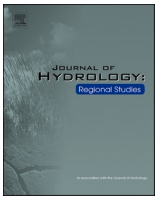
Journal of Hydrology: Regional Studies 62 (2025) 102774

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

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



# Modeling hydrological functioning of karst aquifer systems in Slovenia using geomorphological features and random forest algorithm

Mitja Janža [a,*], Valter Hudovernik [b], Luka Serianz [a], Andrej Stroj [c]

[a] Geological Survey of Slovenia, Dimičeva ulica 14, Ljubljana 1000, Slovenia
[b] Faculty of Computer and Information Science, University of Ljubljana, Večna pot 113, Ljubljana 1000, Slovenia
[c] Croatian Geological Survey, Milana Sachsa 2, Zagreb 10 000, Croatia

ARTICLE INFO

*Keywords:*
Karst aquifer
Random forest
Machine learning
Ungauged catchment
Spring discharge
Recession curve analysis

ABSTRACT

*Study region:* Slovenia
*Study focus:* This study investigates the relationship between the hydrological functioning of karst aquifer systems and the geomorphological characteristics of their catchments. It is based on the analysis of discharge time series from 15 karst springs. Hydrograph analysis of these time series was used to estimate 11 hydrological parameters that characterize aquifer system functioning. A spatial analysis of morphological, geological, and hydrological data was carried out to assess 7 lumped geomorphological features of the catchments. These features (independent variables) and hydrological parameters (dependent variables) were used to develop random forest models for predicting the hydrological functioning of karst springs.
*New hydrological insights for the region:* The developed methodological approach provides a basis for improved characterization and prediction of the hydrological functioning of ungauged karst systems. Groundwater availability in these systems is largely controlled by aquifer retention capacity and spring discharge variability. These characteristics can be inferred from hydrological parameters that can be predicted using the developed random forest models. Feature importance analysis indicated that catchment area, cave density, and slope gradient are the most important geomorphological features for predicting the hydrological characteristics of spring discharge.

## 1. Introduction

Karst aquifers represent a vital source of drinking water for nearly 10 % of the global population and extend across approximately 14 % of the Earth's ice-free surface (Stevanovicć, 2019; Goldscheider et al., 2020). Due to their distinctive hydrogeological properties, they are particularly susceptible to contamination and difficult to monitor and manage effectively. These challenges underscore the need for a better understanding of karst system behavior, especially in regions where water supply heavily depends on these aquifers. In Slovenia, for example, karst covers nearly half of the national territory, and about 50 % of the drinking water originates from karst sources (Gostinčar and Stepišnik, 2023; Kovačič et al., 2012). Improving knowledge of karst hydrodynamics is therefore crucial for ensuring long-term protection and sustainable use of water resources in such areas.

* Corresponding author.
*E-mail addresses:* mitja.janza@geo-zs.si (M. Janža), valter.hudovernik@gmail.com (V. Hudovernik), luka.serianz@geo-zs.si (L. Serianz), astroj@hgi-cgs.hr (A. Stroj).

https://doi.org/10.1016/j.ejrh.2025.102774
Received 5 June 2025; Received in revised form 29 August 2025; Accepted 11 September 2025
Available online 19 September 2025

Karst aquifers are widely recognized as some of the most heterogeneous and complex groundwater flow systems (Ford and Williams, 2007; Goldscheider and Drew, 2007; Jourde and Wang, 2023; Verbovšek, 2008). They typically discharge through a limited number of large springs. The discharge, turbidity, and chemistry of water from karst springs provide information about processes occurring in the groundwater basin along the flow paths that are generally inaccessible. Due to the high complexity and heterogeneity of groundwater flow in karst aquifers, their hydrogeologic characterization is typically based on the analysis of the discharge time series of the main spring outlets of the aquifer system (e.g., Fiorillo, 2014; Kovács et al., 2005; Kovács and Sauter, 2008; Maillet, 1905). Consequently, the spring discharge of a karst system is considered as the basic information in karst hydrogeology (Cinkus et al., 2021; Dufoyer et al., 2019; Mangin, 1984). The discharge of a karst spring is commonly divided into two components. The baseflow, which comprises the slow contribution from the saturated zone and fast flow which is typically a response to rainfall and represents the fast contribution from the infiltration zone. The ratio between these two components is an important parameter for characterizing the vulnerability of the karst aquifer to pollution, its water availability and its resilience to droughts (Padilla et al., 1994). Groundwater flow in karst aquifers occurs through two primary pathways: concentrated turbulent flow in conduits and diffuse laminar flow in fractures, fissures, and the low-permeability rock matrix (Liedl et al., 2003).

Several studies have proposed classification frameworks as a fundamental tool for understanding of karst aquifer hydrodynamics (e.g., Cinkus et al., 2021; Hobbs and Smart, 1986; Kiraly, 1994; Kovács, 2021; Mangin, 1975). These classifications are typically based on key hydrodynamic indicators derived from spring discharge data, hydrograph analysis, and other hydrogeological parameters. Cinkus et al. (2021) identified the most relevant indicators of karst hydraulics for the proposed classification using multivariate analyses of spring discharge data. Traditionally, the hydrological properties of karst aquifers have been studied using karst spring recession hydrographs, which provide an insight into the storage and drainage characteristics of the aquifer (e.g., Brenčič, 2000; Kovács, 2021; Kovács and Perrochet, 2008; Kovács et al., 2005; Kresic and Bonacci, 2010; Li et al., 2016; Malík and Vojtková, 2012; Mangin, 1975; Olarinoye et al., 2022; Xu et al., 2018). Recently, machine learning models have shown great potential for analyzing environmental factors and hydrological processes for use in water resources management (e.g., Aguilera et al., 2022; Fernandes et al., 2024; Gaertner, 2023; Gomez et al., 2024; Koch et al., 2024; Patra et al., 2024). These methods are particularly valuable for studying nonlinear and nonstationary behaviors, which are characteristic of karst aquifers (Zhang et al., 2022; Zhou and Zhang, 2023). By integrating machine learning techniques with traditional hydrogeological approaches, it is possible to gain deeper insights into spring discharge dynamics, improve aquifer characterization and develop more effective water resource management strategies.

To gain a deeper understanding of the relationship between the physical characteristics of karst systems and their hydrological functioning, and to support the development of predictive models for ungauged catchments, advanced machine learning techniques such as the Random Forest (RF) algorithm (Breiman, 2001) are increasingly being employed (Dasgupta et al., 2024). RF is a robust, regularized algorithm that requires minimal parameter tuning, making it particularly suitable for applications with limited datasets (Biau and Scornet, 2016). Hydrological processes are often difficult to interpret due to non-linear interactions among climatic and hydrological variables, as well as the complexity of parameter relationships. As a result, the RF algorithm has gained widespread use in water resources research in recent years (Tyralis et al., 2019). This includes applications focused on the prediction of flow parameters at ungauged sites (Booker and Snelder, 2012; Booker and Woods, 2014; Brunner et al., 2018), streamflow forecasting (Bond and Kennard, 2017; Fang et al., 2024; Li et al., 2019; Masrur Ahmed et al., 2021; Moges et al., 2024; Pham et al., 2021;), and groundwater hydrology and characterization (Fiedler, 2024; Naghibi et al., 2017; Rahmati et al., 2016; Rodriguez-Galiano et al., 2014; Singh et al., 2018; Wang et al., 2018; Xu et al., 2017).

This study investigates the relationship between catchment-scale geomorphological characteristics and the hydrological functioning of karst aquifer systems. Karst-related surface features and terrain characteristics, derived from publicly available regional-scale geospatial datasets, are analyzed and the potential of using these surface geomorphological indicators to infer karst aquifer hydrodynamics is evaluated. Discharge time series of 15 karst springs in Slovenia and the geomorphological and geological characteristics of their catchments are analyzed. The analysis is based on spring discharge time series analysis and RF modeling. The approach was developed to support the assessment of groundwater availability in karst aquifers, with a particular focus on applications in ungauged catchments. Groundwater availability in these systems is largely controlled by internal flow and storage dynamics, particularly the aquifer's retention capacity and the variability of spring discharge.

**Table 1**
Description of geomorphological parameters (independent variables).

| Parameter | Unit | Abbreviation | Description and source of data |
|---|---|---|---|
| Catchment area | $km^2$ | AREA | The area of the delineated catchment the spring. |
| Mean slope gradient | degrees (°) | Slope | The mean slope gradient derived from the Digital Elevation Model with a grid size 5 × 5 m (GURS, 2022). |
| Karst caves density | No. of karst caves/$km^2$ | Caves | The number of karst caves per square kilometer of the catchment area, derived from the cave cadaster using the locations of cave entrances (JZS, 2022). |
| Permanent streams density | $km^{-1}$ | STREAMS | The length of permanent streams per unit area of the catchment (DRSV, 2022). |
| Portion of karst rock outcrops | % | Karst_ro | The ratio of the area of outcropping lithological units susceptible to karstification to the total area of the catchment. (Novak et al., 2014). |
| Volume of dolines | m | V_doline | The volume of dolines per unit area of the catchment (Mihevc and Mihevc, 2021). |
| Active faults density | $km^{-1}$ | Faults | The total linear length of mapped active faults (Atanackov et al., 2021) within a catchment, normalized by the catchment area. |

## 2. Methodology

### 2.1. Workflow

The study was based on the selection of karst springs in Slovenia, which are the main outlets of karst aquifer systems and have continuous discharge measurements with daily resolution. A geomorphological analysis of the selected catchments was then conducted, assessing parameters such as surface area, mean slope gradient, karst cave density, density of permanent streams, portion of karst outcrops, volume of dolines, and density of active faults (Table 1). This analysis used spatial datasets covering the entire set of spring catchments to ensure consistent data resolution and quality across the study region. Subsequently, the spring discharge time series were analyzed using basic statistical methods and recession curve analysis to assess parameters characterizing the hydrological behavior of the studied karst aquifer systems. These parameters were used as dependent variables (targets), while the geomorphologic characteristics of the catchments, assessed in the previous step, were used as independent variables (features) for training the RF models. The performance of the RF models was validated using the cross-validation technique and the importance of each geomorphological variable in predicting the hydrological parameters of the karst aquifer system was determined.

### 2.2. Study area - selected karst aquifer systems

The selection of the karst aquifer systems analyzed in the study is based on several criteria. Springs were selected that are the main outlets of karst aquifers and have discharge time series with daily resolution, covering more than one year of continuous measurements. The catchment areas of the springs are mainly composed of carbonate rocks and the discharge measured at the gauging stations is close to the spring outlet, thereby minimizing the influence of surface water flow or tributaries. Based on these criteria, 15 karst aquifer systems were selected and named after the corresponding spring or river (Fig. 1, Table 2).

### 2.3. Geomorphological analysis of karst spring catchment areas

The catchment boundaries of the selected karst systems were delineated based on the results of previous hydrogeological studies (Janža et al., 2011; Mezga, 2014; Pavlič et al., 2008; Prestor et al., 2004). A GIS spatial analysis of morphological, geological, and hydrological data (Fig. 2) was then conducted to assess seven lumped parameters that characterize the catchments (Table 1).

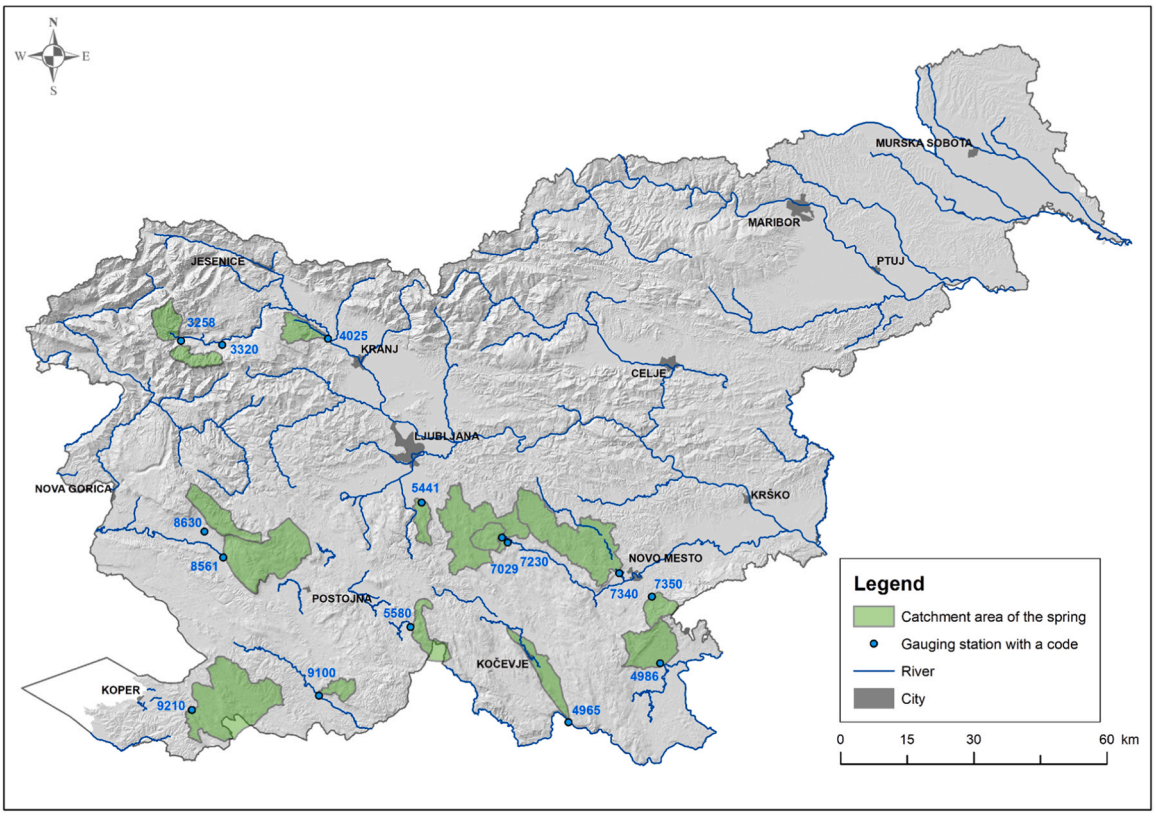



**Fig. 1.** The locations of selected karst aquifer system catchments and the corresponding gauging stations.

**Table 2**
Values of lumped parameters characterizing the geomorphological properties of karst aquifer system catchments (independent variables).

| Gauging Station ID | Aquifer system | Catchment area (km²) | Mean slope gradient (°) | Karst caves density (no. of karst caves/ km²) | Permanent streams density (km/km²) | Share of karst rocks (%) | Volume of dolines (m³/ m²) | Active faults density (km/ km²) |
|---|---|---|---|---|---|---|---|---|
| 3258 | SAVICA | 43.31 | 25.04 | 4.59 | 0.11 | 88.09 | 0.02 | 0 |
| 3320 | BOHINJSKA BISTRICA | 35.15 | 30.23 | 3.02 | 0.04 | 75.81 | 4.37 | 0 |
| 4025 | LIPNICA | 40.41 | 16.81 | 3.34 | 1.11 | 64.86 | 3.95 | 0.09 |
| 4965 | BILPA | 61.88 | 8.24 | 1.91 | 0.33 | 60.72 | 1.15 | 0.18 |
| 4986 | KRUPA | 93.98 | 10.19 | 1.23 | 0.03 | 87.83 | 1.4 | 0.13 |
| 5441 | IŽICA | 28.34 | 12.87 | 0.56 | 0.49 | 33.12 | 0.95 | 0.76 |
| 5580 | VELIKI OBRH PRI LOŽU | 49.89 | 12.96 | 0.5 | 0.03 | 78.84 | 0.45 | 0.31 |
| 7029 | KRKA | 206.97 | 9.41 | 1.23 | 0.58 | 59.85 | 0.69 | 0.18 |
| 7230 | POLTARICA | 33.27 | 10.78 | 2.1 | 0 | 97.23 | 1.34 | 0.09 |
| 7340 | PREČNA | 214.14 | 9.32 | 0.96 | 0.12 | 65.66 | 0.86 | 0.06 |
| 7350 | TEŽKA VODA | 34 | 12.12 | 0.58 | 0.01 | 35.96 | 1.58 | 0 |
| 8561 | VIPAVA | 179.47 | 15.46 | 2.36 | 0.14 | 88.88 | 0.69 | 0.11 |
| 8630 | HUBELJ | 68.89 | 18.97 | 1.53 | 0 | 93.54 | 2.18 | 0.26 |
| 9100 | ILIRSKA BISTRICA | 23.92 | 12.67 | 0.75 | 0.08 | 98.97 | 0.8 | 0 |
| 9210 | RIŽANA | 244.68 | 11.86 | 2.28 | 0.26 | 71.41 | 0.62 | 0.21 |
| **Minimum** | | **23.92** | **8.24** | **0.50** | **0.00** | **33.12** | **0.02** | **0.00** |
| **Maximum** | | **244.68** | **30.23** | **4.59** | **1.11** | **98.97** | **4.37** | **0.76** |
| **Median** | | **49.89** | **12.67** | **1.53** | **0.11** | **75.81** | **0.95** | **0.11** |
| **Average** | | **90.55** | **14.46** | **1.80** | **0.22** | **73.38** | **1.40** | **0.16** |

### *2.4. Analysis of spring discharge time series*

Karst spring discharge time series were obtained from the Slovenian Environment Agency, which provides online access to its datasets (ARSO, 2022). The time series datasets for the selected karst springs cover time spans from 20 months to a maximum of 20 years (Table 3, Appendix 1).

The spring discharge time series were analyzed using the XLKarst tool (Bailly-Comte et al., 2023), developed in Visual Basic for use with Microsoft Excel. To identify the hydrological functioning of the selected karst aquifer systems, we computed basic statistics and analyzed recession curves to assess the parameters of the spring discharge recession models:

- basic statistics: mean, median, minimum (Min), maximum (Max);
- the coefficient of variation (CV) - the ratio of the standard deviation of a time series to itsmean (Flora, 2004; Mangin, 1975);
- the spring variability coefficient (SVC) - the ratio between percentile 0.1 and 0.9 (Flora, 2004),
- the base flow index (BFI) - the volume ratio of baseflow to the total flow, the separation of baseflow is based on the digital filter of Lyne and Hollick (1979) according to the standard approach of Ladson et al. (2013) using a filter coefficient value of 0.91 and employing three filter passes;
- the ratio $\sigma_{250}/\sigma$ - where $\sigma_{250}$is the standard deviation of the long-term processes assessed using a 250-day moving average filter—serves as an indicator of the system's regulation capacity, i.e., a measure of its long-term versus short-term response (Bailly-Comte et al., 2023);
- regulation time (RT) - corresponds to the half of the spectral density function for the 0 frequency, computed as the discrete Fourier transform of the autocorrelation function using a Tukey lag window (Jenkins and Watts, 1968), with a truncation point m = 125 days of the correlation function, providing an estimate of the duration of the system's long-term response (Mangin, 1984);
- the memory effect (ME) - corresponds to the lag for which the autocorrelation function falls below 0.2, indicating the average response time of the system (Mangin, 1984);
- composite parameter KA - calculated as (σ250/σ)*(median/mean) and reflects both the system's regulation capacity and the uniformity of its discharge recession.

#### *2.4.1. Recession curve analysis*

An analysis of the karst spring discharge recession curves was applied to separately characterize flood recession, influenced by recharge processes, and baseflow recession. This approach follows a conceptual model introduced by Mangin (1975) and considers a two-component recession curve, consisting of fast-flow (mainly through karst conduits) and slow-flow (mainly through the fissure matrix of the carbonate rock) recessions. The results of this analysis help to interpret the recharge dynamics through the infiltration zone and the hydraulic functioning of the storage compartments in the karst aquifer system.

The analysis began with the automatic selection of recessions using the XLKarst tool (Bailly-Comte et al., 2023). The selected recessions were reviewed, inconsistent recessions were removed, and the recession parameters of the remaining selections were

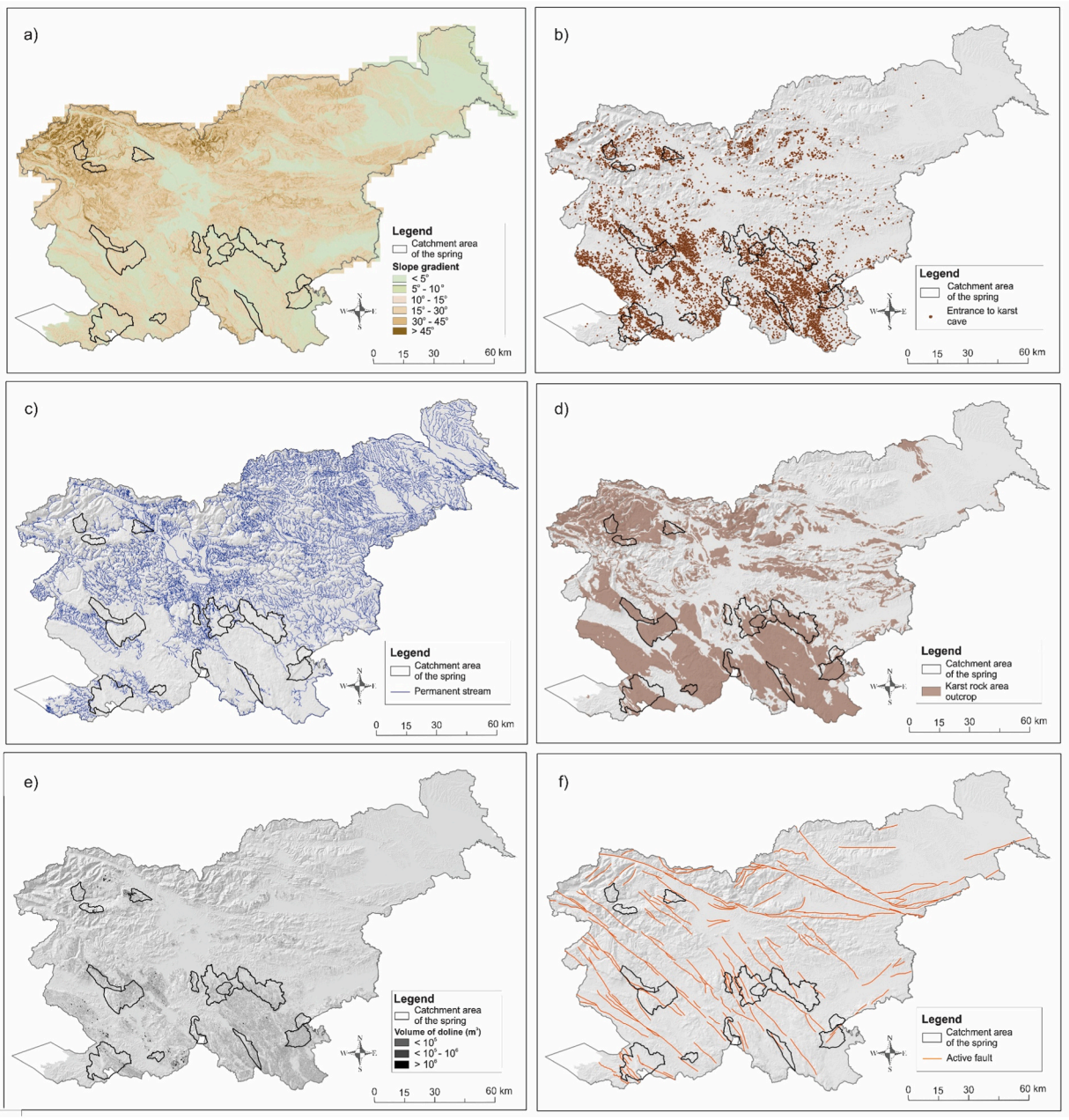



**Fig. 2.** The graphical representation of the spatial data used to assess the geomorphological parameters of the selected kart spring catchment areas; a) Slope gradient, b) Entrances to karst caves, c) Permanent streams, d) Karst rock area, e) Volume of dolines, f) Active faults.

adjusted to improve the fit to the theoretical recession model. The fitted recession curves are presented in Appendix 1. Based on the fitted models, two aquifer parameters were calculated for all selected karst systems: the infiltration delay (i), which describes the rate of the fast-flow response, and the regulating capacity (k), which reflects the maximum system storage volume relative to its average flow. These parameters were then used to classify the selected karst systems into five domains. The classification scheme (El-Hakim and Bakalowicz, 2007; Mangin, 1975) distinguishes karst aquifer systems based on their hydrodynamic functioning.

### *2.5. Random forest modeling*

RF is an ensemble method based on bootstrap aggregation or bagging of decision trees. Bootstrapping is a resampling method based on sampling with replacement from the empirical distribution of the data and is commonly used to estimate the distribution of a parameter of interest (Efron, 1977). In terms of the bias-variance tradeoff, a single decision tree is a highly variable estimator but has a

**Table 3**
Calculated basic statistics and assessed parameters of the spring hydrograph recession models (dependent variables).

| Gauging Station ID | 3258 | 3320 | 4025 | 4965 | 4986 | 5441 | 5580 | 7029 | 7230 | 7340 | 7350 | 8561 | 8630 | 9100 | 9210 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Date from | 01.01. 2016 | 01.01. 2010 | 01.01. 2011 | 01.01. 2014 | 01.01. 2008 | 01.01. 2014 | 01.01. 2010 | 01.01. 2000 | 01.01. 2008 | 01.01. 2000 | 01.01. 2010 | 01.01. 2010 | 01.01. 2017 | 13.06. 2016 | 01.01. 2010 |
| Date to | 11.12. 2020 | 11.12. 2020 | 11.12. 2021 | 19.08. 2015 | 11.12. 2009 | 11.12. 2015 | 11.12. 2015 | 01.02. 2015 | 11.12. 2015 | 11.12. 2020 | 11.12. 2015 | 11.12. 2020 | 11.12. 2020 | 11.12. 2020 | 11.12. 2020 |
| Mean ($m^3/s$) | 4.53 | 3.10 | 2.34 | 3.42 | 3.29 | 2.01 | 2.01 | 7.56 | 1.13 | 4.06 | 0.80 | 6.14 | 2.93 | 1.59 | 4.27 |
| Median ($m^3/s$) | 1.93 | 1.31 | 0.89 | 0.80 | 2.30 | 1.27 | 0.92 | 4.56 | 0.62 | 2.59 | 0.46 | 2.66 | 1.47 | 0.76 | 1.46 |
| Min ($m^3/s$) | 0.05 | 0.12 | 0.29 | 0.25 | 0.69 | 0.55 | 0.19 | 0.90 | 0.07 | 0.91 | 0.07 | 0.81 | 0.50 | 0.18 | 0.27 |
| Max ($m^3/s$) | 84.73 | 63.38 | 44.59 | 30.23 | 32.03 | 12.88 | 30.12 | 102.4 | 8.13 | 44.87 | 8.72 | 81.43 | 35.66 | 6.63 | 63.39 |
| SVC (-) | 22.69 | 16.09 | 13.03 | 30.96 | 5.88 | 5.36 | 18.35 | 9.19 | 17.78 | 5.92 | 11.65 | 13.23 | 10.04 | 13.16 | 31.72 |
| CV (-) | 1.72 | 1.68 | 1.75 | 1.71 | 1.10 | 1.00 | 1.53 | 1.23 | 1.18 | 1.04 | 1. 20 | 1.39 | 1.36 | 0.97 | 1.59 |
| ME (d) | 5.00 | 3.00 | 5.00 | 10.00 | 9.00 | 12.00 | 16.00 | 16.00 | 17.00 | 18.00 | 17.00 | 6.00 | 6.00 | 26.00 | 15.00 |
| RT (d) | 6.93 | 7.10 | 9.75 | 11.61 | 19.10 | 11.77 | 18.10 | 22.25 | 22.56 | 26.07 | 27.45 | 13.21 | 11.00 | 28.63 | 19.50 |
| δ250/δ (%) | 21.00 | 22.00 | 27.00 | 31.00 | 39.00 | 33.00 | 38.00 | 41.00 | 42.00 | 45.00 | 47.00 | 32.00 | 32.00 | 46.00 | 39.00 |
| BFI (-) | 0.44 | 0.43 | 0.43 | 0.30 | 0.61 | 0.60 | 0.47 | 0.56 | 0.50 | 0.64 | 0.58 | 0.46 | 0.49 | 0.59 | 0.41 |
| KA (-) | 0.09 | 0.07 | 0.21 | 0.17 | 0.23 | 0.28 | 0.27 | 0.14 | 0.16 | 0.22 | 0.13 | 0.27 | 0.25 | 0.47 | 0.10 |

low bias. Bagging allows us to reduce the variance of the model while preserving the desirable low-bias property. Each tree in the RF is constructed independently based on a bootstrap sample of data, and the final model prediction is the aggregate of the individual predictions. RF combines a collection of simple and easily interpretable decision tree models and aggregates their predictions to obtain a robust ensemble model.

In the bagging process, each tree is constructed based on a subset of data, leaving an unused out-of-bag (OOB) subset of data that can be used to estimate the importance of each independent variable. Drop-column importance (DCI) is a simple and easily interpretable method that can be used to estimate the importance of each feature or independent variable (Parr et al., 2018). First, the performance of the model with all features is calculated. Then, one feature is removed, the model is retrained, and its performance is compared with that of the original model. The more important the variable is, the greater the decrease in model performance.

For each dependent variable, an RF model was developed, based on 100 random trees with a selected set of geomorphological parameters (Table 1) as features or independent variables. The model was fitted using the squared error loss function.

To validate the performance of the models, the root mean square error (RMSE) was calculated. Due to the limited size of the dataset, the repeated leave-one-out cross-validation (LOOCV) procedure was used. In the LOOCV procedure, the model is fitted several times, each time using a different training dataset. The dataset is split into a training set and a test set, using all but one observation in the training set. Models are developed using only data from the training set, and the predicted value of the one observation not included in the model is used to estimate error of the models. The LOOCV procedure was repeated ten times for each observation using different model seeds (karst aquifer system), and the average RMSE across all tests was calculated.

To assess the predictive performance of the RF models, they were compared to two baseline models: an average prediction model (APM) and a linear regression model (LRM). These simple baseline models served as a reference point for validation. The APM approach is widely used, often indirectly through the use of the coefficient of determination (R-Squared) as a metric (e.g., Gaertner, 2023). When using the mean squared error (MSE) - a suitable evaluation rule in the context of empirical risk minimization - the mean predictor is the best possible model based solely on the values of the dependent variable. To test whether the RMSE of the RF models was statistically significantly lower than the RMSE of the APM baseline models, a one-tailed *t*-test was performed, and p-values were estimated.

To better understand the dependence of the RF model error on dataset size a simple simulation experiment was performed. The procedure involved randomly sampling 10,000 datasets of size N with replacement, training an RF model on each sample, and obtaining predictions for the observations held out from the dataset. Additionally, predictions were obtained using the APM constructed from the N observations for comparison. The performance of both models was then assessed using data not included in the sampled training sets. The root mean squared error (RMSE) was computed for each dataset size, and the process was repeated for $N = 2,\ldots,14$.

#### 2.5.1. Feature importance

The importance of the features (independent variables) in the RF model was estimated by computing the out-of-bag drop-column importance (OOBDCI). The performance of the model with the selected variables was computed, then one variable was removed, the model was retrained and its performance was compared with the original model. To estimate the variance of importance, the procedure was repeated on 100 bootstrap samples.

Occasionally, OOBDCI assigns negative importance to certain variables, indicating that they degrade model performance. To address this issue, a simple greedy feature selection algorithm was implemented. Starting with an initial set of features, the importance of each feature was assessed, and the feature with the lowest importance was iteratively removed until only features with positive importance (i.e., those that improve model performance) remained. This approach allowed for the selection of a subset of features in the RF models for each dependent variable. Details of the greedy feature selection algorithm are provided in Appendix 2.

#### 2.5.2. Partial dependence and accumulated local effect plots

To examine the relationships between the important features and the dependent variable, partial dependence plots (PDP) were used. The PDP shows the marginal effect of a feature on the predicted outcome of a machine learning model (Friedman, 2001). It is illustration of a model's predictions change with a single feature, holding others constant. The partial dependence is expressed as:

$$f_S(x_S) = E_{X_C}[f(x_S, X_C)] = \int f(x_S, X_C) dP(X_C)$$

where $f$ is the RF model, $x_S$ the variable of interest, and $X_C$ the rest of the independent variables. Partial dependence was computed for $S$ by marginalizing the RF model output over the distribution of the variables in set $C$, so that the function shows the relationship between the variable $S$ and model predictions. By integrating out the other variables, a function that depends solely on variable SSS is obtained. In practice, partial dependence is estimated using the Monte Carlo method (Molnar, 2022):

$$f_S(x_S) = \frac{1}{N} \sum_{i=1}^{N} f\left(x_S, x_C^i\right)$$

The partial dependence function is constructed in a way that assumes the $x_S$and $X_C$ are not correlated. If the independent variables are correlated, new data points can be generated with values that are unlikely in reality (Molnar, 2022). An alternative to PDP, which also works when independent variables are correlated, is accumulated local effect (ALE) plots (Apley and Zhu, 2020).

## 3. Results

### 3.1. Geomorphological analysis of karst spring catchment areas and analysis of spring discharge time series

Results of GIS spatial analysis of the morphological, geological and hydrological data in the areas of selected karst spring catchments are summarized in Table 2.

Calculated basic statistics and assessed parameters of the karst spring hydrograph recession models are presented in Table 3. The results of the analyzed parameters indicate notable differences in karst spring characteristics, as reflected by a wide range of parameter values (e.g., ME, derived from the autocorrelation function, ranging from 3 to 26 days). These variations highlight the complexity of individual spring dynamics and the importance of the proposed approach in capturing and interpreting these differences.

### 3.2. Recession curve analysis

The selected karst aquifer systems are classified into four domains according to the karst aquifer classification (Mangin, 1975; El-Hakim and Bakalowicz, 2007), all reflecting the karst character (Fig. 3). No aquifer system was characterized as poorly karstified (porous and fissured aquifers, domain 5). Most aquifer systems (seven) are in a class characterized by a karst conduit system that is more developed in the upper part than in parts close to the spring, and by delayed recharge due to either non-karstic terrain or snow or sediment cover (domain 2). Four aquifer systems belong to the class of systems with a well karstified infiltration zone and an extensive conduit network ending in a flooded phreatic zone (domain 4). Three aquifer systems are characterized as intensely karstified systems in both the infiltration and phreatic zones, with a well-developed conduit system directly connected to the spring (domain 3), and one aquifer system is characterized as a complex karst system, extensive and composed of several subsystems (domain 1).

### 3.3. Random forest modeling

RF models developed use catchment geomorphological features to predict 11 dependent variables that indicate the hydrological functioning of aquifer systems. Validation of the predictive performance of the RF models, expressed by the RMSE, and the comparison with the results of the APM and LRM are presented in Table 4. The statistical significance of the differences in RMSE between the RF models and the baseline models was assessed by estimating p-values. Due to the weaker performance of the RF model in predicting SVC compared to the APM, this model was excluded from further analysis.

Fig. 4 shows the errors of the APM and RF models of four dependent variables using different numbers of observations (karst aquifer systems).

#### 3.3.1. Variable importance

Table 5 shows the importance of the features (in columns) in the RF models of the dependent variables (in rows) estimated by computing the OOBDCI. Relatively higher values reflect greater importance of the variable in the RF model.

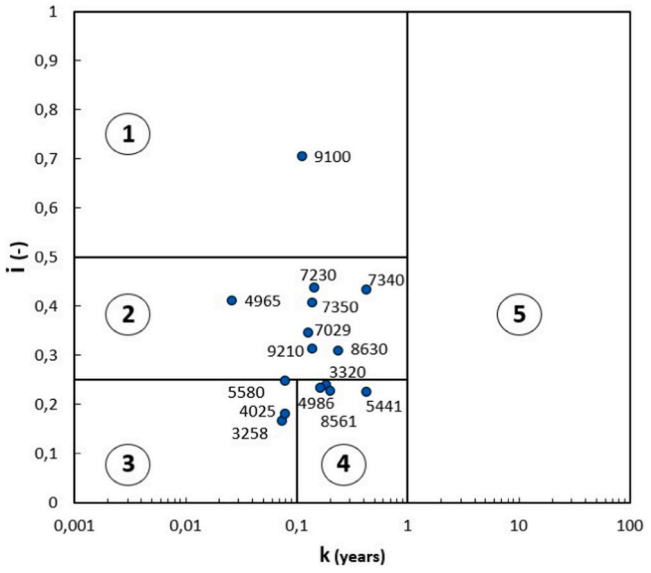



**Fig. 3.** Classification of selected karst aquifer systems based on the recession curves analysis, proposed by Mangin (1975) and modified by El-Hakim and Bakalowicz (2007).

**Table 4**
Comparison of RMSE of random forest models (RFM), and baseline models - linear regression models (LRM) and average prediction models (APM), and corresponding statistical significance level (p-value). The RFMs are ordered according to their predictive performance, estimated in comparison to the APM and expressed with the p-value.

| Dependent variable | RMSE | | | p-value |
|---|---|---|---|---|
| | RFM | LRM | APM | |
| δ250/δ (%) | 0.045 | 0.054 | 0.078 | 4.1E−15 |
| BFI (-) | 0.069 | 0.095 | 0.092 | 3.0E−14 |
| CV (-) | 0.221 | 0.260 | 0.274 | 7.1E−13 |
| Min ($m^3/s$) | 0.255 | 0.267 | 0.302 | 1.2E−12 |
| RT (d) | 5.076 | 5.047 | 7.099 | 2.0E−12 |
| Max ($m^3/s$) | 22.777 | 31.894 | 29.108 | 5.2E−12 |
| KA (-) | 0.058 | 0.070 | 0.070 | 6.9E−12 |
| ME (d) | 4.796 | 5.674 | 6.255 | 1.0E−11 |
| Mean ($m^3/s$) | 1.477 | 1.538 | 1.781 | 2.9E−10 |
| Median ($m^3/s$) | 1.037 | 1.110 | 1.041 | 3.0E−01 |
| SVC (-) | 8.431 | 12.648 | 7.969 | 1.0E+ 00 |

### *3.3.2. Partial dependence and accumulated local effect plots*

In interpreting the relationships between the most influential features and the corresponding dependent variables, results of the two methods were used. The marginal effect of most influential features on the predicted dependent variables (Table 5) is presented with partial dependence (Fig. 5) and accumulated local effect plots (Appendix 3).

## 4. Discussion

### *4.1. Geomorphological analysis of karst spring catchment areas and recession curve analysis*

For this study, 15 karst aquifer systems were selected based on the characteristics of spring discharge gauging stations, their catchments, and data availability. In this sample, both the time span of measurements and their reliability vary, which is an important source of uncertainty in the analysis of the recession curves and affects the representativeness of the input data used. The inclusion of shorter series increased the number of samples, which contributes to the robustness of the RF model and helps to reduce the prediction error (as shown in Fig. 4). This trade-off between time series length and sample size was a necessary compromise given the available data. The selected data range captures a variety of hydrological conditions that occur throughout the year, although in shorter time

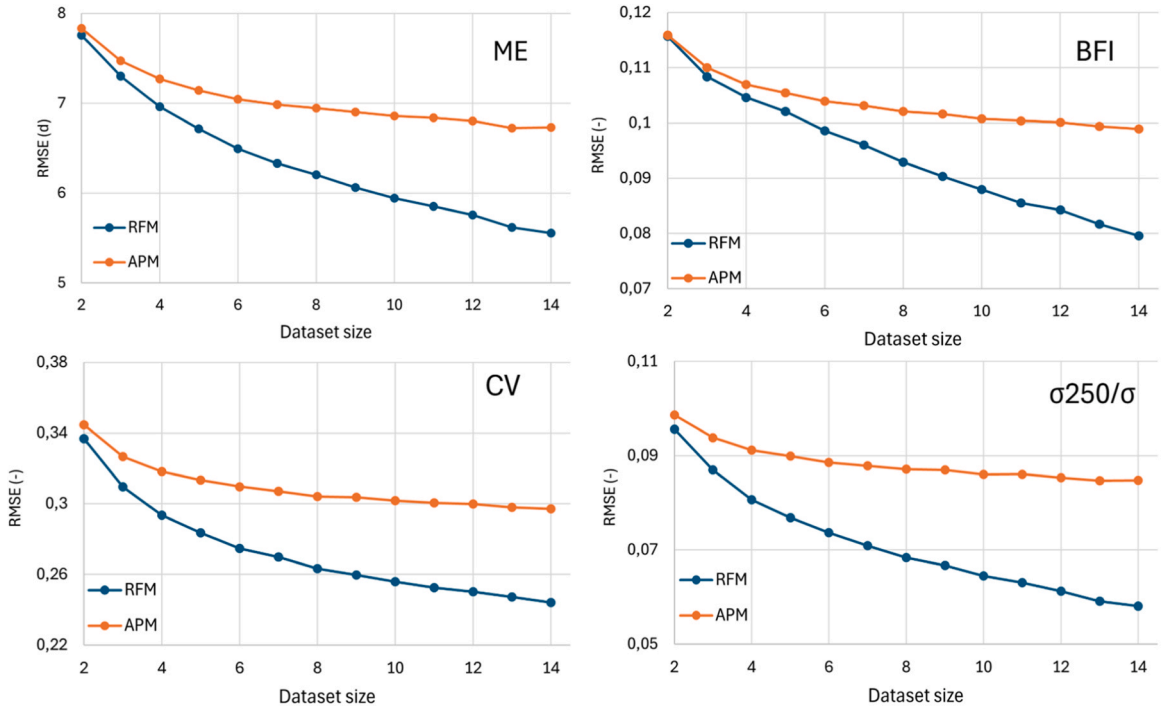



**Fig. 4.** Average RMSE for the APM and RF models for four dependent variables (σ250/σ, BFI, CV and ME) using different dataset sizes.

**Table 5**
Importance of features in RF models. Higher values indicate greater influence of a feature in predicting the target variable (i.e., greater predictive importance). The most important feature in each model is shown in bold. "N/A" indicates that the feature was not included in the RF model for the corresponding dependent variable.

| Dependent variable | AREA | CAVES | KARST_RO | V_DOLINE | FAULTS | SLOPE | STREAMS |
|---|---|---|---|---|---|---|---|
| Mean ($m^3/s$) | **0.570** | N/A | 0.006 | 0.053 | 0.043 | 0.087 | N/A |
| Median ($m^3/s$) | **0.263** | 0.015 | N/A | 0.023 | N/A | 0.031 | 0.027 |
| Min ($m^3/s$) | **0.078** | 0.002 | N/A | N/A | N/A | N/A | N/A |
| Max ($m^3/s$) | N/A | N/A | N/A | **9.628** | N/A | 4.658 | N/A |
| CV (-) | 0.005 | **0.059** | 0.010 | 0.006 | N/A | N/A | N/A |
| ME (d) | N/A | 0.066 | 0.494 | 0.224 | 0.255 | **1.064** | 0.104 |
| RT (d) | N/A | 0.071 | N/A | N/A | **0.384** | 0.343 | N/A |
| δ250/δ (%) | N/A | 0.002 | N/A | N/A | 0.003 | **0.007** | N/A |
| BFI (-) | 0.001 | **0.054** | 0.005 | 0.003 | N/A | N/A | N/A |
| KA (-) | N/A | **0.018** | 0.001 | N/A | 0.000 | 0.003 | 0.000 |

series, long-term trends and extreme events may not be fully represented. This introduces some uncertainty, which is recognized as a limitation of the study.

Recession curve analysis and classification of the karst aquifer systems (Mangin, 1975; El-Hakim and Bakalowicz, 2007) confirmed the karst character of the selected aquifer systems. The classification of the systems into four domains shows the wide range of hydrological characteristics of the aquifer systems.

### *4.2. Interpretation of parameters indicating the hydrological functioning of karst aquifer systems*

In this study, 11 hydrological parameters (dependent variables) were used to characterize karst aquifer systems. Basic statistics (mean, median, minimum, maximum and standard deviation) provide fundamental information on the characteristics of spring discharge and indirectly on the hydraulic functioning of the karst aquifer system. The minimum and maximum discharge indicate the range of the observed spring discharge. However, they are subject to uncertainties, especially the maximum discharge, which often occurs in very short periods and is difficult to measure. The ratio between the minimum and maximum observed discharge is an indicator of the variability of the spring discharge, but for the same reason it is an unreliable indicator. Therefore, the SVC (ratio between Q90 and Q10) is proposed as a more stable and reliable indicator of spring discharge variability (Flora, 2004). In general, relative discharge variability (i.e., SVC) is higher in less regulated karst aquifer systems characterized by very rapid infiltration and water flow through the subsurface. This is most commonly a characteristic of aquifer systems developed in highly karstified but compact rocks, where conduit porosity is very high, while matrix and fracture porosity is low. As a result, these systems have limited water retention capacity and exhibit high transmissivity and low storativity. The standard deviation is influenced by the magnitude of the spring discharge. The CV is calculated as the ratio of the standard deviation to the mean (Mangin, 1975; Meinzer, 1923). This normalization facilitates comparison between karst systems of different area sizes. The CV is correlated with the SVC and serves as an additional indicator of the variability of spring discharge, which generally increases when the regulation capacity of the aquifer system decreases.

In addition to the variability of spring discharge, the duration of the response of spring discharge to infiltration pulses is a decisive factor in describing the regulatory capacity of the system. The ME (Mangin, 1984) is derived from the characteristics of the autocorrelation function of the discharge time series. Another indicator proposed by Mangin (1984) is the RT, which is based on the spectral density function and estimates the duration of the long-term response of the system. The RT can also be calculated approximately as $125 \times (\sigma 250/\sigma)^2$ (Bailly-Comte et al., 2023). The indicators ME, RT and σ250/σ are strongly correlated and together reflect the inertia of the system, i.e., how fast the discharge decreases during the recession after a rainfall pulse. In general, systems characterized by a higher storage capacity, often associated with high matrix and fracture porosity or soil cover, exhibit a slower flow recession during dry periods. The BFI and the ratio between median and mean discharge indicate the uniformity of the spring discharge hydrograph recession curve, as extreme values that deviate strongly affect the mean but not the median discharge. Both indicators are also related to CV and SVC, as springs with more uniform recessions (higher BFI and median/mean) tend to have less pronounced total discharge variations due to the absence of a strong discharge increase after rainfall. A weak distinction between fast and base flow on the spring discharge hydrograph indicates the absence of one or more factors that facilitate rapid infiltration and groundwater flow. These factors include well-developed epikarst and/or concentrated surface water inflows (e.g., sinkholes and ponors), and an integrated network of highly permeable karst conduits connecting the infiltration zone to the spring(s).

However, the rate of discharge decline during the recession (indicated by ME, RT and σ250/σ) should be distinguished from the uniformity of the spring recession curve (represented by BFI and median/mean). Therefore, we introduce a composite parameter, KA, which considers both system inertia (rate of discharge decrease) and recession uniformity. It is defined as the product of two statistical indicators: σ250/σ and median/mean. In this way, KA serves as a comprehensive measure of the regulatory capacity of the karst system, which is crucial for the management of water resources.

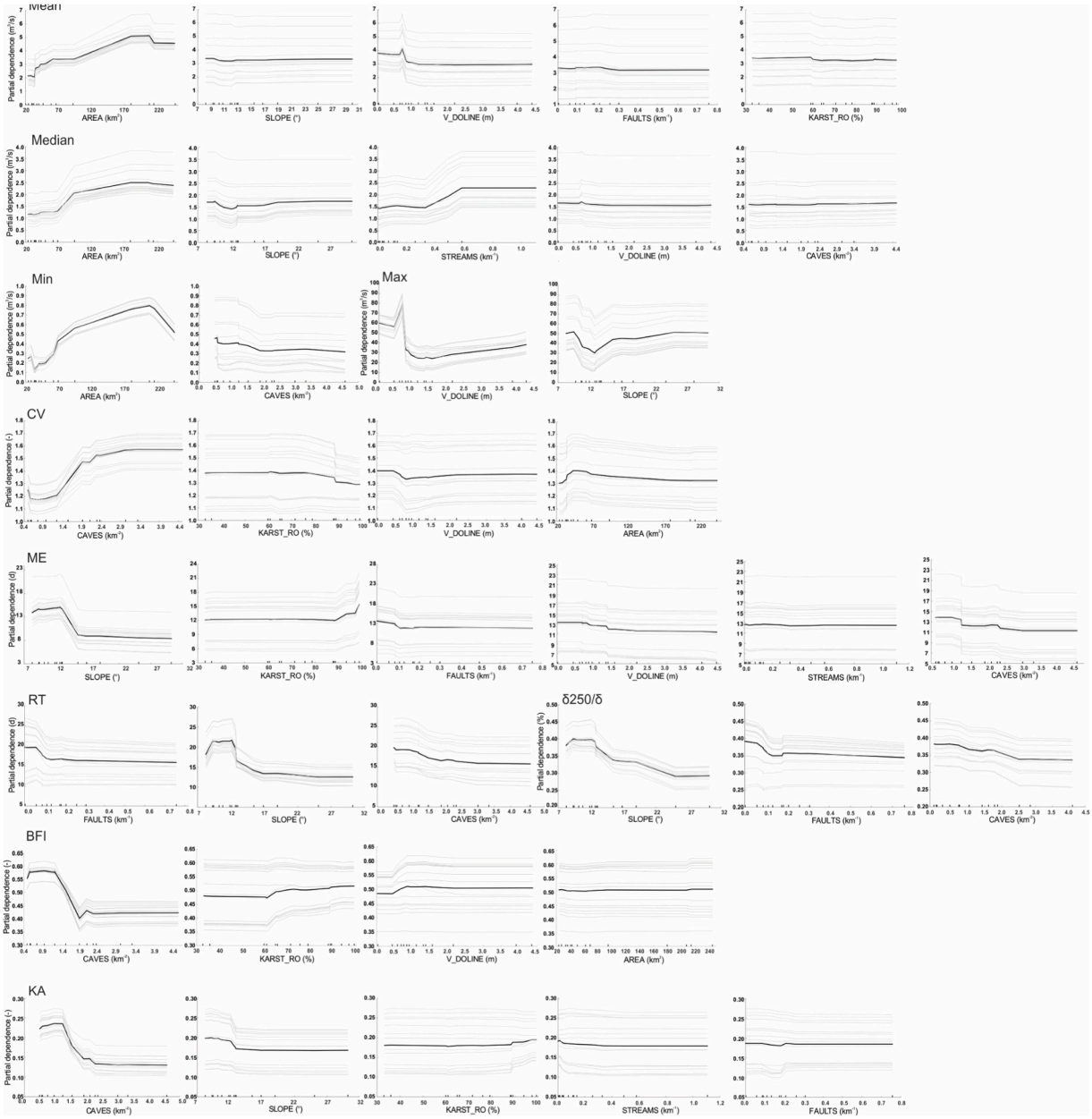



**Fig. 5.** PDPs for the RF models illustrate the marginal effects of the most influential features on the predicted dependent variables (shown in Table 5). The x-axis represents the range of values for the features, while the y-axis shows the predicted outcome of the model. Gray lines represent the relationships between individual observations (aquifer systems) and the selected feature, while the black line represents the average partial dependence across all observations.

### *4.3. Random forest modeling*

Random forest modeling was used to predict the hydrological properties of the karst aquifer system based on the geomorphological characteristics of the catchment. The evaluation of the predictive performance of the RF models compared to the baseline APM models shows that the RF models, except for the dependent variable SVC, have a lower RMSE (Table 4). At a p-value threshold of 0.05, the RF models perform statistically significantly better than the corresponding APMs in predicting 10 out of 11 dependent variables (except SVC). The RF models also perform better than the LRMs. Only for the dependent variable RT is the RMSE of the LRM lower than that of the RF model.

To counter the risks of overfitting arising from the limited sample size and to provide a robust assessment of the RF model's generalizability, a rigorous evaluation protocol based on LOOCV was implemented. LOOCV is widely recognized as a robust method

for estimating a predictive model's generalization error (Vehtari et al., 2017; Zhang and Yang, 2015). Although LOOCV is generally computationally demanding, the small size of the dataset made it feasible in this case. The method enables a comprehensive evaluation of the model's predictive ability on previously unseen data.

Due to the stochastic nature of the RF algorithm, even with a fixed number of trees, the evaluation was further enhanced by using repeated LOOCV with different random seeds. This approach allowed for the assessment of variability in model performance across different initializations, thereby providing stronger insight into the model's generalization capabilities and its susceptibility to overfitting. The adopted evaluation protocol effectively identified overfitting, as demonstrated by the model for SVC prediction. In this instance, the protocol estimated the model's generalization error to be worse than that of the baseline (mean predictor), indicating that this model is not reliable for predictive purposes. The available dataset may not have provided sufficient information for the model to accurately predict SVC from the independent variables. Consequently, the model may have overfit the training data, potentially being influenced by a small number of highly influential observations. Given the strong correlation between SVC and CV, and the more robust performance of the CV model, the latter is recommended for use in predicting spring discharge variability. Similarly, the prediction of the Mean is more reliable than that of the Median, whose performance is not significantly better than that of the APM baseline model. Use of the model's prediction of Mean is therefore recommended as an estimator of characteristic spring discharge in this context.

#### 4.3.1. *Dataset size*

The dependence of the RMSE on the dataset size of the APM and RF models shows a similar pattern for all four dependent variables analyzed (Fig. 4). The RMSE of the APM converges to the true standard deviation as the dataset size increases. On the other hand, the RMSE of the RF models decreases continuously, even at the maximum available dataset size. The gain in predictive performance with increasing data size is significantly greater for RF models compared to APMs. This indicates that the predictive performance of RF models can be improved by increasing the dataset size. Therefore, the small sample size of 15 aquifer systems is a strong limitation to better demonstrate the applicability of the presented approach. These results provided an estimate of how model error decreases with dataset size, highlighting the benefits of acquiring larger datasets and demonstrating the potential for improved effectiveness in similar problems with more observations.

#### 4.3.2. *Feature importance and interpretation of their effects*

The OOBDCI procedure highlighted the area of the spring catchment (AREA) as by far the most important feature for predicting mean, median and minimum spring discharge. This is an expected result and confirms the ability of the methodology to identify the most important physical drivers. PD and ALE plots (Fig. 5 and Appendix 3) show that with increasing AREA increase also values of these dependent variables. Generally, values of dependent variables increase up to a catchment area size of 240 $km^2$, then start to decrease. This trend is most evident for minimum discharge. Since the pattern is influenced by one observation, spring with the largest catchment area, it is most likely an anomaly related to the uncertainty in the delineation of the catchment area of this complex karst aquifer system. This uncertainty is further linked to the dependence of groundwater flow direction on hydrological conditions, including possible flow bifurcation (Janža, 2010; Serianz and Prestor, 2022).

In contrast to RF models of the basic statistics mentioned above, the most important features in RF model for predicting maximum discharge are the volume of dolines (V_DOLINE) and the mean slope gradient (SLOPE). However, increasing values of these features do not have a uniform effect on maximum discharge (Fig. 5); the correlation is negative at lower values and reverses at higher values. The uncertainties associated with measuring maximum discharges, as mentioned above, make interpreting the non-uniform relationships identified by the model challenging and potentially unreliable. Due to the limited dataset, other important factors (e.g., variations in soil type and thickness, rock permeability, and epikarst development) may also influence these relationships and act as potential confounding variables.

The absence of catchment area among the important features for predicting maximum spring discharge may indicate that the karst springs studied have catchment areas that vary depending on hydrological conditions, a phenomenon known as bifurcation. This is a common phenomenon in karst aquifer systems, which often do not discharge through a single outlet during floods, but through several outlets or springs acting as overflows (Kresic and Bonacci, 2010). As the aquifer is predominantly recharged by direct infiltration of precipitation through open fractures and conduits, this leads to progressive saturation of the system from the uppermost karst features downwards. During intense rainfall events, the upper conduit network and cavities are the first to become saturated, with water subsequently infiltrating at different rates into the deeper karst zones. This results in a highly dynamic and heterogeneous flow regime, which is influenced by the structural and hydrodynamic properties of the karst system.

RF models for predicting σ250/σ and RT use the same three features. For σ250/σ, the most important feature is SLOPE. Their negative correlation (Fig. 5) indicates that steeper slopes facilitate higher hydraulic gradients and faster groundwater flow, as well as the concentration of surface water flows and more concentrated infiltration into the subsurface. This results in faster drainage of catchments, which leads to less flow regulation and a faster decline of spring discharge during recession. The identification of the same important features for predicting σ250/σ and RT, which are highly correlated hydrologic indicators, confirms the ability of the RF models to consistently identify the main geomorphological drivers for the respective hydrological characteristics. The most important feature for predicting RT is FAULTS (followed by SLOPE), and FAULTS also ranks second for σ250/σ. This feature is negatively correlated with RT and σ250/σ. The presence of active faults likely facilitates higher terrain karstification (Dantas et al., 2025; De la Torre et al., 2020; Häuselmann et al., 1999; Herold et al., 2000), which promotes faster water flow and results in lower σ250/σ and RT values. CAVES is the third most important feature for predicting σ250/σ and RT. Cave entrances are generally more numerous in less fractured limestones, which have thicker stratification and greater resistance to surface erosion. This resistance contributes to the

preservation of cave entrances (Stroj and Paar, 2019). In contrast, dolomites and highly fractured limestones are more susceptible to erosion, often resulting in cave entrances being covered by accumulated sediment. Lower spatial density of fractures reduces water storage capacity in the epikarst and deeper zones, leading to lower water retention and faster infiltration, which accelerates groundwater flow through the subsurface. This is reflected in reduced σ250/σ, RT and ME values. A similar interpretation can be applied to other dependent variables strongly influenced by CAVES (e.g., negative correlation with BFI and KA, and positive correlation with CV).

The hydraulic parameter ME is correlated with σ250/σ and RT. The most important feature for predicting ME is SLOPE and its effect is similar to that in models of σ250/σ and RT (Fig. 5). The influence of other features (KARST_RO, FAULTS, V_DOLINE, WATER_C and CAVES) is less pronounced.

RF models for predicting CV and BFI use the same features. The most important feature is CAVES, followed by KARST_RO, V_DOLINE, and AREA. CV and BFI are negatively correlated and the features have opposite effects on them (Fig. 5), which confirms the consistency of the modeling results. A higher density of karst cave entrances, as explained above, indicates more compact karst rocks with lower water storage capacity and faster flow response of the system. The portion of karst rock outcrops in the catchment (KARST_RO) is negatively correlated with CV (and positively affects BFI). This may be because diffuse infiltration through the epikarst is the dominant process in catchments composed predominantly of karst rocks. Diffuse infiltration slows down flow concentration in the karstic subsurface, thereby attenuating the spring's response. In contrast, surface flows from low-permeable (non-karst) parts of the catchment, typically torrential in nature (Carlier et al., 2018), converge and rapidly sink upon reaching the karst terrain (allogenic recharge), resulting in a quicker, more pronounced spring response and, consequently, reduced regulation of spring discharge. The influence of V_DOLINE and AREA is less pronounced. Trends of V_DOLINE at lower values (positively correlated with BFI and negatively correlated with CV) probably also indicate less allogenic recharge. The expected same trend related to better flow regulation with increasing catchment area and associated longer water retention time is not well pronounced.

KA is a composite hydraulic parameter that characterizes both the uniformity of the flow regime and the inertia of the aquifer system. RF model for predicting this parameter is based on the features CAVES, SLOPE, and KARST_RO. CAVES is by far the most important feature in the model (Table 5), which emphasizes the importance of this geomorphologic feature in interpreting the hydrologic properties of the karst aquifer system. The feature STREAMS is included in only three RF models, predicting the dependent variables Median, ME, and KA. In all of these models, its importance is relatively low and generally among the lowest compared with other features. This suggests that the length of permanent streams in the spring catchment is a less informative indicator of karstification and spring discharge variability than the other geomorphological features considered.

While the relationships between most of the independent geomorphological variables (features) and the dependent variables or hydrological parameters of karst aquifer system can be explained, there are differences in the identification and ranking of the most important features for some of the well-correlated hydrological parameters (e.g., ME compared to σ250/σ and RT). This is likely due to the relatively small dataset. The uncertainties associated with the determination of the geomorphological parameters and the monitoring accuracy of the discharge of individual springs are also factors that influence the results of RF models.

The study is limited to the available data, which are indirect surface indicators of karst aquifer systems, whose hydrological behavior is primarily controlled by the development and connectivity of karst conduit networks. However, such detailed subsurface information is typically restricted to well-studied systems and was not available for all of the aquifers included in this study. Including only those aquifer systems with detailed subsurface data would have further reduced the already limited sample size and compromised the representativeness of the analysis. For this reason, the study focuses on parameters that could be obtained consistently across the full set of aquifers.

Selected geomorphological parameters have certain constraints in terms of hydrological relevance. However, the use of machine learning methods, as applied in this study, requires a range of sample characteristics that must be derived from data available at the regional level. Some of these indicators, particularly karst cave density, are influenced by human factors. Despite its anthropocentric nature, karst cave density emerged as an important feature in RF modeling karst aquifer hydrological functioning. This importance likely reflects the relatively comprehensive cave mapping in Slovenia, which is supported by a well-established speleological exploration community. In regions without such a tradition of systematic exploration, karst caves density may not hold the same relevance.

## 5. Conclusion

This study presents a novel application of RF modeling to link geomorphological catchment features with the hydrological functioning of karst aquifers. The approach enables improved characterization and prediction of hydrological behavior in ungauged karst systems, with key hydrological parameters inferred from the developed models. Feature importance analysis identified catchment area (AREA), cave entrance density (CAVES), slope gradient (SLOPE) and fault density (FAULTS) as the most important geomorphological features for predicting the hydrologic properties of karst spring discharge.

Compared to traditional modeling approaches (APM and LRM), the RF framework demonstrates higher predictive performance, highlighting the potential of machine learning tools to derive hydrological information from available geomorphological data. These results advance understanding of karst systems by quantifying the relative contribution of geomorphological features, going beyond descriptive assessments and simple analyses of individual features used in previous studies.

The developed RF models provide a practical tool for predicting the hydrological behavior of ungauged karst catchments in areas with reliable geomorphological data. This capability has practical application for water resource management and supports the prioritization and planning of research activities at the regional scale, particularly in areas where long-term monitoring is limited or unavailable.

Future improvements could be achieved by incorporating additional karst catchments and regional-scale datasets (e.g., properties and thickness of soil and the unsaturated zone, and epikarst development), which would further enhance model performance and extend the applicability of this approach to broader karst regions.

## CRediT authorship contribution statement

**Valter Hudovernik:** Writing – review & editing, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Mitja Janža:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Andrej Stroj:** Writing – review & editing, Methodology, Investigation. **Luka Serianz:** Writing – review & editing, Methodology, Investigation, Conceptualization.

## Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT and DeepL in the writing process to improve the readability and language of the manuscript. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

## Funding

This work was supported by the Slovenian Water Agency (project Development of a decision support system for groundwater use) and by the Slovenian Research and Innovation Agency through the research program Groundwater and Geochemistry (P1–0020).

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The authors thank Rada Peternel Rikanović for collecting and processing spatial data and Nika Pišek Szillich for assistance in preparing the figures.

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ejrh.2025.102774.

## Data availability

The code and data used in this study are available at https://zenodo.org/records/17150604.

## References

Aguilera, H., Guardiola-Albert, C., Moreno Merino, L., Baquedano, C., Díaz-Losada, E., Robledo Ardila, P.A., Durán Valsero, J.J., 2022. Building inexpensive topsoil saturated hydraulic conductivity maps for land planning based on machine learning and geostatistics. Catena 208, 105788. https://doi.org/10.1016/j.catena.2021.105788.

Apley, D., Zhu, J., 2020. Visualizing the effects of predictor variables in black box supervised learning models. J. R. Stat. Soc. Ser. B (Stat. Methodol. ) 82. https://doi.org/10.1111/rssb.12377.

ARSO, 2022. Hydrological data archive - daily data [dataset]. Slovenian Environment Agency (Agencija RS za okolje), https://vode.arso.gov.si/hidarhiv/pov_arhiv_tab.php.

Atanackov, J., Jamšek Rupnik, P., Jež, J., Celarc, B., Novak, M., Milanič, B., Markelj, A., Bavec, M., Kastelic, V., 2021. Database of active faults in Slovenia: compiling a new active fault database at the junction between the Alps, the dinarides and the pannonian basin tectonic domains. Front. Earth Sci. 9. https://doi.org/10.3389/feart.2021.604388.

Bailly-Comte, V., Ladouche, B., Charlier, J.B., Hakoun, V., Maréchal, J.C., 2023. XLKarst, an excel tool for time series analysis, spring recession curve analysis and classification of karst aquifers. Hydrogeol. J. 31 (8), 2401–2415. https://doi.org/10.1007/s10040-023-02710-w.

Biau, G., Scornet, E., 2016. A random forest guided tour. Test 25 (2), 197–227. https://doi.org/10.1007/s11749-016-0481-7.

Bond, N.R., Kennard, M.J., 2017. Prediction of hydrologic characteristics for ungauged catchments to support hydroecological modeling. Water Resour. Res. 53, 8781–8794, 2017.

Booker, D.J., Snelder, T.H., 2012. Comparing methods for estimating flow duration curves at ungauged sites. J. Hydrol. 434–435, 78–94.

Booker, D.J., Woods, R.A., 2014. Comparing and combining physically-based and empirically-based approaches for estimating the hydrology of ungauged catchments. J. Hydrol. 508, 227–239.

Breiman, L., 2001. Random forests. Mach. Learn 45 (1), 5–32. https://doi.org/10.1023/A:1010933404324.

Brenčič, M., 2000. Hidrogeološka analiza velikih kraških izvirov v Sloveniji. Doctoral Dissertation Thesis, University of Ljubljana, Ljubljana, 357 pp.

Brunner, M.I., Furrer, R., Sikorska, A.E., Viviroli, D., Seibert, J., Favre, A.C., 2018. Synthetic design hydrographs for ungauged catchments: a comparison of regionalization methods. Stoch. Environ. Res. Risk Assess. 32, 1993–2023, 2018.
Carlier, C., Wirth, S.B., Cochand, F., Hunkeler, D., Brunner, P., 2018. Geology controls streamflow dynamics. J. Hydrol. 566, 756–769. https://doi.org/10.1016/j.jhydrol.2018.08.069.
Cinkus, G., Mazzilli, N., Jourde, H., 2021. Identification of relevant indicators for the assessment of karst systems hydrological functioning: proposal of a new classification. J. Hydrol. 603. https://doi.org/10.1016/j.jhydrol.2021.127006.
Dantas, T.B., Striglio, G., La Bruna, V., Bezerra, F.H.R., Balsamo, F., Araujo, R.E.B., Furtado, C.P.Q., Rabelo, J.G., Lima, R.S., Vasconcelos, D.L., Ibanez, D.M., Lima-Filho, F.P., 2025. Fault-induced karst features: insights from the Poço verde fault tip in the northwestern portion of potiguar basin, Brazil. J. Struct. Geol. 192, 105341. https://doi.org/10.1016/j.jsg.2025.105341.
Dasgupta, R., Das, S., Banerjee, G., Mazumdar, A., 2024. Revisit hydrological modeling in ungauged catchments comparing regionalization, satellite observations, and machine learning approaches. HydroResearch 7, 15–31. https://doi.org/10.1016/j.hydres.2023.11.001.
De la Torre, B., Mudarra, M., Andreo, B., 2020. Investigating karst aquifers in tectonically complex alpine areas coupling geological and hydrogeological methods. J. Hydrol. X 6, 100047. https://doi.org/10.1016/j.hydroa.2019.100047.
DRSV, 2022. Hydrography [dataset]. Slovenian Water Agency (Direkcija za vode), https://podatki.gov.si/dataset/hidrografija?resource_id=798e79bb-9262-4e97-bb0d-6763140211e8.
Dufoyer, A., Massei, N., Lecoq, N., Marechal, J.-C., Thiery, D., Pennequin, D., David, P.-Y., 2019. Links between karst hydrogeological properties and statistical characteristics of spring discharge time series: a theoretical study. Environ. Earth Sci. 78 (14). https://doi.org/10.1007/s12665-019-8411-0.
Efron, B., 1977. The efficiency of cox's likelihood function for censored data. J. Am. Stat. Assoc. 72 (359), 557–565. https://doi.org/10.1080/01621459.1977.10480613.
El-Hakim, M., Bakalowicz, M., 2007. Significance and origin of very large regulating power of some karst aquifers in the Middle East. Implication on karst aquifer classification. J. Hydrol. 333 (2-4), 329–339. https://doi.org/10.1016/j.jhydrol.2006.09.003.
Fang, W., Ren, K., Liu, T., Shang, J., Jia, S., Jiang, X., Zhang, J., 2024. An evaluation of random forest based input variable selection methods for one month ahead streamflow forecasting. Sci. Rep. 14, 29766. https://doi.org/10.1038/s41598-024-81502-y.
Fernandes, V.J., de Louw, P.G.B., Bartholomeus, R.P., Ritsema, C.J., 2024. Machine learning for faster estimates of groundwater response to artificial aquifer recharge. J. Hydrol. 637, 131418. https://doi.org/10.1016/j.jhydrol.2024.131418.
Fiedler, M., 2024. Using random forest regression to model the spatial distribution of concentrations of selected metals in groundwater in forested areas of the wielkopolska national park, Poland. Forests 15, 2191. https://doi.org/10.3390/f15122191.
Fiorillo, F., 2014. The recession of spring hydrographs, focused on karst aquifers. Water Resour. Manag. 28 (7), 1781–1805. https://doi.org/10.1007/s11269-014-0597-z.
Flora, S.P., 2004. Hydrogeological characterization and discharge variability of springs in the Middle Verde River watershed, Central Arizona, Northern Arizona University, 237 pp.
Ford, D., Williams, P., 2007. Karst Hydrogeology and Geomorphology. John Wiley & Sons Ltd, p. 562. https://doi.org/10.1002/9781118684986.
Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29 (5), 1189–1232.
Gaertner, B., 2023. Geospatial patterns in runoff projections using random forest based forecasting of time-series data for the mid-Atlantic region of the United States. Sci. Total Environ. 912, 169211. https://doi.org/10.1016/j.scitotenv.2023.169211.
Goldscheider, N., Chen, Z., Auler, A.S., Bakalowicz, M., Broda, S., Drew, D., Hartmann, J., Jiang, G., Moosdorf, N., Stevanovic, Z., Veni, G., 2020. Global distribution of carbonate rocks and karst water resources. Hydrogeol. J. 28 (5), 1661–1677. https://doi.org/10.1007/s10040-020-02139-5.
Goldscheider, N., Drew, D. (Eds.), 2007. Methods in Karst Hydrogeology: IAH: International Contributions to Hydrogeology, 26. CRC Press. https://doi.org/https://doi.org/10.1201/9781482266023.
Gomez, M., Nölscher, M., Hartmann, A., Broda, S., 2024. Assessing groundwater level modelling using a 1-D convolutional neural network (CNN): linking model performances to geospatial and time series features. Hydrol. Earth Syst. Sci. 28 (19), 4407–4425. https://doi.org/10.5194/hess-28-4407-2024.
Gostinčar, P., Stepišnik, U., 2023. Extent and spatial distribution of karst in Slovenia. Acta Geogr. Slov. 63 (1), 111–129. https://doi.org/10.3986/AGS.11679.
GURS, 2022. Digital elevation model [dataset]. Surveying and Mapping Authority (Geodetska uprava Republike Slovenije), https://podatki.gov.si/dataset/digitalni-model-visin.
Häuselmann, P., Jeannin, P.-Y., Bitterli, T., 1999. Relationships between karst and tectonics: case-study of the cave system north of lake thun (Bern, Switzerland). Geodin. Acta 12 (6), 377–388. https://doi.org/10.1016/S0985-3111(99)00104-7.
Herold, T., Jordan, P., Zwahlen, F., 2000. The influence of tectonic structures on karst flow patterns in karstified limestones and aquitards in the Jura mountains, Switzerland. Eclogae Geol. Helv. 93, 349–362.
Hobbs, S.L., Smart, P.L., 1986. Characterization of carbonate aquifers: a conceptual base., Proceed 9th Int Speleol Congress, Barcelona, pp. 43-46.
Janža, M., 2010. Hydrological modeling in the karst area, Rižana spring catchment, Slovenia. Environ. Earth Sci. 909–920. https://doi.org/10.1007/s12665-009-0406-9.
Janža, M., Meglič, P., Šram, D., Prestor, J., Skaberne, D., Rozman, D., 2011. Improvement of hydrogeological conceptual and geological model (A.2.6). Final report. Projekt INCOME-LIFE07 ENV/SLO/000725. Geological Survey of Slovenia, Ljubljana.
Jenkins, G.M., Watts, D.G., 1968. Spectral Analysis and its Applications. Holden-Day, San Francisco.
Jourde, H., Wang, X., 2023. Advances, challenges and perspective in modelling the functioning of karst systems: a review. Environ. Earth Sci. 82 (17), 396. https://doi.org/10.1007/s12665-023-11034-7.
JZS, 2022. Cave Registry [dataset]. Speleological Association of Slovenia (Jamarske zveze Slovenije), https://www.jamarska-zveza.si/index.php/foreigners/cave-registry.
Kiraly, L., 1994. Groundwater flow in fractures rocks: models and reality, 14th Mintrop seminar über Interpretationsstrategien in Exploration und Produktion. Ruhr Universität Bochum, Bochum, pp. 1-21.
Koch, J., Kim, H., Tirado-Conde, J., Hansen, B., Møller, I., Thorling, L., Troldborg, L., Voutchkova, D., Højberg, A.L., 2024. Modeling groundwater redox conditions at national scale through integration of sediment color and water chemistry in a machine learning framework. Sci. Total Environ. 947, 174533. https://doi.org/10.1016/j.scitotenv.2024.174533.
Kovačič, G., Ravbar, N., Petrič, M., Kogovšek, J., 2012. Latest research on karst waters in Slovenia and their significance. Geogr. Vestn. 84, 65–75.
Kovács, A., 2021. Quantitative classification of carbonate aquifers based on hydrodynamic behaviour. Hydrogeol. J. 29 (1), 33–52. https://doi.org/10.1007/s10040-020-02285-w.
Kovács, A., Perrochet, P., 2008. A quantitative approach to spring hydrograph decomposition. J. Hydrol. 352 (1), 16–29. https://doi.org/10.1016/j.jhydrol.2007.12.009.
Kovács, A., Sauter, M., 2008. Modelling karst hydrodynamics. In: Martin, J.B., White, W.B. (Eds.), Frontiers of Karst Research. Karst Waters Institute, pp. 13–26.
Kovács, A., Perrochet, P., Király, L., Jeannin, P.-Y., 2005. A quantitative method for the characterisation of karst aquifers based on spring hydrograph analysis. J. Hydrol. 303 (1), 152–164. https://doi.org/10.1016/j.jhydrol.2004.08.023.
Kresic, N., Bonacci, O., 2010. Spring discharge hydrograph, in: Kresic, N., Stevanovic, Z. (Eds.), Groundwater Hydrology of Springs. Butterworth-Heinemann, Boston, pp. 129-163. https://doi.org/https://doi.org/10.1016/B978-1-85617-502-9.00004-9.
Ladson, A.R., Brown, R., Neal, B., Nathan, R., 2013. A standard approach to baseflow separation using the lyne and hollick filter. Aust. J. Water Resour. 17 (1), 25–34. https://doi.org/10.7158/W12-028.2013.17.1.
Li, G., Goldscheider, N., Field, M.S., 2016. Modeling karst spring hydrograph recession based on head drop at sinkholes. J. Hydrol. 542, 820–827. https://doi.org/10.1016/j.jhydrol.2016.09.052.
Li, X., Sha, J., Wang, Z.-L., 2019. Comparison of daily streamflow forecasts using extreme learning machines and the random forest method. Hydrol. Sci. J. 64, 1857–1866.

Liedl, R., Sauter, M., Hückinghaus, D., Clemens, T., Teutsch, G., 2003. Simulation of the development of karst aquifers using a coupled continuum pipe flow model. Water Resour. Res. 39 (3). https://doi.org/10.1029/2001WR001206.
Lyne, V., Hollick, M., 1979. Stochastic time-variable rainfall-runoff modelling. Institute of Engineers Australia Barton, Australia, pp. 89-93.
Maillet, E.T., 1905. Essais d'hydraulique souterraine et fluviale. A. Hermann, Paris, 218pp.
Malík, P., Vojtková, S., 2012. Use of recession-curve analysis for estimation of karstification degree and its application in assessing overflow/underflow conditions in closely spaced karstic springs. Environ. Earth Sci. 65 (8), 2245–2257. https://doi.org/10.1007/s12665-012-1596-0.
Mangin, A., 1975. Contribution à l'étude hydrodynamique des aquifères karstiques. Univ. é De. Dijon. 267. https://hal.science/tel-01575806.
Mangin, A., 1984. Pour une meilleure connaissance des systèmes hydrologiques à partir des analyses corrélatoire et spectrale. J. Hydrol. 67 (1), 25–43. https://doi.org/10.1016/0022-1694(84)90230-0.
Masrur Ahmed, A.A., Deo, R.C., Feng, Q., Ghahramani, A., Raj, N., Yin, Z., Yang, L., 2021. 2021. Deep learning hybrid model with Boruta-Random forest optimiser algorithm for streamflow forecasting with climate mode indices, rainfall, and periodicity. J. Hydrol. 599, 126350. https://doi.org/10.1016/j.jhydrol.2021.126350.
Meinzer, O.E., 1923. Outl. GroundWater Hydrol. Défin. 494, 10.3133/wsp494.https://pubs.usgs.gov/publication/wsp494.
Mezga, K., 2014. Natural Hydrochemical Background and Dynamics of Groundwater in Slovenia. Doctoral Dissertation Thesis, University of Nova Gorica, Nova Gorica, p. 226.
Mihevc, A., Mihevc, R., 2021. Morphological characteristics and distribution of dolines in Slovenia, a study of a lidar-based doline map of Slovenia. Acta Carsol. 50 (1). https://doi.org/10.3986/ac.v50i1.9462.
Moges, D.M., Virro, H., Kmoch, A., Cibin, R., Rohith, R.A.N., Martínez-Salvador, A., Conesa-García, C., Uuemaa, E., 2024. Streamflow prediction with Time-Lag-Informed random forest and its performance compared to SWAT in diverse catchments. Water 16, 2805. https://doi.org/10.3390/w16192805.
Molnar, C., 2022. Interpretable Machine Learning: A Guide for Making Black Box Models Explainable. Independently published.
Naghibi, S.A., Ahmadi, K., Daneshi, A., 2017. Application of support vector machine, random forest, and genetic algorithm optimized random forest models in groundwater potential mapping. Water Resour. Manag 31, 2761–2775.
Novak, M., Bavec, M., Trajanova, M., 2014. Lithological Map (of Slovenia). Geological Survey of Slovenia.
Olarinoye, T., Gleeson, T., Hartmann, A., 2022. Karst spring recession and classification: efficient, automated methods for both fast- and slow-flow components. Hydrol. Earth Syst. Sci. 26 (21), 5431–5447. https://doi.org/10.5194/hess-26-5431-2022.
Padilla, A., Pulido-Bosch, A., Mangin, A., 1994. Relative importance of baseflow and quickflow from hydrographs of karst spring. Groundwater 32 (2), 267–277. https://doi.org/10.1111/j.1745-6584.1994.tb00641.x.
Investigathor, Parr, T., Turgutlu, K., Csiszar, C., Howard, J. 2018. Beware default random forest importances. https://explained.ai/rf-importance/index.html (accessed 17 October 2023).
Patra, S.R., Chu, H.-J., Aman, M.A., 2024. Utilizing deep learning to investigate the impacts of climate change on groundwater dynamics and pumping variability. Sci. Total Environ. 957, 177784. https://doi.org/10.1016/j.scitotenv.2024.177784.
Pavlič, U., Trišič, N., Souvent, P., 2008. Ocena prispevnih zaledij izbranih kraških izvirov: interno poročilo (Assessment of catchments of selected karst springs: Internal report). Slov. Environ. Agency Ljubl. (https://meteo.arso.gov.si/met/sl/watercycle/observation_sites/springs/).
Pham, L.T., Luo, L., Finley, A., 2021. Evaluation of random forests for Short-Term daily streamflow forecasting in Rainfall- and Snowmelt-Driven watersheds. Hydrol. Earth Syst. Sci. 25, 2997–3015, 2021.
Prestor, J., Urbanc, J., Meglič, P., Lapanje, A., Rajver, D., Hribernik, K., Šinigoj, J., Strojan, M., Bizjak, M., 2004. Nacionalna baza hidrogeoloških podatkov za opredelitev teles podzemne vode republike slovenije (National Database of Hydrogeological Data for Defining Groundwater Bodies of the Republic of Slovenia). Geol. Surv. Slov.
Rahmati, O., Reza, H., Melesse, A.M., 2016. Application of GIS-based data driven random forest and maximum entropy models for groundwater potential mapping: a case study at mehran region, Iran. Catena 137, 360–372. https://doi.org/10.1016/j.catena.2015.10.010.
Rodriguez-Galiano, V., Mendes, M.P., Garcia-Soldado, M.J., Chica-Olmo, M., Ribeiro, L., 2014. Predictive modeling of groundwater nitrate pollution using random forest and multisource variables related to intrinsic and specific vulnerability: a case study in an agricultural setting (Southern Spain). Sci. Total Environ. 476–477, 189–206. https://doi.org/10.1016/j.scitotenv.2014.01.001.
Serianz, L., Prestor, J., 2022. Hydrogeological conceptualization of Rižana spring karst system, in: Brenčič, M., Žvab Rožič, P., Torkar, A. (Eds.), Making groundwater in the Danube region visible: 5th IAH CEG Conference Slovenian Committee of International Association of Hydrogeologist - SKIAH, Rogaška Slatina, pp. 50.
Singh, N.K., Emanuel, R.E., Nippgen, F., McGlynn, B.L., Miniat, C.F., 2018. The relative influence of storm and landscape characteristics on shallow groundwater responses in forested headwater catchments. Water Res. Res. 54, 9883–9900.
Stevanović, Z., 2019. Karst waters in potable water supply: a global scale overview. Environ. Earth Sci. 78 (23). https://doi.org/10.1007/s12665-019-8670-9.
Stroj, A., Paar, D., 2019. Water and air dynamics within a deep vadose zone of a karst massif: observations from the lukina jama–Trojama cave system (−1,431 m) in Dinaric karst (Croatia). Hydrol. Process 33 (4), 551–561. https://doi.org/10.1002/hyp.13342.
Tyralis, H., Papacharalampous, G., Langousis, A., 2019. A brief review of random forests for water scientists and practitioners and their recent history in water resources. Water 11, 910. https://doi.org/10.3390/w11050910.
Vehtari, A., Gelman, A., Gabry, J., 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Stat. Comput. 27 (5), 1413–1432. https://doi.org/10.1007/s11222-016-9696-4.
Verbovšek, T., 2008. Hydraulic conductivities of fractures and matrix in slovenian carbonate aquifers. Geologija 51 (2), 245–255. https://doi.org/10.5474/geologija.2008.025.
Wang, X., Liu, T., Zheng, X., Peng, H., Xin, J., Zhang, B., 2018. Short-term prediction of groundwater level using improved random forest regression with a combination of random features. Appl. Water Sci. 8, 125. https://doi.org/10.1007/s13201-018-0742-6.
Xu, B., Ye, M., Dong, S., Dai, Z., Pei, Y., 2018. A new model for simulating spring discharge recession and estimating effective porosity of karst aquifers. J. Hydrol. 562, 609–622. https://doi.org/10.1016/j.jhydrol.2018.05.039.
Xu, T., Valocchi, A.J., Ye, M., Liang, F., 2017. Quantifying model structural error: efficient Bayesian calibration of a regional groundwater flow model using surrogates and a data-driven error model. Water Res. Res. 53, 4084–4105, 2017.
Zhang, Y., Yang, Y., 2015. Cross-validation for selecting a model selection procedure. J. Econ. 187 (1), 95–112. https://doi.org/10.1016/j.jeconom.2015.02.006.
Zhang, Y., Tian, Y., Li, Y., Wang, D., Tao, J., Yang, Y., Lin, J., Zhang, Q., Wu, L., 2022. Machine learning algorithm for estimating karst rocky desertification in a peak-cluster depression basin in southwest guangxi, China. Sci. Rep. 12 (1), 19121. https://doi.org/10.1038/s41598-022-21684-5.
Zhou, R., Zhang, Y., 2023. Linear and nonlinear ensemble deep learning models for karst spring discharge forecasting. J. Hydrol. 627. https://doi.org/10.1016/j.jhydrol.2023.130394.