

Primljen / Received: 10.7.2017.

Ispravljen / Corrected: 25.10.2017.

Prihvaćen / Accepted: 12.12.2017.

Dostupno online / Available online: 10.5.2018.

# Methodology for developing hydrological discharge model for small catchments

## Authors:



Ivana Sušanji, MCE  
University of Rijeka  
Faculty of Civil Engineering  
[isusanj@uniri.hr](mailto:isusanj@uniri.hr)



Prof. Nevenka Ožanić, PhD. CE  
University of Rijeka  
Faculty of Civil Engineering  
[nozanic@uniri.hr](mailto:nozanic@uniri.hr)



Prof. Barbara Karleuša, PhD. CE  
University of Rijeka  
Faculty of Civil Engineering  
[barbara.karleusa@uniri.hr](mailto:barbara.karleusa@uniri.hr)

Subject review

**Ivana Sušanji, Nevenka Ožanić, Barbara Karleuša**

## Methodology for developing hydrological discharge model for small catchments

Detailed steps of a hydrological model implementation methodology, as based on application of the artificial neural network (ANN) in small catchments, is presented in this paper. The motivation for this paper arises from the established lack of precise steps and procedures in the existing ANN-based methodologies for development of hydrological models. The implementation of hydrological discharge model based on ANN in a small catchment of Slani potok served as a basis for the development of detailed procedural steps for this methodology.

### Key words:

methodology, hydrological discharge model, artificial neural network (ANN), small catchment

Pregledni rad

**Ivana Sušanji, Nevenka Ožanić, Barbara Karleuša**

## Metodologija za implementaciju hidrološkog modela otjecanja na malim slivovima

U ovom se radu opisuju detaljni koraci metodologije za implementaciju hidrološkog modela otjecanja na malim slivovima temeljenog na primjeni umjetne neuronske mreže (UNM). Motivacija za izradu ovoga rada proizlazi iz utvrđenog nedostatka preciznih koraka i procedure u postojećim metodologijama za implementaciju hidroloških modela temeljenih na upotrebi umjetnih neuronskih mreža. Implementacija hidrološkog modela otjecanja temeljenog na umjetnoj neuronskoj mreži na malom slivu Slanog potoka poslužila je kao temelj za utvrđivanje detaljnih proceduralnih koraka metodologije.

### Ključne riječi:

metodologija, hidrološki model otjecanja, umjetna neuronska mreža (UNM), mali sliv

Übersichtsarbeit

**Ivana Sušanji, Nevenka Ožanić, Barbara Karleuša**

## Methodologie zur Implementierung des hydrologischen Abflussmodells an kleinen Wasserbecken

In dieser Abhandlung werden die detaillierten Schritte der Methodologie zur Implementierung des hydrologischen Abflussmodells an kleinen Wasserbecken beschrieben, basierend auf der Anwendung künstlicher neuronaler Netze. Die Motivation für die Ausarbeitung dieser Abhandlung geht aus dem festgestellten Mangel an präzisen Schritten und Prozeduren in den bestehenden Methodologien für die Implementierung hydrologischer Modelle basierend auf der Anwendung künstlicher neuronaler Netze hervor. Die Implementierung hydrologischer Abflussmodelle basierend auf der Anwendung künstlicher neuronaler Netze am kleinen Wasserbecken des Slani potok (*Salzbach*) diente als Grundlage für die Festlegung detaillierter prozeduraler Schritte der Methodologie.

### Schlüsselwörter:

Methodologie, hydrologisches Abflussmodell, künstliches neuronales Netzwerk, kleines Wasserbecken

## 1. Introduction

Development of a hydrological prediction model requires a number of technologies and areas of expertise that normally include several elements, such as the long-term monitoring and collection of existing data, data analysis, and model development, validation and evaluation [1].

Problems related to the development of hydrological prediction rainfall-runoff models currently arise from their diversity, various approaches to modelling process, their complexity, use of various model development technologies and model validation and evaluation procedures, etc. Hydrological prediction models are usually prepared for specific large catchments, and they cannot be used anywhere else. Such models cannot be applied in, for example, small catchments, where the solution-finding process and prediction time are more sensitive. It can be assumed that it is difficult to predict hydrological variables in small catchments with short rainfall response periods [2]. From the hydrological point of view, it is difficult to determine the boundary between the small and large catchments based on the catchment size only. In fact, it is necessary to take into account factors influencing conditions of runoff from the catchment. The uniform distribution of precipitation in the catchment, as well as participation of the entire area in the runoff process, is therefore the best way to define a small catchment [3].

In this paper, the focus is placed on the development of an accurate model development methodology using parametric (data-driven) models, such as the artificial neural networks (ANN), in order to develop a hydrological rainfall-runoff prediction model for small catchments. The collection of meteorological and hydrological data is necessary for the development of such model.

The ANN can be classified as a parametric model that is generally lumped because rainfall-runoff processes are treated as a "black box" with inputs and outputs [4, 5]. A novel, simple and appealing solution to complex hydrological processes is derived from a large number of existing studies and ANN-based hydrological rainfall-runoff models. The lack of hydrological prediction model implementation in small catchments, and the absence of an accurate model development methodology, is also established using the developed model.

## 2. Methods

The basic methodology for the hydrological ANN development is used in order to determine precise steps of the methodology. Using the main guidelines given in the basic methodology, the model is developed and implemented in the case study focusing on the Slani potok catchment area. The results of successful implementation of the ANN prediction model in the

small catchment area of Slani potok are described in detail in a doctoral thesis [6].

The ANN is a massively parallel distributed processor that has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects: knowledge is acquired by the network through a learning process and interneuron connection strengths, known as synaptic weights, are used to store the knowledge [7].

The reason why the ANN model has been chosen for use in a small catchment is its fast and efficient response in the prediction of hydrological variables. These ANN characteristics are very important if a short rainfall runoff response prediction of a small catchment is considered.

The ANN use in the field of hydrology can be considered relatively new since the first application was noted in papers published by Daniell in [8], French et al. in [9], and Hall and Minns in [10]. Since then, the ANN has been widely used in hydrology, and especially in the design of models that describe the relationship between rainfall and runoff. A contribution to that field was made by Halff et al. in by implementation of the ANN in order to predict hydrographs. Raman and Sunilkumar developed in [11] a model for the estimation of monthly precipitation. Since then, the ANN use in the field of hydrology has become quite frequent. One of the most important ANN uses in hydrology involves the first use of ANN for flood forecasting, as proposed by Openshaw et al. in [12], See and Openshaw in [13], Dawson and Wilby in [14], Maier and Dandy in [15], and others. They successfully presented runoff models for various mostly large catchments, while Jayawardena and Fernando in [16] were the first that applied the ANN in small catchments [17, 18].

In general, the basic steps of the methodology for the development of hydrological models are similar to basic steps of the methodology for the development of the ANN-based models. Although Maier and Dandy in [19], and Maier et al. in [20], highlighted good predicting properties of the ANN models, they emphasized at the same time the lack of systematic approach to the development of the ANN-based models in hydrology. Due to the lack of methodology for the development of hydrological models based on the ANN, in many cases the number of network inputs, data division, choice of architecture, activation functions, and the number of hidden neurons, have been selected at random [21]. In addition, there is also an inconsistency in the model quality assessment procedure, where quality measures are not selected appropriately. Basic general guidelines for the ANN-based development of hydrological models are provided by Dawson and Wilby in [4], Maier et al. in [20], and Matić et al. in [22]. All these methodologies for the ANN model development are quite similar, and they describe the same basic steps without guidelines that can provide precise and uniform implementation of the ANN.

## 2.1. Implementation of artificial neural network

ANNs are considered to be a mathematical data-driven type of model. In order to implement an ANN model, an appropriate quantity of input and output data sets has to be available. Data sets have to be divided into three sub-sets of data:

- training data set
- validation data set
- evaluation data set.

The first data set is used in training process when developed ANN model is adjusting the weight coefficient in order to link input and output data. Validation and evaluation data sets are used in the model quality assessment process. The ANN has the property of universal approximation, meaning that it has the ability to approximate any input to the output data with relative accuracy. Although it has this ability, it does not mean that it is able to respond accurately to the input data that was not included in the training process. That is why, before ANN model is accepted as a system model, it is necessary to assess its generalization properties through the process of validation and evaluation.

### 2.1.1. Architecture of artificial neural network

The artificial neuron node is the main microstructural component of the ANN. An artificial neuron node can be defined by mathematical expressions [7] shown in Eq. (1) and Eq. (2).

$$o_k = \varphi(v_k) \quad (1)$$

$$v_k = \sum_{n=0}^m (w_k \cdot x_k)_n \quad (2)$$

where  $o_k$  is the neuron node response in the  $k$ -th epoch of the calculation,  $v_k$  is the sum of products of the weight coefficients  $w_k$ ,  $x_k$  is the input data in the  $k$ -th epoch of the calculation, and  $\varphi$  is the activation function of the neuron node. The ANN model implementation involves:

- selection of an adequate ANN architecture and training algorithm
- the ANN training procedure.

The implementation starts with selection of the ANN mesostructure, which refers to the type of network (architecture) with which the model will be built [20]. The most common types of networks used in the development of rainfall-runoff models are:

- multilayer perceptron (MLP)
- radial basis function (RFB)
- self-organizing map (SOP)
- support vector machines (SVMs) [7, 21].

The MLP architecture is the best choice for the data-driven prediction model development [17]. The MLP architecture can be described as a static feed forward neuron network that consists of a minimum of three layers - input layer, hidden layer and output layer. Every layer consists of neurons that are connected by activation functions. Activation functions can be:

- linear
- limited linear
- unipolar sigmoid
- bipolar sigmoid
- hyperbolic tangent, etc. [21, 14].

Their purpose is to direct data through network layers, from the input layer to the output layer. The numbers of neurons in the input and output layers are defined by the number of selected data, whereas the number of neurons in the hidden layer should be optimized to avoid model overfitting, defined as the loss of generalization ability [21].

The MLP architecture was introduced by Werbos in 1974 in his PhD thesis [23]. Its final form was introduced by Rumelhart, Hinton and Williams in [24], who also presented applications of the MLP architecture and a description of its success in prediction, classification and association related to real problems. Model of MLP architecture is shown in Figure 1.

For the purpose of predictive hydrological ANN model development, it is important that the input layer consists of the data with minimum ten delay steps, and the output layer with the prediction time step. Delay steps can be defined as the input data from previous time steps.

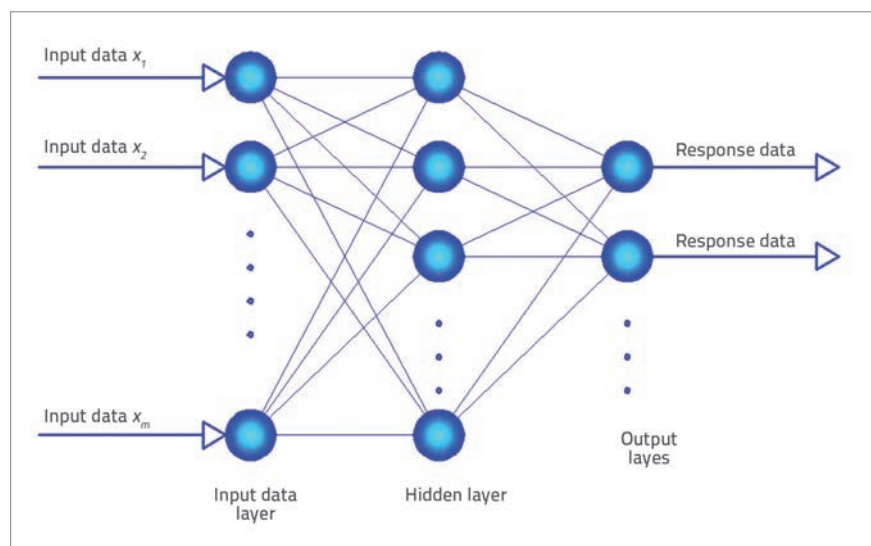


Figure 1. Multilayer Perceptron (MLP) model [17]

Because the output data from the network in one iteration of calculation will have errors, which are a function of the target output and model response in the output layer, an algorithm for determining the change  $\Delta w_k$  of the weight coefficient  $w_k$  is needed. These algorithms are known as training algorithms because they optimize input data in each following epoch, which reduces the error in the output layer with respect to the target output. The optimization of the weight coefficient can be defined as shown in Eq. (3) [7]:

$$w_{k+1} = w_k + \Delta w_k \quad (3)$$

where  $w_{k+1}$  is the weight coefficient in the  $k+1$ th epoch, and  $\Delta w_k$  is the change determined by the training algorithm. Training algorithms can be divided into three groups:

- first-order local algorithms (error backpropagation (generalized delta rule))
- second-order local algorithms (Newton algorithm, quasi-Newton algorithm, Levenberg-Marquardt (LM) algorithm)
- global algorithms (genetic algorithm, simulated annealing and evolutionary programming) [21, 25].

The LM algorithm is the fastest and most appropriate for training simpler structures [26] under the MLP architecture, and it was specially developed for the ANN training. Because of such characteristics, this algorithm is proposed for the development of ANN models for small catchments. Using the second-order local algorithms, the change measure  $\Delta w_k$  is obtained from the squared approximation of the error function, which is represented by the Hessian matrix. Algorithms that avoid solving the Hessian matrix, such as the LM algorithm, are used because the Hessian matrix typically cannot be used in ANN training, and because it is not in compliance with appropriate conditions, and is thus unsolvable. The LM algorithm [27], which is a special combination of the Gauss-Newton and error backpropagation algorithms, uses a conjugate gradient method by introducing the Jacobian matrix instead of the Hessian matrix. The change measure  $\Delta w_k$  can be defined as shown in Eq. (4).

$$\Delta w_k = -\left(\mathbf{J}^T \cdot \mathbf{J} + \mu \cdot \mathbf{I}\right)^{-1} \times \mathbf{J}^T \cdot \mathbf{e} \quad (4)$$

where  $\mathbf{J}$  is the Jacobian matrix of the error vector  $\mathbf{e}$  with respect to the weight coefficients in the  $k$ -th epoch of the calculation,  $\mathbf{J}^T$  is the transpose of the Jacobian matrix, and  $\mu$  is a scalar representing the learning rate.

At the end of every calculating epoch, the Sum Squared Error (SSE) is calculated as in Eq. (5) [5]:

$$SSE = \sum_{k=1}^n (e_k)^2 = \sum_{k=1}^n (d_k - o_k)^2 \quad (5)$$

where  $e_k$  is the error in  $k$ -th epoch of the calculation,  $d_k$  is the target value, and  $o_k$  is the response model value in  $k$ -th epoch of the calculation.

Depending on the increase or decrease in the SSE, the learning rate scalar  $\mu$  changes through every epoch of the calculation by dividing or multiplying by a constant factor (for example,  $\beta$  in the range [0,1]) to make the LM algorithm more similar to the Gauss-Newton error backpropagation algorithm, and to increase the training speed. If the sum of squared errors increases, the learning rate scalar  $\mu$  will be multiplied by a constant quantity  $\beta$ , and the LM algorithm will be more similar to the Gauss-Newton algorithm; otherwise, it will be more similar to the backpropagation algorithm [21].

### 3. Results and discussion

The case study area of the Slani potok catchment (Republic of Croatia) is used as the basis for development of precise ANN model methodology steps. The model development is described in detail in the doctoral thesis [6], while only a short description and results are presented in this paper.

As a part of the bigger Dubračina river catchment, Slani potok is considered to be a small catchment occupying approximately 2 square kilometres in area. The catchment ranges from 50 to 700 m a.s.l. in altitude, with an average slope of 22 %. The Slani potok catchment can be divided into two areas: lower part (0,9 km<sup>2</sup>), which is formed of flysch sediments (mainly siltstone), and the upper part, which is a karstic plateau. The lower part contributes more significantly to the surface runoff compared to the upper part. Continuous data monitoring points of hydrological and meteorological variables were established in 2012 for the model development purposes. Water levels in the Slani Potok creek waterbed are measured by a Mini Diver pressure probe (manufactured by Schlumberger Water Services) at the mouth of the Slani Potok creek, where it joins the Dubračina River. Meteorological parameters were measured using a Vantage Pro 2 meteorological station (manufactured by Davis Instruments Corporation) with the measurement frequency interval of two minutes. After three years of data collection, data from 2013 were selected as the representative data set because of significant rainfall. The data set consists of 132.772 samples, and includes meteorological and hydrological variables. The following meteorological parameters were selected as input data for the ANN model development:

- rain
- rain rate
- air temperature
- humidity
- air pressure
- solar radiation (insolation)
- wind direction
- wind speed.

River water levels were used as output prediction data. Selected data have shown direct or indirect connection with the catchment runoff process.

For the purposes of the ANN model implementation, the data is divided into the training, validation and evaluation data sets, i.e. 70 % (92 948 samples) for training, 15 % (19 912 samples) for validation, and 15 % (19 912 samples) for evaluation. After the data has been prepared, the MLP mesostructure is used and the processing is conducted in the MATLAB software (MathWorks, Natick, Massachusetts, US) so as to develop the data-driven ANN model for the small catchment with the sigmoid and linear activation function trained by LM algorithm. The input layer consists of the input data matrix with the minimum of ten delay steps, in order to provide data about past meteorological conditions to the model. The hidden layer consists of ten neurons and the output layer consists of the model prediction data. The model is developed for three prediction steps:

- S15 ( $\Delta t = 30$  minutes)
- S30 ( $\Delta t = 60$  minutes = 1 hour)
- S60 ( $\Delta t = 120$  minutes = 2 hours).

After that, the training process is validated and evaluated by visual and quantitative evaluation.

Schematic representation of the ANN water level prediction model is given in Figure 2. The results of the validation and evaluation process are presented in Table 1. According to the validation and evaluation of the models, they have been classified as "very good" (assessment validation and evaluation criteria [6]) for all prediction steps. The developed model served as the basis for future research aimed at defining methodology steps in implementation of the ANN models in small catchments. A lack of precise methodologies and procedures for the development of hydrological models based on ANNs was established based on the developed model. More precisely, the steps of the methodology for development of ANN-based hydrological models have to be defined in order to improve the model quality in general, and to equalize procedures in the use of ANNs for the hydrological models at small catchments. Therefore, the detailed methodology for development of hydrological model in small catchment is described below.

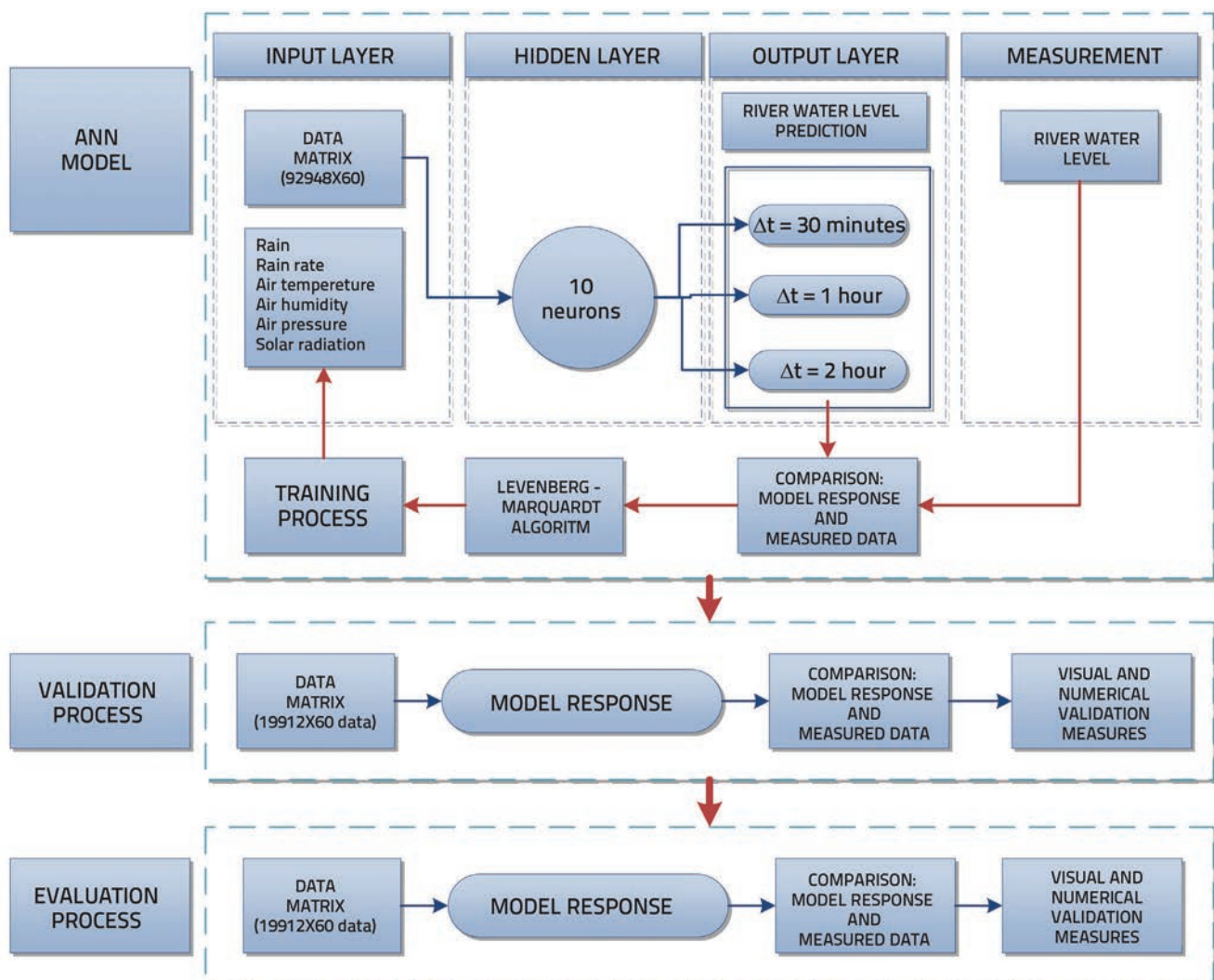


Figure 2. ANN water level prediction model

Table 1. Results of numerical validation and evaluation by ANN water level prediction model

Prediction step	Validation		Evaluation			
	MSE	r <sup>2</sup>	MSE	MSRE	CE	r <sup>2</sup>
S15 (t + 30 minutes)	0.603	0.96	5.737	0.0003	0.833	0.902
S30 (t + 60 minutes)	1.15	0.94	9.359	0.0005	0.728	0.849
S60 (t + 120 minutes)	1.391	0.932	11.656	0.0007	0.661	0.809

Mean Squared Error (MSE); Coefficient of determination (r<sup>2</sup>); Mean Squared Relative Error (MSRE); Coefficient of efficiency (CE)

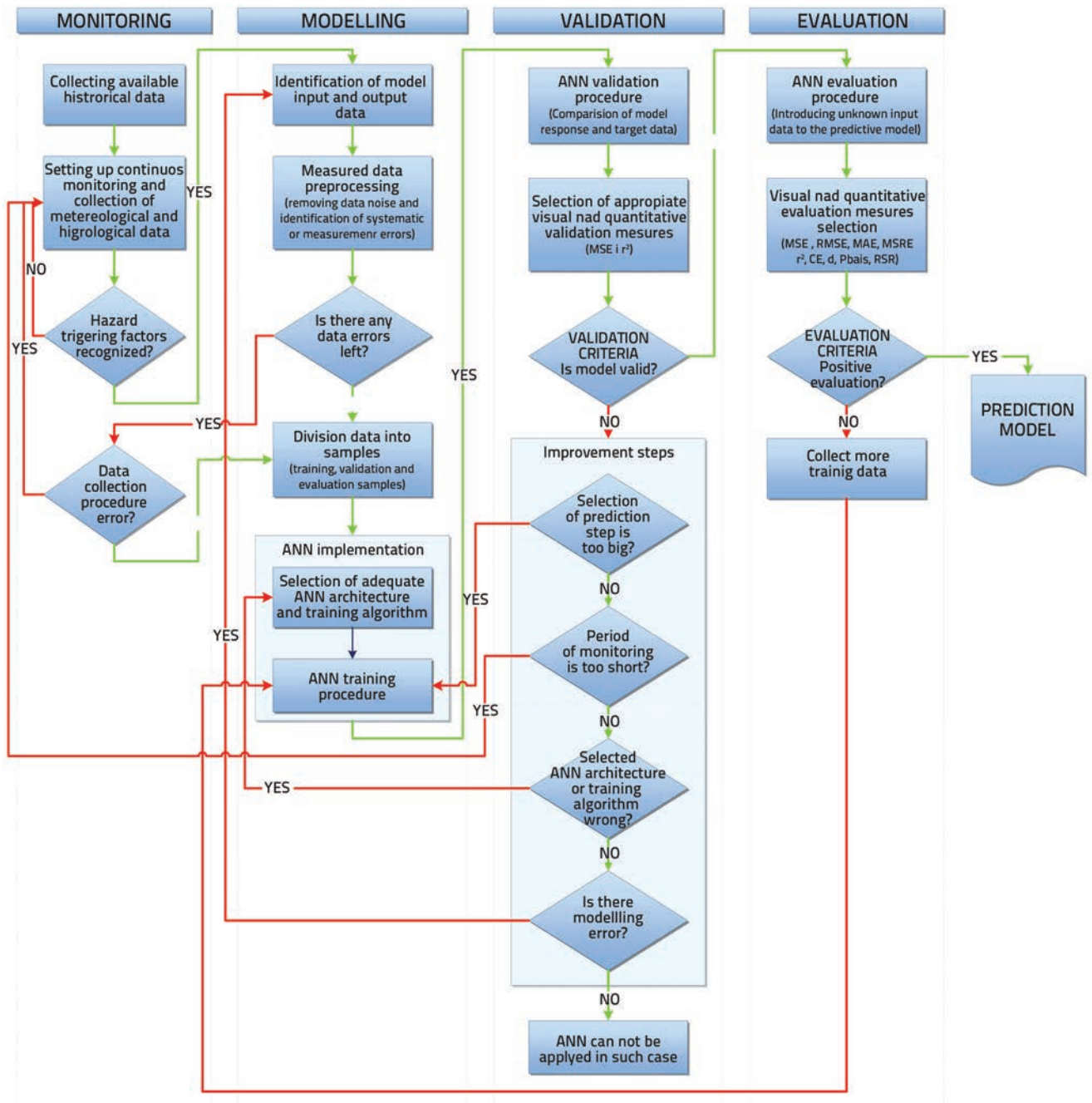


Figure 3. ANN predictive model development flowchart for small catchments

### 3.1. Methodology for development of hydrological model

The ANN model for development of the hydrological rainfall - runoff model for the small catchment of Slani Potok was implemented according to the general methodology suggested by Maier et al. (2010) [22]. The development of the model provided precise procedural steps for model development. The methodology for hydrological model development in small catchments was defined as a result of research and development of the model. The methodology, shown as a flowchart in Figure 3, is divided into four main groups of steps:

- measurement
- modelling
- validation
- evaluation.

Each step of the methodology is explained in detail.

## 3.2. Monitoring

### 3.2.1. Collecting available historical data

At the very beginning, it is important to collect all available historical data related to the case area by conducting preliminary research. Preliminary research includes collection of all historical research data (climatological, geological, hydrogeological, and hydrological data) and cartographic maps related to the catchment under study. It is then necessary to establish the current state of the catchment area and of the watercourse itself, and to analyse physical properties of the catchment and runoff characteristics. This procedure allows the researcher to perceive the problem and define further research.

### 3.2.2. Setting up continuous measurement and collection of meteorological and hydrological data

Based on previous research and collection of available historical data, it is possible to pinpoint the location of the measuring instruments and the time step for collecting meteorological or hydrological data on the catchment. At least two automatic weather stations and one profile in the watercourse for continuous measurement of water levels are necessary at small catchments. The installation of two meteorological stations is necessary to establish uniform distribution of precipitation on the catchment surface, so as to check whether the catchment under study is a small one. Measurements of meteorological and hydrological data in a small basin need to be carried out continuously with an appropriate time step from  $t = 1$  minute to maximum  $t = 10$  minutes in order to include any variability in meteorological and hydrological data that commonly occurs in a short period of time at small catchments. An appropriate time step for data collection can be determined by previous comparative experimental measurements with a different time step of data collection and

by comparing their accuracy. The data collection is very important for the ANN model implementation, because the model can be developed only if the measured data exist.

### 3.2.3. Detection of triggering factors

After establishment of continuous meteorological and hydrological data collection, it is also necessary to continuously monitor the diversity of collected data. In the hydrological research, the length of a collected data series and their significance is very important. In general, the 30-year long data series is considered to be sufficiently long for relevant data analysis. On the other hand, when meteorological and hydrological data are measured at small catchments for a short period of several years, the main goal is to collect the data that are as diverse as possible. The data significant for development of a hydrological rainfall runoff model are the data that are connected with the runoff process, such as the data on the rain, rain rate, air temperature, air humidity, air pressure, solar radiation, etc. It is desirable that a set of continuously collected data consists of at least ten significant precipitation episodes from different annual seasons in order to analyse characteristics of runoff. Measurements should be carried out until the moment when the triggering factors are recognized, and the influence of particular meteorological variables on the runoff process are defined.

## 3.3. Modelling

### 3.3.1. Identification of model input and output

Once triggering factors are identified, it is necessary to determine the input and output data of the model. It is important to select the data that are physically related or that affect the output data of the model. For example, the precipitation is associated with the runoff process causing the watercourse level increase. In order to develop the hydrological runoff model, the input data are represented by the meteorological data (rain, rain rate, air temperature, air humidity, air pressure, insolation, etc.) while the output data are hydrological data (water level or discharge).

### 3.3.2. Measured data preprocessing

The data collected by measuring instruments and selected for model development must be preprocessed. The data preprocessing involves elimination of errors and the so-called "noise" of the data. Data errors can be either systematic or accidental. They must be removed from the data set. When the time step of data collection is short ( $t = 2$  min), the "noise" or disturbance is visible in the form of minimum oscillations around the mean value of the measured data. Data noise also needs to be corrected, as well as measurement errors, in order to avoid unnecessarily

weighting of the model and decreasing its accuracy. Eliminating noise in data can be accomplished by means of a local non-linear regression method (Loess method) at the interval of 10 data samples.

**3.3.3. Additional data errors**

If errors in collected data still exist despite comprehensive data preprocessing, then the measurement procedure is not appropriate or there is a procedural error. Therefore, it is advisable to review the procedure under which the measurements were established. Establishment of new measurements of some meteorological and/or hydrological data is recommended. If additional measurement errors and data noise are not detected, the prepared data can be sent to the next step of the procedure, in which they are divided into the sets.

**3.3.4. Dividing data into sets**

The prepared data, which are to constitute the input and output data of the ANN model, should be divided into sets. To develop an ANN model, it is necessary to divide the data into three sets of data as explained in Section 3.3. The data is proportioned as follows: 70 % of total data for training purposes, 15 % of total data for validation purposes, and 15 % of total data for evaluation purposes, as is usual in the development of this type of models. Once the data sets are established, the implementation of the ANN can be conducted. The number of total data samples can vary according to the time collection step and period of time in which the measurement is carried out. The model validation and evaluation process will reveal whether the data set is significantly long and varied.

**3.3.5. Implementation of artificial neural network**

The implementation of the ANN on the previously preprocessed and divided data begins with the selection of: ANN architecture, training algorithm to train the model at a predetermined prediction time, and program package for model development. The ANN implementation is presented via flowchart in Figure 4. The first step in the procedure involves adequate selection of the

ANN architecture. Although various architectures, such as Multiple Layer Perceptron (MLP), Radial Base Grid (RFB), Self-Organizing Maps (SOM) and others, can be used in hydrological models, the MLP is currently the most widely used architecture [5]. The MLP architecture is therefore also recommended for development of a prediction model for small catchments as the initial choice, which can subsequently be changed depending on further model development and model quality. If the MLP architecture is chosen, the next step in the ANN implementation involves definition of the input, hidden, and output layers, setting of the input data matrix, selection of neurons in the hidden layer, and choosing time step prediction of output data. Since the aim is to develop prediction model for a small catchment, it is important to include at least ten previously measured data in each line of the input layer matrix, which will enable the model to gain insight in the training process into meteorological conditions over a given time period. The training capacity of the model depends on the type of activation function and the number of neurons. For the development of a hydrological prediction model for small catchments, it is advisable to use in the hidden layer a minimum of 10 neurons that are connected with the input and output layers by nonlinear activation functions. Using a

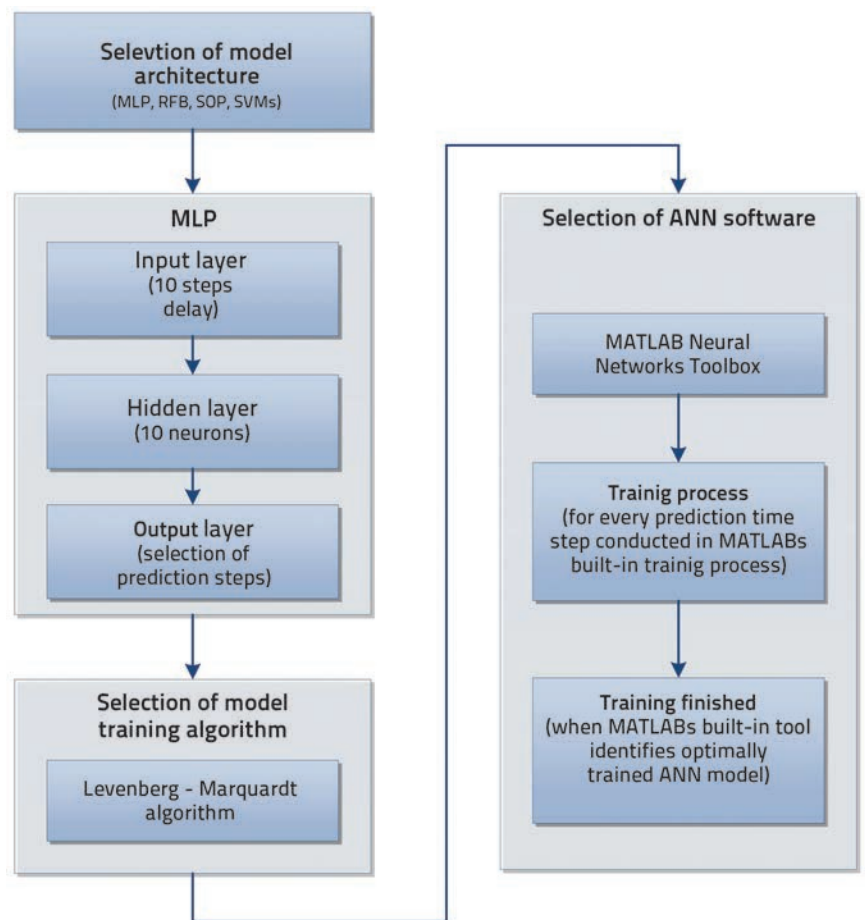


Figure 4. ANN implementation flowchart



larger number of neurons does not necessarily increase the quality of the model, and so the number of neurons should be selected with caution. Increasing the number of neurons when a significant amount of data is implemented in the model can lead to a slowdown in the training process. There is no direct rule for selecting the number of neurons, and so it is suggested to use 10 neurons in the initial phase of model development, which can be corrected depending on the subsequent progress and quality of the model.

Choosing the right training algorithm, i.e. the algorithm that will change optimization parameter, is very important from the standpoint of an efficient and fast training of the model. Algorithms of the first and second order are used in the development of the ANN model. Second-order algorithms are better suited to non-linear problems that occur in the case of hydrological models. The already mentioned Levenberg-Marquardt algorithm is characterized by its high speed and precision of adapting to non-linear problems. Precisely for these reasons, the use of the Levenberg-Marquardt algorithm is suggested for the hydrological prediction model training in small catchments.

After defining the model architecture, the training algorithm and the prediction step is chosen, and a program package in which the model will be programmed is selected. Nowadays, a large number of ready-made program packages are available, which have a predefined ANN architecture and the associated training algorithms, and where it is not possible to have an insight into the whole process of the model's work, and individual elements cannot be altered. Therefore, it is better to use program packages in which the model can fully be programmed. One of the program packages that have this option is the MATLAB (MathWorks) software that has distinct ANN elements such as architectural types and elements, training algorithms, training methods, validation and model evaluation, so that each element of the architecture can be defined during the model programming process. Therefore, it is recommended to use either this program package or other packages that have programming capabilities.

After the ANN model has been prepared, the set of data prepared for the training process (70 % of the total data) is incorporated in the model. In the training process, the optimization parameter is influenced and modified by the training algorithm through the iterations of the calculation. The goal of the training process is to optimize the model in order to avoid the state of overtraining, which would lead to the loss of generalization properties of the model. The way in which the number of iterations needed for optimum training of the model is calculated has not been clearly defined. One of the safe ways is to stop the training process when the SSE of the model response and measured data, calculated at the end of iteration, reaches an amount that does not result in an increase of the SSE calculated in the validation process.

### 3.4. Validation

#### 3.4.1. Artificial neural network validation procedure

The validation procedure is defined as the model's quality response after the model completes the training process. The developed model should be verified with a set of input and output data that have been earlier prepared for that purpose (15 % of data) in order to compare model response to the measured data [26, 28].

#### 3.4.2. Selection of adequate visual and numerical quality measures

The response given by the model based on the validation data set has to be evaluated graphically and by applying numerical quality measures. Visual comparison of model responses and measured data can help in identification of model errors that then have to be evaluated by numerical model quality measures. For the purposes of developing a hydrological rainfall runoff model for a small catchment, it is advisable to use at least one absolute (Mean Square Error (MSE)) and one relative (Coefficient of Determination ( $r^2$ )) numerical model quality measure.

#### 3.4.3. Validation criteria

The results of numerical model quality measures taken during the validation process need to be appraised according to each of the numerical model quality criteria used in the validation. There is a whole range of distinct model quality measures but not all of them have clear assessment criteria. It is very important to approach the results of the model quality measures assessment carefully, objectively, and critically.

If the model shows good results according to validation criteria, it may be possible to move on to the next step in model development. If the model does not show good results according to the validation criteria, several steps can be taken to improve the model. The first improvement possibility is to shorten prediction step of the model, after which the entire model training and model validation procedure have to be repeated. If this reduction of prediction step does not result in model improvement, it is possible that an insufficiently large data set was used for the design of the model, especially in the training process. This can be solved by collecting more data, after which the entire procedure should be repeated. If a larger set of data in the model training process does not improve quality of the validation model, another ANN architecture and/or training algorithm can be chosen, and the procedure can then be repeated. The last possible model improvement option is to check the modelling process in order to find possible errors in a particular step of the procedure, such as the wrong input and output selection, inadequate data pre-processing, etc. If the quality improvement procedure does not lead to improvement of the model, it is only possible to state that, in this case, the ANN model can not be applied.

### 3.5. Evaluation

#### 3.5.1. Evaluation of artificial neural network model

The evaluation procedure is defined as the quality of model response to the data set that has not been used in the training or validation process. The developed model is evaluated by means of a set of input and output data earlier prepared for that purpose (15 % of data), so that the model response can be compared to the measured data [26, 28].

#### 3.5.2. Selection of adequate visual and numerical quality measures

Procedure used in the process of model validation is similar to the one applied in the evaluation process. The model response data are compared to the measured data. The difference is in the number of numerical quality measures used. For the purposes of developing a hydrological model, it is desirable to use a larger number of absolute model quality measures such as the Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and a number of relative measures such as the Mean Squared Relative Error (MSRE), Coefficient of determination ( $r^2$ ), Coefficient of Efficiency (CE), Index of Agreement (d), Percentage to BIAS (PBIAS), and Root Mean Squared Error TO STANDARD DEVIATION (RSR) [5].

#### 3.5.3. Evaluation criteria

The results of numerical model quality measures taken during the evaluation process need to be appraised according to each of the numerical model quality criteria used, just like in the process of validation. If the model shows good results according to evaluation criteria, it can be concluded that the model has been successfully developed. Otherwise, it is recommended to collect more data for model training.

### 4. Conclusion

Although the lack of accurate, systematic, and uniform approach to the development of hydrological ANN models has been felt for many years, no significant progress has so far been made in this field. Also, the ANN models have rarely been used to predict hydrological variables in small catchments. Therefore, this paper presents precise steps in the methodology for hydrological model development based on the use of the artificial neural network (ANN) in small catchments, as based on the research and experience gained in the development of a hydrologic ANN model for the Slani Potok catchment area.

As the proposed methodology is based on a single case study, some problems may be experienced during implementation in other catchments. Thus, the methodology needs to be implemented in others small catchments to determine its general applicability to such catchments. Precise steps of the methodology presented in this paper will enable future researchers to repeat the same procedural steps for model development in other small catchments.

If this methodology is implemented at other small catchments, the improvement of the model quality comparisons and possible methodology improvements can also be anticipated.

### Acknowledgements

The research for this paper was conducted within the bilateral international Croatian-Japanese project "Risk identification and Land-Use Planning for Disaster Mitigation of Landslides and Floods in Croatia", as well as in the scope of scientific projects "Water Resources Hydrology and Floods and Mud Flow Risks Identification in the Karstic Area" (13.05.1.1.03) and "Development of New Methodologies in Water and Soil Management in Karstic, Sensitive and Protected Areas" (13.05.1.3.08), both financed by the University of Rijeka.

### REFERENCES

- [1] Krzhizhanovskaya, V.V., Shirshov, G.S., Melnikova, N.B.: Flood early warning system: design, implementation and computational modules, *Procedia Computer Science*, 4 (2011), pp. 106-115, <https://doi.org/10.1016/j.procs.2011.04.012>
- [2] Collier, C.G.: Flash flood forecasting: what are the limits of predictability?, *Quarterly Journal of the Royal Meteorological Society*, 133 (2007) 622, pp. 3-23, <https://doi.org/10.1002/qj.29>
- [3] Žugaj, R.: Hidrologija, University of Zagreb, Faculty of mining, geology and petroleum engineering Zagreb, 2000.
- [4] Dawson, C.W., Wilby, R.L.: Hydrological modelling using artificial neural networks, *Progress in Physical Geography*, 25 (2001) 1, pp. 80-108, <https://doi.org/10.1177/030913330102500104>
- [5] Abraham, R., Kneale, P. E., See, L. M.: *Neural networks for hydrological modelling*, Taylor & Francis Group plc, London, UK, 2004.
- [6] Sušan, I.: Development of the hydrological rainfall-runoff model based on artificial neural network in small catchments, Ph.D. Thesis, University of Rijeka, Faculty of Civil Engineering, Rijeka, 2016.

- [7] Haykin, S.: *Neural Networks: A Comprehensive Foundation*, Macmillan, New York, 1994.
- [8] Daniell, T. M.: *Neural networks. Applications in hydrology and water resources engineering, Australia*: In National Conference Publication-Institute of Engineers, 1991.
- [9] French, M.N., Krajewski, W.F., Cuykendall, R.R.: Rainfall forecasting in space and time using a neural network. *Journal of hydrology*, 137 (1992) 1-4, pp. 1-31, [https://doi.org/10.1016/0022-1694\(92\)90046-X](https://doi.org/10.1016/0022-1694(92)90046-X)
- [10] Hall, M.J., Minns, A.W.: *Rainfall-runoff modelling as a problem in artificial intelligence: experience with a neural network*, Cardiff, England: BHS 4th National Hydrology Symposium, pp. 5-51, 1993.
- [11] Raman, H., Sunilkumar, N.: Multivariate modelling of water resources time series using artificial neural networks. *Hydrological Sciences Journal*, 40 (1995) 2, pp. 145-163.
- [12] Openshaw, S., Openshaw, C.: *Artificial Intelligence in Geography*, John Wiley & Sons, London, 1997.
- [13] See, L., Openshaw, S.: Applying soft computing approaches to river level forecasting. *Hydrological Sciences Journal*, 44 (1999) 5, pp. 763-778.
- [14] Dawson, C.W., Wilby, R.L.: A comparison of artificial neural networks used for river forecasting. *Hydrology and Earth System Sciences Discussions*, 3 (1999) 4, pp. 529-540, <https://doi.org/10.5194/hess-3-529-1999>
- [15] Maier, H.R., Dandy, G.C.: Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental modelling & software*, 15 (2000) 1, pp. 101-124, [https://doi.org/10.1016/S1364-8152\(99\)00007-9](https://doi.org/10.1016/S1364-8152(99)00007-9)
- [16] Jayawardena, A.W., Fernando, D.A.K.: Use of radial basis function type artificial neural networks for runoff simulation. *Computer Aided Civil and Infrastructure Engineering*, 13 (1998) 2, pp. 91-99, <https://doi.org/10.1111/0885-9507.00089>
- [17] Abraham, R., Kneale, P.E., See, L.M.: *Neural networks for hydrological modelling*, Taylor & Francis Group plc, London, UK, 2004.
- [18] Govindaraju, R.S., Rao, A.R.: *Artificial neural networks in hydrology*, Water science and technology library, Vol. 36, Springer Science & Business Media, 2013.
- [19] Maier, H.R., Dandy, G.C.: Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental Modelling & Software*, 15 (2000), pp. 101-124, [https://doi.org/10.1016/S1364-8152\(99\)00007-9](https://doi.org/10.1016/S1364-8152(99)00007-9)
- [20] Maier, H.M., Jain, A., Dandy, G.C., Sudheer, K.P.: Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. *Environmental Modelling & Software*, 25 (2010), pp. 891-909, <https://doi.org/10.1016/j.envsoft.2010.02.003>
- [21] Matić, P.: *Short-term forecasting of hydrological inflow by use of the artificial neural networks*, Ph.D. Thesis, University of Split, Faculty of Electrical Engineering, Mechanical Engineering And Naval Architecture, Split, 2014.
- [22] Matić, P., Bego, O., Goić, R.: A systematic approach to a time series neural model development for river flow forecasting. *International review of automatic control*, 5 (2012) 3, pp. 367-372.
- [23] Werbos, P.J.: *Beyond regression: new tools for prediction and analysis in the behavioural sciences*, Ph.D. Thesis, Harvard University, Cambridge, Mass., 1974.
- [24] Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. *Nature*, 323 (1986) 6188, pp. 533-536, <https://doi.org/10.1038/323533a0>
- [25] Tingleff, O., Madsen, K., Nielsen, H.B.: *Methods for non-linear least squares problems. Lecture Note in Computer Science 02611, Optimization and Data Fitting, Informatics and Mathematical Modelling*, 2<sup>nd</sup> edn., Technical University of Denmark, 2004.
- [26] Demuth, H.B., Beale, M.H.: *Neural Network Toolbox™ User's Guide*, The MathWorks, INC., 2004.
- [27] Marquardt, D.W.: An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11 (1963) 2, pp. 431-441, <https://doi.org/10.1137/0111030>
- [28] Hagan, M.T., Demuth, H.B., Beale, M.H., De Jesús, O.: *Neural network design* (20), Boston, USA: PWS publishing company, 1996.