

Assessing water resources systems' dynamic resilience under hazardous events via a genetic fuzzy rule-based system

Visnja Simic ^{a,*}, Milan Stojkovic^b, Nikola Milivojevic^c and Nikola Bacanin^a

^a Faculty of Science, University of Kragujevac, Kragujevac, Serbia

^b Institute for Artificial Intelligence R&D Serbia, Novi Sad, Serbia

^c Jaroslav Černi Water Institute, Belgrade, Serbia

*Corresponding author. E-mail: visnja@kg.ac.rs

 VS, 0000-0001-9644-8640

ABSTRACT

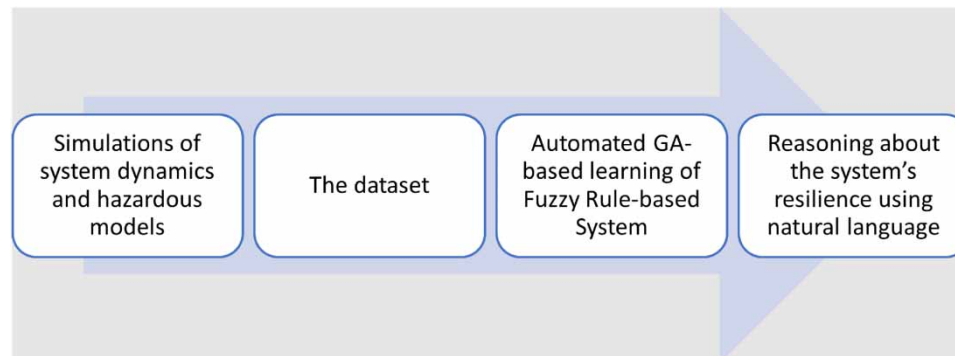
In this paper, the use of a novel genetic fuzzy rule-based system (FRBS) is proposed for assessing the resilience of a water resources system to hazards. The proposed software framework generates a set of highly interpretable rules that transparently represent the causal relationships of hazardous events, their timings, and intensities that can lead to the system's failure. This is achieved automatically through an evolutionary learning procedure that is applied to the data acquired from system dynamics (SD) and hazard simulations. The proposed framework for generating an explainable predictive model of water resources system resilience is applied to the Piroć water resources system in the Republic of Serbia. The results indicate that our approach extracted high-level knowledge from the large datasets derived from multi-model simulations. The rule-based knowledge structure facilitates its common-sense interpretation. The presented approach is suitable for identifying scenario components that lead to increased system vulnerability, which are very hard to detect from massive raw data. The fuzzy model also proves to be a satisfying fuzzy classifier, exhibiting precisions of 0.97 and 0.96 in the prediction of low resilience and high rapidity, respectively.

Key words: fuzzy rule-based system, genetic algorithm, water system resilience

HIGHLIGHTS

- A FRBS was proposed for assessing the resilience of a water resources system to hazardous events and system element failures.
- An evolutionary learning procedure based on a genetic algorithm (GA) was used to generate a set of linguistic rules from data.
- Fuzzy rules provide a means for transparent reasoning about a system's resilience using natural language.

GRAPHICAL ABSTRACT



INTRODUCTION

Dam and reservoir systems are critical infrastructures and play important roles in hydropower generation, irrigation, water supply, flood protection, and water quality control. Nowadays, dams are, on average, 50 years old, which makes them highly prone to structural damage. Global climate change is expected to lead to dramatic changes in precipitation patterns and an increased frequency and intensity of extreme weather events. As a result, this can lead to unfavorable working conditions outside the designed envelopes of dams and accumulation systems and, consequently, to functional failure. Another common cause of disturbance is earthquakes that induce landslides and pose challenges to hazard and risk assessment (Fan *et al.* 2019). Failures of large dams due to disturbances caused by undesirable events are of serious concern, putting asset owners in need of assessing the water systems' dynamic resilience, i.e., its capability to recover after the disturbance event (Simonovic & Arunkumar 2016; Khatri 2022). Many frameworks have been proposed for analysing and assessing the resilience of water systems (Nikolopoulos *et al.* 2019; Behboudian & Kerachian 2021; Liu *et al.* 2021; Roni *et al.* 2022). Water system safety risk assessment requires the identification of different hazardous events and analysis of their interactions, as well as the spatial and temporal evolution of hazards that may lead to system failure (De Angeli *et al.* 2022). To prevent the disruption of water system functionality, it is essential to identify and highlight multi-hazard interactions that can jeopardize the water system services and safety. To analyse the response of the system to various operating conditions and estimate its resilience, a system dynamics (SD) simulation model's outputs can be used (King *et al.* 2017; Ignjatović *et al.* 2021). A wide range of disturbance scenarios can be generated by combining different hazardous events and system failures and their various intensities, starting points, and durations (King & Simonovic 2020). Those input scenarios are then simulated by SD models to quantify the risks related to flooding protection and hydro-energy generation under environmental hazards and unfavourable conditions (system element failures).

The disturbance scenarios alongside the simulations' results of those scenarios yield a large dataset. That dataset should be analysed to identify and highlight those interactions between system states and hazards that lead to low resilience. Analysis of such large data presented in a tabular form (using sorting and filtering) is practically impossible, while a graphical display of data reveals only part of the useful information. The right way to access big data is automated knowledge extraction from it. That knowledge reveals the system's states and hazards (and their combinations and timings) that should be considered when assessing system resilience. Discovering interpretable knowledge from data and finding potentially useful patterns in data are complex processes. Knowledge processing for effective decision-making requires the direct use of computers (Simonovic 2020) and especially the methods of artificial intelligence (AI), namely its subfield of machine learning (ML).

Employing ML techniques to discover knowledge from data generated through computational simulations has rapidly spread in the field of hydro science (Xu & Liang 2021; Zounemat-Kermani *et al.* 2021; Stojković *et al.* 2023). Hybrid ML and optimization techniques were successfully utilized for hydrological streamflow forecasting (Ibrahim *et al.* 2022), prediction of water resource demand, and their optimal allocation (Men *et al.* 2019; Li *et al.* 2021; Wu *et al.* 2021; Zhang & Zhang 2021). A long short-term memory (LSTM) deep learning model with parameters optimized by the Ant Lion Optimizer was used for streamflow time-series prediction, resulting in remarkable accuracy (Yuan *et al.* 2018; Latif & Ahmed 2021). Least square support vector machine (LSSVM) models and their hybrid versions were found to be highly accurate (Adnan *et al.* 2020; Ikram *et al.* 2022a). In Adnan *et al.* (2020), authors showed that LSSVM and multivariate adaptive regression splines (MARS) can provide more accurate streamflow predictions in comparison to optimally pruned extreme learning machine (OP-ELM) and M5Tree-based models. The MARS model's ability to predict monthly streamflow was also compared to the group method of data handling-neural network (GMDH-NN) and the dynamic evolving neural-fuzzy inference system (DENFIS) (Adnan *et al.* 2021a). The accuracy of ML models for streamflow prediction was successfully improved by using a covariance matrix adaptation evolution strategy for tuning control parameters (Ikram *et al.* 2022b). Liu *et al.* (2021) successfully applied a support vector machine model optimized by the modified Grey Wolf Algorithm to improve the accuracy of the evaluation of the resilience of the water resource system in the irrigation areas.

However, even with their remarkable advancement, a key limitation to the use of the most prominent ML models is that they often lack transparency and interpretability. Those techniques generate black-box models, which do not provide an explanation for the decisions they will take and cannot help in understanding the dependencies between adverse events and the system's resilience. Additionally, such models cannot be extended with a human experience. For AI to be trusted, greater transparency can be provided by the means of explainable AI (XAI) systems (Adadi & Berrada 2018). One way to ensure explainability is by creating a fuzzy rule-based system (FRBS) that is intrinsically understandable and comprehensible

(Chimatapu *et al.* 2018; Mencar & Alonso 2019; Fernandez *et al.* 2019). The FRBS is based on fuzzy logic introduced by Zadeh (1965). The main part of the FRBS is a set of IF-THEN rules expressed in natural language, which serve as a part of the knowledge base (KB) for fuzzy logic-based inferencing. Another important part of the KB is the database of the membership functions (MFs), which enable the linguistic representation of numeric values. The data generated through the numerous simulations of disturbance scenarios can be used for learning MFs and fuzzy rules. The process of learning fuzzy rules can be viewed as finding the set of linguistic rules that, based on the data, best represents a relationship between input and output. Searching for the best set of rules is guided by the optimization of a given performance metric. Having the optimization task in mind, the automatic learning of fuzzy rules can be performed by means of a genetic algorithm (GA) (Holland 1992). A GA is a powerful evolutionary algorithm that is widely used as a global search technique for finding near-optimal solutions in complex search spaces. To date, various approaches had been proposed for GA-based learning of fuzzy rules (Herrera & Magdalena 1997), and many of them have been successfully employed in complex real-world systems modelling (Harp *et al.* 2009; Zanganeh 2017).

Fuzzy systems are often used in hydro science. The majority of the research is focused on the use of the adaptive neuro-fuzzy inference system which is a mixed fuzzy-neural approach (ANFIS) (Adnan *et al.* 2021b; Jain *et al.* 2022; Vakili & Mousavi 2022). Zanganeh (2017) combines subtractive clustering, GA, and ANFIS to generate and optimize input fuzzy sets and fuzzy rules for the prediction of wind-driven wave parameters. The fuzzy logic-based approach has also proven successful in water system safety risk assessment (Duhalde *et al.* 2018; Fu *et al.* 2018; Ribas *et al.* 2021). Duhalde *et al.* (2018) apply fuzzy sets to identify areas with a high vulnerability that were not detected by using traditional approaches. Ribas *et al.* (2021) partition the input space to linguistic variables based on literature findings and construct the rule base using expert knowledge. Jacquin & Shamseldin (2006) develop a rainfall-runoff model using the Takagi–Sugeno fuzzy inference system. For generating the FRBS from data, they use a two-stage constrained optimization procedure, involving an evolutionary algorithm and simplex search. Van der Heijden & Haberlandt (2015) introduce a fuzzy rule-based metamodel for the simulation of monthly nitrate loads to replace a heavy process-based model. The authors carried out the training of the fuzzy rule systems with simulated annealing, while the fuzzy sets for describing input variables were allocated through a statistical procedure. Sedighkia *et al.* (2021) use expert-defined fuzzy inference systems as part of a coupled knowledge-based system–optimization model to assess the environmental flow downstream of the reservoirs as one of the important water resource systems. Despite the popularity of the fuzzy approach in solving problems of the aquatic environment, to the best of the authors' knowledge, GA-based learning of FRBS has not yet been used to assess water resource system resilience.

The present study proposes a general software framework for support in identifying the system's states, external events, and their temporal dependencies that affect the reduced dynamic resilience of the system. Two properties of dynamic resilience are considered: robustness and rapidity. The proposed framework automatically extracts the knowledge from data and represents it by fuzzy rules. The rules express the cause-and-effect relationship between the system state and hazardous events on one side and robustness and rapidity on the other. Linguistic terms are used to qualitatively describe system states, events, and resilience measures. We aim to overcome the problem of perceiving mutual connections and influences between inputs and outputs from a large amount of data. The main goal is to provide a comprehensible FRBS that enables a user to easily grasp the causal relationships of events, their timings, and intensities that can lead to the system's failure. To generate a compact set of highly interpretable linguistic rules, we use an evolutionary learning procedure based on the GA. The data-driven GA-based generation of FRBS does not rule out further refinement using experts' opinions but rather allows expertise to be incorporated into the FRBS naturally and transparently by translating it into fuzzy IF-THEN rules. The obtained FRBS can also be used as the metamodel for predicting the system's resilience under given conditions avoiding costly simulations to reason about the resilience of the water resource system. The FRBS enables the fuzzy classification of disturbance scenarios according to how they affect system resilience in terms of qualitatively described system robustness and rapidity. These two properties of dynamic resilience were previously used to assess the water system's dynamic resilience under hazardous events (Stojković *et al.* 2023). Predicting resilience properties is considered a regression problem and is solved using an artificial neural network (ANN), which accurately reproduces key dynamic resilience parameters. However, the ANN is a black-box predictive model, which does not provide a transparent explanation of how the variables are jointly related to each other to reach a final prediction. Our aim is to unbox the reasoning process and assess the dynamic resilience under hazardous events transparently. We extract the high-level compact knowledge from the low-level large data. In contrast to the ANN-based approach, we gain a comprehensible KB for reasoning about the system's resilience using natural language and thus facilitate the analysis of large data gathered from simulations. The method that we propose is universally applicable

to any water system for which data are available regarding the system's states and external events that can affect the resilience of a system.

The rest of the paper is organized as follows: the second section gives a short description of the water resource system model and the case study, followed by an explanation of the FRBS and GA-based learning of fuzzy rules from data in more detail. The third section describes the process of generating our FRBS. The results and discussion are presented in fourth section followed by the concluding remarks.

METHODS

Water resource system model

To mimic the non-linear behaviour of the water resources system, a system dynamic model is used (Stojkovic & Simonovic 2019; King & Simonovic 2020) alongside the earthquake (Rakić *et al.* 2022) and flood dynamics models to simulate the impact of the temporal coincidence of unexpected hazardous events on the system. To include the dependence of the system's operation on the inflow and the system's initial state, the variable initial water content in the reservoir at the time of system failure is also generated (Ivetić *et al.* 2022). Simulation inputs are varied to determine system behaviour under various conditions by analysing a wide range of simulation outcomes depending on the initial reservoir levels, inflows, timing, and intensities of hazards. To measure the flood-related risk of the water system based on the output from SD simulations, dynamic resilience is used as a time-dependent parameter since it overperforms the static risk measures such as system reliability or vulnerability (Ignjatović *et al.* 2021). Dynamic resilience captures the system's robustness and rapidity (Simonovic & Arunkumar 2016). Rapidity describes the recovery time of the water resources system forced by an external or internal hazard, whereas robustness explains a maximal reduced capacity of the water system in terms of delivering the service requirements. To provide an explainable framework for assessing the resilience of water systems, the FRBS is developed using the data gathered through the simulations of SD and hazard models of the Pirot water resources system in the Republic of Serbia.

Case study

The Pirot water resources system, located in the Republic of Serbia, extends over a flood-prone area and includes the Zavoj reservoir at the Visočica river, which is hydraulically connected by a pressure tunnel with a hydropower plant (HPP) Pirot. The HPP Pirot conveys the HPP outflows from the Visočica river to the Nišava river. The Pirot water system mitigates floods at the Nišava river, generates hydropower, and controls downstream water quality by regulating the outflows from the reservoir over the low-flow seasons. The management of the Pirot water system depends on the actual volume of water stored in the reservoirs, inflows, and energy demand (Ignjatović *et al.* 2021). The main characteristics of the water system used in this study are listed in Table 1.

The dataset generated from the outputs of simulations of the SD and hazards models of the Pirot water system contains 900 records, with six input and two output variables. The input variables are the following: the flood hydrograph peak value (Q_{max}), the corresponding return period (T), the initial water volume in the reservoir (V_{init}), the temporal distance between the flood peak and the earthquake start time (t_{dist}), the earthquake duration (t_e), and the normalized value of the maximum seismic acceleration (α). The outputs are numerical values: robustness and rapidity. In Table 2, a look at the first five rows of the dataset is given. It can be noticed that the data entries are in the form of scalars, with Q_{max} obtained as the peak value of the flow time-series (hydrograph).

Fuzzy rule-based system

The most powerful form of conveying information that people have about a real-world problem that requires reasoning is natural language (Ross 2005). Fuzzy logic and FRBSs as the most important area of its application successfully utilize this

Table 1 | The characteristics of the Zavoj reservoir in the Pirot water resources system

Reservoir	Year built	Drainage area (km ²)	Annual inflows (m ³ /s)	Active volume (10 ⁶ m ³)	Flood storage volume (10 ⁶ m ³)	Minimal operational level (m.a.s.l.)	Spillway capacity (m ³ /s)
Zavoj	1990	571	6.2	140	5.5	568	1,820

Table 2 | First five rows of the dataset

	Q_{max}	T	V_{init}	α	t_{dist}	t_e	Robustness	Rapidity
0	2592.41	4414.84	1.67×10^8	0.05	6.00	95.00	0.01	1117.00
1	875.31	195.49	1.75×10^8	0.34	3.00	65.00	0.80	10.00
2	1737.83	1400.35	1.59×10^8	0.40	2.00	60.00	0.04	259.00
3	3287.09	8728.54	1.53×10^8	0.67	7.00	33.00	0.03	387.00
4	1278.08	579.57	1.63×10^8	0.54	4.00	45.00	0.07	139.00

power. In the FRBS, knowledge of the modelled system as well as the interactions and relationships that exist between its parts are presented using fuzzy sets and fuzzy logic (Cordoon *et al.* 2001). Fuzzy sets (Zadeh 1965) enable the linguistic representation of numerical variables through their membership degree to a certain linguistic term. The gradual transition of the membership in the range from 0 to 1 describes vagueness and ambiguity that is common in natural language and allows direct human interaction (Ishibuchi *et al.* 2004). The basic concepts of fuzzy sets theory are not included here but can be found in Ross (2005). The FRBS consists of two main components: KB and the fuzzy inference system. KB is composed of the rule base (RB) and the database of the MFs used to model the linguistic terms. The fuzzy inference system uses fuzzy logic to determine the conclusions that can be inferred considering the information stored in the KB and the user input. Both MFs and IF-THEN rules can be designed by experts or extracted from data.

Generating the FRBS and using it to assess the water resources system resilience

In the present research, we have used the method for the estimation of MFs from data presented in Bhatt *et al.* (2012) and the GA-based FRBS learning method presented in Yuan & Zhuang (1996) for generating KB. For defining the linguistic terms, the authors assume trapezoidal and triangle MFs. The algorithm for learning the MFs from the data presented in Bhatt *et al.* (2012) determines the MF's parameters using the data clustered by fuzzy c-means (FCM) clustering (Ross 2005) – cluster matrix and centre vector. The algorithm approximates as many MFs, as there are clusters in the clustered data. The detailed description of the algorithm will be omitted here as it can be found in Bhatt *et al.* (2012).

To generate fuzzy rules for the KB, we use genetic algorithms (Goldberg 2006), which have proven as a robust and powerful mechanism when it comes to solving challenging optimization problems. They mimic the natural evolution process by modifying the set of potential solutions called the population, through selection, crossover, and mutation of individuals (chromosomes). GA starts with the randomly chosen initial population. To select the best candidates for reproduction, each chromosome in the current population must be evaluated and assigned a fitness value. An offspring population is created from the selected parents in the current population by applying a crossover operator with a certain probability. According to a predefined probability, the mutation operator then alters some of the offspring to increase the variability of a population and prevent premature convergence to a local optimum. The offspring population replaces the current population, and the processes of selection, crossover, and mutation for evolving new generations are repeated in a loop until a stopping condition is met.

The three most popular approaches in GA-based learning of FRBS are the Pittsburgh approach (Smith 1980), the Michigan approach (Booker *et al.* 1989), and iterative rule learning (Cordón & Herrera 1997). The Pittsburgh approach encodes an entire set of rules in a single chromosome, whereas the Michigan approach and iterative rule learning use separate chromosomes to encode each rule in the KB. Yuan & Zhuang's (1996) method for learning the FRBS utilized in this research uses the Michigan approach. The method is presented in detail in Yuan & Zhuang (1996), and only the key points will be addressed here as part of a detailed description of the general workflow of developing and using the FRBS to assess the resilience of a complex water resources system.

A general workflow of developing and using the explainable FRBS for assessing the resilience of a complex water resources system to hazards is depicted in Figure 1.

The workflow includes the following four phases:

Phase 1. The dataset is split into the training and test sets using an 80–20% split, respectively. The data for these subsets are selected randomly to preserve generality.

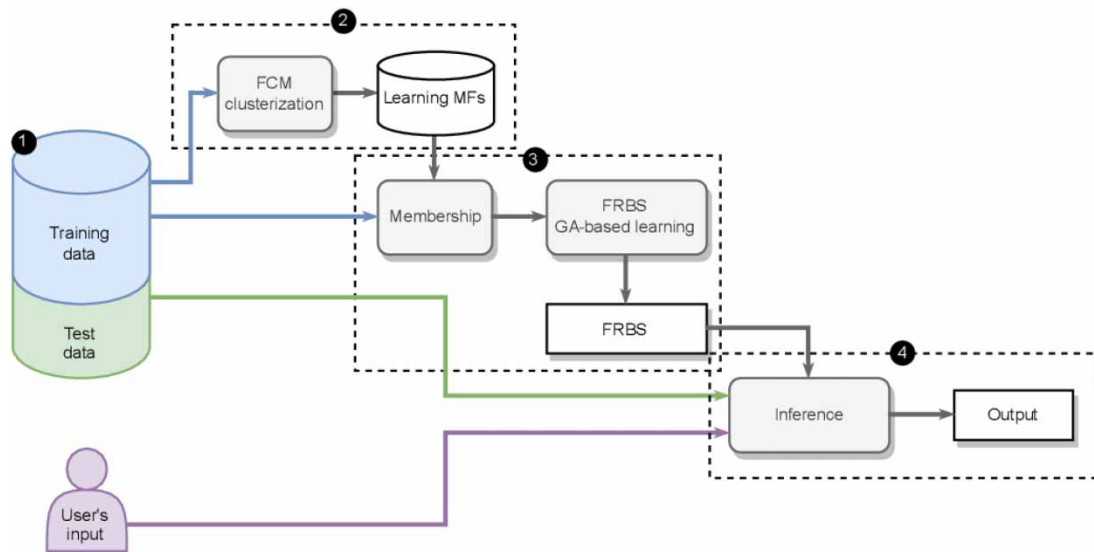


Figure 1 | A general workflow of developing and using the FRBS for assessing the resilience of a complex water resources system.

Phase 2. The training data are used for fuzzy partitioning of input and output space that is creating the database of the MFs, which model the linguistic terms for all inputs and all outputs. For example, the following three linguistic terms can be used for describing a value of the flood hydrograph peak (Q_{max}): *Low*, *Medium*, or *High*. Each linguistic term is determined by a corresponding fuzzy set, i.e., by its MF. To determine MFs from data, the data must first be clustered. The FCM clustering algorithm (Bezdek 1982) generates cluster centres from the data and assigns membership degrees of each numeric (crisp) value to each fuzzy cluster. FCM allows data to belong to more than one cluster at the same time. Interpretability of clustered data is improved by representing it with triangular, trapezoidal, or Gaussian MFs, rather than representing it as a matrix of membership values. The algorithm for learning MFs from clustered data (Bhatt *et al.* 2012) is used for generating trapezoidal MFs.

Phase 3. Obtained MFs are used for determining the membership degree of the crisp values of the training data (both input and output variables) to each fuzzy partition. For example, according to the MFs for the fuzzy sets *Low*, *Medium*, and *High*, a crisp Q_{max} value of 2,500 m³/s could, at the same time, belong to the fuzzy set *Medium* and the fuzzy set *High* to a degree of 0.4 and 0.7, respectively. By determining the memberships of the crisp training data to fuzzy partitions, we transform our training dataset, such that each numeric input (output) value is replaced with its membership degrees to the fuzzy sets that correspond to the linguistic terms used for describing the input (output) variable.

The transformed dataset is utilized for learning the FRBS through the evolutionary learning procedure based on GA, which consists of the steps depicted in Figure 2.

Instead of the initial population consisting entirely of randomly chosen chromosomes, the proposed method uses the binary encoding of the training dataset as part of the initial population. Combining randomly generated rules with the rules generated from training instances provides a diverse initial population with rather specific rules converted from training data, and the additional knowledge (not covered by training examples) contained in randomly generated individuals. A rule is coded as one chromosome consisting of segments that correspond to either an input variable in the condition part of the rule or an output variable in the conclusion part of the rule. Each segment consists of binary genes corresponding to the linguistic term for the inputs or outputs. For example, the chromosome (0 1 0, 1 0 0, 1 0 0, 0 0 1 0, 1 0 0, 1 0 0; 1 0 0 0, and 0 0 1), where commas separate conditions and a semicolon separates the IF and the THEN part of the rule translates to the rule IF (Q_{max} is *Medium*) AND (T is *Short*) AND (V_{init} is *Low*) AND (α is *Medium*) AND (t_{dist} is *Short*) AND (t_e is *Short*) THEN (Robustness is *Low*) AND (Rapidness is *High*). A training example is converted to a chromosome by encoding a membership degree to 1 if it is greater or equal to 0.5 and 0 otherwise.

After creating the initial population in Step 1 of the algorithm, the fitness of all rules from the population is determined in Step 2. Each rule is evaluated based on three criteria: the accuracy, the coverage, and the competitiveness measure in terms of

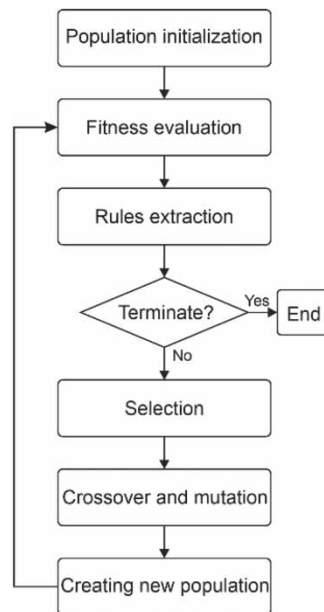


Figure 2 | The steps of the genetic evolution process.

the contribution of each rule to the population. The accuracy of the rule indicates the degree to which the condition of the rule implies the conclusion of the rule. The coverage refers to the portion of the training dataset that is covered by the specific rule, meaning that the conditional part of the rule corresponds to the input of the training example. The larger the coverage, the more general the rule is. The third criterion for rule evaluation is the rule's contribution to determining the correct conclusion for each example in the training set. Further details on fitness evaluation are beyond the scope of this paper and can be found in [Yuan & Zhuang \(1996\)](#).

In Step 3, rule extraction is performed by first selecting the rules with accuracy above the desired accuracy level specified by the user. From this candidate set, the best rules are further selected one by one based on three criteria in sequential order: accuracy, coverage, and fitness. From all the rules with accuracy above the predefined value, the rules with the highest accuracy are first extracted. If more rules have approximately the same accuracy, the one with the largest coverage is selected. If several rules have approximately the same coverage, the one with the highest fitness value is extracted and put into the final set of rules. All the training examples that are correctly classified by the extracted rule are removed from the training set. The process of moving the rules from the population into the final set of rules is performed until all the examples are removed from the training set or there is no rule remaining in the current population. Next, if the stopping criterion is not satisfied, the procedure advances to Step 4, where parents are selected proportionally to their fitness. The two parents can crossover if they have the same conclusion part of the chromosome. The crossover operator in Step 5 exchanges the randomly selected segments between parent chromosomes to generate two child chromosomes. The mutation operator generates a new chromosome by modifying the genes of one or more segments in an existing chromosome. An offspring generated through crossover and mutation cannot survive if it represents a rule that is covered by any rule in the population.

In Step 6, a new generation is formed through the replacement of old, weak members with new offspring. After Step 6, the algorithm returns to Step 2 to evaluate the new generation.

Phase 4. The FRBS created in Phase 3 represents the knowledge in the form of fuzzy rules. The rule base and the MFs can be used to perform the Mamdani fuzzy inferencing. A set of six numeric values of input variables defines one scenario in which the water resources system and hazards have the characteristics determined by the values of the input variables. The FRBS, through the process of Mamdani inferencing, assigns each scenario to a certain fuzzy set corresponding to one of the linguistic variables that describe robustness and to the one fuzzy set that describes rapidity. Crisp input values intersect the antecedent MFs of certain rules at some membership levels. Due to the overlapping of the fuzzy sets, several rules are applicable at the same time for one specific scenario. The minimum of all membership values determines matched rules' consequent MFs. Aggregation (union) of conclusions of all matched rules yields the overall conclusion in the form of a fuzzy set,

e.g., the linguistic value of an output. Once the FRBS is learned and tested, it can be used for inferencing with new, user-provided data for assessing the system's resilience.

RESULTS AND DISCUSSION

The proposed procedure for GA-based learning of the FRBS from data was implemented in the Python programming language, and the Pirot water system simulation dataset was used for training and testing. As previously mentioned, the data has six inputs, two outputs, and 900 records (Table 2). Figure 3 visualizes the relationships between each input and the output. The scatterplots at the first line show how robustness depends on Q_{max} , T , V_{init} , α , t_{dist} , and t_e , respectively. The scatterplots at the second line depict the dependence between each input and the rapidity. The relationships between the inputs and outputs are quite non-linear. The data were split into the training and test sets using an 80–20% split, respectively. The data for these subsets are selected randomly to preserve generality.

In Phase 2 of the abovementioned procedure, we used the FCM algorithm to cluster the values of each input and output variable and obtained the MFs of all linguistic terms used for describing the inputs and outputs. The inputs, outputs, corresponding universal set, and the terms used for partitioning the input and output universes are presented in Table 3. In Figure 4, a graphical representation is given for each fuzzy set that corresponds to a specific linguistic term used for describing the values of the input and output variables.

The crisp values of inputs and outputs for each data record were mapped into corresponding memberships to fuzzy sets (Phase 3). For example, the Q_{max} crisp value of $1,500 \text{ m}^3/\text{s}$ has the membership degree of 0.25 to the fuzzy set *Low* and 0.8 to the fuzzy set *Medium*. Considering the memberships of all six inputs and two outputs to each of the corresponding fuzzy sets, each record of the dataset was coded as an array of 26 real numbers. Each element of the array corresponds to a membership degree of input (output) to one of the fuzzy sets used for the linguistic description of the input's (output's) intensity. After determining the memberships of the crisp training data to fuzzy partitions, we have noticed that the training dataset is highly imbalanced with approximately 75% of records related to the *Low* robustness of the system, 16% of *Medium-low*, 6% of *Medium*, and just 3% of records with *High* robustness. This is the consequence of numerous simulations being dedicated to analysing situations that could lead to the Pirot system failure under hazards. The distribution of data records according to rapidity is 20% *High*, 32% *Medium*, and 48% *Low* rapidity.

The GA was further applied to generate rules from the training data. The initial population consisted of 720 binary-encoded training examples and 280 randomly generated individuals. In each generation, 20% of the population was selected for reproduction and the mutation rate was 0.02. The parameters used in calculating the coverage and accuracy of the rules are the same as in Yuan & Zhuang (1996). The number of rules extracted from the population decreased through generations. The final rule base comprised 240 rules. Table 4 presents the part of the fuzzy rule system obtained in Phase 3, for the Pirot case study. Each rule includes six inputs and two outputs in the binary notation of GA chromosomes.

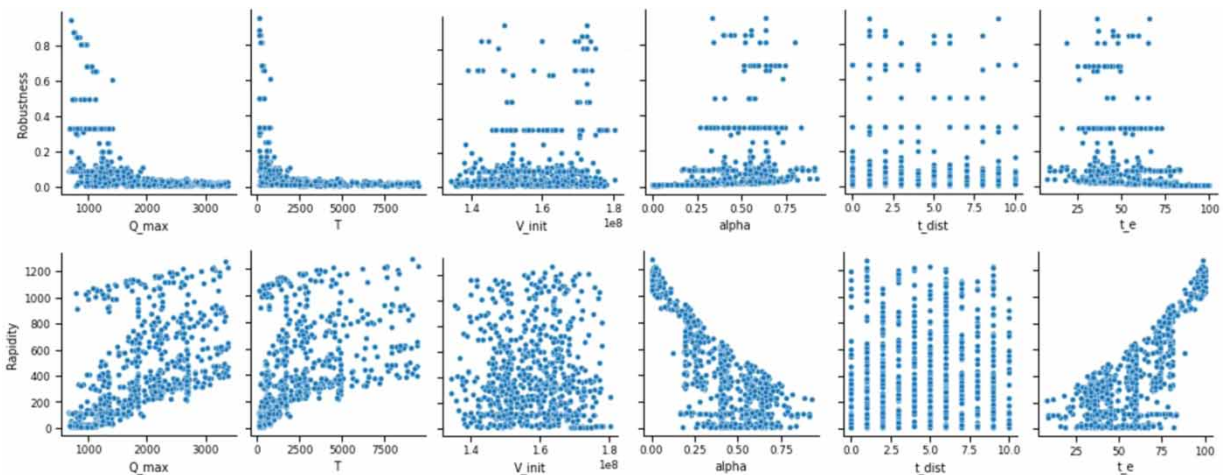


Figure 3 | Pairwise scatterplots of input and output data.

Table 3 | Input and output variables and the main characteristics of corresponding MFs

Inputs	Description	Universal set	Terms	Outputs	Universal set	Terms
Q_{max}	The flood hydrograph peak value	[500, 4,500]	Low Medium High	Robustness	[0, 1]	Low Medium- Low Medium High
T	The return period	[10, 10,000]	Short Medium Long	Rapidity	[10, 1,300]	Low Medium High
V_{init}	The initial water volume in the reservoir	$[133 \times 10^6, 190 \times 10^6]$	Low Medium High			
α	The normalized value of the maximum seismic acceleration	[0, 1]	Small Medium- Small Medium Large			
t_{dist}	The temporal distance between the flood peak and the earthquake start time	[0, 12]	Short Medium Long			
t_e	The earthquake duration	[0, 100]	Short Medium Long			

Fuzzy rules can be translated into verbal expressions and thus provide a means for reasoning about a system’s resilience using natural language. This gives the fuzzy model transparency, which popular black-box models do not have, and thus an advantage over them since it enables a user to easily grasp the causal relationships of events, their timings, and intensities that can lead to the system’s failure. Since there is no objective criterion to measure transparency, here we can only show how to interpret the fuzzy rules from Table 3 and thus demonstrate transparency in the specific example.

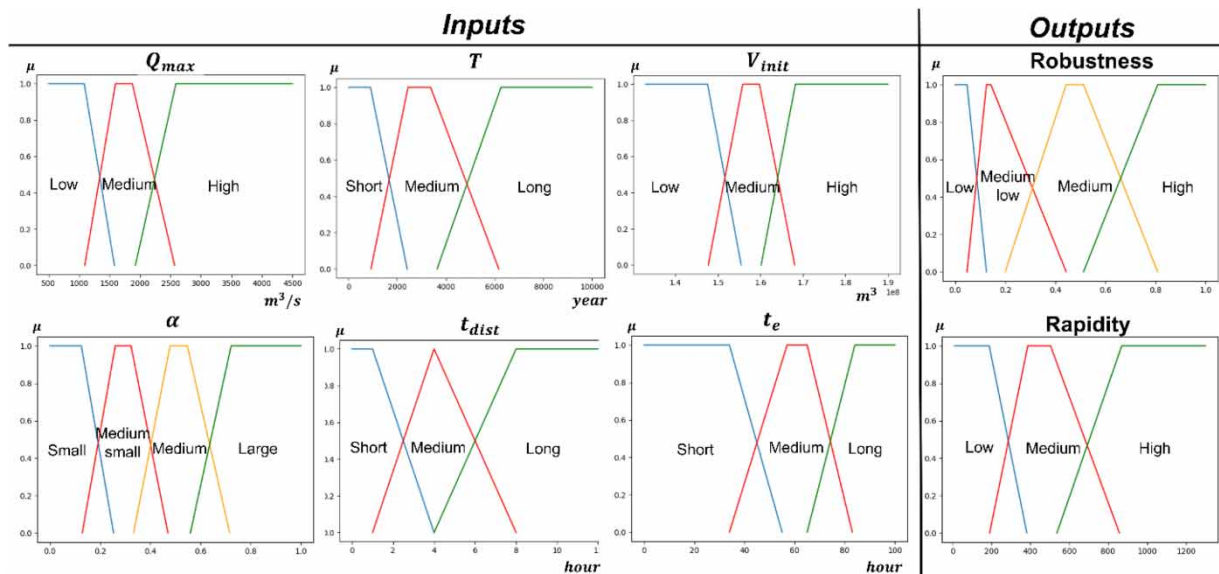


Figure 4 | Fuzzy sets for the linguistic description of inputs and outputs obtained from the data.

Table 4 | Examples of the fuzzy rules obtained via the GA

No.	Q_{max}			T			V_{init}			α			t_{dist}			t_e			Robustness			Rapidly				
	L	M	H	Sh	M	Lg	L	M	H	S	MS	M	La	Sh	M	Lg	Sh	M	Lg	L	ML	M	H	L	M	H
R1	1	0	0	1	0	0	0	0	1	0	0	1	0	0	0	1	1	0	0	0	0	0	1	1	0	0
R2	1	0	0	1	0	0	1	0	0	0	0	1	0	0	1	0	0	1	0	0	0	1	0	0	1	0
R3	0	0	1	0	1	0	1	0	0	1	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	1
R4	0	1	0	1	0	0	1	0	0	0	0	1	0	0	1	0	1	0	0	0	0	1	0	0	1	0
R5	0	0	1	0	0	1	0	1	0	0	0	0	1	0	0	1	1	0	0	1	0	0	0	0	0	1

L, Low; M, Medium; H, High; Sh, Short; Lg, Long; S, Small; MS, Medium-Small; ML, Medium-Low; La, Large.

R1: IF (Q_{max} is Low) AND (T is Short) AND (V_{init} is High) AND (α is Medium) AND (t_{dist} is Long) AND (t_e is Short) THEN (Robustness is High) AND (Rapidly is Low)

R2: IF (Q_{max} is Low) AND (T is Short) AND (V_{init} is Low) AND (α is Medium) AND (t_{dist} is Medium) AND (t_e is Medium) THEN (Robustness is Medium) AND (Rapidly is Medium)

R3: IF (Q_{max} is High) AND (T is Medium) AND (V_{init} is Low) AND (α is Small) AND (t_{dist} is Short) AND (t_e is Long) THEN (Robustness is Low) AND (Rapidly is High)

R4: IF (Q_{max} is Medium) AND (T is Short) AND (V_{init} is Low) AND (α is Medium) AND (t_{dist} is Medium) AND (t_e is Short) THEN (Robustness is Medium) AND (Rapidly is medium)

R5: IF (Q_{max} is High) AND (T is Long) AND (V_{init} is Medium) AND (α is Large) AND (t_{dist} is Long) AND (t_e is Short) THEN (Robustness is Low) AND (Rapidly is High)

The examples of obtained rules indicate that the proposed approach successfully extracted high-level knowledge from the large amount of numerical data gathered by simulating numerous scenarios using the SD model and hazard models.

The fuzzy KB provides useful insights into the system’s vulnerability. It enables the analysis of potential combinations of intensities and durations of indicators and events that were most likely to result in significant safety impacts. The bubble chart in Figure 5 visualizes how often each of the input values occurs in rules with low robustness and high or medium rapidly as an outcome. It is obvious that in the majority of scenarios, which lead to an unfavourable outcome, the earthquake

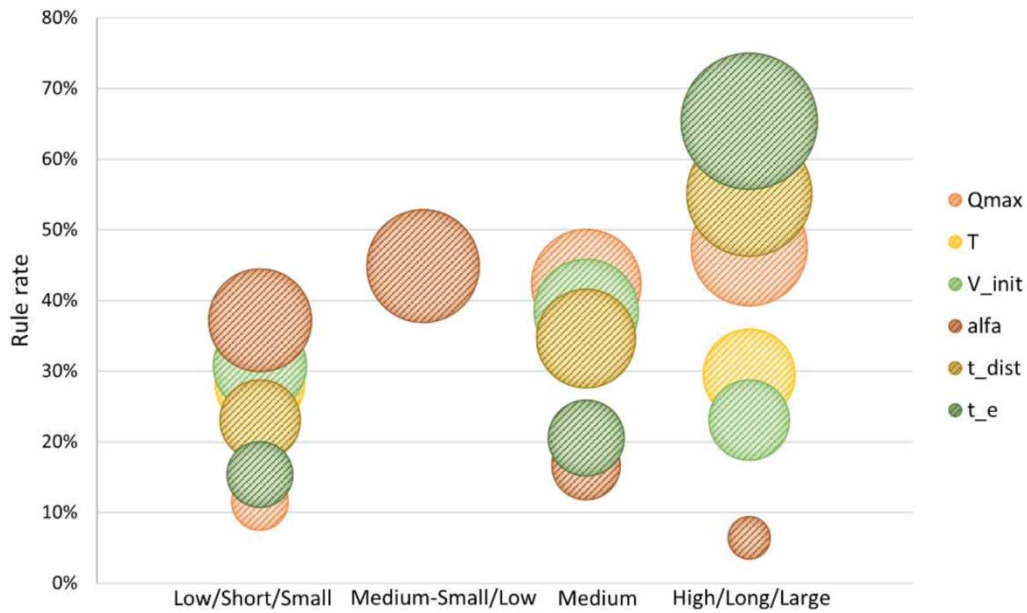


Figure 5 | The occurrences of each input’s linguistic value in rules with low robustness and high or medium rapidly as an outcome. The size of a bubble is determined by the number of rules that contain the specific linguistic value of an input.

duration (t_e) is long. Medium and high peak values of the flood hydrograph (Q_{max}) have an almost equal share, while low Q_{max} is rarely present in critical scenarios. Unexpectedly, low and medium initial water volume in the reservoir (V_{init}) are more frequent in those scenarios than high volume, but they are usually combined with long earthquake durations or high values of Q_{max} . All three values for T , as well as for t_{dist} , are almost equally represented in the scenarios that lead to the system disruption, while high α values do not play a significant role in low system resilience.

The Python Scikit-Fuzzy fuzzy logic toolbox (Scikit-Fuzzy) was used to implement the obtained FRBS and enable Mamdani inferencing with the test data that contained 182 records. First, we evaluated the proposed solution by comparing the expected output's assignment to one of the fuzzy sets used for partitioning the output universe with the predicted output's assignment. Tables 5(a) and 5(b) show the confusion matrices for robustness and rapidity, respectively. According to the confusion matrix, 17 out of 157 scenarios that lead to low robustness were classified as medium-low. Six out of 15 scenarios that should be medium-low were misclassified, five as ones with low robustness, and one as medium robust. For the eight scenarios that yield medium robustness, three were incorrectly classified as medium-low, whereas one high robust example was classified as a medium. Having in mind that adjacent fuzzy sets overlap, misclassifying an example into a fuzzy set that is adjacent to the correct one is not exactly a wrong judgment. The confusion matrix for the rapidity shows results almost similar to the ones for the robustness, with the exception that one example with low rapidity was misclassified as high, which is not the fuzzy set adjacent to the fuzzy set *Low*. Table 6 lists the values of the metrics resulting from the confusion matrices. Since the data were imbalanced, we did not calculate the accuracy. As expected, precision, recall, and f_1 -score have very good values when it comes to the examples with low robustness and with low or high rapidity because they prevailed in the training set.

The results prove that the fuzzy KB extracted from data using the proposed software framework can serve for the comprehensible identification of potential combinations of events that were most likely to result in significant safety impacts. Moreover, the evaluation of the classification model reveals that the obtained FRBS may serve as an explainable fuzzy predictor.

Table 5 | Confusion matrix for (a) the robustness and (b) the rapidity

(a) Robustness

Actual	Predicted			
	Low	Medium-Low	Medium	High
Low	140	17	0	0
Medium-Low	5	9	1	0
Medium	0	3	5	0
High	0	0	1	1

(b) Rapidity

Actual	Predicted		
	Low	Medium	High
Low	90	7	1
Medium	6	15	1
High	0	10	52

Table 6 | Results on the test set

Robustness	Precision	Recall	f_1 -score	Rapidity	Precision	Recall	f_1 -score
L	0.97	0.89	0.93	L	0.94	0.92	0.93
ML	0.31	0.60	0.41	M	0.47	0.68	0.56
M	0.71	0.62	0.67	H	0.96	0.84	0.90
H	1.00	0.50	0.67				

CONCLUSIONS

This study aimed to propose a general software framework for assessing the resilience of a water resources system to hazardous events using a FRBS. The proposed methodology includes four phases to create an FRBS from data. The data were obtained as a result of numerous simulations of the SD model and hazard models, which introduce the temporal coincidence of unfortunate events. As the case study area, we used the Pirot water resources system in the Republic of Serbia. However, the framework applies to any water system for which there is available data regarding the indicators that can affect the resilience of a system.

The main part of the framework is the evolutionary learning procedure based on a GA that is used for learning the FRBS. The framework also includes determining the fuzzy partitioning of input and output spaces, fuzzification of numerical input and output values, and evolutionary learning of the fuzzy rules. By learning the rules, the high-level compact knowledge is extracted from the low-level large data. In this way, we gain a transparent and comprehensible KB for reasoning about the system's resilience using natural language. Once the fuzzy rule-based model is trained, it can be used independently of the original SD model and hazard models, avoiding costly simulations to analyse the resilience of the water resources system. Additionally, the fuzzy KB can be upgraded using experts' opinions, which can be naturally and transparently translated to fuzzy IF-THEN rules. Knowledge represented in the form of fuzzy rules is used for identifying the events (or system's states), their characteristics, and temporal relations that are most likely to result in significant safety impacts. It enables the detection of scenario components to which the system is sensitive. In the case of the Pirot water system, rule analysis showed that the system's resilience is most threatened by the duration of an earthquake rather than its intensity. One weakness of the proposed framework is the duration of the GA-based learning procedure. The computational cost of the fitness evaluation is extremely high due to the immense training dataset against, which all the rules in the population are evaluated. This can be overcome by the parallel execution of the GA, which will be the aim of our future work.

The FRBS obtained for the Pirot use case was also tested as the predictor of the system's resilience under given conditions. The FRBS was used for fuzzy classification of disturbance scenarios according to how they affect system resilience in terms of qualitatively described system robustness and rapidity. The proposed solution has been evaluated using the test dataset. The expected output's assignment to one of the fuzzy sets used for partitioning the output universe was compared to the predicted output's assignment. This resulted in confusion matrices, precision, recall, and f_1 -score for both robustness and rapidity resilience measures. All obtained results were satisfying and showed that it is possible to learn an explainable fuzzy KB from data via GA and use it for highlighting potential combinations of events that were most likely to result in significant safety impacts. With a sufficiently large KB, the proposed solution can be used to guide the further generation of risk scenarios, so rather than simulate numerous randomly generated scenarios, one can use fuzzy analysis to direct the investigation toward combinations of states and events that pose a risk to system resilience. Hence, this approach constitutes a very promising contribution toward bringing great predictive power to assess the resilience of complex water and infrastructure systems to hazardous events and benefits water management practitioners.

ACKNOWLEDGEMENTS

The authors express their gratitude to the Science Fund of the Republic of Serbia for the support through the project of the PROMIS call, 6062556, DyRes_System: 'Dynamics resilience as a measure for risk assessment of the complex water, infrastructure and ecological systems: Making a context'.

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

- Adadi, A. & Berrada, M. 2018 Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). *IEEE Access* 6, 52138–52160.

- Adnan, R. M., Liang, Z., Heddarn, S., Zounemat-Kermani, M., Kisi, O. & Li, B. 2020 Least square support vector machine and multivariate adaptive regression splines for streamflow prediction in mountainous basin using hydro-meteorological data as inputs. *Journal of Hydrology* **586**, 124371.
- Adnan, R. M., Liang, Z., Parmar, K. S., Soni, K. & Kisi, O. 2021a Modeling monthly streamflow in mountainous basin by MARS, GMDH-NN and DENFIS using hydroclimatic data. *Neural Computing and Applications* **33** (7), 2853–2871.
- Adnan, R. M., Mostafa, R. R., Islam, A. R. M. T., Kisi, O., Kuriqi, A. & Heddarn, S. 2021b Estimating reference evapotranspiration using hybrid adaptive fuzzy inferencing coupled with heuristic algorithms. *Computers and Electronics in Agriculture* **191**, 106541.
- Behboudian, M. & Kerachian, R. 2021 Evaluating the resilience of water resources management scenarios using the evidential reasoning approach: the Zarrinehrud river basin experience. *Journal of Environmental Management* **284**, 112025.
- Bezdek, J. C. 1982 *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, NY.
- Bhatt, R. B., Narayanan, S. J., Paramasivam, I. & Khalid, M. 2012 Approximating fuzzy membership functions from clustered raw data. In: *2012 Annual IEEE India Conference (INDICON)*. IEEE, Kochi, Kerala, India, pp. 487–492.
- Booker, L. B., Goldberg, D. E. & Holland, J. H. 1989 Classifier systems and genetic algorithms. *Artificial Intelligence* **40** (1–3), 235–282.
- Chimatapu, R., Hagrass, H., Starkey, A. & Owusu, G. 2018 Explainable AI and fuzzy logic systems. In: D. Fagan, C. Martín-Vide, M. O'Neill & M. A. Vega-Rodríguez (eds). *International Conference on Theory and Practice of Natural Computing*. Springer, Cham, pp. 3–20.
- Cordón, O. & Herrera, F. 1997 A three-stage evolutionary process for learning descriptive and approximate fuzzy-logic-controller knowledge bases from examples. *International Journal of Approximate Reasoning* **17** (4), 369–407.
- Cordon, O., Herrera, F., Hoffmann, F. & Magdalena, L. 2001 Genetic fuzzy systems: evolutionary tuning and learning of fuzzy knowledge bases. *Advances in Fuzzy Systems–Applications and Theory* **19**.
- De Angeli, S., Malamud, B. D., Rossi, L., Taylor, F. E., Trasforini, E. & Rudari, R. 2022 A multi-hazard framework for spatial-temporal impact analysis. *International Journal of Disaster Risk Reduction* **73**, 102829.
- Duhalde, D. J., Arumí, J. L., Oyarzún, R. A. & Rivera, D. A. 2018 Fuzzy-based assessment of groundwater intrinsic vulnerability of a volcanic aquifer in the Chilean Andean Valley. *Environmental Monitoring and Assessment* **190** (7), 1–14.
- Fan, X., Scaringi, G., Korup, O., West, A. J., van Westen, C. J., Tanyas, H., Hovius, N., Hales, T. C., Jibson, R. W., Allstadt, K. E. & Zhang, L. 2019 Earthquake-induced chains of geologic hazards: patterns, mechanisms, and impacts. *Reviews of Geophysics* **57** (2), 421–503.
- Fernandez, A., Herrera, F., Cordon, O., del Jesus, M. J. & Marcelloni, F. 2019 Evolutionary fuzzy systems for explainable artificial intelligence: why, when, what for, and where to? *IEEE Computational Intelligence Magazine* **14** (1), 69–81.
- Fu, X., Gu, C. S., Su, H. Z. & Qin, X. N. 2018 Risk analysis of earth-rock dam failures based on fuzzy event tree method. *International Journal of Environmental Research and Public Health* **15** (5), 886.
- Goldberg, D. E. 2006 *Genetic Algorithms*. Pearson Education, India.
- Harp, D. R., Taha, M. R. & Ross, T. J. 2009 Genetic-fuzzy approach for modeling complex systems with an example application in masonry bond strength prediction. *Journal of Computing in Civil Engineering* **23** (3), 193–199.
- Herrera, F. & Magdalena, L. 1997 Genetic fuzzy systems: a tutorial. *Tatra Mountains Mathematical Publications (Slovakia)* **13**, 93–121.
- Holland, J. H. 1992 *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. MIT Press, Cambridge, MA.
- Ibrahim, K. S. M. H., Huang, Y. F., Ahmed, A. N., Koo, C. H. & El-Shafie, A. 2022 A review of the hybrid artificial intelligence and optimization modelling of hydrological streamflow forecasting. *Alexandria Engineering Journal* **61** (1), 279–303.
- Ignjatović, L., Stojković, M., Ivetić, D., Milašinović, M. & Milivojević, N. 2021 Quantifying multi-parameter dynamic resilience for complex reservoir systems using failure simulations: case study of the Pirot Reservoir System. *Water* **13** (22), 3157.
- Ikram, R. M. A., Dai, H. L., Ewees, A. A., Shiri, J., Kisi, O. & Zounemat-Kermani, M. 2022a Application of improved version of multi verse optimizer algorithm for modeling solar radiation. *Energy Reports* **8**, 12063–12080.
- Ikram, R. M. A., Goliatt, L., Kisi, O., Trajkovic, S. & Shahid, S. 2022b Covariance matrix adaptation evolution strategy for improving machine learning approaches in streamflow prediction. *Mathematics* **10** (16), 2971.
- Ishibuchi, H., Nakashima, T. & Nii, M. 2004 *Classification and Modeling with Linguistic Information Granules: Advanced Approaches to Linguistic Data Mining*. Springer Science & Business Media, Berlin, Heidelberg, New York.
- Ivetić, D., Milašinović, M., Stojković, M., Šotić, A., Charbonnier, N. & Milivojević, N. 2022 Framework for dynamic modelling of the dam and reservoir system reduced functionality in adverse operating conditions. *Water* **14** (10), 1549.
- Jacquín, A. & Shamseldin, A. 2006 Development of rainfall–runoff models using Takagi–Sugeno fuzzy inference systems. *Journal of Hydrology* **329** (1–2), 154–173.
- Jain, A., Rallapalli, S. & Kumar, D. 2022 Cloud-based neuro-fuzzy hydro-climatic model for water quality assessment under uncertainty and sensitivity. *Environmental Science and Pollution Research* **29** (43), 65259–65275.
- Khatri, K. B. 2022 Current state and future direction for building resilient water resources and infrastructure systems. *Eng* **3** (1), 175–195.
- King, L. M. & Simonovic, S. P. 2020 A deterministic Monte Carlo simulation framework for dam safety flow control assessment. *Water* **12** (2), 505.
- King, L. M., Simonovic, S. P. & Hartford, D. N. D. 2017 Using system dynamics simulation for assessment of hydropower system safety. *Water Resources Research* **53** (8), 7148–7174.
- Latif, S. D. & Ahmed, A. N. 2021 Application of deep learning method for daily streamflow time-series prediction: a case study of the Kowmung River at Cedar Ford, Australia. *International Journal of Sustainable Development and Planning* **16** (3), 497–501.

- Li, Y., Xie, J., Jiang, R. & Yan, D. 2021 Application of edge computing and GIS in ecological water requirement prediction and optimal allocation of water resources in irrigation area. *PLoS One* **16** (7), e0254547.
- Liu, D., Li, M., Ji, Y., Fu, Q., Li, M., Faiz, M. A., Ali, S., Li, T., Cui, S. & Khan, M. I. 2021 Spatial-temporal characteristics analysis of water resource system resilience in irrigation areas based on a support vector machine model optimized by the modified gray wolf algorithm. *Journal of Hydrology* **597**, 125758.
- Men, B., Wu, Z., Liu, H., Hu, Z. & Li, Y. 2019 Improved grey prediction method for optimal allocation of water resources: a case study in Beijing in China. *Water Supply* **19** (4), 1044–1054.
- Mencar, C. & Alonso, J. M. 2019 Paving the way to explainable artificial intelligence with fuzzy modeling. In: R. Fullér, S. Giove & F. Masulli (eds). *Fuzzy Logic and Applications*. WILF 2018. Lecture Notes in Computer Science 11291. Springer, Cham.
- Nikolopoulos, D., van Alphen, H. J., Vries, D., Palmen, L., Koop, S., van Thienen, P. & Makropoulos, C. 2019 Tackling the ‘New normal’: a resilience assessment method applied to real-world urban water systems. *Water* **11** (2), 330.
- Rakić, D., Stojković, M., Ivetić, D., Živković, M. & Milivojević, N. 2022 Failure assessment of embankment dam elements: case study of the Pirot Reservoir System. *Applied Sciences* **12** (2), 558.
- Ribas, J. R., Severo, J. C. R., Guimarães, L. F. & Perpetuo, K. P. C. 2021 A fuzzy FMEA assessment of hydroelectric earth dam failure modes: a case study in Central Brazil. *Energy Reports* **7**, 4412–4424.
- Roni, M. S., Mosier, T., Feinberg, T. D., McJunkin, T., Toba, A. L., Boire, L. D. & Parvania, M. 2022 Integrated water-power system resiliency quantification, challenge and opportunity. *Energy Strategy Reviews* **39**, 100796.
- Ross, T. J. 2010 *Fuzzy Logic with Engineering Applications*, 3rd edn. John Wiley & Sons, UK.
- Scikit-Fuzzy. Available from: <https://scikit-fuzzy.readthedocs.io/en/latest/overview.html> (accessed 12 July 2022).
- Sedighkia, M., Datta, B., Abdoli, A. & Moradian, Z. 2021 An ecohydraulic-based expert system for optimal management of environmental flow at the downstream of reservoirs. *Journal of Hydroinformatics* **23** (6), 1343–1367.
- Simonovic, S. P. 2020 Systems approach to management of water resources – toward performance based water resources engineering. *Water* **12** (4), 1208.
- Simonovic, S. P. & Arunkumar, R. 2016 Comparison of static and dynamic resilience for a multipurpose reservoir operation. *Water Resources Research* **52** (11), 8630–8649.
- Smith, S. F. 1980 *A Learning System Based on Genetic Adaptive Algorithms*. PhD Thesis, University of Pittsburgh, Pennsylvania, USA.
- Stojkovic, M. & Simonovic, S. P. 2019 System dynamics approach for assessing the behaviour of the Lim Reservoir system (Serbia) under changing climate conditions. *Water* **11** (8), 1620.
- Stojković, M., Marjanović, D., Rakić, D., Simić, V., Milivojević, N. & Trajković, S. 2023 Assessment of water resources system resilience under hazardous events using system dynamic approach and artificial neural networks. *Journal of Hydroinformatics*. <https://doi.org/10.2166/hydro.2023.069>.
- Vakili, S. & Mousavi, S. M. 2022 Rainfall–runoff modeling using adaptive neuro-fuzzy inference system (ANFIS) and genetic algorithm (GA). *Water Supply* **22** (10), 7460–7475.
- Van der Heijden, S. & Haberlandt, U. 2015 A fuzzy rule based metamodel for monthly catchment nitrate fate simulations. *Journal of Hydrology* **531**, 863–876.
- Wu, J., Wang, Z. & Dong, L. 2021 Prediction and analysis of water resources demand in Taiyuan City based on principal component analysis and BP neural network. *AQUA – Water Infrastructure, Ecosystems and Society* **70** (8), 1272–1286.
- Xu, T. & Liang, F. 2021 Machine learning for hydrologic sciences: an introductory overview. *Wiley Interdisciplinary Reviews: Water* **8** (5), e1533.
- Yuan, Y. & Zhuang, H. 1996 A genetic algorithm for generating fuzzy classification rules. *Fuzzy Sets and Systems* **84** (1), 1–19.
- Yuan, X., Chen, C., Lei, X., Yuan, Y. & Muhammad Adnan, R. 2018 Monthly runoff forecasting based on LSTM–ALO model. *Stochastic Environmental Research and Risk Assessment* **32** (8), 2199–2212.
- Zadeh, L. A. 1965 Fuzzy sets. *Information and Control* **8** (3), 338–353.
- Zanganeh, M. 2017 Simultaneous optimization of clustering and fuzzy IF-THEN rules parameters by the genetic algorithm in fuzzy inference system-based wave predictor models. *Journal of Hydroinformatics* **19** (3), 385–404.
- Zhang, F. & Zhang, Y. 2021 A multi-objective optimization prediction approach for water resources based on swarm intelligence. *Earth Science Informatics* **14** (1), 457–468.
- Zounemat-Kermani, M., Batelaan, O., Fadaee, M. & Hinkelmann, R. 2021 Ensemble machine learning paradigms in hydrology: a review. *Journal of Hydrology* **598**, 126266.

First received 14 July 2022; accepted in revised form 17 February 2023. Available online 2 March 2023