

KarstID : an R Shiny application for the analysis of karst spring discharge time series and the classification of karst system hydrological functioning

Guillaume Cinkus, Naomi Mazzilli, Hervé Jourde

▶ **To cite this version:**

Guillaume Cinkus, Naomi Mazzilli, Hervé Jourde. KarstID : an R Shiny application for the analysis of karst spring discharge time series and the classification of karst system hydrological functioning. Environmental Earth Sciences, 2023, 82 (6), pp.1-19. 10.1007/s12665-023-10830-5 . hal-04018696

HAL Id: hal-04018696
https://hal.science/hal-04018696

Submitted on 7 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

KarstID: An R Shiny application for the analysis of karst spring discharge time series and the classification of karst systems hydrological functioning

Guillaume Cinkus[1], Naomi Mazzilli[2] and Hervé Jourde[1]

[1]HydroSciences Montpellier (HSM), Univ. Montpellier, CNRS, IRD, 34090 Montpellier, France

[2]UMR 1114 EMMAH (AU-INRAE), Université d'Avignon, 84000 Avignon, France

Corresponding author

Name: Guillaume Cinkus

Email address: guillaume.cinkus@umontpellier.fr

Postal address: Guillaume Cinkus, HydroSciences Montpellier, 300 avenue du professeur Emile Jeanbrau, 34090, Montpellier, France

ORCID IDs

Guillaume Cinkus: 0000-0002-2877-6551

Naomi Mazzilli: 0000-0002-9145-5160

Hervé Jourde: 0000-0001-7124-4879

Abstract

Karst spring discharge time series analyses are often used to gain preliminary insights into the hydrological functioning of a karst system. KarstID is an R Shiny application that facilitates the completion of such analyses and allows the identification of karst system hydrological functioning. The application permits (i) to perform statistical, recession curves, classified discharges and signal (simple correlational and spectral) analyses; (ii) to calculate relevant indicators representative of distinct hydrological characteristics of karst systems, (iii) to classify karst systems hydrological functioning; and (iv) to compare the results to a database of 78 karst systems. The KarstID software is free, open source, and actively developed on a developer community platform. The user-friendly installation and launch make it especially accessible even for non-programmers, therefore KarstID can be used for both research and educational purposes. The application and its user manual are both available on the French SNO KARST website (https://sokarst.org/en/softwares-en/karstid-en/).

Keywords

karst, hydrological functioning, classification, time series analysis, hydrogeology, R Shiny.

Acknowledgements

This article is published in Environmental Earth Sciences 82(6). https://doi.org/10.1007/s12665-023-10830-5.

The French Ministry of Higher Education and Research is thanked for the thesis scholarship of G. Cinkus. This application is developed within the framework of (i) the KARST observatory network (www.sokarst.org) initiated by the INSU/CNRS, which aims to strengthen knowledge-sharing and promote cross-disciplinary research on karst systems; and (ii) the KARMA project (Karst Aquifer Resources availability and quality in the Mediterranean Area,

http://karma-project.org/). The DREAL Provence Alpes-Côtes d'Azur (PACA) is also acknowledged for providing Fontaine de Vaucluse's data.

The manuscript was written with the Rmarkdown framework (Allaire et al., 2021; Xie et al., 2020, 2018), using R (R Core Team, 2021) and knitr (Xie, 2021). KarstID software uses functions from the following packages: data.table (Dowle and Srinivasan 2021), DT (Xie et al., 2021), minpack.lm (Elzhov et al., 2016), padr (Thoen, 2021), plotly (Sievert, 2020), shiny (Chang et al., 2021), shinyFeedback (Merlino and Howard, 2020), shinyhelper (Mason-Thom, 2019), shinyjs (Attali, 2020), waiter (Coene, 2021), zoo (Zeileis and Grothendieck, 2005), and readxl, readr, dplyr, stringr, magrittr, ggplot2, lubridate (Wickham et al., 2019). The KarstID package was developed and structured using devtools (Wickham et al., 2021), usethis (Wickham and Bryan, 2021) and testthat (Wickham, 2011) packages.

Software availability

- Name of software: KarstID
- Type of software: R Shiny application embedded into an R package
- Developer: Guillaume Cinkus
- Contact information: guillaume.cinkus@umontpellier.fr
- Year first available: 2021
- Software required: R & Any web browser
- Availability and cost: Free
- Program language: R (Requires version 4.0.0 or later)
- Package size: 4.6 Mo
- Code repository: https://github.com/busemorose/KarstID
- License: Creative Commons Attribution 4.0 International

1 Introduction

Around 9% of the world's population is dependent on karst water resources for drinking water (Stevanović, 2019). Karst systems are heterogeneous media with high contrasts in porosity and permeability, inducing a high variability in infiltration and internal flow processes (Bakalowicz 2005; Ford and Williams 2007). With the increasing demand for water, the characterisation of the functioning of karst aquifers become a major challenge for water resource management (Bakalowicz 2011). Among the numerous methods to study karst aquifers, analyses of spring discharge time series (recession curves, signal, statistics) are the most accessible as they only require the monitoring of spring discharge. Therefore, they are generally used as a preliminary step for characterising the hydrological functioning of karst systems, and subsequently for developing and designing hydrological models. Also, many authors declined these analyses into classifications for differentiating karst systems, with recession curves (Bonacci 1993; Cinkus et al. 2021; Dewandel et al. 2003; Fiorillo 2014; Kullman 2000; Malík and Vojtková 2012; Mangin 1975; Soulios 1991), signal (Gárfias-Soliz et al. 2010; Mangin 1984), hydrograph (Kovács 2021; Zhang et al. 2020) and statistical indicators (Flora 2004; Hakoun et al. 2022; Rashed 2012; Springer et al. 2008).

Recession curves analysis has been widely developed over the past century (Barnes 1939; Boussinesq 1903; Coutagne 1948; Drogue 1972; Horton 1933; Kullman 2000; Maillet 1905; Mangin 1975; Padilla et al. 1994). It consists in calibrating numerical models on a selection of recession curves from a hydrograph and interpreting the parameters of the equations. Signal analyses —originally developed by Box and Jenkins (1976), Brillinger (1975), and Jenkins and Watts (1968)— were introduced in karst hydrology by Mangin (1984). Their purpose is to characterise the temporal structure of hydrological signals, which allows deducing information on the inertia of a karst system (Jeannin and Sauter 1998; Kovács 2003; Larocque et al. 1998). Statistical analyses include distribution indicators such as mean, standard deviation and quantiles, but also cumulative frequency curves (Malík 2015; Mangin 1971). Numerous studies across the world are based on these analyses for characterising karst systems properties, determining the hydrodynamic parameters of aquifers and providing information on flow dynamics (Guo et al. 2021; Lorette et al. 2018; Malík et al. 2021; Nurkholis et al. 2019; Sağır et al. 2020; Vrsalović et al. 2022; Zerouali et al. 2021).

The completion of these analyses often requires a meticulous reading of the literature and appropriate programming skills. The application of statistical and signal analyses (e.g. simple correlational and spectral analyses) is generally done using or writing specific code functions. The recession curves analysis requires to (i) select and isolate several parts of the discharge time series, (ii) calibrate a recession model over each recession curve, and (iii) calculate indicators of functioning from the model's parameters values. These operations can be tedious —especially for long time series— and are subjects to errors in selection, calibration and indicators calculation. For these reasons, several authors proposed powerful toolboxes and software to facilitate the completion of the recession curves analysis (Arciniega-Esparza et al. 2017; Carlotto and Chaffe 2019; Gregor and Malík 2016; Posavec et al. 2017), one of them also including statistical and signal analyses (BRGM 2022).

This paper presents an application (KarstID) that provides the user a toolbox for both the analysis of karst spring discharge time series and the characterisation of karst systems hydrological functioning. KarstID is distinguishable from other software because (i) it supports multiple analyses of discharge time series (statistical, recession curves, simple correlational and spectral, classified discharges) and automatic calibration of recession model; (ii) it proposes a classification of karst systems hydrological functioning (according to the proposal of Cinkus et al. (2021)) and a comparison of the results to a database of 78 karst systems; and (iii) it is free, open source and actively developed on a developer community platform. KarstID is built with the R Shiny framework (Chang et al. 2021) and is embedded into an R package (R Core Team 2021), which make the installation and launch easy even for non-programmers.

2 Software overview

The links to the user guide, the source code and the git repository are available on the French SNO KARST (Service National d'Observation du Karst) website (https://sokarst.org/en/softwares-en/karstid-en/). The user guide provides guidelines for the installation and launch, as well as a technical, in-depth and visual description of all the features of the application. The source code includes the data and functions used (i) for applying the analyses, (ii) for generating the plots, (iii) for managing the application and (iv) for building the R package. Users can start

discussions or raise issues in the git repository, as well as propose new code or modify existing code with Pull Requests.

2.1 Workflow

First, the user has to load an appropriate dataset using the "Data import" tab. The second step is to apply four different methods for analysing the hydrological functioning of the system (Figure 1). Two of these analyses (statistical and classified discharges analyses) do not need any actions from the user and the results of these analyses are directly displayed in their respective tabs. Two complementary analyses (recession curves and signal analyses) require the user to select curves and/or define parameters for the functions (i.e. recession model, autocorrelation function). The completion of the recession curves analysis automatically launches the third step which is the classification of the hydrological functioning of the system according to the methodology proposed by Cinkus et al. (2021). The user can then appreciate the results of the classification and compare the various hydrological characteristics of the analysed karst system with the ones of 78 karst systems located worldwide.

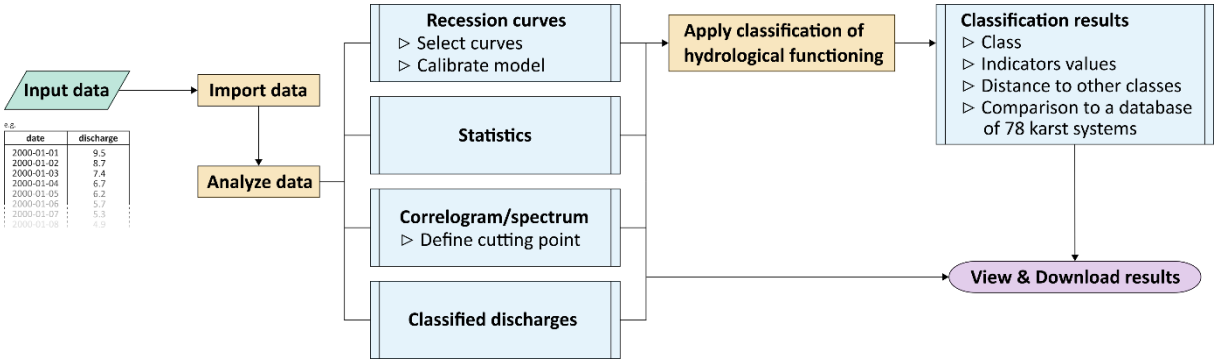


Fig. 1 Synthetic workflow of the KarstID application. Green, yellow, blue, and purple boxes represent, respectively, (i) input data, (ii) action within KarstID, (iii) available analyses within KarstID, and (iv) output data

2.2 Data import

The "Data import" tab allows the user to load a karst spring discharge time series into KarstID. The raw data can be either a plain text or an Excel file, and must have only two columns referring to date and discharge, respectively. KarstID supports date and date-time format for the date column, and numeric format for the discharge column. The application proposes several features to minimise the preprocessing of the data: it is thus possible to (i) skip rows, (ii) select a specific Excel sheet, (iii) use a header or not, (iv) define the decimal mark, (v) define the delimiter and (vi) specify the date format. The user can give a name to the dataset, which will be used when displaying or downloading results.

The user can choose to interpolate missing discharge values and specify the maximum gap that will be covered. The interpolation is performed with the *na.spline(method = "monoH.FC")* R function from the *zoo* package (Zeileis and Grothendieck 2005). The method is particularly well suited for the interpolation of small gaps, but users must be careful when using it for large gaps. Critical gap length cannot be specified a priori since it depends on both the time step on the time series and the hydrological behaviour of the investigated system. The user can also choose to either (i) keep all missing values in the time series or (ii) keep only the longest part of the time series without missing values.

After defining the import options and starting the importation, the application will (i) look for missing date entries and fill the blanks if necessary (adapted to the time step of the time series), (ii) interpolate missing discharge values, (iii) perform a daily or hourly mean over the discharge time series, and (iv) display a hydrograph on the same page. The interpolation and daily/hourly mean are realised according to the user-defined options. Note that the hourly mean can only be applied if the initial time step of the time series is at an hourly time step or less.

2.3 Methods

Four different methods are proposed in KarstID for analysing karst spring discharge time series. The methods can be applied independently of each other in their respective tabs.

2.3.1 Statistical analyses

Statistical analyses of spring discharge provide fundamental information about the hydrological functioning of a system. In KarstID, the following indicators are automatically calculated over the discharge time series when a dataset is imported: mean, maximum, minimum, standard deviation, 0.1 quantile (Q10), 0.9 quantile (Q90), Coefficient of Variation (CV) and Spring Variability Coefficient (SVC). The coefficient of variation corresponds to the ratio between the standard deviation σ and the mean μ of the values:

$$CV = \frac{\sigma}{\mu}$$

The SVC, which corresponds to the proposal of a "characteristic discharge" by Netopil (1971), is the ratio between Q90 (value that is exceeded 10% of the time) and Q10:

$$SVC = \frac{Q90}{Q10}$$

The statistical indicators appear in a table below the hydrograph in the "Data import" tab. The number of missing discharge values are given in the last column of the table. The statistical analyses can be performed even if there are missing discharge values in the discharge time series.

2.3.2 Recession curves analysis

Recession curves correspond to the periods when the discharge gradually decreases without replenishment of water (Toebes and Strang 1964). The analysis of recession curves can be used to assess groundwater storage and gain insights into the hydrological functioning of an aquifer (Drogue 1972; Forkasiewicz and Paloc 1967; Kovács 2003; Krešić 2007; Kullman 2000; Malík 2006; Malík and Vojtková 2012; Mangin 1975). Generally, a recession curve can be divided into (i) an influenced regime or quickflow component, and (ii) a non-influenced regime or baseflow component. Usually, the influenced regime results from the fast infiltration of the precipitation through large fractures and conduits, while the non-influenced regime results from slow infiltration through a less transmissive media such as a porous matrix (Mangin 1975). Numerous recession models exist whose indicators of functioning and interpretation differ.

To date, after analyses of the various aforementioned methods, Mangin's recession model (Mangin 1975) was identified as the most informative model (Cinkus et al. 2021). Accordingly, KarstID only propose Mangin's recession model to identify relevant indicators necessary for classifying karst systems hydrological functioning (see section 2.4). Mangin's model is a two-equations recession model that requires the manual definition of an inflexion point for distinguishing between influenced and non-influenced regimes:

$$Q_t = Q_{R0}e^{-\alpha t} + q_0\frac{1-\eta t}{1+\varepsilon t}$$

With Q_t the discharge at time t, α the recession coefficient, Q_{R0} the baseflow extrapolated at $t = 0$, q_0 the influenced discharge corresponding to the difference between Q_0 (discharge at $t = 0$) and Q_{R0}, η a constant characterising the speed of infiltration ($\eta = 1/t_i$, with t_i the beginning of the non-influenced regime) and ε a constant characterising the concavity of the influenced part of the recession curve.

The Mangin's recession model is widely used as several indicators can be calculated for characterising the hydrological functioning of a karst system. The indicator k gives information about the capacity of a system to store and release recharge water, and is calculated as follows:

$$k = \frac{V_{DYN}}{V_{an}}$$

With V_{an} the yearly mean volume of water discharged at the spring. The dynamic volume V_{DYN} corresponds to the integral of the exponential function of the recession model:

$$V_{DYN} = \int_0^{\infty} Q_i \, e^{-\alpha t} dt = \frac{Q_i}{a}$$

The indicator i can be used to characterise the capacity of a system to dampen the precipitation signal, and corresponds to the discharge generated by the influenced regime two days after the flood peak:

$$i = \frac{1 - 2\eta}{1 + 2\varepsilon}$$

The "Recession curves analysis" tab allows to perform the recession curves analysis. The selection of recession curves is done with the cursor using the graphical interface. The retained recession curves appear in a recap table where they can be selected to apply Mangin's model. The user has to define the inflexion point of the curves, based on his knowledge and experience. The recession model is calibrated with the nonlinear least squares *nlsLM()* function from the *minpack.lm* package (Elzhov et al. 2016), which minimises the squared sum of the residuals between observed and simulated discharges. The Root Mean Square Error (RMSE) between observed and simulated discharges is displayed below the recession model plot and helps to appreciate the performance of the model. Once a recession model is calibrated and validated, the indicators of functioning are calculated. They appear in the recap table when saved by the user.

The user can choose to remove spikes on the recession curves, which usually correspond to the system's response to small precipitation events and can be considered as noise for the modelling. Recession curves analysis can be performed even if there are missing discharge values in the discharge time series.

2.3.3 Simple correlational and spectral analyses

Simple correlational and spectral analyses are used to study the frequency content of a signal (Massei et al. 2006) by calculating the autocorrelation function and the associated spectrum with a Fourier transform. Mangin (1984) first applied these signal analyses to karst hydrology and proposed three indicators of karst hydrological functioning: the memory effect, the regulation time and the cut-off frequency. These three indicators mainly help to characterise the inertia of a karst system and its capacity to filter unitary impulse (Larocque et al. 1998; Marsaud 1997; Massei et al. 2006). The autocorrelation r_k and autocovariance C_k functions are calculated as follows:

$$r_k = \frac{C_k}{C_0}$$

$$C_k = \frac{1}{n} \sum_{1}^{n-k} (x_i - \bar{x}) \, (x_{i+k} - \bar{x})$$

With n the length of the series, m the maximum possible shift (usually $m < n/3$), k the shift (between 0 and m), $\bar{x}$ the mean of the series, and x_i and x_{i+k} the ith and the (i+t)th elements of the series, respectively. The spectrum s_f is derived from the autocorrelation function:

$$s_f = 2 \left[1 + 2 \sum_{k=1}^{m} D_k \, r_k cos(2\pi f k) \right]$$

With f the frequency ($f = j/2m$ at daily time step) and D_k a weighting function to ensure that s_f is not biased (Mangin 1984):

$$D_k = \frac{1 + cos\left(\frac{\pi k}{m}\right)}{2}$$

The correlogram is represented as the plot of r_k against k, and the spectrum of sf against f. The memory effect corresponds to the value of k for a r_k of 0.2, which can be read on the correlogram or calculated from the data. The regulation time corresponds to the value of the integral of the spectrum between 0 and $+\infty$, i.e. the maximum value of the the spectrum divided by 2.

The "Simple correlational and spectral analyses" tab displays the results of the simple correlational and spectral analyses. The user can define the cutting point m, which correspond to the maximum shift possible for the calculation. The cut-off frequency is not displayed as it results from a visual, subjective assessment of the spectrum. Simple correlational and spectral analyses cannot be performed if there are any missing discharge values in the discharge time series. Appendix A presents a comparison of the results obtained by Mangin (1984) and those calculated with KarstID, although the databases are different as the ones used in Mangin (1984) are unavailable.

2.3.4 Analysis of classified discharges

The analysis of classified discharges provides information on flow dynamics within a system by analysing the distribution of the discharges at the spring. For most authors, classified discharges are equivalent to the empirical cumulative function of discharge (Stevanović 2015). Mangin (1971) proposed a variant based on the assumption that the distribution of the discharges can be approximated by a half-normal distribution. From this perspective, classified discharges refer to the quantile-quantile graph of observed discharges quantiles against quantiles of the half-normal distribution. Homogeneous hydrological functioning should be outlined by a straight line in the classified discharge plot. Interpretation of Mangin's classified discharges thus consists of assessing the discontinuities of the curve and to relate them to changes in the hydrological functioning (e.g. activation of overflow springs, storage and release of water, leakage to another aquifer or miscalibration of the gauging station). The repartition function corresponding to the cumulative probability density regarding the standard normal distribution is calculated as follows:

$$P(X \leq z) = \frac{1}{2}\left[1 + erf\left(\frac{z}{\sqrt{2}}\right)\right]$$

For a half-Gaussian distribution:

$$P(X \leq z) = erf\left(\frac{z}{\sqrt{2}}\right)$$

The "Analysis of classified discharges" tab displays the results of both analyses of classified discharges. No user action is needed for the calculation. Analyses of classified discharges can be performed even if there are missing discharge values in the discharge time series.

2.4 Classification

In KarstID, it is possible to characterise a karst system after the methodology proposed by Cinkus et al. (2021) and compare the results with 78 karst systems located worldwide, the discharge of which being extracted from different database (Banque Hydro (Banque Hydro 2021), SNO KARST (Jourde et al. 2018), WoKaS (World Karst Spring hydrograph; Olarinoye et al. 2020). This dataset covers a wide diversity of karst hydrological functioning (from very reactive to inertial responses) with data from 17 countries in 12 different climatic conditions, according to the Köppen-Geiger classification (Cinkus et al. 2021). The classification allows characterising karst systems hydrological functioning according to 6 classes based on 3 indicators of functioning (Table 1). The indicators are derived from the results of the analysis of at least two recession curves. The draining of the capacitive function α_{mean} is calculated by averaging the α parameters of the recession models. The capacity of dynamic storage k_{max} corresponds to the maximum value of k among the analysed recession curves. The variability of the hydrological functioning IR corresponds to the difference between the maximum and minimum of the i distribution:

$$IR = i_{max} - i_{min}$$

Table 1 Indicator thresholds and corresponding characterisation of hydrological functioning for each class

Class	k_{max}	α_{mean}	IR	Capacity of dynamic storage	Draining of the capacitive function	Variability of the hydrological functioning
C1	≤ 0.4	≥ 0.03	≥ 0.25	Poor	Fast	Substantial
C2	≤ 0.4	≥ 0.03	< 0.25	Poor	Fast	Low
C3	≤ 0.4	< 0.03	≥ 0.25	Poor	Moderate	Substantial
C4	≤ 0.4	< 0.03	< 0.25	Poor	Moderate	Low
C5	> 0.4	/	≥ 0.25	Noticeable	Slow	Substantial
C6	> 0.4	/	< 0.25	Noticeable	Slow	Low

The "Classification" tab highlights the results obtained for the analysed karst system and summarises the values of the various indicators considered for the classification (Figure 6). A flowchart thus indicates how the system is classified according to the values of the indicators of functioning. The associated text section (i) describes the hydrological functioning of the system according to its class, (ii) displays the indicators values and (iii) shows the distance to other classes. A 3D scatter plot shows the investigated system (highlighted in red) alongside 78 other karst systems, with each axis corresponding to one indicator of functioning. Results from statistical, recession curves, simple correlational and spectral analyses, as well as indicators of functioning of all 78 systems also appear in a recap table. By default, the systems in the table are ordered by increasing distance to the investigated system. The user can select a system in the table to highlight (in yellow) its position on the 3D scatter plot.

3 Test case

Fontaine de Vaucluse is a karst spring located South-East of France. Its recharge area is estimated to be about 1160 km^2 (Ollivier et al. 2019), resulting in one of the highest karst spring interannual mean discharge in Europe (17.5 $m^3.s^{-1}$ over the 1966–2018 period).

Fontaine de Vaucluse's daily discharges over the 2013–2019 period (amounting to 1923 observations) are provided in KarstID as a test dataset. After importation using the "load test dataset" button, the hydrograph is loaded on the import page (Figure 2). The statistical indicators and number of missing discharge values are displayed in the table below the plot. For this period, Fontaine de Vaucluse's interannual mean discharge is about 15.9 $m^3.s^{-1}$ with no missing discharge values. The maximum observed discharge (about 67.1 $m^3.s^{-1}$) and the 90th quantile of observed discharges (about 32.0 $m^3.s^{-1}$) show that the discharge at the spring can be and stay very high during wet periods. The minimum observed discharge and the 10th quantile of observed discharges are relatively close (about 3.3 and 5.6 $m^3.s^{-1}$, respectively), highlighting a slow and consistent release of water from storage during dry periods. The coefficient of variation (72.3%) and SVC (5.7) are average and correspond to a "moderate" and "balanced" discharge variability, respectively (Flora 2004; Springer et al. 2008). The moderate discharge variability and the fact that the discharge can attain very high values can be related to a strong karstification of a part of the system. Using cross-correlation analyses between precipitation and discharge, Ollivier et al. (2015) found a transfer time between 1 and 6 days, indicating a somewhat reactive response of the system to precipitation events.

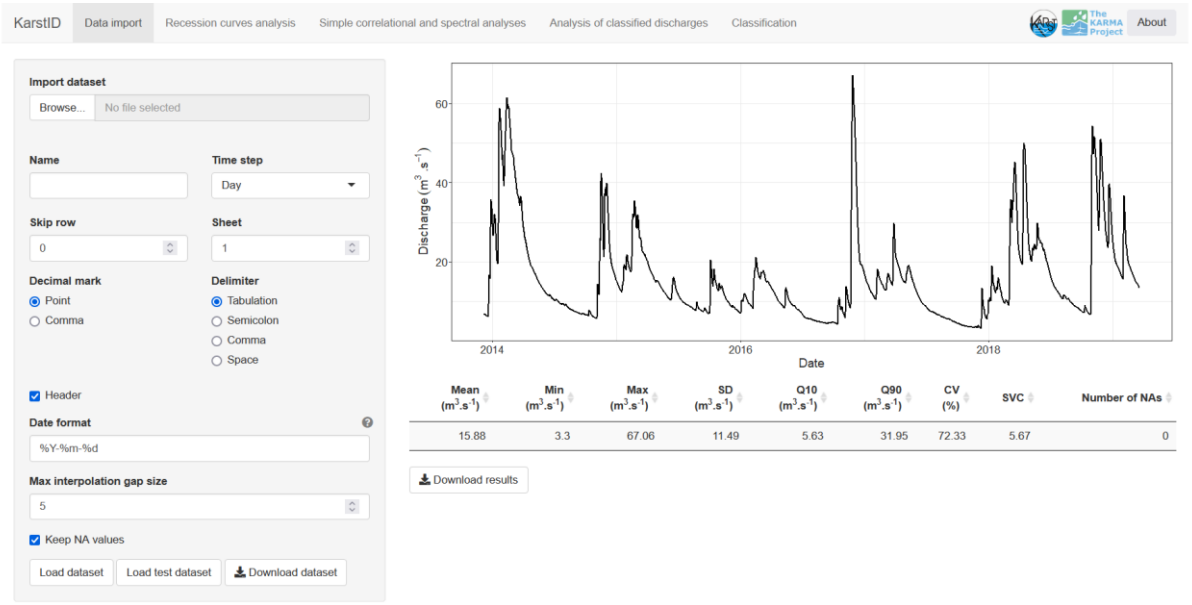


Fig 2 Import and statistical analyses tab. Left pane is dedicated to data import (section 2.2). Right part presents the hydrograph and the results of statistical analyses (section 2.3.1)

The autocorrelation function of discharge (Figure 3) declines slowly and steadily, reaching 0 at 117 days. The memory effect and the regulation time are of about 56.0 and 44.0 days, respectively. These values testify of a significant capacity of filtration of the precipitation signal, which relates to the overall organization of flows in the system (Jeannin and Sauter 1998). The noticeable dampening of the recharge in the Fontaine de Vaucluse karst system can be related to the very large dimensions of its recharge area and unsaturated zone (Ollivier et al. 2019) or the characteristics of the Urgonian limestones (Carrière et al. 2016).

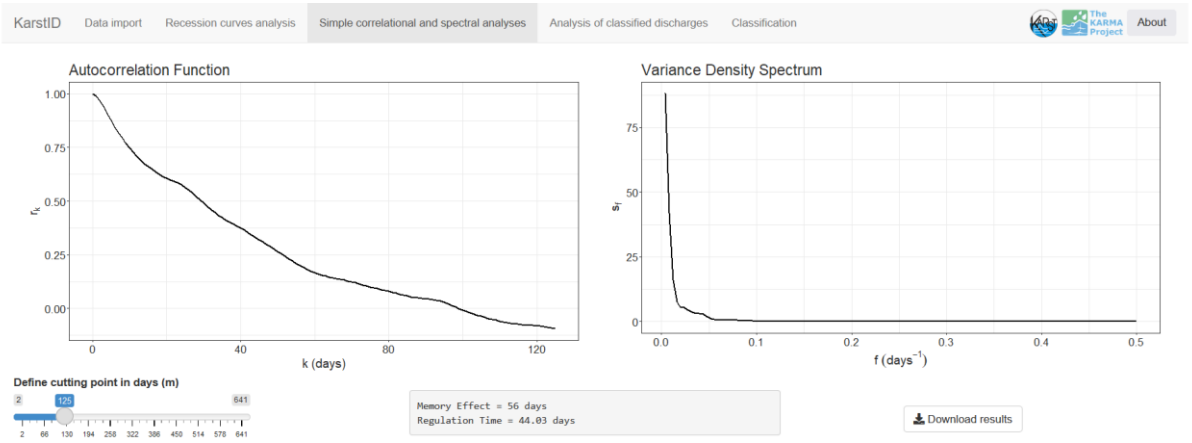


Fig. 3 Simple correlational and spectral analyses tab. Left and right graphs present the autocorrelation function and the variance density spectrum, respectively (section 2.3.3)

The analysis of classified discharges (Figure 4) according to the methodology proposed by Mangin (1975) hints that there are flow properties changes beyond 20 $m^3.s^{-1}$ (less steep slope following the inflexion point). This discontinuity reflects the overflow threshold of the upper spring pool (Mangin 1975). The other inflexion point, occurring at 57.5 $m^3.s^{-1}$, can be related to several hydrological processes: activation of an overflow, temporary storage of water or leakage to another aquifer. It can be also due to a miscalibration of the gauging station or uncertainties on the water level-discharge calibration curve.

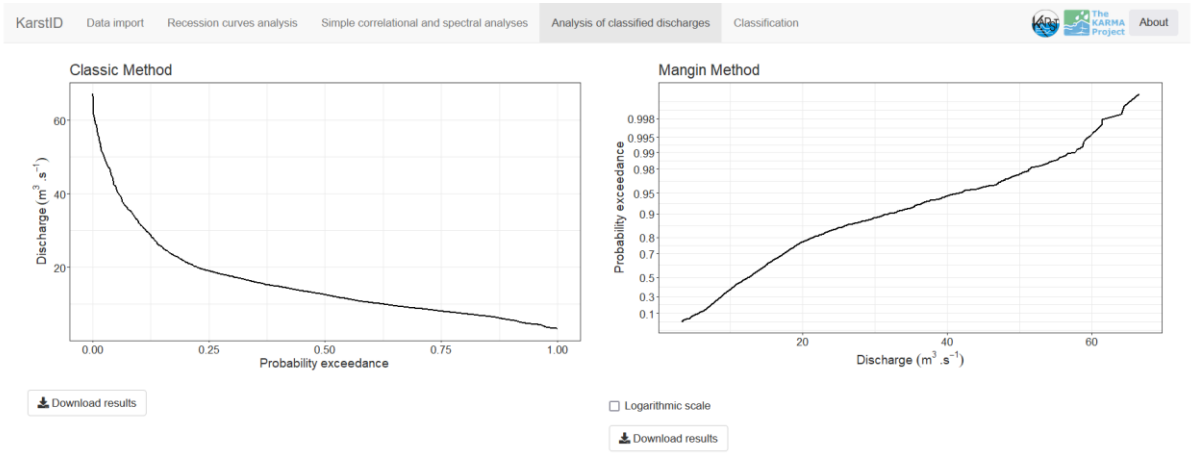


Fig. 4 Analysis of classified discharges tab. Left and right graphs present the empirical cumulative function of discharge and the Mangin classified discharges, respectively (section 2.3.4)

Three recession curves were selected and applied Mangin's recession model to identify relevant indicators (Figure 5). The recession curves were chosen according to the following criteria to ensure a maximum relevance of the analysis and its results: (i) the peak discharge must be at least one tenth of the maximum discharge of the time series, (ii) there must be little or no untimely peaks during the recession, and (iii) the recession must include both influenced and non-influenced regimes. The inflexion points (n.b. "breakpoint" in the application) were defined manually based on expert knowledge and RMSE values. The indicators k, i, and α are then calculated for each recession curve and appear in the recap table.

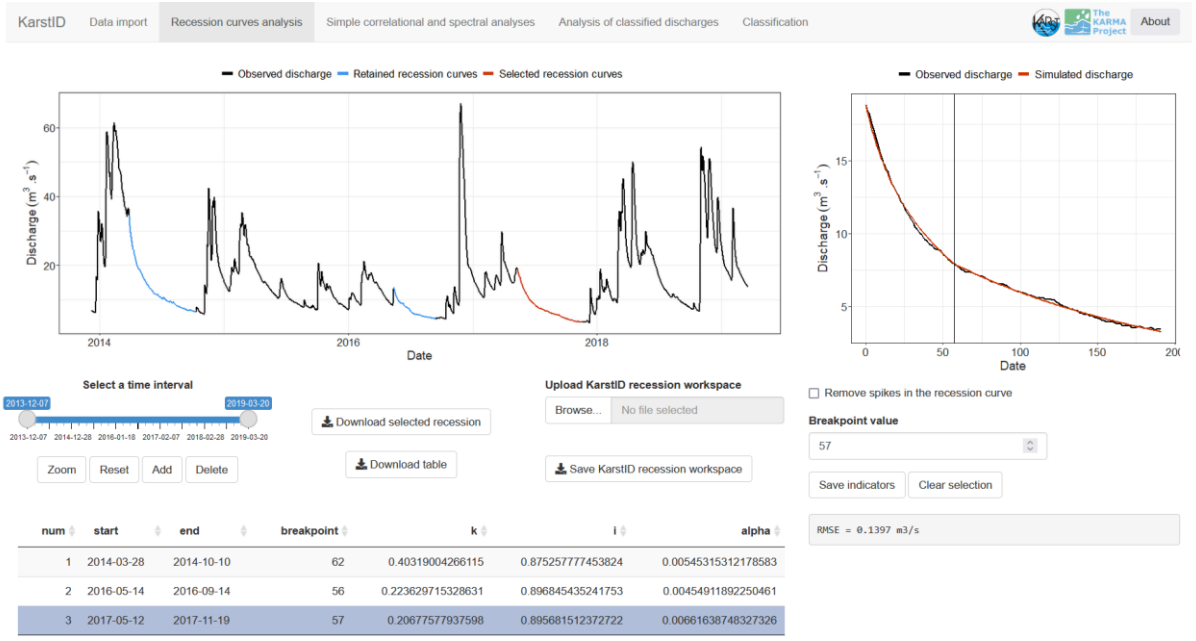


Fig. 5 Recession curves and modelling tab. The left graph presents the studied time series and retained/selected recession curves. The right graph displays the selected recession with the Mangin recession model. The table shows the details of each recession curves and their corresponding indicators values (section 2.3.2)

Fontaine de Vaucluse is classified C6 with a k_{max} of 0.403, an α_{mean} of 0.006 and an IR of 0.022 (Figure 6). This class characterises a system with noticeable capacity of dynamic storage, slow draining of the capacitive function and low variability of hydrological functioning. Fontaine de Vaucluse is considered close to the C4 class with a distance of about 0.8% (normalised Euclidean distance in the three-dimensional criteria space), meaning that C4 characteristics can also be considered in the interpretation. It highlights that the system may have a capacity of

dynamic storage and a draining of the capacitive function in-between C4 and C6 characteristics. Fontaine de Vaucluse's class is also far from the classes C3 and C5 with distances of about 91.4%, which is due to the very low *IR* (0.022). The low variability of hydrological functioning and noticeable capacity of dynamic storage assigned to this system can be also related to the large extent of the recharge area and the thick unsaturated zone. Local variability of hydrological functioning may thus be mitigated as a consequence of the spatial averaging (indirectly inducing a strong filtration of the precipitation signal). The dampening of the rainfall-discharge relationship and the noticeable capacity of dynamic storage may also be related to particular hydrological behaviour of Urgonian limestones (Carrière et al. 2016). By looking at the page 1 of the database table, users can find other karst systems with similar hydrological functioning: e.g. Taillade, PR_0005 and IE_0018. These systems are highlighted in yellow on the 3D scatter plot, alongside the investigated system highlighted in red. Studying the characteristics of other similar systems may help to support the interpretation of the investigated system. Note that Fontaine de Vaucluse also appears in the table but here corresponds to the permanent entry of the database, which results from the analysis of the whole discharge time series.

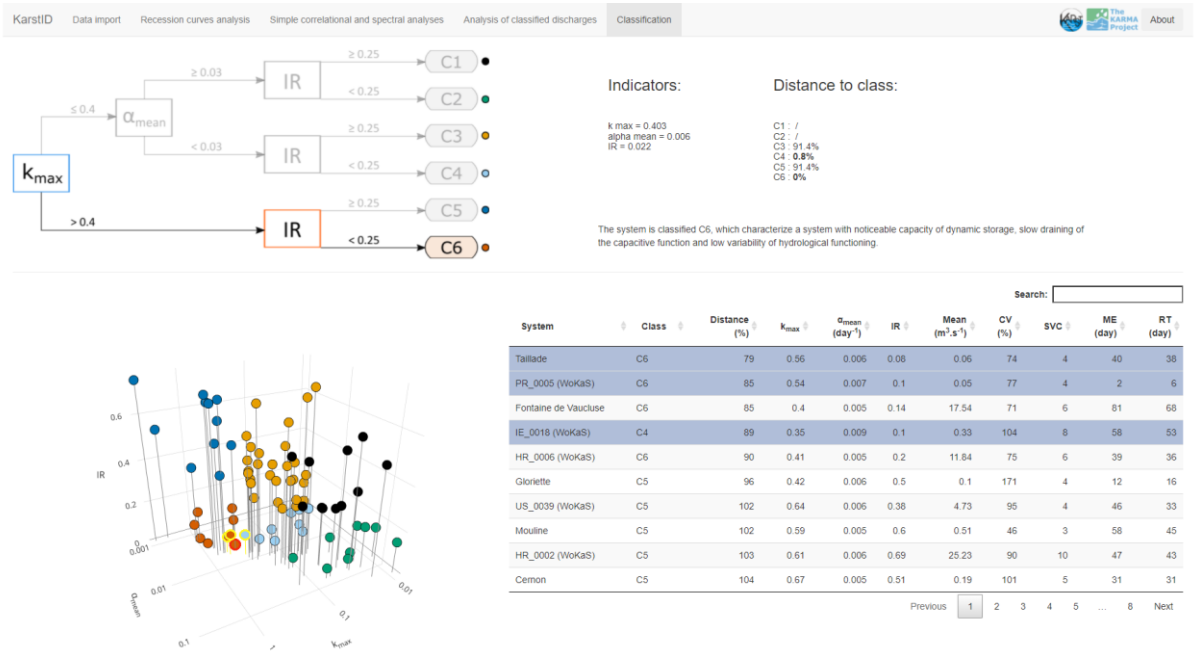


Fig. 6 Classification tab. The top part shows the classification flowchart and its associated text description: indicator values and distance to other classes (section 2.4). The bottom part presents the classification of the 78 karst systems in a 3D scatter plot with the values of key indicators in the associated table

4 Conclusion

KarstID can be seen as a useful tool for gaining preliminary insights into the hydrological functioning of a karst system. The application supports different methods for analysing discharge time series and proposes a classification of karst systems hydrological functioning. It is also possible to compare the results with a database of 78 karst systems located worldwide. KarstID is free, open source, and available on a developer community platform, which allow potential interaction between users and developers for improving software efficiency or adding new features. Other than the installation of R and R packages, no programming skills are required to use the application. KarstID could therefore also be relevant for occasional users or educational purpose. Future developments of the application include (i) a continuous consideration of feature requests and bug reports to improve user experience, (ii) the proposition of additional recession models (Drogue 1972; Kullman 2000; Padilla et al. 1994), and (iii) the addition of other discharge time series analyses (e.g. wavelet analyses).

Appendix

A. Comparison of indicators values from the correlational and spectral analyses between the original publication of Mangin (1984) and recent work of Cinkus et al. (2021) – calculated with KarstID. The exact time series used by Mangin (1984) are unavailable so different – and more recent – time series were used by Cinkus et al. (2021).

System	Indicator	Mangin (1984)	Cinkus et al. (2021)
Aliou	Memory effect (days)	4	5
	Regulation Time (days)	14	11
Baget	Memory effect (days)	15	18
	Regulation Time (days)	22.5	24

References

Allaire J, Xie Y, McPherson J, et al (2021) Rmarkdown: Dynamic documents for r. https://cran.r-project.org/package=rmarkdown. Last accessed: 2022-10-01

Arciniega-Esparza S, Breña-Naranjo JA, Pedrozo-Acuña A, Appendini CM (2017) HYDRORECESSION: A Matlab toolbox for streamflow recession analysis. Comput Geosci 98:87–92. https://doi.org/10.1016/j.cageo.2016.10.005

Attali D (2020) Shinyjs: Easily Improve the User Experience of Your Shiny Apps in Seconds. R package version 2.0.0. https://CRAN.R-project.org/package=shinyjs

Bakalowicz M (2005) Karst groundwater: A challenge for new resources. Hydrogeol J 13:148–160. https://doi.org/10.1007/s10040-004-0402-9

Bakalowicz M (2011) Management of Karst Groundwater Resources. In: van Beynen PE (ed) Karst Management. Springer Netherlands, Dordrecht, pp 263–282

Banque Hydro (2021) French Ministry of Ecology, Energy, Sustainable Development, Archive of hydrological data. Available online: http://hydro.eaufrance.fr/. Last accessed: 2022-10-01

Barnes BS (1939) The structure of discharge-recession curves. Trans AGU 20:5. https://doi.org/10.1029/TR020i004p00721

Bonacci O (1993) Karst springs hydrographs as indicators of karst aquifers. Hydrol Sci J 38:51–62. https://doi.org/10.1080/02626669309492639

Boussinesq J (1903) Sur un mode simple d'écoulement des nappes d'eau d'infiltration à lit horizontal, avec rebord vertical tout autour lorsqu'une partie de ce rebord est enlevée depuis la surface jusqu'au fond. C R Acad Sci 137:5–11

Box GEP, Jenkins GM (1976) Time Series Analysis: Forecasting and Control, Revised Edition. Holden Day, San Franscisco and Düsseldrof and Johannesburg etc.

BRGM (2022) XLKarst : une application Excel pour la caractérisation hydrodynamique des systèmes karstiques. https://www.brgm.fr/en/software/xlkarst-excel-application-hydrodynamic-characterisation-karst-systems. Last accessed: 2022-10-01

Brillinger D (1975) The Identification of Point Process Systems. Ann Probab 3:909–924. https://doi.org/10.1214/aop/1176996218

Carlotto T, Chaffe PLB (2019) Master Recession Curve Parameterization Tool (MRCPtool): Different approaches to recession curve analysis. Comput Geosci 132:1–8. https://doi.org/10.1016/j.cageo.2019.06.016

Carrière SD, Chalikakis K, Danquigny C, et al (2016) The role of porous matrix in water flow regulation within a karst unsaturated zone: An integrated hydrogeophysical approach. Hydrogeol J 24:1905–1918. https://doi.org/10.1007/s10040-016-1425-8

Chang W, Cheng J, Allaire J, et al (2021) Shiny: Web Application Framework for R. R package version 1.6.0. https://CRAN.R-project.org/package=shiny. Last accessed: 2022-10-01

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 603:127006. https://doi.org/10.1016/j.jhydrol.2021.127006

Coene J (2021) Waiter: Loading Screen for 'Shiny'. R package version 0.2.3. https://CRAN.R-project.org/package=waiter. Last accessed: 2022-10-01

Coutagne A (1948) Étude générale des variations de débit en fonction des facteurs qui les conditionnent. La Houille Blanche 134–146. https://doi.org/10.1051/lhb/1949025

Dewandel B, Lachassagne P, Bakalowicz M, et al (2003) Evaluation of aquifer thickness by analysing recession hydrographs. Application to the Oman ophiolite hard-rock aquifer. J Hydrol 274:248–269. https://doi.org/10.1016/S0022-1694(02)00418-3

Dowle M, Srinivasan A (2021) Data.table: Extension of 'data.frame'. R package version 1.14.0. https://CRAN.R-project.org/package=data.table. Last accessed: 2022-10-01

Drogue C (1972) Analyse statistique des hydrogrammes de decrues des sources karstiques statistical analysis of hydrographs of karstic springs. J Hydrol 15:49–68. https://doi.org/10.1016/0022-1694(72)90075-3

Elzhov TV, Mullen KM, Spiess A-N, Bolker B (2016) Minpack.lm: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK, Plus Support for Bounds. R package version 1.2-1. https://CRAN.R-project.org/package=minpack.lm

Fiorillo F (2014) The Recession of Spring Hydrographs, Focused on Karst Aquifers. Water Resour Manage 28:1781–1805. https://doi.org/10.1007/s11269-014-0597-z

Flora SP (2004) Hydrogeological Characterization and Discharge Variability of Springs in the Middle Verde River Watershed, Central Arizona. PhD thesis, Northern Arizona University

Ford D, Williams P (2007) Karst Hydrogeology. In: Karst Hydrogeology and Geomorphology. John Wiley & Sons, Ltd, pp 103–144

Forkasiewicz MJ, Paloc H (1967) Régime de tarissement de la foux-de-la-vis (Gard) étude préliminaire. La Houille Blanche 29–36. https://doi.org/10.1051/lhb/1967002

Gárfias-Soliz J, Llanos-Acebo H, Martel R (2010) Time series and stochastic analyses to study the hydrodynamic characteristics of karstic aquifers. Hydrol Process 24:300–316. https://doi.org/10.1002/hyp.7487

Gregor M, Malík P (2016) User manual for HydroOffice RC 4.0 tool. Online only, 35 pp. http://hydrooffice.org. Last accessed: 2022-10-01

Guo Y, Wang F, Qin D, et al (2021) Hydrodynamic characteristics of a typical karst spring system based on time series analysis in northern China. China Geol 4:433–445. https://doi.org/10.31035/cg2021049

Hakoun V, Bailly-Comte V, Charlier J-B, et al (2022) Definition of new indicators for the characterization and classification of karst aquifers using discharge time series. In: Eurokarst 2022

Horton RE (1933) The Role of infiltration in the hydrologic cycle. Trans AGU 14:446. https://doi.org/10.1029/TR014i001p00446

Jeannin P-Y, Sauter M (1998) Analysis of karst hydrodynamic behaviour using global approaches: A review. Bull Hydrogeol 31–48

Jenkins GM, Watts DG (1968) Spectral Analysis and its Applications. Louvain Econ Rev 36:554. https://doi.org/10.1017/S0770451800043062

Jourde H, Massei N, Mazzilli N, et al (2018) SNO KARST: A French Network of Observatories for the Multidisciplinary Study of Critical Zone Processes in Karst Watersheds and Aquifers. Vadose Zone J 17:180094. https://doi.org/10.2136/vzj2018.04.0094

Kovács A (2003) Geometry and hydraulic parameters of karst aquifers: A hydrodynamic modeling approach. PhD thesis, Neuchâtel University

Kovács A (2021) Quantitative classification of carbonate aquifers based on hydrodynamic behaviour. Hydrogeol J 29:33–52. https://doi.org/10.1007/s10040-020-02285-w

Krešić N (2007) Hydrogeology and Groundwater Modeling, 2e edition. CRC Press, Boca Raton (Fla.) and London and New york

Kullman E (2000) Nové metodické prístupy k riešeniu ochrany a ochranných pásiem zdrojov podzemných vôd v horninových prostrediach s krasovopuklinovou a puklinovou priepustnosťou. Podzemná voda ISSN 1335–1052:31–41

Larocque M, Mangin A, Razack M, Banton O (1998) Contribution of correlation and spectral analyses to the regional study of a large karst aquifer (Charente, France). J Hydrol 205:217–231. https://doi.org/10.1016/S0022-1694(97)00155-8

Lorette G, Lastennet R, Peyraube N, Denis A (2018) Groundwater-flow characterization in a multilayered karst aquifer on the edge of a sedimentary basin in western France. J Hydrol 566:137–149. https://doi.org/10.1016/j.jhydrol.2018.09.017

Maillet ET (1905) Essais d'hydraulique souterraine et fluviale. A. Hermann, Paris

Malík P (2006) Assessment of regional karstification degree and groundwater sensitivity to pollution using hydrograph analysis in the Velka Fatra Mountains, Slovakia. Environ Geol 51:707–711. https://doi.org/10.1007/s00254-006-0384-0

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 65:2245–2257. https://doi.org/10.1007/s12665-012-1596-0

Malík P (2015) Evaluating Discharge Regimes of Karst Aquifer. In: Stevanović Z (ed) Karst Aquifers and Engineering. Springer International Publishing, Cham, pp 205–249

Malík P, Švasta J, Bajtoš P, Gregor M (2021) Discharge recession patterns of karstic springs as observed in Triassic carbonate aquifers of Slovakia. Hydrogeol J 29:397–427. https://doi.org/10.1007/s10040-020-02276-x

Mangin A (1971) Etude des débits classés d'exutoires karstiques portant sur un cycle hydrologique. Ann spéléol 26:283–329

Mangin A (1975) Contribution à l'étude hydrodynamique des aquifères karstiques. PhD thesis, Université de Dijon

Mangin A (1984) Pour une meilleure connaissance des systèmes hydrologiques à partir des analyses corrélatoire et spectrale. J Hydrol 67:25–43. https://doi.org/10.1016/0022-1694(84)90230-0

Marsaud B (1997) Structure et fonctionnement de la zone noyée des karsts à partir des résultats expérimentaux. PhD thesis, Université Paris XI Orsay

Mason-Thom C (2019) Shinyhelper: Easily Add Markdown Help Files to 'shiny' App Elements. R package version 0.3.2. https://CRAN.R-project.org/package=shinyhelper. Last accessed: 2022-10-01

Massei N, Dupont JP, Mahler BJ, et al (2006) Investigating transport properties and turbidity dynamics of a karst aquifer using correlation, spectral, and wavelet analyses. J Hydrol 329:244–257. https://doi.org/10.1016/j.jhydrol.2006.02.021

Merlino A, Howard P (2020) shinyFeedback: Display User Feedback in Shiny Apps. R package version 0.3.0. https://CRAN.R-project.org/package=shinyFeedback. Last accessed: 2022-10-01

Netopil R (1971) Ke Klasifikaci pramenu podle variability vydatnasti (The classification of water springs based on the basis of the variability of yields). Sbornik-Hydrological Conference, Papers 22:145–150

Nurkholis A, Adji TN, Haryono E, et al (2019) Time series analysis application for karst aquifer characterisation in Pindul Cave karst system, Indonesia. Acta Carsologica 48: https://doi.org/10.3986/ac.v48i1.6745

Olarinoye T, Gleeson T, Marx V, et al (2020) Global karst springs hydrograph dataset for research and management of the world's fastest-flowing groundwater. Sci Data 7:59. https://doi.org/10.1038/s41597-019-0346-5

Ollivier C, Danquigny C, Mazzilli N, Barbel-Perineau A (2015) Contribution of Hydrogeological Time Series Statistical Analysis to the Study of Karst Unsaturated Zone (Rustrel, France). In: Andreo B, Carrasco F, Durán JJ, et al. (eds) Hydrogeological and Environmental Investigations in Karst Systems. Springer, Berlin, Heidelberg, pp 27–33

Ollivier C, Chalikakis K, Mazzilli N, et al (2019) Challenges and Limitations of Karst Aquifer Vulnerability Mapping Based on the PaPRIKa Method to a Large European Karst Aquifer (Fontaine de Vaucluse, France). Environments 6:39. https://doi.org/10.3390/environments6030039

Padilla A, Pulido-Bosch A, Mangin A (1994) Relative Importance of Baseflow and Quickflow from Hydrographs of Karst Spring. Ground Water 32:267–277. https://doi.org/10.1111/j.1745-6584.1994.tb00641.x

Posavec K, Giacopetti M, Materazzi M, Birk S (2017) Method and Excel VBA Algorithm for Modeling Master Recession Curve Using Trigonometry Approach. Groundwater 55:891–898. https://doi.org/10.1111/gwat.12549

R Core Team (2021) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. Last accessed: 2022-10-01

Rashed KA (2012) Assessing degree of karstification: A new method of classifying karst aquifers. Sixteenth International Water Technology Conference (IWTC)

Sağır Ç, Kurtuluş B, Razack M (2020) Hydrodynamic Characterization of Mugla Karst Aquifer Using Correlation and Spectral Analyses on the Rainfall and Springs Water-Level Time Series. Water 12:85. https://doi.org/10.3390/w12010085

Sievert C (2020) Interactive Web-Based Data Visualization with R, plotly, and shiny. Chapman and Hall/CRC Florida, 2020. https://cran.r-project.org/package=plotly. Last accessed: 2022-10-01

Soulios G (1991) Contribution à l'étude des courbes de récession des sources karstiques: Exemples du pays Hellénique. J Hydrol 124:29–42. https://doi.org/10.1016/0022-1694(91)90004-2

Springer AE, Stevens LE, Anderson DE, et al (2008) A comprehensive springs classification system: Integrating geomorphic, hydrogeochemical, and ecological critera. Aridland Springs in North America. Ecol Conserv 49–75

Stevanović Z (ed) (2015) Karst Aquifers and Engineering. Springer International Publishing, Cham

Stevanović Z (2019) Karst waters in potable water supply: A global scale overview. Environ Earth Sci 78:662. https://doi.org/10.1007/s12665-019-8670-9

Thoen E (2021) Padr: Quickly get datetime data ready for analysis. https://CRAN.R-project.org/package=padr. Last accessed: 2022-10-01

Toebes C, Strang DD (1964) On recession curves - Recession Equations. J Hydrol 3:2–15

Vrsalović A, Andrić I, Buzjak N, Bonacci O (2022) Karst Lake's Dynamics Analysis as a Tool for Aquifer Characterisation at Field Scale, Example of Cryptodepression in Croatia. Water 14:830. https://doi.org/10.3390/w14050830

Wickham H (2011) Testthat: Get Started with Testing. R J. 3 (1), 5-10. R Journal 3:5–10

Wickham H, Averick M, Bryan J, et al (2019) Welcome to the tidyverse. J Open Source Softw 4:1686. https://doi.org/10.21105/joss.01686

Wickham H, Bryan J (2021) Usethis: Automate Package and Project Setup. R package version 2.0.1. https://CRAN.R-project.org/package=usethis. Last accessed: 2022-10-01

Wickham H, Hester J, Chang W (2021) Devtools: Tools to Make Developing R Packages Easier. R package version 2.4.2. https://CRAN.R-project.org/package=devtools. Last accessed: 2022-10-01

Xie Y, Allaire JJ, Grolemund G (2018) R markdown: The definitive guide. Chapman and Hall/CRC, Boca Raton, Florida

Xie Y, Dervieux C, Riederer E (2020) R markdown cookbook. Chapman and Hall/CRC, Boca Raton, Florida

Xie Y (2021) Knitr: A General-Purpose Package for Dynamic Report Generation in R. R package version 1.33. https://cran.r-project.org/package=knitr. Last accessed: 2022-10-01

Xie Y, Cheng J, Tan X (2021) DT: A Wrapper of the JavaScript Library 'DataTables'. R package version 0.19. https://CRAN.R-project.org/package=DT. Last accessed: 2022-10-01

Zeileis A, Grothendieck G (2005) Zoo: S3 Infrastructure for Regular and Irregular Time Series. J Stat Softw 14:1–27. https://doi.org/10.18637/jss.v014.i06

Zerouali B, Chettih M, Alwetaishi M, et al (2021) Evaluation of Karst Spring Discharge Response Using Time-Scale-Based Methods for a Mediterranean Basin of Northern Algeria. Water 13:2946. https://doi.org/10.3390/w13212946

Zhang R, Chen X, Zhang Z, Soulsby C (2020) Using hysteretic behaviour and hydrograph classification to identify hydrological function across the "hillslopedepressionstream" continuum in a karst catchment. Hydrol Process 34:3464–3480. https://doi.org/10.1002/hyp.13793

Statements and Declarations

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Authors Contributions

G.C., N.M. and H.J. worked on the methodology and underlying sciences; G.C. programmed, coded and developed the application; G.C. wrote the original draft; G.C., N.M. and H.J. reviewed and edited the original draft.