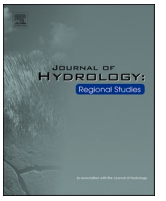
Journal of Hydrology: Regional Studies 65 (2026) 103407

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

journal homepage: www.elsevier.com/locate/ejrh



Towards adaptive stage-flow rating curve for large lowland river streams on the lower Tisza River with backwater impacts using deep learning and copula approach

Milan Stojković [a,b,*], Milan Dotlić [a], Luka Vinokić [a], Zoran Kapelan [c,f], Slobodan Kolaković [d], Veljko Prodanović [a,e]

[a] The Research and Development Institute for Artificial Intelligence of Serbia, Fruškogorska 1, Novi Sad 21000, Australia
[b] Faculty of Civil Engineering, University of Novi Sad, Kozaračka 2a, Subotica 24000, Serbia
[c] Delft University of Technology, Department of Water Management, Stevinweg 1, Delft 2628CN, Netherlands
[d] Faculty of Technical Science, Department of Civil Engineering, University of Novi Sad, Trg Dositeja Obradovića 6, Novi Sad 21000, Serbia
[e] University of South Wales (UNSW), School of Civil and Environmental Engineering, Sydney, NSW 2052, Australia
[f] Faculty of Civil Engineering, University of Belgrade, Belgrade, Serbia

ARTICLE INFO

Keywords:
Backwater impacts
Lowland river stream, synthetic generation
Kolmogorov–Arnold Network
Multilayer perception
Support vector regression
Adaptive rating curve
Tisza River

ABSTRACT

Study region and rationale: The case study focuses on two hydrological stations on the lower Tisza River (Serbia and Hungary), located on a large lowland river stream strongly influenced by the downstream Novi Bečej reservoir (Serbia), where backwater effects and long-term hydraulic variability pose significant challenges for accurate flow estimation.
Methods and data: Conventional approaches that assume a stable stage-flow relationship fail to capture rating curve complexity. To address these limitations, this study introduces a joint machine learning (ML)–copula framework in which ML-based rating models are developed and verified on measured data and stochastically generated synthetic stage-flow pairs using a Gumbel copula. The framework integrates traditional power-law regression with Support Vector Regression (SVR), Multilayer Perceptron (MLP), and Kolmogorov–Arnold Networks (KAN), and evaluates uncertainty through confidence intervals and performance metrics (MAE, RMSE, MAPE, R^2, PICP).
Main results and conclusions: ML models outperform classical power regression across low, mean, and high flows, with SVR, MLP, and KAN achieving RMSE ≈ 78–163 m^3 /s compared to RMSE ≈ 80–173 m^3 /s for power regression. Under synthetic Gumbel-generated datasets, KAN maintains performance comparable to SVR (RMSE ≈ 129–212 m^3/s) and preserves stable behavior across flow regimes, avoiding the underprediction observed in MLP. Consequently, KAN demonstrates the robustness necessary for adaptive stage-flow rating curve estimation under changing hydraulic conditions.

1. Introduction

Reliable river flow estimation plays a pivotal role in water management, serving multiple critical functions. It enables assessment of

* Corresponding author at: The Research and Development Institute for Artificial Intelligence of Serbia, Fruškogorska 1, Novi Sad 21000, Serbia.
E-mail address: stojkovic@ivi.ac.rs (M. Stojković).

https://doi.org/10.1016/j.ejrh.2026.103407
Received 14 January 2026; Received in revised form 1 March 2026; Accepted 28 March 2026
Available online 4 April 2026

design flood magnitudes for dams with significant regulating capacity by estimating flow return periods (Domínguez and Arganis, 2012), informs crest elevation and freeboard requirements for embankments along flood-prone sections (Alhamid et al., 2024), supports preparation of input data for hydrological and hydrodynamic models in forecasting systems (Yu et al., 2025), and facilitates long-term hydrological regime analyses by comparing observed and simulated flows under future climate scenarios (Stojković et al., 2020).

The measurement of water level (stage) is relatively straightforward and does not require expensive equipment, whereas flow measurement is far more demanding, requiring skilled personnel, time-intensive procedures, and specialized instruments (Jain and Singh, 2019). Stage records are used to estimate flow through a stage–flow rating curve (hereafter referred to as the rating curve) derived from systematic paired measurements under varying hydrological conditions. These empirical data establish a mathematical relationship that enables continuous conversion of stage to discharge. To ensure long-term validity, the U.S. Geological Survey routinely revisits each hydrological station roughly every six weeks, performing direct flow measurements, including during high flows, to update the rating curve (USGS, 2011).

Rating curves typically rely on a conventional power-law model assuming stationary channel geometry and hydraulic conditions (Jain and Singh, 2019; Rantz et al., 1982). However, this neglects processes like sedimentation, erosion, vegetation growth, and channel degradation, which progressively reduce accuracy (Church, 2015; Rojas et al., 2020). Such morphological changes dynamically alter the values in rating curves, necessitating periodic recalibration or fully dynamic flow equations designed to capture non-stationary behavior more reliably (Dottori et al., 2009). Backwater effects further challenge reliability, as observed pairs depend on downstream boundary conditions, especially near reservoirs. Installing a downstream station (“twin gage”) to estimate water-surface slope can correct backwater influences (Plavšić 2019), but is impractical in lowland, meandering rivers with minimal gradients (Hidayat et al., 2011). Recent efforts therefore consider translating flow into water level via volume routing or estimating backwater storage by balancing surface slopes (Hellmers and Fröhle, 2022). This, however, limits the proposed conceptual approach, which is designed for mesoscale lowland river basin, and makes it difficult to apply to large low-slope river basins, where backwater effects extend over long distances. Moreover, hydrodynamic models explicitly simulate downstream influences such as variable reservoir levels (Liro et al., 2022), though real-time use remains difficult due to continuous data requirements (Brunner et al., 2021). Consequently, most existing approaches for estimating rating curves for low-sloping large rivers are either overly simplified (hence not very accurate) or require extensive input data which is difficult and/or time-consuming, i.e. expensive to collect. All this limits the applicability and generalization capability of these approaches.

More recently, above challenges have been addressed by using the ML methods. These methods have emerged as a promising new technology particularly given their increasing adoption in recent hydrological practice (Vinokić et al., 2025; Dodig et al., 2024) and their potential value in developing practical tools for rating curve estimation (Rozos et al., 2022). Among the ML methods, those based on regression modeling have shown effectiveness (Rozos et al., 2022, He et al., 2023; Vishwakarma et al., 2023; Kisi et al., 2024; Dasgupta et al., 2025) but only a limited number of studies have explored temporal dependencies, often with poorer performance compared to conventional statistical methods (Maghrebi and Vatanchi, 2024). Early applications of neural networks demonstrated less effective performance relative to traditional techniques (Ajmera and Goyal, 2012), but the advent of deep learning has significantly improved flow estimation accuracy (He et al., 2023; Vishwakarma et al., 2023; Kisi et al., 2024; Dasgupta et al., 2025). In fact, improvements in predictive reliability exceeding 50% over conventional methods have been reported (Dasgupta et al., 2025). The spectrum of developed methods for rating curve estimation ranges from single ML to ensemble approaches, such as Support Vector Regression (SVR), Random Forests (RF), eXtreme Gradient Boosting (XGBoost), and Model Trees, and more recent advances like symbolic regression (Dasgupta et al., 2025).

While considerable progress has been made in estimating flow and stage relationships using ML-based models, some key limitations remain unexplored. In particular, the evolution of rating curves under headwater impacts and changing river morphology are still insufficiently addressed (Rojas et al., 2020; Hellmers and Fröhle, 2022). The existing models assume a stationary dependence between stage and flow, which constrains their accuracy when applied to unseen hydrological conditions. As a consequence, their predictive reliability decreases when new measurements are influenced by backwater effects.

There is, therefore, a need to propose new approaches for rating curve assessment and better understanding of related uncertainties. This is crucial for supporting water managers in flow estimation, under changing river conditions which are common in highly regulated river systems. In response to these challenges, the present research is guided by the following objectives:

- To develop a novel ML-based regression framework capable of estimating rating curves that could adapt to hydrological changes. The proposed models are trained and validated using historical datasets affected by downstream regulation (e.g., reservoir influence), thereby demonstrating their robustness and reliability under variable hydraulic conditions.
- To validate the developed ML-based rating curve using stochastically generated flows-stage pairs derived from a site-specific bivariate copula distribution. This validation aims to assess the ML model’s performance and generalization ability for hydrological events not represented in the observed data indicating its adaptive potential.
- To estimate confidence intervals for the derived rating curve using unseen observed data and testing it on stochastically generated data, thereby providing a quantitative measure of model reliability across different flow regimes (low, medium and high flows).

The joint ML–Copula framework offers a novel approach for addressing variability in records and reliability in empirical rating curves, enhancing both scientific understanding and practical application in river flow estimation.

2. Methods

2.1. Case study

With a total length of approximately 1400 kilometers and a drainage area of 157,186 km^2, the Tisza River represents the largest tributary of the Danube (ICPDR, 2008; Spatial and plan, 2023). It originates from the confluence of the Black and White Tisza near Rakhiv, Ukraine, and flows through Ukraine, Hungary, and Serbia, with brief crossings into Romania and Slovakia. The annual precipitation across the Tisza River Basin ranges from 500 to 1600 mm, with the lowest values recorded in the southwestern plains and the highest in the northwestern Carpathians. The mean multi-annual flow of the Tisza river at its confluence with the Danube is approximately 830 m^3 /s. The minimum recorded flow is 80 m^3 /s, while the maximum observed flow reached 3718 m^3 /s (Spatial and plan, 2023).

The lower Tisza river is recognized as a flood-prone area due to its geomorphological characteristics, particularly the low channel gradients in the downstream river sections and high precipitation sums in the upstream parts. The first major organized efforts to construct flood defense infrastructure along the Tisza River began in the 1860s, with most of the infrastructure completed prior to World War I (Kolaković et al., 2018). Historical design standards prescribed clayey cohesive material for embankment construction, with crest elevations positioned one meter above the historically highest water levels. Lately, a significant intervention at the river stream with implication on hydrological regime was the construction of the Novi Bečej dam in 1979, downstream from the Serbian–Hungarian border (Spatial and plan, 2023). The dam was designed for a flood flow of 4650 m^3 /s and operates at a normal reservoir level ranging from 74.50 to 75.50 m a.a.s.l. Its primary purposes are upstream water level regulation for navigation, irrigation, and drought mitigation.

The hydrological behavior of the Tisza River is substantially influenced by sediment transport and bank erosion processes. Erosion along the lower Tisza, especially at concave riverbank sections, remains a persistent geomorphic feature that intensifies during high-flow conditions (Kiss et al., 2019). Recent studies highlight that lateral channel migration is a suitable proxy for reconstructing long-term sediment budgets in low-gradient environments (Kiss et al., 2024). Between 1976 and 2017, erosional processes have been intensified, transforming the Tisza River into a net sediment source, with the annual sediment budget ranging from −0.03 to −0.08 m^2/m/year (Kiss et al., 2024). The hydraulic structures, namely the Novi Bečej dam have considerably impacted these processes. Flow regulation has resulted in increased sediment deposition in the upstream rive reach and progressive bed degradation downstream (Stipić et al., 2021; Mohsen et al., 2021). Broader analyses of the low Tisza river suggest that sediment deficits below dams can lead to long-term disconnection of floodplains, deterioration of aquatic habitats, and compromised infrastructure resilience (Jánosi et al., 2025).

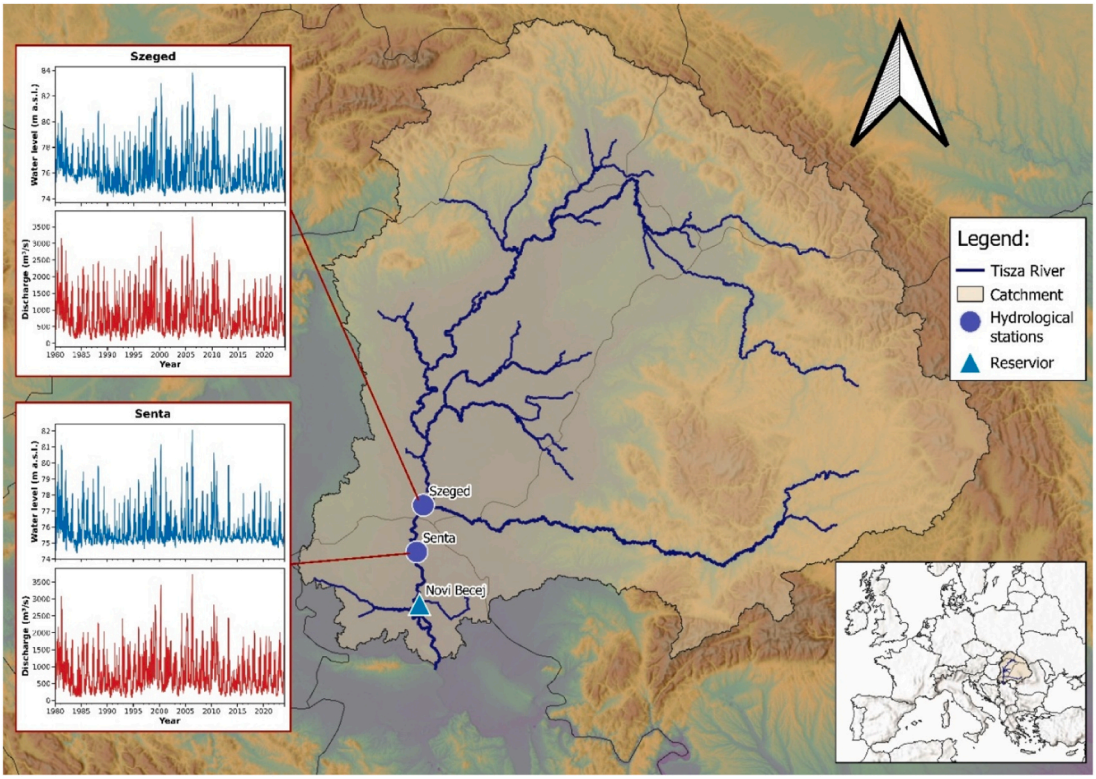


Fig. 1. Daly water levels and flows for hydrological stations at the lower Tisza river (left) alongside with their locations (right).

This study utilizes data from two hydrological stations (h.s.) situated along the lower Tisza River: Szeged in Hungary and Senta in Serbia. The geographic locations of the stations are illustrated in Fig. 1. The Szeged and Senta h.s. drain catchment areas of 141,715 km^2 and 138,858 km^2, respectively. Both stations lie within the backwater zone influenced by the Novi Bečej dam, situated approximately 110.2 km downstream of Szeged and 60.1 km downstream of Senta.

Both stations have continuous long-term records of daily flows and water levels (stages) spanning from 1931 to 2023. To ensure consistency with the hydraulically regulated regime, only data recorded after the construction of the Novi Bečej dam (i.e. post-1979) are used in this study. However, to assess the stationarity of the extreme flow series, homogeneity and trend analyses are conducted over the full observed period (1931–2023), as shorter time series can yield unreliable test results (Stojković et al., 2014). Applying the Mann–Kendall trend test at both stations returns p-values of 0.484 for Senta and 0.902 for Szeged, both well above the 0.05 significance threshold. Although a mild downward tendency in peak flows is noted, no statistically significant trends are detected, confirming the temporal stability of the flow regime and the reliability of the learned rating curves.

2.2. Adaptive rating curve model

This research presents a joint ML-Copula based framework for estimating rating curve by accounting for human-introduced changes. Fig. 2 summarizes the four-phase workflow used to develop the framework: (1) recorded data collection, preprocessing, and outlier identification using a weighted regression approach; (2) development, testing, and verification of regression and ML models to derive the rating curve; (3) verification of these models using stochastically generated events from a bivariate Copula; and (4) estimation of confidence interval of rating curve with the recorded and stochastically generated data. The performance of a traditional power-law regression model is compared with several models that use the following ML methods: support vector regression - SVR, multilayer perceptron - MLP, Kolmogorov-Arnold network - KAN.

2.2.1. Dataset processing

Daily flows (Q) and water stage/level (H) records are compiled, quality-checked, and synchronized to a common daily time step, within the period from 1980 to 2023 which corresponds to the operational phase of the Novi Bečej dam reservoir. The dataset is divided into two subsets: 1980–2010 for model training and 2011–2023 for model verification. The Szeged station record has 0.2% missing values, whereas the Senta station record is complete. To infill the Szeged data gaps, both time series of water level and flow are normalized, and values from Senta are transferred to Szeged using the multi-nonlinear regression method (Stojković et al., 2014). The harmonized series are then merged into a paired Q–H dataset for subsequent modeling and evaluation.

After recorded dataset is completed, outlier detection is performed. Outliers are defined as observations whose absolute standardized residual from the regression fit exceeds 3 standard deviations (Chatterjee and Hadi, 2006). Because the relationship is nonlinear, the regression function is estimated using locally weighted regression (LOESS). LOESS (Loader, 1999) is chosen because it captures complex nonlinear structure without assuming a parametric form. Moreover, it is resilient to local irregularities through distance-based weighting and yields a smooth deterministic component that can be separated from stochastic variability (Stojković et al., 2015, 2017).

The LOESS smoothing parameter is expressed as the width of a temporal window containing n_w observations, centered at time t where $t_i \in [t - (n_w - 1)/2, t + (n_w - 1)/2]$. Within this window, each observation at time t_i is assigned a weight w_i via the tricube kernel (Loader, 1999):

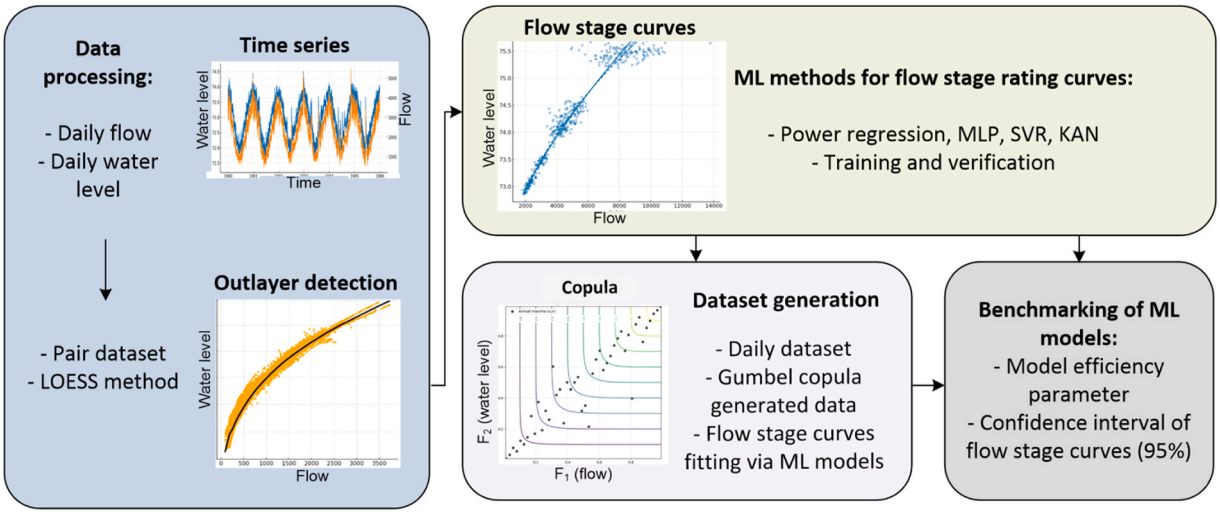


Fig. 2. A workflow of proposed ML-Copula based framework for rating curve estimation for the large low-slope river streams.

$$w_i(Q) = \left[1 - \left|\frac{Q - Q_i}{d(Q)}\right|^3\right]^3, i = 1, 2, ..., n_w. \quad (1)$$

where $w_i$is the smoothing kernel function; w_i are the local weights assigned to each time series member, with Q denoting the flow values of the member being smoothed, Q_i the neighboring member at step $i = 1,2,\ldots, n_w$, and d is the temporal distance between Q_i and Q. With LOESS friction of 0.3 and a tricube weighting function applied, it results in 132 pairs (0.8%) of outliers at Senta and 374 pairs (2.3%) of outliers at Szeged.

2.2.2. Rating curve models

To estimate the rating curve, we employ several baseline models: power regression model, MLP, and SVR, that have been used in this domain for a long time due to their ability to capture nonlinearity in flow and stage relationship (Sivapragasam and Muttil, 2005). Recent advances in ML further motivate complementing these baselines with a KAN, a neural architecture grounded in the Kolmogorov–Arnold representation theorem in which each connection implements a learnable univariate function, rather than relying on fixed activations or piecewise-constant tree splits (Vinokić et al. 2025, Liu et al., 2024).

2.2.3. Power regression

Flows are converted from the observed stages using the rating curve power equation. In a river system under steady flow conditions when the energy grade line, and the channel bed slope are all parallel to each other, a hydraulic equation for flow estimation can be written as (Kumar and Sen, 2024):

$$Q = a(h + h_0)^c \quad (2)$$

where Q is the flow, h is the water depth, and h_0 is in offset while a and c are power law coefficient and exponent related to the type of river section.

2.2.4. Support vector regression

Support vector regression (SVR) is a powerful kernel-based learning method known for its high predictive accuracy and strong generalization capabilities. When equipped with the Radial Basis Function (RBF) kernel, SVR becomes a universal approximator capable of modeling any continuous function on a compact set (Wang et al., 2004). The RBF kernel enables nonlinear mapping of input data into a high-dimensional feature space, allowing SVR to capture complex relationships while maintaining a balance between model complexity and fitting error through regularization.

For the SVR with an RBF kernel, two key hyperparameters must be specified: the parameter γ (gamma), which defines the inverse of the radius of influence of individual training samples, and the regularization parameter C, which controls the trade-off between model complexity and training error. Since the optimal values of these parameters are problem-dependent and typically unknown a priori, a parameter search is required. A grid search approach is employed, which is computationally feasible given the low dimensionality of the hyperparameter space and provides a reliable means of identifying parameter combinations (Hastie et al., 2005).

2.2.5. Multilayer perception (fully connected feedforward Neural Network)

A fully connected feedforward neural network, also known as multilayer perceptron (MLP), represents a powerful and versatile tool for regression tasks. Theoretically, it has been established (Hornik at al. 1989) that multilayer feedforward networks are universal function approximators, capable of representing any measurable function to an arbitrary degree of accuracy.

Although rigorous hyperparameter optimization and architectural tuning can be done using grid search, random search, or Bayesian optimization, such approaches are computationally expensive (Pedregosa et al., 2011). Therefore, the hyperparameters are determined through manual tuning guided by empirical performance on a validation set, following standard model selection practice. Empirical performance is assessed by computing multiple forecast evaluation metrics, including the Nash–Sutcliffe efficiency, on the validation set. Hyperparameter configurations are selected based on their ability to consistently maximize validation performance across performance metrics.

2.2.6. Kolmogorov-Arnold neural network

This neural network is based on the Kolmogorov-Arnold representation theorem, which states that continuous multivariate function can be decomposed into a finite set of univariate functions and their combinations. Mathematically, KAN is:

$$f(x_1, \ldots, x_n) \approx \sum_{q=0}^{2n} \Phi_q \left(\sum_{p=1}^{n} \phi_{q,p}(x_p) \right), \quad (3)$$

where functions ϕ_q and ϕ_p are learned from data.

KAN bridges classical curve fitting and deep learning with an architecture that represents the Q–H curves as a sum of learnable spline functions (Panahi et al., 2025), which yields to a smooth model with local control. As a result, KAN complements SVR and neural networks by offering flexible structured function learning that aligns with the physics of rating curves (Cui et al., 2025).

2.2.7. Flow and stage data generation via copula

Generation of Q-H data pairs used for the training and verification of above models is performed via a bivariate Copula framework

(Nelsen, 2006). In the first step, the marginal distributions of daily flows and stages are fitted under empirical distribution separately for each month.

To capture the observed upper-tail dependence of flow and water level time series, a Gumbel copula is adopted (Grimaldi and Serinaldi, 2006; Genest and Favre, 2007). This choice is motivated by its ability to represent extended right tails and positive tail dependence, which are typical for flood extremes. While alternative copula families, such as the Clayton copula, which emphasizes lower-tail dependence, or the Frank copula, which assumes symmetric tail behavior could generate synthetic datasets with different distributional properties, their use would likely underrepresent the joint extremes most relevant to rating curve estimation under flood conditions (Nelsen, 2006).

A mathematical expression of Gumbel copula C_θ is as follows (Nelsen, 2006):

$$C_{\theta(u,v)} = \exp\left(-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{\frac{1}{\theta}}\right), \quad \theta \geq 1 \tag{4}$$

where $u = F_1(Q)$ and $v = F_2(H)$ are respectively flow and water levels scores, while θ is Gumbel copula dependence.

2.2.8. Rating curve confidence intervals

Confidence intervals for the models mentioned above are derived from the residuals between the modeled stage and flow relationship and the corresponding observed or simulated flow and stage pairs. Only the held-out verification set is used, and the sample is approximately expanded to population level using a simple residual-tails bootstrap, thereby estimating uncertainty intervals entirely on data never seen during training. These intervals are defined separately for different hydrological regimes – low, mean, and high flows using quantile classification (Richter et al., 1996). The low-water regime covers the range from the minimum observed flows up to the first quartile (q_1), the mean-water regime extends from the first (q_1) to the third quartile (q_3), while the high-water regime spans from the third quartile (q_3) to the maximum observed flow.

Let $y_i$be flow, $\widehat{y}_i$ the model estimate, and $e_i = y_i - \widehat{y}_i$ the residual. For a regime $k \in \{\text{low, mean, high}\}$, let $\mathcal{I}_k$ index verification samples in that regime, with confidence level $1-\alpha$ and B bootstrap resamples. For each bootstrap $b = 1, \ldots, B$, (re)sample with replacement indices $\mathcal{I}_k^{(b)} \subset \mathcal{I}_k$ and form residuals $\{e_j^{(b)} : j \in \mathcal{I}_k^{(b)}\}$. Regime-specific residual quantiles are computed as follows:

$$q_{k,\alpha/2}^{(b)} = \text{quantile}_{\alpha/2}\left(e_j^{(b)}\right) \tag{5}$$

$$q_{k,1-\alpha/2}^{(b)} = \text{quantile}_{1-\alpha/2}(\{e_j^{(b)}\}) \tag{6}$$

The regime-conditioned PI for a prediction $\widehat{y}_i$ is then calculated by adding aggregated bootstraps of lower and upper quantiles (via median) $\widetilde{q}_{k,\alpha/2}$and $\widetilde{q}_{k,1-\alpha/2}$ (Lee and Scholtes, 2014):

$$\widehat{y}_i^L = \widehat{y}_i + \widetilde{q}_{k,\alpha/2} \tag{7}$$

$$\widehat{y}_i^U = \widehat{y}_i + \widetilde{q}_{k,1-\alpha/2} \tag{8}$$

2.3. Model performance metrics

Standard performance metrics are employed for model training and for verification on unseen data. Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) are selected to illustrate average predictive error over the entire rating curve. In addition, the maximum absolute error (MAE) is reported to highlight extreme deviations that are critical during high-water conditions. All metrics are computed on Q-H observations where lower values indicate better predictive performance:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|e_i| \tag{9}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}e_i^2} \tag{10}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{e_i}{y_i}\right| \tag{11}$$

where n denotes the number of time-series observations, i is indexes the time step, and e_i is the error computed as the difference between the observed and predicted value at time step i.

Additional metric showcasing predictive accuracy R^2, where a value of one indicates a perfect match between model prediction and observation, is used:

$$R^2 = \frac{\sum_{i=1}^{n} e_i^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{12}$$

where $\bar{y}$ denotes the mean observed value within the sample.

A standard metric for evaluating the coverage of confidence intervals, the prediction interval coverage probability (PICP), is used too. This metric quantifies the proportion of data points that fall within the predicted uncertainty bounds, as defined by:

$$PICP = \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}(y_i \in [\hat{y}_i^L, \hat{y}_i^U]), \tag{13}$$

where Ndenotes the total number of samples, $\hat{y}_i^L$ and $\hat{y}_i^U$represent the lower and upper prediction bounds for the i-th observation, respectively, and $\mathbb{I}(\cdot)$is the indicator function, equal to 1 when the condition is true and 0 otherwise.

3. Results

3.1. Flow rating curve models

Fig. 3 and Table 1 present the performance of the tested models for the Senta and Szeged stations, showing both the observed and modelled stage-flow relationships and the corresponding performance metrics. The dataset (1980–2023) is divided into two subsets: 1980–2010 for training and 2011–2023 for verification. Both datasets exhibit a strong dependence between stage and flow, with a

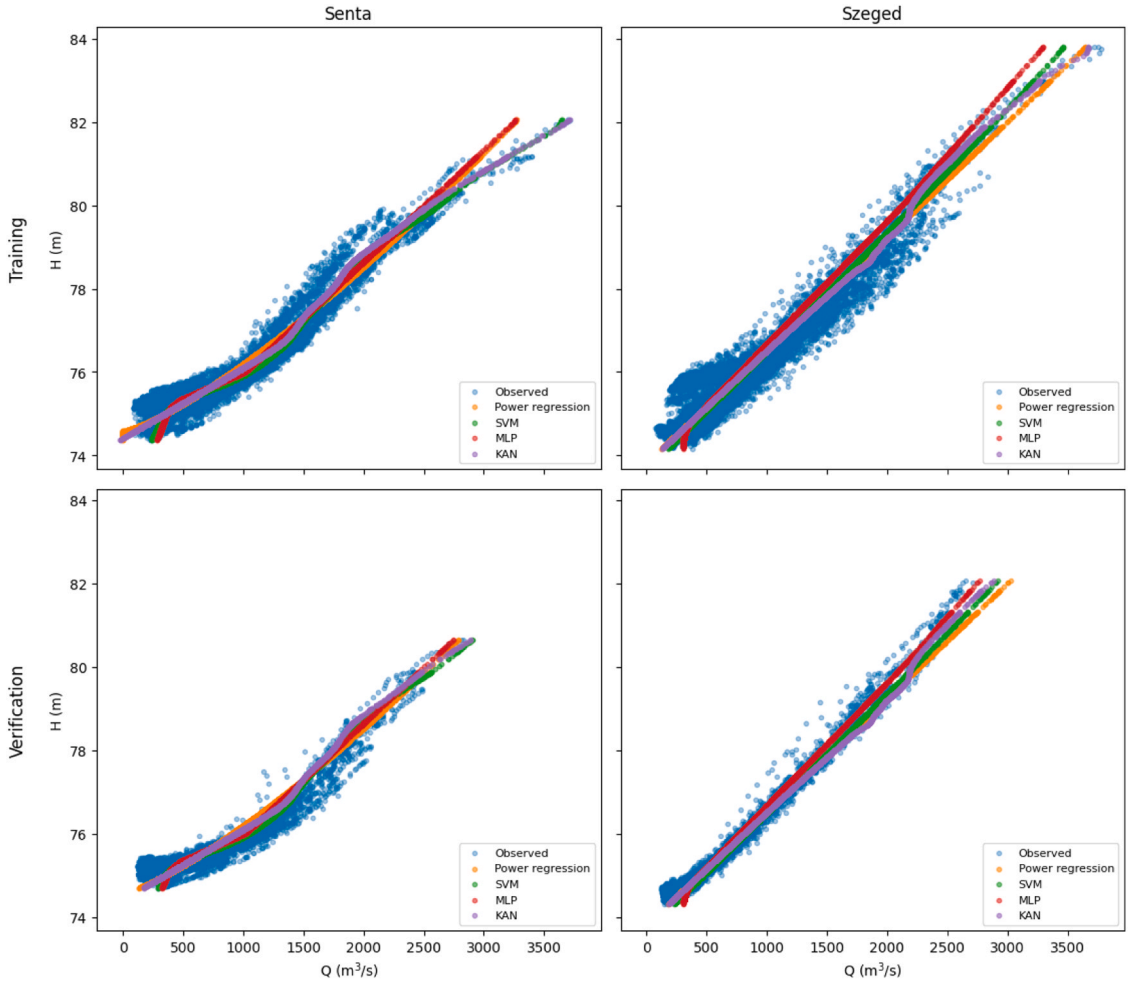


Fig. 3. Observed and modelled rating curves for Senta and Szeged during the training and verification phases. The observed data (blue points) are compared with four modelling approaches (Power regression, SVR, MLP, and KAN).

Table 1
Model performance metrics for the Senta and Szeged during the verification phase under observed and copula-generated datasets: Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Coefficient of Determination (R^2), Root Mean Square Error (RMSE), and Prediction Interval Coverage Probability (PICP). Best results are shown in **bold** text.

	Station	Model	MAE (m3/s)	MAPE (%)	R^2 (-)	RMSE (m^3/s)	PICP (-)
Verification – observed dataset	**Senta**	Power	140.9	37.41	0.887	173.1	0.771
		SVR	**114.4**	**31.65**	**0.922**	**143.8**	**0.820**
		MLP	120.8	31.84	0.915	150.2	0.762
		KAN	132.3	35.72	0.899	163.3	0.782
	Szeged	Power	**55.8**	**14.06**	0.976	80.2	0.940
		SVR	64.9	19.04	0.972	86.8	0.913
		MLP	59.6	17.93	**0.977**	**78.0**	0.914
		KAN	56.1	14.26	0.977	78.5	**0.940**
Verification – Copula generated datasets	**Senta**	Power	159.2	30.76	0.847	216.8	0.714
		SVR	**151.9**	**29.22**	**0.854**	**211.4**	**0.719**
		MLP	155.9	29.68	0.849	214.9	0.693
		KAN	154.3	29.80	0.854	211.7	0.706
	Szeged	Power	93.7	**15.16**	0.951	129.3	**0.679**
		SVR	92.2	17.26	**0.953**	**125.8**	0.667
		MLP	100.1	18.75	0.947	134.2	0.650
		KAN	**92.2**	15.22	0.952	128.8	0.669

larger spread at higher flows, particularly at Senta station. The rating curve at this station displays a higher degree of nonlinearity than at Szeged station, reflecting the different hydraulic characteristics of the river Tisza at these locations (Fig. 3). The spread of point measurements in both rating curves results from the backwater influence of the Novi Bečej reservoir, which affects the geographically closer Senta station more strongly than the upstream Szeged.

As it can be seen from Table 1 and Fig. 3, all models show a strong correspondence with the observed Q-H data at both river stations. At Senta station, the Power regression rating curve model provides a fit that captures the overall rating curve trend but respectively underestimates the lowest and overestimates the upper stages. It demonstrates satisfactory verification performance across different flow regimes (Table 1), as indicated by the performance metrics ($R^2 = 0.887$, $RMSE = 173.1 m^3/s$, $MAPE = 37.4\%$). However, a visual inspection of Fig. 3 reveals a pronounced bias at extreme flow conditions. In addition, the SVR model aligns most closely with the observed data across the entire flow range, demonstrating high flexibility in representing nonlinear segments and achieving the best overall metrics at Senta ($R^2 = 0.922$, $MAE = 114.5 m^3/s$, $RMSE = 143.8 m^3/s$). The MLP model also reproduces the Q-H relationship with high accuracy ($R^2 = 0.915$), local deviations are observed at extreme flow conditions, as illustrated in Fig. 3, which is consistent with the increased RMSE (150.2 m^3/s). The KAN model achieves a good fit across all flow regimes, with particularly strong performance in the transition zones between moderate and high flows (Fig. 3). It yields a competitive error metrics on the verification dataset (R^2=0.899; RMSE = 163.3 $m^3\ s^{-1}$), confirming its ability to effectively capture nonlinearities in the rating curve, even in the river sections influenced by hysteresis. At Szeged station (Fig. 3), for the observed Q-H data, all models perform well since the relationship in the measured points does not possess the pronounced nonlinearity (for all models R^2 exceed 0.97). The performance metrics from Table 1 suggest that the MLP slightly surpasses other models ($R^2 = 0.977$, $RMSE = 78.0 m^3/s$), while KAN and Power regression exhibit nearly identical accuracy ($R^2 \approx 0.977$, $PICP \approx 0.94$). The SVR model, though slightly less accurate ($R^2 = 0.972$), is also performing well on the verification dataset. However, discrepancies between observed and modelled points at high flow conditions are evident for the power regression and SVR model (Fig. 3), whereas satisfactory matching is achieved by the KAN and MLP models.

Residual plots for all models are presented in Figure S2 and Figure S3 (Supplementary material), providing further insight into the systematic behavior of each model across different flow regimes. At low and medium flows, all models exhibit considerable residual scatter, with power regression and SVR showing the widest spread and the most pronounced systematic bias. At high flows, the differences between models become particularly evident, reinforcing the visual discrepancies noted in Fig. 3. Power regression and SVR systematically underestimate large flows, with negative residuals becoming especially pronounced at Szeged, where they reach values of 600 m^3 /s. MLP improves upon these models but still retains a tendency to underestimate peak flow values. KAN, in contrast, is the only model that consistently produces residuals near zero at high flows for both stations, supporting its performance at extreme conditions observed in Fig. 3, and avoiding the negative bias present in the other models.

Having examined the predictive performance of the models, it is worth addressing the generalization behavior of KAN and MLP during training. Despite the substantially larger parameterization of KAN due to the replacement of scalar weights with learned univariate functions, no empirical evidence of overfitting is observed (Figure S1, Supplementary material). As shown in Figure S1, both training and validation losses decrease rapidly and stabilize at nearly identical values, with the generalization gap remaining minimal across both study locations. The MLP baseline exhibits a similar convergence pattern, with no systematic advantage in validation performance despite its smaller parameter count. This suggests that the functional parameterization of KAN introduces implicit structural regularization that mitigates overfitting, even under a relatively limited training window for a large river system.

In addition to the results obtained using observed datasets, Table 1 also summarizes the performance metrics for Copula-generated synthetic datasets at both stations. The Gumbel copula is employed to provide a statistical expansion of the observed data range while preserving the upper-tail dependence structure. It generates synthetic flow-stage pairs that are statistically consistent with the observed marginal distributions and their correlation. It is important to note that downstream reservoir operations predominantly

affect the low and medium flow regimes at the Tisza river, whereas high flows are not significantly impacted, as the homogeneity analysis indicates stationarity in the annual flow series. Consequently, the copula-based approach treats hysteresis at the sites of Senta and Szeged stations as statistical variability inherent in the backwater-affected data, rather than incorporating physically informed dynamics through explicit covariates (e.g., gate operations, reservoir water levels).

For the Senta station, all models exhibit comparable performance on the synthetic data, with R^2 values ranging between 0.847 and 0.854 and RMSE values between 151.9 m^3 /s and 159.2 m^3 /s. The SVR and KAN models slightly outperform the power regression and MLP in the terms of R^2 (0.854), indicating a capability to reproduce joint dependence embedded in the Copula-generated samples. At the Szeged station, higher overall performance is observed across all models, primarily due to the lower degree of nonlinearity in the simulated Q–H pairs, with R^2exceeding 0.94 and RMSE values remaining below 101 $m^3\ s^{-1}$. The power regression, SVR, and KAN models demonstrate nearly identical performance, while the MLP shows increased errors.

3.2. Flow and stage synthetic data generation via copula

Fig. 4 illustrates the dependence between the observed daily flows and water level (black points) at the Senta and Szeged hydrological stations, along with synthetic data generated using the Gumbel copula (red points). The estimated Gumbel copula parameters indicate a pronounced positive dependence between the variables, stronger at Szeged ($\Theta = 4.69$), where backwater effects from the Novi Bečej reservoir are less pronounced. The higher Θ value at Szeged reflects a tighter coupling between flow and water level, suggesting more stable hydraulic conditions and a consistent rating relationship. Conversely, the slightly lower Θ value at Senta (3.16) implies greater variability in the rating curve, influenced by the operational regime of the downstream reservoir.

3.3. Confidence intervals

Confidence intervals for the flow-stage curves are estimated using both the observed verification dataset (which is not included in model training) and the Gumbel-copula–based synthetic datasets (Fig. 5 and Fig. 6, respectively). The results are presented via confidence intervals (CIs) of the predicted rating curves for the Senta and Szeged stations under different flow regimes (low flows, medium, and high flows). The shaded bands shown in above two figures represent 95% confidence intervals derived from residuals across low, medium and high flow regimes.

As shown in Fig. 5, the confidence intervals at the Szeged station are narrower than those at Senta, indicating a more stable Q-H relationship. This can be attributed to the stronger backwater influence at Senta, which is located closer to the Novi Bečej dam and therefore exhibits increased hydraulic variability. The patterns observed in Fig. 5 are also reflected in the PICP values reported in Table 1. At Szeged, the KAN model achieves a PICP of 0.940, whereas CI coverage at the Senta station is lower, with the best performance obtained by the SVR model (PICP = 0.820).

The transition of confidence intervals across different flow regimes (low, medium, and high flows) for the observed dataset exhibits station-dependent behaviour (Fig. 5). At the Senta station, the confidence intervals are wider, while their transition is smoother, reflecting increased variability of Q-H pairs mainly within the medium- and high-flow domains, which can be attributed to the influence of the Novi Bečej dam. In contrast, at the Szeged station, the impact of the Novi Bečej dam is predominantly manifested in the high-flow regime, leading to a pronounced spread in the observed rating curves and a subsequent widening of the confidence intervals across all models. In terms of low flows, KAN is the only model that provides consistent estimates across both stations (Fig. 5), preserving the shape of the rating curve more effectively than the other models. Consistent with this behaviour, confidence interval asymmetry is more pronounced at the Szeged station for all models except the MLP, whereas at the Senta station asymmetry is less evident.

When evaluating the synthetic Q-H data (Fig. 6), all models successfully reproduce the trends of the rating curves, reflecting the behaviour observed for the measured dataset (Fig. 5). Moreover, the width of the confidence intervals does not increase substantially

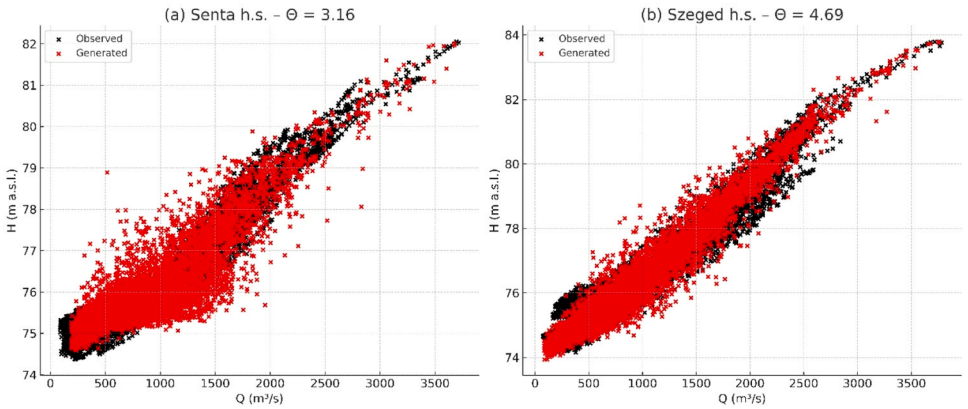


Fig. 4. Observed (black) and Gumbel copula–generated (red) flow–water level pairs at Senta (Θ = 3.16) and Szeged (Θ = 4.69) stations.

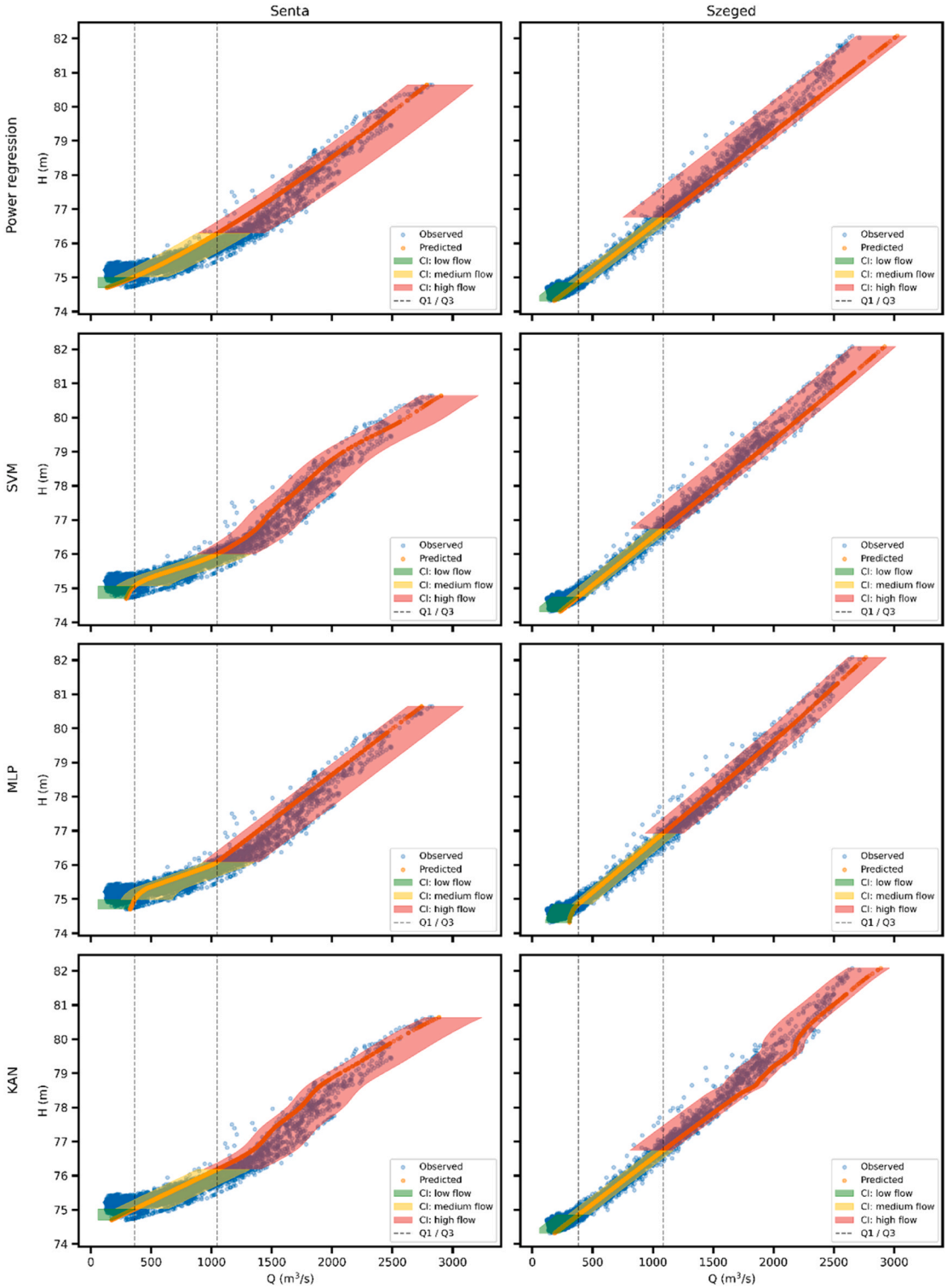


Fig. 5. Confidence intervals of modelled rating curves (power regression, SVR, MLP, KAN) for the Senta and Szeged stations using the observed verification datasets (1980–2023). Shaded bands represent 95% confidence intervals conditioned on model structure and residual-based uncertainty formulation across low, medium, and high flow regimes.

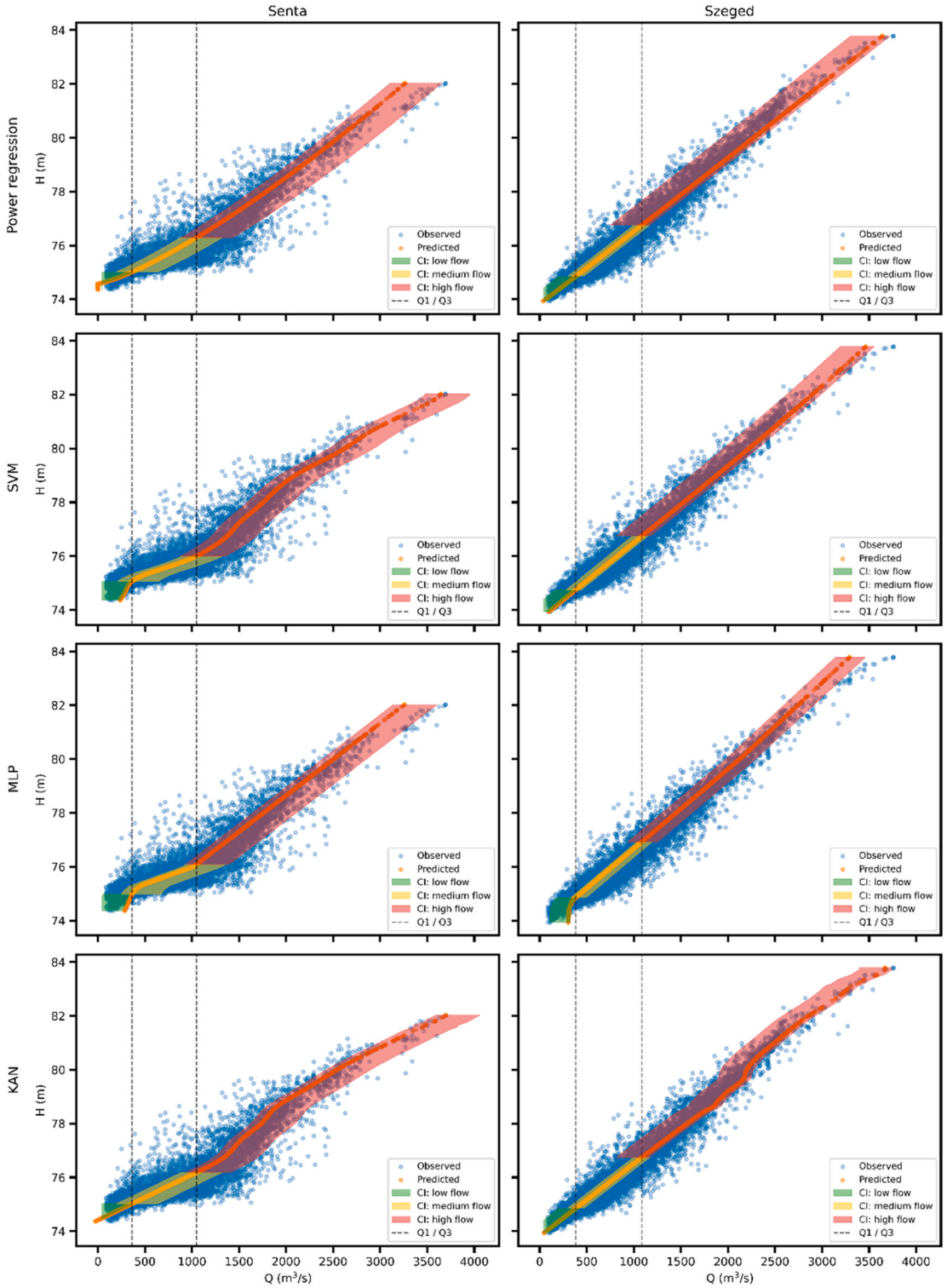


Fig. 6. Confidence intervals of rating curve estimates (Power Regression, SVR, MLP, and KAN) for the Senta and Szeged stations derived from Gumbel copula-generated synthetic datasets. Shaded bands represent 95% confidence intervals conditioned on model structure and residual-based uncertainty formulation across low, medium, and high flow regimes.

for the synthetic datasets compared to the verification dataset, indicating that the predictive performance remains robust even when the models are exposed to distribution-extended, synthetic Q–H pairs. Consistent with the results obtained from the observed data, the confidence intervals are wider at the Senta station than at Szeged, while the transition of the CIs across the medium- to high-flow regimes is smoother at Senta than at Szeged.

Among all evaluated models, only the KAN maintains a stable representation in the low-flow domain for the generated dataset at both stations (Fig. 6), in agreement with its performance on the observed data (Fig. 5). Nevertheless, the coverage of the CIs is, as expected, reduced at both Senta and Szeged, reaching values of 0.719 and 0.679, respectively (Table 1). This reduction primarily reflects that the bootstrap-based CIs are estimated on the empirical data distribution, whereas copula sampling extends the joint distribution beyond observed regimes.

However, the key sought difference between the synthetic and observed datasets is the distributional extension in the high-flow domain. Specifically, the copula-based synthetic dataset spans a substantially wider magnitude range (Fig. 6), than those recorded during the verification period, during which maximum flows and stages remain moderate (Fig. 5). Under such distributional expansion, nominal 95% bootstrap intervals may become under-dispersed because they reflect residual variability learned within the observed data manifold. Within this extended range, the KAN model, unlike MLP and, to the same extent, SVR, successfully captures the highest synthetic values without underestimation. This capability is reflected in its performance metrics under extreme synthetic flow conditions, with $R^2 = 0.854$ at Senta and $R^2 = 0.952$ at Szeged (Table 1). Such behaviour represents an advantage when extrapolating toward high-flow conditions. Notably, even though the associated CI coverage for copula-generated dataset is reduced, when restricting evaluation to the upper tail (Q above the 95th percentile), coverage increases PICP ranges from 0.75 to 0.77 and from 0.78 to 0.89 across models at Senta and Szeged, respectively), indicating that most under-coverage in the full synthetic set is driven by low- and medium-flow samples rather than the targeted high-flow extremes.

4. Discussion

The comparative analysis of the modelling approaches (SVR, MLP, and KAN) reveals clear differences in their ability to represent nonlinear and hydraulically complex rating curves at stations influenced by backwater effects. These findings reinforce previous research demonstrating that classical power-law formulations often fail under conditions of hydraulic nonstationarity (Ajmera and Goyal, 2012; Rozos et al., 2022), whereas state-of-the-art ML methods are more effective in capturing localized nonlinearities (Dasgupta et al., 2025; Nguyen et al., 2022; Vishwakarma et al., 2023). Moreover, the application of a physically based approach for deriving a rating curve relationship (e.g. "twin-gauge") is hampered in the presented case study, as the backwater effects of the Novi Bečej reservoir influence the regime of both hydrological stations. The water surface slope in the considered river reach is extremely low (1.8×10^{-5}), further precluding the application of such an approach.

The results indicate that the novel deep learning approach (KAN) can outperform both classical regression and commonly applied ML models (SVR and MLP), particularly across the hydrological conditions (Table 1), and more importantly within the low- and high-flow regimes (Fig. 6). The KAN model achieves R^2 values ranging from 0.94 to 0.98 respectively for Senta and Szeged, together with high confidence-interval coverage (the PICP values reached 0.94). In contrast, for small river basins characterized by natural hydrological regimes, reported R^2 values for classical linear regression typically range between 0.86 and 0.92 (Kumar and Sen, 2024). Under comparable hydrological conditions, MLP models have achieved R^2 values of up to 0.96 (Rozos et al., 2022), whereas XGBoost models generally exhibit more moderate predictive skill, with R^2 values of approximately 0.70 (Baruah et al., 2025). Similar performance levels have been observed for alternative ML techniques, where SVM achieved $R^2 = 0.792$, Random Subspace $R^2 = 0.937$, and M5-pruned models $R^2 = 0.841$ (Dinesh Kumar Vishwakarma et al., 2023). More recently, hybrid deep ML architectures, such as Transformer- Convolutional neural network combination, have demonstrated R^2 approaching 0.98 (Feizi et al., 2025). However, these results have been obtained for small catchments (~169 km^2) under semi-arid climatic conditions. Similarly, hybrid physics-informed machine-learning frameworks, incorporating additional model-correction terms, have suggested R^2 values of up to 0.98, while relying on considerably more complex architectures (Langhammer et al., 2025).

The proposed framework, which integrates data preprocessing techniques, outlier detection, and machine-learning models, and is evaluated through two testing phases using observed and statistically generated data (Section 2.2), is well suited for practical engineering applications. Its main advantage lies in its applicability to large and low-slope river basins, where backwater effects significantly influence the rating curve (Hidayat et al., 2011; Rojas et al., 2020; Kiss et al., 2024). The framework is inherently adaptive in terms of robustness to data spread. However, the ability to adapt to morphological changes would require the implementation of a sliding window approach for ML model training, which is not applied in this study due to the short recorded period. In the presented methodology, the model is trained using point-by-point mapping of stage-flow values for the post-dam period.

Building on this, the framework is designed to model the flow–stage relationship on a multi-annual basis, using daily observations from the post-dam period. Within this temporal scope, it captures the overall variability of the rating curve but does not explicitly model or detect interannual variability. Interannual shifts in the stage–flow relationship can have a significant impact on high-flow estimation, highlighting the importance of accounting for such dynamics in operational applications (Kona and Bhowmik, 2025). Gradual shifts driven by progressive morphological adjustments, vegetation encroachment, or long-term sediment redistribution are therefore not resolved within the current formulation, and may contribute to additional uncertainty in flow observations (Di Baldassarre and Montanari, 2009).

Among the tested ML models, SVR consistently achieves the highest PICP values, indicating that its kernel-based formulation provides strong resilience to rating-curve variability and produces conservative uncertainty envelopes (He et al., 2023; Maghrebi and Vatanchi, 2024). However, this robustness is accompanied by wider CIs, as observed for both measured and copula-generated datasets

(Figs. 5 and 6). In contrast, KAN exhibits a more balanced behavior although its PICP values are lower than those of SVR, they remain higher than those of MLP across both observed and synthetic datasets, which is consistent with its locally adaptive functional representation (Liu et al., 2024; Panahi et al., 2025; Cui et al., 2025). MLP shows the lowest PICP values for both observed and synthetic Q-H pairs, suggesting that its fixed activation functions limit its ability to represent uncertainty under hydraulic variability. For the copula-generated datasets, reduced PICP should be interpreted as a coverage limitation arising under distributional extension. Accordingly, caution is required when interpreting 95% intervals under distribution-extended flows. Taken together, the findings indicate that the KAN model represents a well-balanced compromise between model complexity and predictive reliability for rating-curve estimation. In the presented study, extrapolation is performed up to the 100-year return-period stage levels, corresponding to 83.80 m a.s.l. at Szeged and 82.06 m a.s.l. at SentaPredictions beyond these levels should be treated with caution, as the model is not constrained by physical hydraulic principles and its behavior in more extreme ranges remains unverified.

A further limitation of the proposed framework concerns the scope of uncertainty quantification. In the present study, uncertainty is characterized primarily through residual-based confidence intervals, without explicitly accounting for measurement uncertainty. Measurement uncertainty in river flow is substantially influenced by hydraulic conditions and gauging methodology, and can propagate nonlinearly into flow estimates (McMillan et al., 2012). The magnitude of such measurement uncertainty can be significant enough to affect estimated flows, underscoring the importance of its explicit incorporation into robust rating curve estimation (Bhowmik et al., 2020).

In addition, particular attention should be given to the optimization of KAN for rating curve estimation, as replacing scalar weights with learned univariate functions substantially increases the number of parameters and may introduce training instability (Ji et al., 2024). While KANs offer strong universality guarantees, these relate primarily to the existence of a solution rather than to efficient learnability or robust generalization. Consequently, in practical settings, KANs may exhibit overfitting and do not consistently demonstrate advantages over well-tuned MLPs.

5. Conclusion

In this paper, a joint ML–Copula framework is presented to capture rating curve while accounting for reservoir-induced backwater effects and long-term morphological adjustments. By combining LOESS-based outlier detection, a Gumbel copula for synthetic data extension, and four modelling approaches (Power regression, SVR, MLP, and KAN), the framework provides reliable estimation of rating curve across a wide range of hydrological regimes. The approach overcomes limitations of traditional methods that assume stationarity, which is often unrealistic for low-slope rivers influenced by hydraulic structures such as reservoirs. The framework is applied to two hydrological stations (Senta in Serbia and Szeged in Hungary) on the low-slope Tisza River, where hydraulic conditions are strongly affected by the downstream Novi Bečej reservoir and long-term morphological changes. Its performance is evaluated under both observed and synthetic flow–stage datasets to assess robustness in a backwater-affected lowland river system.

The results provided support the following conclusions:

- ML rating curve methods developed here (SVR, MLP and KAN) outperform classical power regression across observed conditions, particularly in error reduction and representation of nonlinearities in rating curves. For measured Q–H pairs, these ML models achieve RMSE $\approx$ 78–163 m^3 /s, lower than the RMSE $\approx$ 80–173 m^3 /s of the power-law curve. These improvements reflect the ability of ML models to flexibly approximate complex, hydraulic-influenced nonlinearities, especially in the high-flow domains where classical regressions exhibit bias. Even under synthetic copula-generated datasets, where flows extend beyond maxima in verification dataset, ML models preserve stable accuracy with RMSE increases limited from 20% to 30%.
- The comparative performance of ML models developed in this work highlights trade-offs between accuracy, interval coverage probability, and response to low and high flows. SVR exhibits the highest PICP values (0.72–0.94), indicating strong uncertainty-envelope coverage but at the cost of wider confidence intervals. MLP performs well within the central range but shows underprediction at extreme flows, seen in both RMSE increases and lower PICP (0.65–0.76). The power regression shows asymmetry in confidence intervals, signaling poor adaptability to backwater-impacts. These differences highlight that model selection depends on the desired balance between accuracy and conservativeness.
- KAN model yields overall-best performance with robust rating curve estimates across both observed and synthetic datasets. It maintains smooth transitions between low-, medium-, and high-flow regimes, avoiding the sharp regime breaks and underestimations that appear in MLP and SVR under synthetic flow extremes. It is also the only model that preserves realistic behavior across synthetic dataset, an essential feature for flood reservoir management and ecological flow release, this demonstrates its promise as a next-generation approach for adaptive flow–stage curve estimation under the backwater impacts and morphological changes in low-land river stream.
- The ML-Copula framework provides a scalable and transferable methodological foundation for adaptive rating curve estimation to other rivers experiencing morphological changes, backwater effects, or nonstationary hydraulic regimes. The joint use of ML and copulas offers a flexible structure that can support operational water management, reservoir regulation, and long-term assessment of hydrological conditions.

Future work should expand the ML–Copula framework to include physics-informed components that directly account for backwater dynamics, sediment transport, and evolving channel geometry. Additional research should evaluate model transferability across multiple river systems with varying hydraulic controls to assess generalization capabilities. Furthermore, future studies should systematically examine the influence of different synthetic data characteristics on model performance, and a formal copula sensitivity

analysis is proposed as a valuable direction for further investigation.

CRediT authorship contribution statement

Veljko Prodanović: Writing – review & editing, Writing – original draft, Supervision, Investigation, Conceptualization. **Luka Vinokić:** Formal analysis. **Zoran Kapelan:** Writing – review & editing, Writing – original draft, Supervision, Conceptualization. **Milan Stojković:** Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis. **Milan Dotlić:** Formal analysis. **Slobodan Kolaković:** Project administration, Funding acquisition, Data curation.

Declaration of Generative AI and AI-assisted technologies in the writing process

We confirm that AI tools (e.g., ChatGPT by OpenAI) were used only to assist in the preparation of non-scientific sections of this manuscript, such as language refinement and formatting. The scientific content, including study design, data analysis, interpretation, and conclusions, were entirely conceived, conducted, and validated by the authors. The authors take full responsibility for the integrity and accuracy of the manuscript.

Declaration of Competing Interest

On behalf of the authors, I confirm that.

- this manuscript is original and has not been published previously.
- it is not under consideration for publication elsewhere.
- all authors have approved the manuscript and agree with its submission.

Milan Stojkovic

Acknowledgement

This research was supported by the European Union's Horizon Europe project ARTIFACT under Grant Agreement No. 101159480 Support was also provided by the Interreg IPA Hungary-Serbia project ADAPTisa (HUSRB/23R/11/006), through the associated study on the hydrological database and the analysis of the hydrological regime of the Serbian section of the Tisza River.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ejrh.2026.103407.

Data availability

Data will be made available on request.

References

Ajmera, T.K., Goyal, M.K., 2012. Development of stage–discharge rating curve using model tree and neural networks: an application to peachtree Creek in Atlanta. Expert Syst. Appl. 39 (5), 5702–5710.
Alhamid, A.K., Akiyama, M., He, Z., Firdaus, P.S., Frangopol, D.M., 2024. LRFD methodology for river embankments against non-stationary flooding under climate change. Struct. Saf. 109, 102477.
Baruah, A., Zarrabi, R., Cohen, S., Johnson, J.M., McDermott, R., 2025. Interpretable machine learning for predicting rating curve parameters using channel geometry and hydrological attributes across the United States. Sci. Rep.
Bhowmik, R.D., NG, T.L., Wang, J.P., 2020. Understanding the impact of observation data uncertainty on probabilistic streamflow forecasts using a dynamic hierarchical model. Water Resour. Res. 56 (4).
Brunner, M.I., Slater, L., Tallaksen, L.M., Clark, M., 2021. Challenges in modeling and predicting floods and droughts: a review. Wiley Interdiscip. Rev. Water 8 (3), e1520.
Chatterjee, S., Hadi, A.S., 2006. Regression Analysis by Example, 4Forth ed. Wiley.
Church, M., 2015. Channel stability: morphodynamics and the morphology of rivers. Rivers–physical, fluvial and environmental processes. Springer, pp. 281–321.
Cui, S., Cao, M., Liao, Y., Wu, J., 2025. Physics-informed Kolmogorov–Arnold networks: Investigating architectures and hyperparameter impacts for solving Navier–Stokes equations. Phys. Fluids 37 (3).
Dasgupta, R., Bhar, A., Das, S., Das, R., Banerjee, G., Mazumdar, A., 2025. Novel mathematical expression for dynamic stage-discharge relationship of rivers under flow unsteadiness explored through machine learning models via symbolic regression in PySR. J. Hydrol., 133947
Di Baldassarre, G., Montanari, A., 2009. Uncertainty in river discharge observations: A quantitative analysis. Hydrol. Earth Syst. Sci. 13 (6), 913–921.
Dodig, A., Ricci, E., Kvascev, G., Stojkovic, M., 2024. A novel machine learning-based framework for the water quality parameters prediction using hybrid long short-term memory and locally weighted scatterplot smoothing methods. J. Hydroinformatics 26 (5), 1059–1079.

Domínguez, R., M., Arganis, M.L., J., 2012. Validation of methods to estimate design discharge flow rates for dam spillways with large regulating capacity. Hydrol. Sci. J. 57 (3), 460–478.
Dottori, F., Martina, M.L.V., Todini, E., 2009. A dynamic rating curve approach to indirect discharge measurement. Hydrol. Earth Syst. Sci. 13 (6), 847–863.
Feizi, H., Sattari, M.T., Milewski, A., 2025. Improving stage-discharge relationship modeling accuracy using a hybrid ViT-CNN framework. Sci. Rep. 15 (1), 38031.
Genest, C., Favre, A.C., 2007. Everything you always wanted to know about copula modeling but were afraid to ask. J. Hydrol. Eng. 12 (4), 347–368.
Grimaldi, S., Serinaldi, F., 2006. Asymmetric copula in multivariate flood frequency analysis. Adv. Water Resour. 29 (8), 1155–1167.
Hastie, T., Tibshirani, R., Friedman, J., Franklin, J., 2005. The elements of statistical learning: data mining, inference and prediction. Math. Intell. 27 (2), 83–85.
Hellmers, S., Fröhle, P., 2022. Computation of backwater effects in surface waters of lowland catchments including control structures – an efficient and re-usable method implemented in the hydrological open-source model Kalypso-NA (4.0). Geosci. Model Dev. 15 (3), 1061–1077.
Hidayat, H., Vermeulen, B., Sassi, M.G., Torfs, P.J.J.F., Hoitink, A.J.F., 2011. Discharge estimation in a backwater affected meandering river. Hydrol. Earth Syst. Sci. 15, 2717–2728.
ICPDR, 2008. Analysis of the Tisza River Basin 2007: Initial step toward the Tisza River Basin Management Plan – 2009. ICPDR Secretariat.
Jain, S.K., Singh, V.P., 2019. Engineering Hydrology: An introduction to processes, analysis, and modeling, First ed. McGraw-Hill Education.
Jánosi, I.M., Zsuffa, I., Bíró, T., Lakatos, B.O., Szöllősi-Nagy, A., Hetesi, Z., 2025. The dynamics of lowland river sections of Danube and Tisza in the Carpathian Basin. Front. Earth Sci. 13, 1391458.
Ji, T., Hou, Y., & Zhang, D. (2024). A comprehensive survey on kolmogorov arnold networks (KAN). *arXiv preprint arXiv:2407.11075*.
Kisi, O., Azamathulla, H.M., Cevat, F., Kulls, C., Kuhdaragh, M., Fuladipanah, M., 2024. Enhancing river flow predictions: comparative analysis of machine learning approaches in modeling stage–discharge relationship. Results Eng. 22, 102017.
Kiss, T., Amissah, G.J., Fiala, K., 2019. Bank processes and revetment erosion of a large lowland river: case study of the lower Tisza River, Hungary. Water 11 (6), 1313.
Kiss, T., Tóth, M., Török, G.T., Sipos, G., 2024. Reconstruction of a long-term, reach-scale sediment budget using lateral channel movement data as a proxy: A case study on the lowland section of the Tisza River, 11. Hydrology, Hungary, p. 67.
Kolaković, S., Kolaković, S., Kovács, S., 2018. Upravljanje poplavama na velikim rekama kroz istoriju i nove strategije u svetlu klimatskih promena. Vodoprivreda 50 (1), 21–29.
Kona, K.S., Bhowmik, R.D., 2025. Understanding the interannual variation in river stage-discharge relationship and its impact on flood discharge estimation. ASCE J. Hydrol. Eng. 30 (4).
Kumar, V., Sen, S., 2024. Rating curve development and uncertainty analysis in mountainous watersheds for informed hydrology and resource management. Front. Water 5, 1323139.
Langhammer, J., Šobr, M., Ba, D., 2025. Machine learning model for stage-discharge curve calculation. Auc Geographica 60 (2), 296–315.
Lee, Y.S., Scholtes, S., 2014. Empirical prediction intervals revisited. Int. J. Forecast. 30 (2), 217–234.
Liro, M., Nones, M., Mikuś, P., Plesiński, K., 2022. Modelling the effects of dam reservoir backwater fluctuations on the hydrodynamics of a small mountain stream. Water 14 (19), 3166.
Liu, Y., Wang, S., Vaidya, F., Ruehle, J., Halverson, M., Soljacic, T.Y., Hou, M., & Tegmark, M. (2024). KAN: Kolmogorov–Arnold networks. *arXiv:2404.19756*.
Loader, C., 1999. Local regression and likelihood. Springer.
Maghrebi, M.F., Vatanchi, S.M., 2024. Comparison of different machine learning methods in river streamflow estimation using isovel contours and hydraulic variables. Int. J. River Basin Manag. 1–18.
McMillan, H., Krueger, T., Freer, J., 2012. Benchmarking observational uncertainties for hydrology: Rainfall, river discharge and water quality. Hydrol. Process 26 (26), 4078–4111.
Mohsen, A., Kovács, F., Mezősi, G., Kiss, T., 2021. Sediment transport dynamism in the confluence area of two rivers transporting mainly suspended sediment based on Sentinel-2 satellite images. Water 13 (21), 3132.
Nelsen, R.B., 2006. An Introduction to Copulas, Second ed. Springer.
Nguyen, A.D., Tran, D.H., Nguyen, Q.T., Dang, T.D., 2022. Accurate discharge and water level forecasting using radial basis neural network and genetic algorithm. Environ. Model. Softw. 150, 105338.
Panahi, S., Moradi, M., Bollt, E.M., Lai, Y.C., 2025. Data-driven model discovery with Kolmogorov-Arnold networks. Phys. Rev. Res. 7 (2), 023037.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Duchesnay, É., 2011. Scikit-learn: machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830.
Plavšić, J., 2019. *Inženjerska hidrologija*. Građevinski fakultet. Univerzitet u Beogradu.
Rantz, S.E., Russell, G.M., Osborne, J.L., Williams, G.P., Hirsch, R.M., Cobb, E.D., 1982. Measurement and computation of streamflow. In: Measurement of stage and discharge (Water-Supply Paper 2175)., 1. U.S. Government Printing Office.
Richter, B.D., Baumgartner, J.V., Powell, J., Braun, D.P., 1996. A method for assessing hydrologic alteration within ecosystems. Conserv. Biol. 10 (4), 1163–1174.
Rojas, M., Quintero, F., Young, N., 2020. Analysis of stage–discharge relationship stability based on historical ratings. Hydrology 7 (2), 31.
Rozos, E., Leandro, J., Koutsoyiannis, D., 2022. Development of rating curves: Machine learning vs. statistical methods. Hydrology 9 (10), 166.
Sivapragasam, C., Muttil, N., 2005. Discharge rating curve extension – a new approach. Water Resour. Manag. 19, 505–520.
Spatial plan, 2023. Prostorni plan područja posebne namene brane na Tisi – Nacrt plana. Zavod za Urban. Vojv.
Stipić, M., Budinski, L., Fabian, J., 2021. Matematičko modeliranje tečenja i pronosa nanosa – primjer brane na rijeci Tisi. Građevinar 73 (9), 917–932.
Stojković, M., Ilić, A., Prohaska, S., Plavšić, J., 2014. Multi-temporal analysis of mean annual and seasonal stream flow trends, including periodicity and multiple nonlinear regression. Water Resour. Manag. 28 (12), 4319–4335.
Stojković, M., Kostić, S., Prohaska, S., Plavšić, J., Tripković, V., 2017. A new approach for trend assessment of annual streamflows: a case study of hydropower plants in Serbia. Water Resour. Manag. 31 (4), 1089–1103.
Stojković, M., Plavšić, J., Prohaska, S., Pavlović, D., Despotović, J., 2020. A two-stage time series model for monthly hydrological projections under climate change in the Lim River basin (southeast Europe). Hydrol. Sci. J. 65 (3), 387–400.
Stojković, M., Prohaska, S., Plavšić, J., 2015. Stochastic structure of annual discharges of large European rivers. J. Hydrol. Hydromech. 63 (1), 63–70.
USGS, 2011. USGS stage-discharge relation example. U.S. Department of the Interior. ⟨https://www.usgs.gov/media/images/usgs-stage-discharge-relation-example⟩.
Vinokić, L., Dotlić, M., Prodanović, V., Kolaković, S., Simonovic, S.P., Stojković, M., 2025. Effectiveness of three machine learning models for prediction of daily streamflow and uncertainty assessment. Water Res. X 27, 100297.
Vishwakarma, D.K., Kuriqi, A., Abed, S.A., Kishore, G., Al-Ansari, N., Pandey, K., … & Jewel, A. (2023). Forecasting of stage–discharge in a non-perennial river using machine learning with gamma test. *Heliyon, 9*(5).
Wang, J., Chen, Q., Chen, Y., 2004. RBF kernel based support vector machine with universal approximation and its application. International Symposium on Neural Networks. Springer, pp. 512–517.
Yu, J., Li, Y., Huang, X., Ye, X., 2025. Data quality and uncertainty issues in flood prediction: a systematic review. Int. J. Digit. Earth 18 (1), 2495738.