation; mean slope; area; mean soil depth; mean permeability.

Station	River	Mean Alt. (m a.s.l.)	Mean Slope (°)	Area (km^2)	Mean soil depth (cm)	Mean permeability (cm d^{-1})
Sonceboz	Suze	1036.95	14.77	127.25	49.12	72.09
Euthal, Rüti	Minster	1346.91	22.20	59.13	33.87	54.27
Salmsach, Hungerbühl	Aach	472.47	3.42	47.38	69.57	52.22
Lavertezzo, Campiòi	Verzasca	1655.71	38.49	185.12	17.16	81.56
Magliaso	Magliasina	928.25	28.44	34.38	29.44	96.77
Chiasso, Ponte di Polenta	Breggia	934.40	33.21	47.10	20.87	75.98

sive) cross-validation experiments. Since we obtained only minor improvements over the default values $\lambda_1 = 10^{-3}$, $\lambda_2 = 1$, we consistently used the defaults across all applications. Since we are only interested in the evaluation of the hydrologic response up to a few days after precipitation, we kept the hyperparameter $T_{\max} = 24 \times 10\,\mathrm{h}$. The positions of the knots to get the B-splines basis functions b_ℓ follow an exponentially increasing sequence, starting at 0 with an initial step of 1. After each step, the step size doubles, leading to a pattern where knots are densely spaced at the beginning and become increasingly sparse as values grow. Following this procedure, the value of $T_{\max}$ automatically sets the number of basis functions to $L = 7$ in our case. All models were trained for the full set of predefined epochs using gradient descent with adaptive learning rate, ensuring a strictly decreasing training loss.

In the observed data series, the response to very small rainfall events may be easily hidden by measurement noise and other processes. While these events are not particularly relevant for the hydrologic response, they may corrupt the training phase. Hence it is convenient to set a precipitation in-

Table 2. Overview of the gauging stations, their hydrological regime type, and key hydrological statistics: snow-free months considered in the study; 10th, 50th, and 90th percentiles of snow-free streamflow (Q10, Q50, Q90); runoff ratio (RR); and center of timing (CT, expressed as the Julian day when 50 % of the snow-free discharge has occurred)

Station	Regime type	Snow-free months	Q10 ($\mathrm{mm\,h^{-1}}$)	Q50 ($\mathrm{mm\,h^{-1}}$)	Q90 ($\mathrm{mm\,h^{-1}}$)	RR (–)	CT (day of year)
Sonceboz	Jura-nivopluvial	May–Oct	0.020	0.053	0.174	0.437	194
Euthal, Rüti	Transition nival	Jun–Oct	0.025	0.077	0.435	0.808	221
Salmsach, Hungerbühl	Pluvial	May–Oct	0.009	0.021	0.092	0.317	211
Lavertezzo, Campiòi	Southern nivo-pluvial	Jun–Oct	0.051	0.106	0.429	0.823	218
Magliaso	Southern pluvio-nival	May–Oct	0.018	0.056	0.211	0.524	187
Chiasso, Ponte di Polenta	Southern pluvio-nival	May–Oct	0.003	0.014	0.111	0.296	184

tensity threshold J_{th} and train the model only for events that exceed J_{th}. We trained GAMCR using $J_{\mathrm{th}} = 0.05\,\mathrm{mm\,h^{-1}}$.

The results from ERRA were obtained using the R scripts accessible at the following repository: https://doi.org/10.16904/envidat.529, as specified in Kirchner (2024). The RRD curves were computed considering a maximum lag of 40 h. Initial estimates of precipitation bins were automatically generated by the algorithm, invoking six approximately even ranges, while ensuring a minimum threshold of 40 nonzero values in each precipitation bin. To improve comparability across models, the same precipitation ensembles were used to average the true transfer functions and the GAMCR estimates. Using a coarser input data resolution is beneficial to ERRA when the hydrologic response is long relative to the input temporal resolution (because in such cases, it can be difficult to separate the overprinted effects of input signals at closely spaced lag times). Using a coarser time step helps clarify these impacts. For this reason, after some initial testing, the flashy, base, and damped synthetic input time series are aggregated into 2, 3, and 6 h time steps, respectively.

4 Results

4.1 Model validation

The model was trained (Sect. 2.3) on the synthetic data (Sect. 3.1), which consists of three cases: the reference response, a flashier response, and a more damped response. We validate the model by first computing the hydrologic response in the form of NRF over six quantiles of precipitation intensity, and comparing it against the ERRA estimates and the benchmark generated directly from the model Fig. 6.

As a result of the different aggregation of the input time series for the three synthetic data sets, their precipitation intensities (and thus the bins used in Fig. 6a–c) appear different, although the original hourly input data are the same.

Figure 6 shows that GAMCR accurately estimates the transfer functions on synthetic data, particularly in the flashy and damped scenarios, where their curves nearly overlap with the benchmark. In the base case, (panel a) the peak value and tail of the response are well captured, but the peak timing is systematically early compared to the benchmark. Overall, in these three cases characterized by very different responses (Fig. 6d) ERRA and GAMCR provide generally consistent estimates.

We also computed the peak height, peak lag and runoff volume of the NRF, and explored their relationship with precipitation intensity (Fig. 7). The results highlight GAMCR's ability to accurately estimate key quantities related to the magnitude of the catchment's response (peak height and runoff volume). These statistics are also very consistent with those estimated by ERRA. For both approaches, the flashy case remains the most sensitive for estimation, with GAMCR underestimating runoff volume for intermediate precipitation values (from 10 to 25 $\mathrm{mm\,h^{-1}}$) but accurately capturing peak height. In the base case, GAMCR slightly underestimates runoff volume while maintaining accurate peak height estimates. For the damped scenario, it closely matches ground truth values for peak height and produces nearly overlapping runoff volume estimates. Overall, both approaches show a strong consistency in their peak height and runoff volume estimates across different scenarios.

Despite the models' strong performance in estimating the magnitude of the catchment response, both face challenges in predicting peak lag, though in opposite ways. ERRA tends to produce more variability across the NRFs (as shown by dashed lines with triangles in all the right panels). By contrast, GAMCR tends to produce lag values that are much less variable than the benchmark across different precipitation ranges.

4.2 Estimation of catchment hydrological responses

When applying GAMCR to observed data it is not possible to validate its accuracy in estimating hydrologic response, because the true response is not known. However, it is instructive to compare the modeled vs measured streamflow series, for both the training and test periods. Since the model was not developed for the purpose of reproducing streamflow, its performance should not be compared to hydrologic models

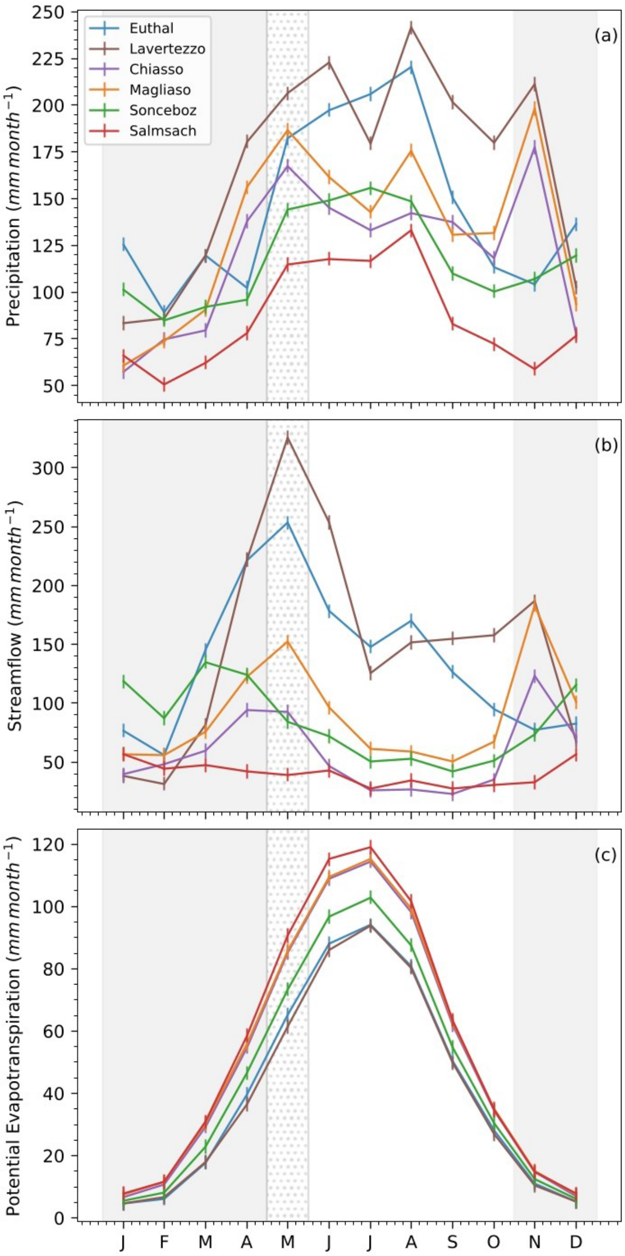

Figure 4. Hydrological regimes, in terms of monthly mean precipitation **(a)**, streamflow **(b)**, and potential evapotranspiration **(c)** for the six sites. The time series were averaged over the complete period of study (2005–2019). The light grey shadowed areas indicate what we considered as snowy periods with potential snow-melt effects on streamflow, including also May for Lavertezzo and Euthal (dotted grey shadowed band).

that are designed to maximize fit, but the simulated hydrograph serves as a valuable diagnostic tool. For example, periods where the modeled hydrograph deviates significantly from the measurements could be flagged as unreliable and excluded from the analysis. To support this evaluation, Fig. 8 shows the streamflow predictions generated by GAMCR.

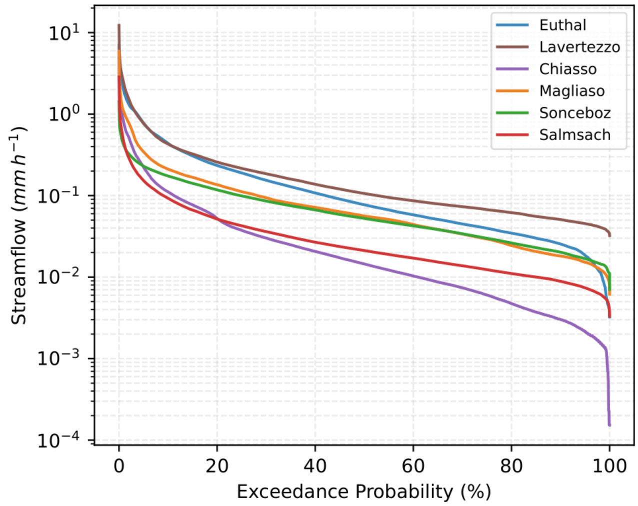

Figure 5. Flow Duration Curves (FDCs) for the six sites during the snow-free period (June–October for Euthal and Lavertezzo and May–October for the other basins).

The agreement between training and test set results varies across basins. Sonceboz and Magliaso show the strongest consistency, with only minor over- or underestimations. Lavertezzo follows a similar pattern but has some overestimated values on training points. Salmsach and Chiasso, by contrast exhibit considerable dispersion and overestimation on the test set, suggesting lower predictive performance. Performance at Euthal is intermediate between these two groups, with overestimation of low test streamflow values. These results suggest that the performance of GAMCR in reproducing streamflow is not directly correlated with the hydrological characteristics of the basins. This is even more visible when looking at the model performance aggregated over streamflow quantiles (Fig. S2), where the fit is consistently good across sites and only a minor underestimation of the lowest flow conditions stands out. Timeseries plots for the test period (Fig. S3) indicate that the temporal dynamics of the predicted hydrograph are appropriate and there are no periods that should be flagged and removed from the analysis.

Figure 9 presents the weighted average RRDs and the peak heights of the NRFs estimated by ERRA and GAMCR for the six sites in the observed dataset. Computations consider all events whose precipitation intensity exceeds $0.5\,\mathrm{mm\,h^{-1}}$. The differences in basin response are also evident in the range of runoff coefficients derived from ERRA, which vary from 0.10 in Sonceboz to 0.64 in Euthal. The results align well with the hydrological regimes and characteristics of the basins (see Table 1 and Figs. 3 and 4). The Sonceboz basin, in particular, shows a very flat runoff-response distribution, which is consistent with the relatively low mean slope, large area, and elongated shape of its basin. These features, along with its moderate permeability and location in the Jura's pluvio-nival region, contribute to the basin's

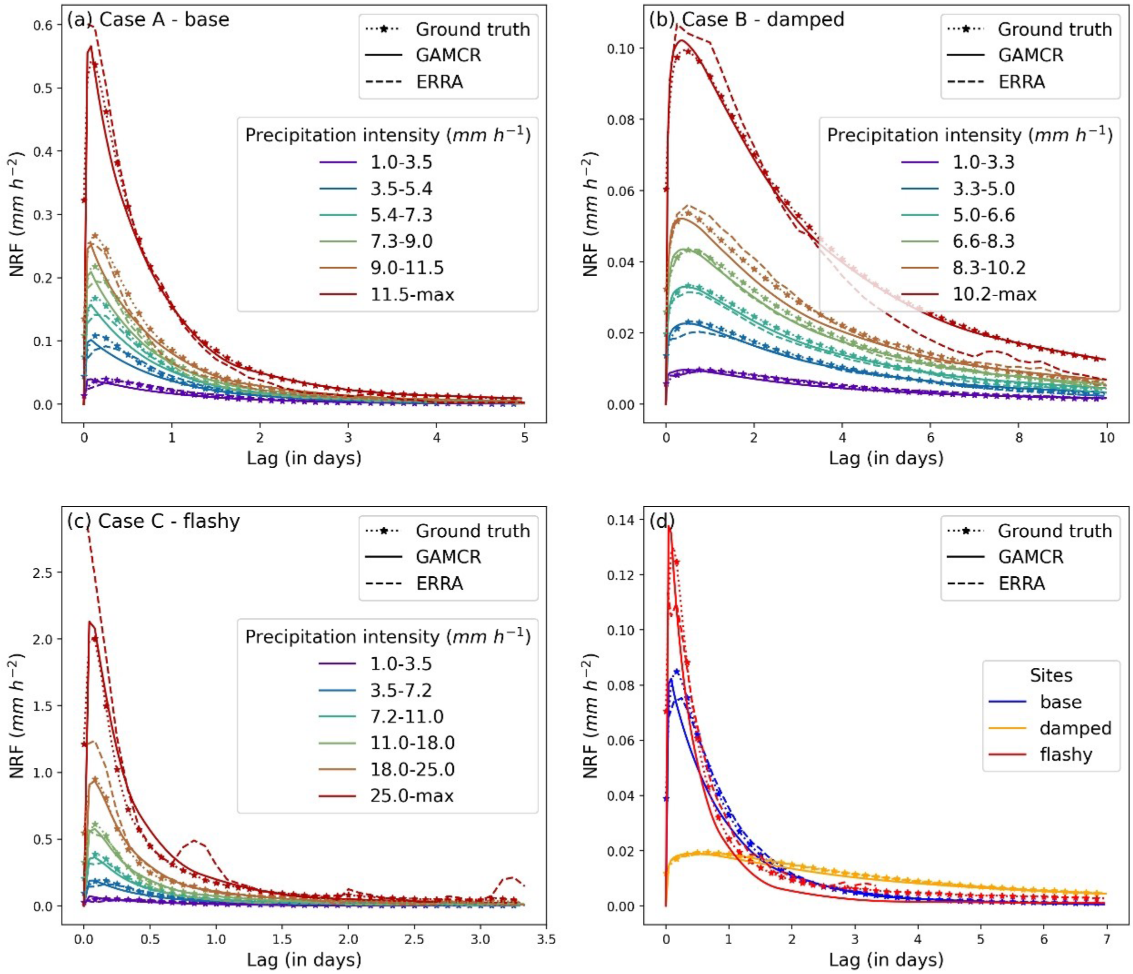

Figure 6. NRFs averaged across different precipitation intensity ensembles from GAMCR, ERRA and the ground truth, for the base **(a)**, damped **(b)**, and flashy **(c)** synthetic time series. Readers should note the different scales of NRFs between flashy, base and damped scenarios. Panel **(d)** combines the overall average NRFs for the three cases in a single plot .

very damped runoff response, reflected in the low runoff coefficient (0.10). A slightly flashier response is observed in the Salmsach basin, which has low mean slope, low permeability, and a pluvial hydrological regime. This results in a damped response, though less damped than Sonceboz's, consistent with its higher runoff coefficient (0.19). The Chiasso and Magliaso basins exhibit similar peak values, but with different response shapes. Despite similarities in altitude and mean slope, Chiasso is larger than Magliaso and has lower permeability, consistent with the larger area under its RRD (i.e. the runoff coefficient, 0.26 versus 0.23 for Magliaso). The flashier response in Magliaso is consistent with its high mean slope, in common with Lavertezzo and Euthal, where the flashiest responses (at least for GAMCR) are observed. Lavertezzo and Euthal are characterized by the highest altitudes, highest annual precipitation and lowest annual potential evapotranspiration. The higher RRD peak for Euthal compared to Lavertezzo is consistent with the lower permeability in the Euthal basin and is also reflected in its larger runoff coefficient (0.64 versus 0.40). Overall, the weighted average RRDs provided by both GAMCR and ERRA are broadly consistent with the distinctive characteristics of each basin. The results shown in Fig. 9 demonstrate the general consistency between the GAMCR and ERRA approaches across the different sites. Only two basins exhibit some discrepancy: the Chiasso and Lavertezzo basins (purple and brown curves, respectively). In both these cases, GAMCR estimates a more pronounced RRD peak than ERRA within the first 7 h, and a slightly lower tail after 10 h. The estimated NRF peaks for different precipitation intensities for these sites (Fig. 9d and f) are consistent between ERRA and GAMCR for most precipitation bins, but deviate slightly for the highest one. Overall, the responses estimated by GAMCR and ERRA are broadly similar, and since the models work very differently, consistency in their estimates increases our confidence in both approaches.

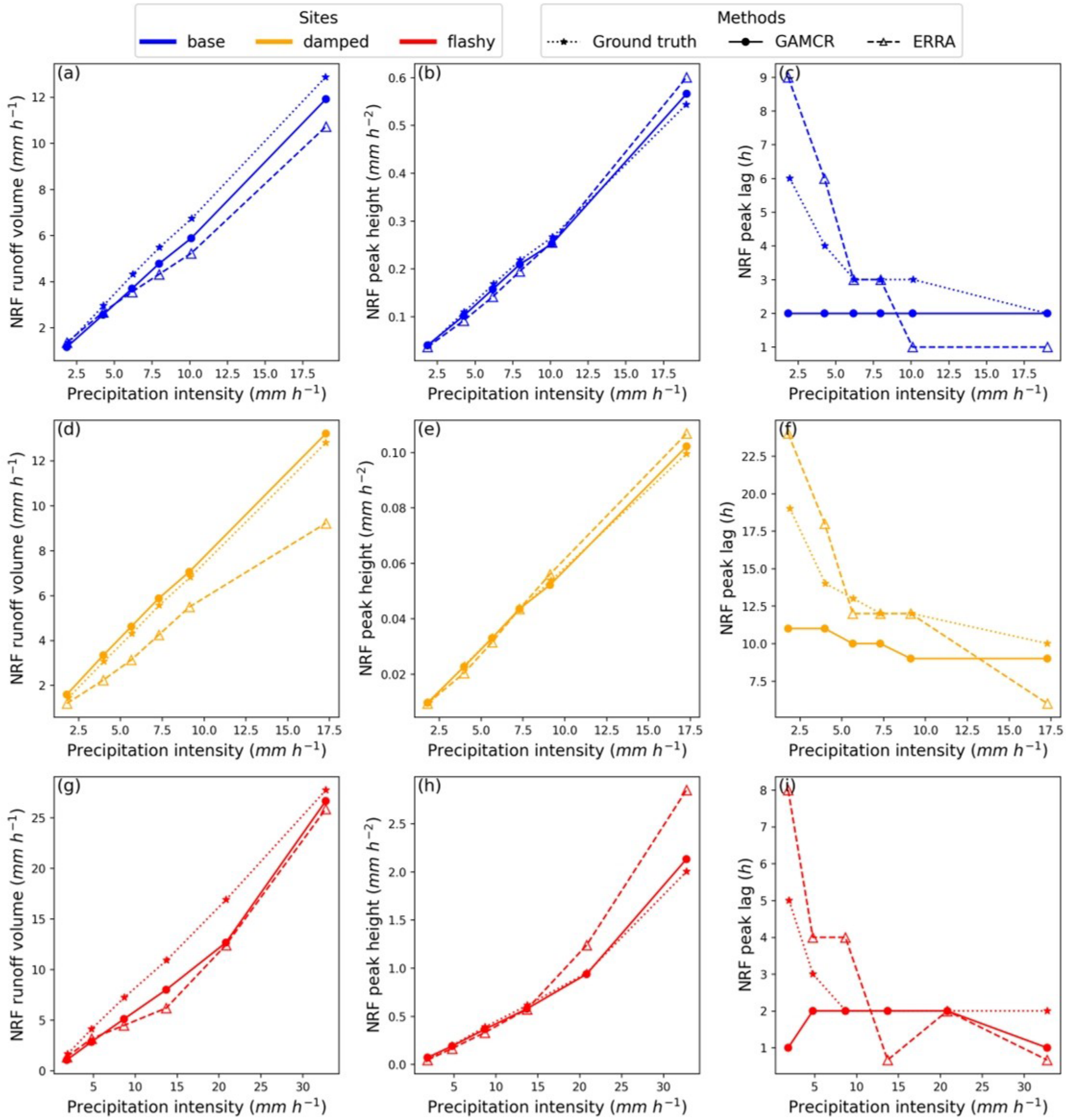

Figure 7. Different statistics computed on NRFs obtained from either GAMCR, ERRA, or the ground truth averaged across different precipitation intensity ensembles for the flashy, base, and damped datasets. Figures **(a)**, **(b)** and **(c)** respectively depict the NRF runoff volume, NRF peak height and the peak lag.

4.3 Effects of precipitation intensity and antecedent wetness

GAMCR can be used to investigate how variations in precipitation intensity and antecedent wetness affect the hydrologic response. Here we explore such effects at the six study sites. To characterize precipitation intensity, we use the same six precipitation intervals defined in Sect. 4.2 above. As a proxy for antecedent wetness, we use the values of streamflow during the timestep prior to the precipitation event under consideration, which we separate into five ranges. We then aggregate the individual response (RRD curves) over each class of precipitation intensity or antecedent wetness. Results are shown in Fig. 10, where we plot the RRD peak height (not to be confused with the peak of the NRF shown in Fig. 9) against precipitation intensity and antecedent wetness.

As Fig. 10a shows, the RRD peak heights do not vary systematically with precipitation intensity. By contrast, Fig. 10b demonstrates clear increasing trends in RRD peak heights with increasing antecedent wetness. Nearly all sites exhibit at least a threefold increase in peak heights across antecedent wetness levels, with the exception of Lavertezzo, which shows a rise in peak heights just for only the last two bins of antecedent wetness. Chiasso, in particular, displays the high-

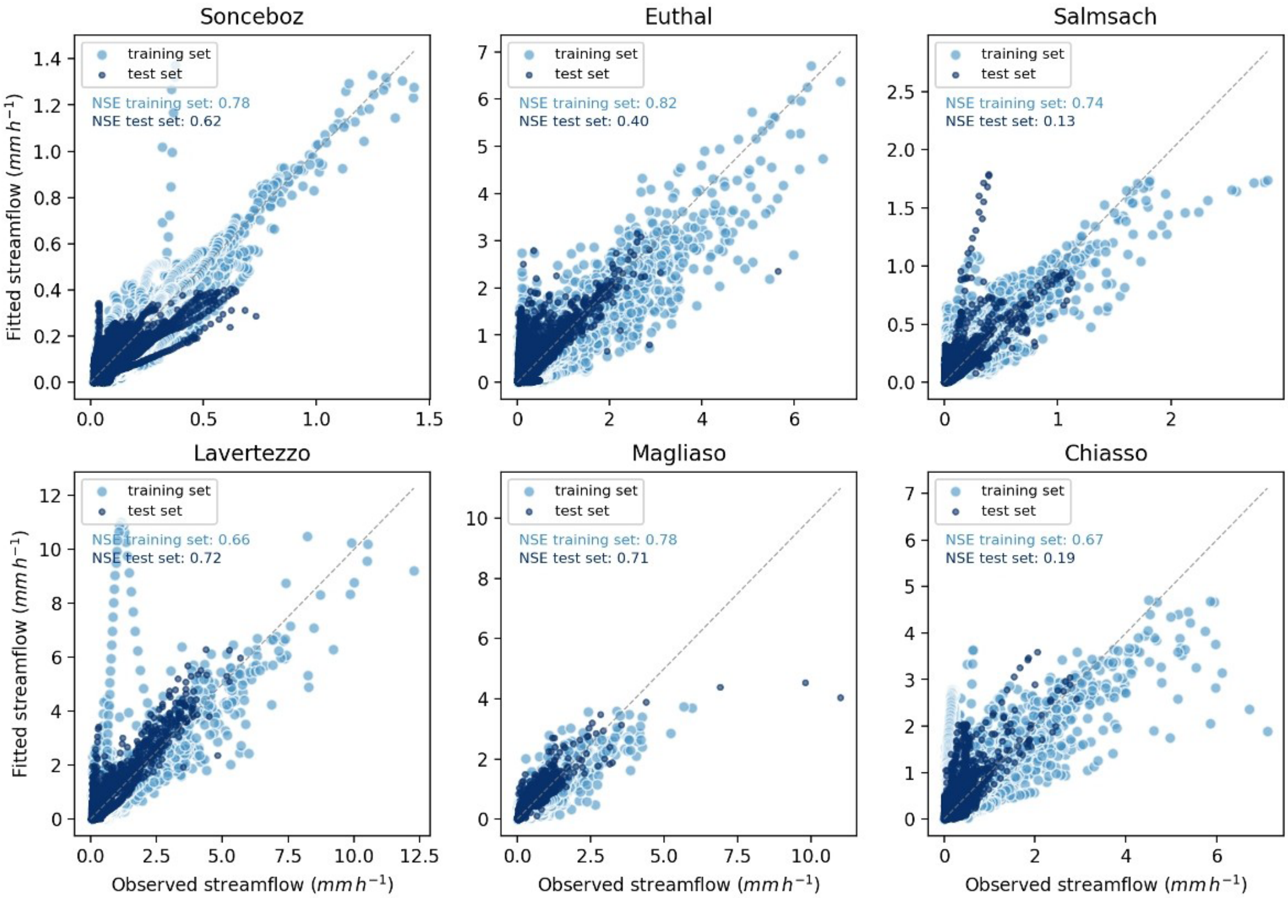

Figure 8. Fitted streamflow estimated through GAMCR for the six investigated sites. The larger light blue dots show measured and fitted discharges during the 2005–2017 snow-free seasons (training set); The smaller dark blue dots indicate measured and predicted discharges during 2018 and 2019 snow-free seasons (test set). 1 : 1 lines are shown in grey.

est variability, with peak heights spanning almost an order of magnitude (from 0.006 to $0.05\,h^{-1}$). Notably, for each site, the highest two antecedent wetness levels are widely separated, leading to a marked increase in RRD peak heights. These findings highlight a clear nonstationary response of the six catchments, strongly influenced by their antecedent wetness.

5 Discussion and Conclusions

We introduced a model based on GAMs to estimate the hydrologic response of watersheds based on precipitation-runoff data. The model was validated against three benchmark synthetic datasets and showed excellent agreement with the response curves of the underlying benchmark model, based only on its input and output time series (Fig. 6). While the accurate reproduction of the individual responses goes beyond the scope of the model, the ensemble responses (RRD and NRF curves) proved accurate. Closer inspection of the statistics of the responses (Fig. 7) showed that GAMCR accurately estimated NRF peak height and volume across different precipitation bins. By contrast, the timing of the NRF peak was generally not very accurate, with GAMCR systematically underestimating the peak lag. While this behavior can likely be improved through a different organization of the basis functions that form the core of the response (Sect. 2.2), GAMCR should currently not be used to estimate the timing of the hydrologic response. Comparisons between GAMCR and ERRA highlight that these two models, despite their very different architectures, provide similar hydrologic responses that closely match the (synthetic) ground truth.

Additionally, we analyzed the runoff response during snow-free periods for six Swiss catchments with diverse climatic and physical characteristics (Sect. 4.2). As the hydrologic response of a catchment is not directly measurable, verifying it in physical basins is challenging. Among the diagnostic tools that help build confidence on the results (beyond the benchmark tests of Sect. 4.1), we verified that the modeled streamflow was generally realistic for both training and test data (Fig. 8) and compared GAMCR's RRD and NRF statistics with those obtained from ERRA (Fig. 9). GAMCR

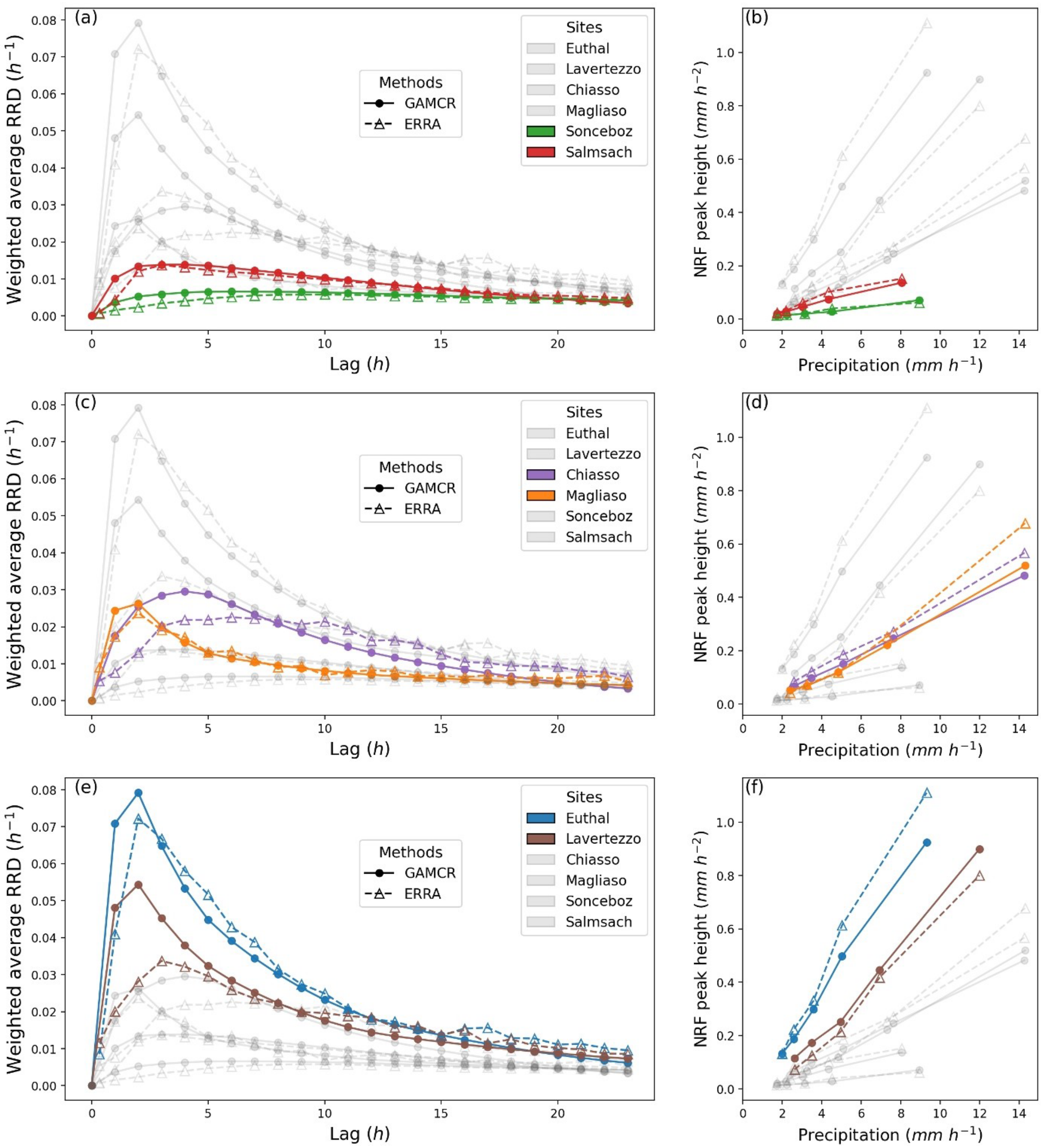

Figure 9. Hydrologic responses and their relationship with precipitation intensity for GAMCR and ERRA. Panels **(a)**, **(c)** and **(e)**: Weighted average RRD, where we keep time points with precipitation intensity above $0.5\,\mathrm{mm\,h^{-1}}$. Panel **(b)**, **(d)** and **(f)**: NRF peak heights against precipitation intensity.

produced results that were closely aligned with ERRA and consistent with the properties of the catchments. For example, the Salmsach catchment, with flatter topography and deeper soils than Euthal, had a slower and less marked average response to rainfall. We conclude that GAMCR is a robust tool to study runoff response behavior in catchments. As such, it enables advanced data-based analyses such as quantifying the effects of precipitation intensity and antecedent wetness on the average response peak (Fig. 10).

Since we have often referred to the ERRA approach (Kirchner, 2024) in our analyses, it is worthwhile to clarify the differences and similarities between ERRA and GAMCR. Both methods aim to estimate runoff response to precipitation based on time series data, and they both can

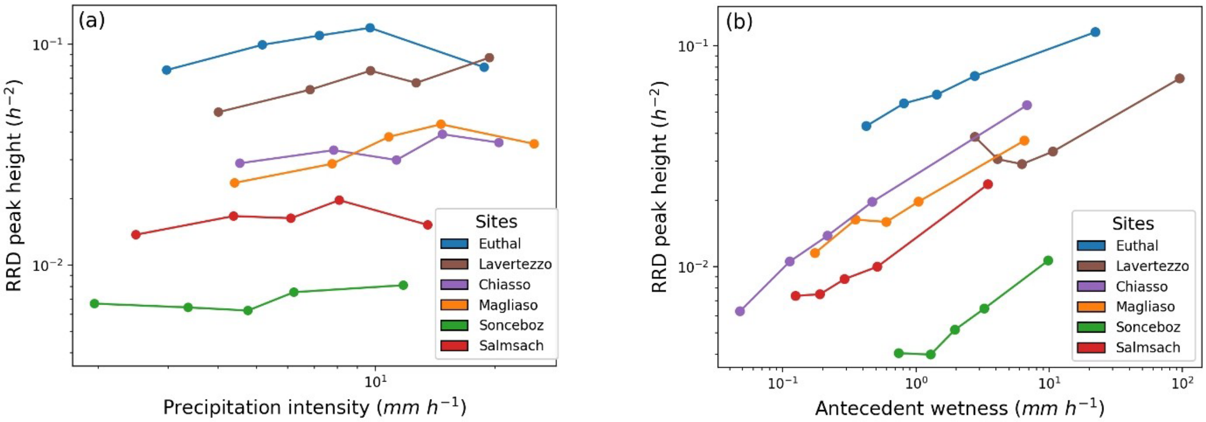

Figure 10. Peak height of RRDs when stratifying with respect to precipitation intensity **(a)** or antecedent wetness **(b)**.

quantify nonlinear and nonstationary runoff responses to precipitation. Both methods also implicitly assume that precipitation intensity and catchment conditions are the main controls on the catchment response. However, the two approaches achieve their (common) objective in radically different ways. ERRA fundamentally works on ensemble responses rather than single events. It extracts information from the entire precipitation-runoff time series or from portions of it that are selected to investigate different periods or conditions (provided that each portion has sufficient data). In contrast, GAMCR estimates the hydrologic response to each individual precipitation event using combinations of spline basis functions, with coefficients determined through machine learning techniques (Sect. 2.2–2.3). These individual responses can then be aggregated to ensemble responses. These different starting points result in different ways to run the models. GAMCR is based on a single training phase to estimate all the responses. Then, users can simply aggregate such responses in various ways as a post-processing phase. Instead, ERRA runs instantly but any sub-setting of the time series (for periods or conditions of particular interest, for example) needs to be specified a priori and the code is re-run for each analysis. The way ERRA and GAMCR are parameterized limits the types of transfer functions they can estimate, embedding specific assumptions about their shape. ERRA produces piecewise linear transfer functions, which might take negative values, especially when the water balance in the data is not maintained. In contrast, GAMCR ensures strictly positive transfer functions and promotes smoothness by relying on smooth basis functions. Other minor operational differences between ERRA and GAMCR include the potential need to aggregate the temporal resolution of the data to improve the estimate (ERRA) and the need for potential evapotranspiration series, along with precipitation and runoff (GAMCR, although in the absence of potential evapotranspiration data the user may simply change the default set of features of GAMCR by removing the ones based on PET). Finally, ERRA not only estimates statistics and responses but also quantifies their uncertainty through standard errors. In contrast, GAMCR currently lacks this capability, highlighting the need for future integration of uncertainty quantification.

These first applications of GAMCR to synthetic and observed data help us identify some current model limitations and encourage further model development. While the peak value and area of the aggregated response functions proved accurate, the timing of the response was not, systematically underestimating the peak lag (Fig. 7). Predicting peak lag is statistically challenging because small changes in lag often cause only minor variations in discharge, making estimation difficult. While larger training datasets can help, model architecture – particularly the choice of basis functions – is crucial. GAMCR's reliance on unimodal basis functions (namely B-splines with irregular knots) may introduce ambiguity, as dense placement at short lags can bias peak selection due to identifiability issues. Future work could focus on learning basis functions directly from data under suitable constraints, allowing models to adapt flexibly to watershed characteristics while preserving meaningful inductive biases.

We also stress that while the model simulates the response to every time step with precipitation, it was only evaluated on its capacity to reproduce ensemble behavior and so we do not recommend using it to evaluate individual responses or predict streamflow time series. Additional features that could be implemented in the future include an uncertainty estimation tool capable of providing accurate uncertainty bounds in the response, and the opportunity to integrate additional data (e.g. soil moisture series from sensors or remote sensing products) while training the model.

Catchments' capacity to mobilize water after storm events is a distinctive feature that is relevant for water resources management and useful to characterize catchment behavior. Quantifying the runoff response to precipitation using data-driven approaches is challenging due to the nonlinear and nonstationary nature of streamflow generation processes. GAMCR addresses these challenges by introducing a robust and flexible framework that leverages spline basis functions and Generalized Additive Models to learn the model's time-variable coefficients. Overall, GAMCR is a modern and effective tool for using the increasingly available rainfall-runoff series to investigate controls on hydrologic responses. Although testing and validation were performed in Swiss catchments, the basins studied span a gradient of physical properties and hydrologic regimes that result in different hydrologic responses. Broader testing of the approach in other climatic regions is a key next step for future research.

Code and data availability. The GAMCR code is available on the Zenodo archive: Duchemin (2025) with https://doi.org/10.5281/zenodo.15180816. Both the synthetic and real data are available on the Zenodo archive: Duchemin et al. (2025) with DOI https://doi.org/10.5281/zenodo.15180911. All the material is published on the FAIR-compliant Zenodo repository: https://doi.org/10.5281/zenodo.15180911 (Duchemin et al., 2025).

Supplement. The supplement related to this article is available online at https://doi.org/10.5194/gmd-18-8663-2025-supplement.

Author contributions. QD developed the GAMCR model and performed the GAMCR computations. MGZ processed all the observational data and performed the ERRA computations. MGF and HS gathered the observed data. GO supervised the model development. PB and JWK conceived the project and supervised this work. JWK developed the ERRA method and the benchmark model. All authors discussed the results and contributed to the final manuscript.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

Financial support. This research has been supported by the Swiss Data Science Center Fifth Call for Data Science Projects, project C21-09).

Review statement. This paper was edited by Charles Onyutha and reviewed by two anonymous referees.

References

Arridge, S., Maass, P., Öktem, O., and Schönlieb, C.-B.: Solving inverse problems using data-driven models, Acta Numerica, 28, 1–174, 2019.

Beven, K.: How far can we go in distributed hydrological modelling?, Hydrol. Earth Syst. Sci., 5, 1–12, https://doi.org/10.5194/hess-5-1-2001, 2001.

Botter, G., Bertuzzo, E., and Rinaldo, A.: Transport in the hydrologic response: Travel time distributions, soil moisture dynamics, and the old water paradox, Water Resources Research, 46, https://doi.org/10.1029/2009WR008371, 2010.

Boyd, S.: Convex optimization, Cambridge UP, https://doi.org/10.1007/978-1-4419-9863-7_1449, 2004.

Bruen, M. and Dooge, J.: Unit hydrograph estimation with multiple events and prior information: I. Theory and a computer program, Hydrological Sciences Journal, 37, 429–443, 1992.

do Nascimento, T. V., Rudlang, J., Höge, M., van der Ent, R., Chappon, M., Seibert, J., Hrachowitz, M., and Fenicia, F.: EStreams: An integrated dataset and catalogue of streamflow, hydro-climatic and landscape variables for Europe, Scientific Data, 11, 879, https://doi.org/10.1038/s41597-024-03706-1, 2024.

Duchemin, Q.: quentin-duchemin/GAMCR: GAMCR package, Zenodo [code], https://doi.org/10.5281/zenodo.15180816, 2025.

Duchemin, Q., Floriancic, M. G., Seybold, H., and Zanoni, M. G.: GAMCR paper: Data, Zenodo [data set], https://doi.org/10.5281/zenodo.15180911, 2025.

Gao, H., Ju, Q., Zhang, D., Wang, Z., Hao, Z., and Kirchner, J. W.: Quantifying Dynamic Linkages Between Precipitation, Groundwater Recharge, and Streamflow Using Ensemble Rainfall-Runoff Analysis, Water Resources Research, 61, e2024WR037821, https://doi.org/10.1029/2024WR037821, 2025.

Gupta, H. V., Wagener, T., and Liu, Y.: Reconciling theory with observations: elements of a diagnostic approach to model evaluation, Hydrological Processes, 22, 3802–3813, https://doi.org/10.1002/hyp.6989, 2008.

Hargreaves, G. H. and Samani, Z. A.: Reference crop evapotranspiration from temperature, Applied Engineering in Agriculture, 1, 96–99, 1985.

Hastie, T., Tibshirani, R., and Friedman, J.: The elements of statistical learning: data mining, inference, and prediction, Springer, https://doi.org/10.1007/978-0-387-84858-7, 2017.

Kirchner, J. W.: A double paradox in catchment hydrology and geochemistry, Hydrological Processes, 17, 871–874, https://doi.org/10.1002/hyp.5108, 2003.

Kirchner, J. W.: Impulse response functions for nonlinear, nonstationary, and heterogeneous systems, estimated by deconvolution and demixing of noisy time series, Sensors, 22, 3291, https://doi.org/10.3390/s22093291, 2022.

Kirchner, J. W.: Characterizing nonlinear, nonstationary, and heterogeneous hydrologic behavior using ensemble rainfall–runoff analysis (ERRA): proof of concept, Hydrol. Earth Syst. Sci., 28, 4427–4454, https://doi.org/10.5194/hess-28-4427-2024, 2024.

Kirchner, J. W., Feng, X., and Neal, C.: Fractal stream chemistry and its implications for contaminant transport in catchments, Nature, 403, 524–527, https://doi.org/10.1038/35000537, 2000.

Kirchner, J. W., Benettin, P., and van Meerveld, I.: Instructive Surprises in the Hydrological Functioning of Landscapes, Annual Review of Earth and Planetary Sciences, 51, 277–299, https://doi.org/10.1146/annurev-earth-071822-100356, 2023.

Knapp, J. L. A., Berghuijs, W. R., Floriancic, M. G., and Kirchner, J. W.: Catchment hydrological response and transport are affected differently by precipitation intensity and antecedent wetness, Hydrol. Earth Syst. Sci., 29, 3673–3685, https://doi.org/10.5194/hess-29-3673-2025, 2025.

Kratzert, F., Nearing, G., Addor, N., Erickson, T., Gauch, M., Gilon, O., Gudmundsson, L., Hassidim, A., Klotz, D., Nevo, S., Shalev, G., and Matias, Y.: Caravan – A global community dataset for large-sample hydrology, Scientific Data, 10, 61, https://doi.org/10.1038/s41597-023-01975-w, 2023.

Mays, L. W.: Water resources engineering, 3rd edn., John Wiley & Sons, Inc., New York, ISBN 978-1-119-59051-4, 2019.

McDonnell, J. J. and Beven, K.: Debates – The future of hydrological sciences: A (common) path forward? A call to action aimed at understanding velocities, celerities and residence time distributions of the headwater hydrograph, Water Resources Research, 50, 5342–5350, https://doi.org/10.1002/2013WR015141, 2014.

McDonnell, J. J., McGuire, K., Aggarwal, P., Beven, K. J., Biondi, D., Destouni, G., Dunn, S., James, A., Kirchner, J. W., Kraft, P., and others: How old is streamwater? Open questions in catchment transit time conceptualization, modelling and analysis, Hydrological Processes, 24, 1745–1754, 2010.

McGuire, K. J. and McDonnell, J. J.: A review and evaluation of catchment transit time modeling, Journal of Hydrology, 330, 543–563, https://doi.org/10.1016/j.jhydrol.2006.04.020, 2006.

MeteoSwiss CombiPrecip: CombiPrecip precipitation data – MeteoSwiss, https://opendatadocs.meteoswiss.ch/d-radar-data/d1-precipitation-radar-products (last access: 28 November 2023), 2005.

Ponce, V. M.: Hydrologic and Environmental Impact of Small and Medium-Size Basins, https://pon.sdsu.edu/protected29/cive445_ponce_chapter05a_lecture.html (last access: 3 September 2025), 1995.

Rigon, R., Bancheri, M., Formetta, G., and de Lavenne, A.: The geomorphological unit hydrograph from a historical-critical perspective, Earth Surface Processes and Landforms, 41, 27–37, 2016.

Sherman, L. K.: Streamflow from rainfall by the unit-graph method, Eng. News Record, 108, 501–505, 1932.

Snyder, W. M.: Hydrograph analysis by the method of least square, in: Proceedings of the American Society of Civil Engineers, vol. 81, 1–25, ASCE, https://books.google.ch/books?id=gDBQ0AEACAAJ (last access: 10 March 2025), 1955.

Soil Conservation Service: Section 4: Hydrology, in: SCS National Engineering Handbook, USDA, Washington D.C., https://books.google.ch/books?id=sjOEf-5zjXgC (last access: 2 April 2025), 1985.

van der Velde, Y., Torfs, P. J. J. F., van der Zee, S. E. A. T. M., and Uijlenhoet, R.: Quantifying catchment-scale mixing and its effect on time-varying travel time distributions, Water Resources Research, 48, https://doi.org/10.1029/2011WR011310, 2012.