Feasibility study on operational use of neural networks in a flash flood early warning system

Estudo de viabilidade do uso operacional de redes neurais em um sistema de alertas antecipados de enxurradas

Glauston Roberto Teixeira de Lima[1]   & Graziela Balda Scofield[1]

[1]Centro Nacional de Monitoramento e Alertas de Desastres Naturais, São José dos Campos, SP, Brasil
E-mails: glauston.lima@cemaden.gov.br (GRTL), graziela.scofield@cemaden.gov.br (GBS)

ABSTRACT

Issuing early and accurate warnings for flash floods is a challenge when the rains that deflagrate these natural hazards occur on very short space-time scales. This article reports a case study in which a neural network-based hydrological model is designed to forecast one hour in advance if the water level in a small mountain watershed with short time to peak, situated in the city of Campos do Jordão in Brazil, will exceed its attention quota. This model can be a powerful auxiliary tool in a flash flood early warning system, since with it decision-making becomes semi-automated, making it possible to improve the warnings advance-accuracy tradeoff. A deep-learning neural network using Exponential Linear Unit activation functions was designed based on 3-years rainfall and water level data from 11 hydrometeorological stations of the National Centre for Monitoring and Early Warning of Natural Disasters. In the training of the neural network, two combinations of input variables were tested. The tuples in the test set were classified through voting with 60 classifiers. The first results obtained in Matlab environment with high percentages of true positives indicate that it is feasible to use the neural model operationally.

Keywords: Flash flood forecast; Machine learning-based hydrological modeling; Natural hazards.

RESUMO

Emitir alertas antecipados e precisos para enxurradas é desafiador quando as chuvas que deflagram esses desastres naturais ocorrem em escala espaço-temporal muito curta. Este artigo relata um estudo de caso no qual um modelo hidrológico baseado em rede neural é projetado para predizer com uma hora de antecedência se o nível de água em uma pequena bacia de montanha com curto tempo de ascensão, situada na cidade de Campos do Jordão no Brasil, excederá sua cota de atenção. O modelo pode aperfeiçoar um sistema de alertas antecipados de enxurradas, pois semiautomatiza a tomada de decisão e melhora o compromisso entre antecedência e precisão dos alertas. Uma rede neural de aprendizado profundo usando funções de ativação do tipo Unidade Linear Exponencial foi projetada com base em dados observados de precipitação e nível de água de 11 estações hidrometeorológicas do Centro Nacional de Monitoramento e Alerta de Desastres Naturais abrangendo um período de 3 anos No treinamento da rede neural, duas combinações das variáveis de entrada foram testadas. As tuplas do conjunto de teste foram classificadas através de votação com 60 classificadores. Resultados preliminares, obtidos em ambiente Matlab, com alto percentual de verdadeiros positivos, indicam que o uso operacional do modelo é viável.

Palavras-chave: Previsão de enxurradas; Modelagem hidrológica baseada em aprendizado de máquina; Desastres naturais.

INTRODUCTION

Every year are recorded in Brazil occurrences of floods in urban areas that cause socioeconomic losses (Banco Mundial, 2012; Haddad & Teixeira, 2015). One of the non-structural measures adopted by the Brazilian government to mitigate the potential impacts of such events was the creation in 2011 of the National Centre for Monitoring and Early Warning of Natural Disasters (Cemaden), which is an applied research Centre on the topic of natural hazards and whose mission includes issuing warnings for previously mapped risk areas, including some urban areas crossed by small watersheds susceptible to hydrological extremes. The warning system maintained by Cemaden needs to be continually improved by the insertion of new operational tools, as data driven hydrological models designed based on machine learning techniques, to assist operators of the situation room in decision making. To support the development of these models, Cemaden has an extensive observational network of hydrometeorological measuring stations that provide rainfall and water level data series. The end users of the warnings issued by Cemaden are the municipality civil defense agents and the warnings help them to decide the actions of the contingency plans in the risk areas monitored by the Centre.

The quality of a warning is checked against some requirements. Among others, two of these requirements are very important in order to help the municipality's civil defense to protect the population and reduce economic and material impacts: advance and accuracy (International Strategy for Disaster Reduction, 2006; International Network for Multi-Hazard Early Warning Systems, 2017). However, a flood can evolve very quickly when is caused by intense convective rains occurring on very short spatial and temporal scales, in which case it's called a flash flood (Kobiyama & Goerl, 2007). Considering the rainfall forecast approaches most commonly used today, the forecast of this type of rain is more accurate the closer to the moment of its occurrence. This makes the two requirements, (advance and accuracy) to conflict and the task of issue a warning for flash floods meeting both becomes very challenging and it is usually only possible to optimize the tradeoff between them.

The use of neural network (NN) targeting different interests in hydrological modeling dates back to the 1990s and since then good solutions based on NN have been obtained, including flood forecast models (Hsu et al., 1995; Dawson & Wilby, 1998, 1999, 2001; Varoonchotikul, 2003; Londhe & Charhate, 2010; Abrahart et al., 2012; Mosavi et al., 2018; Oyebode & Stretch, 2019). Some advantages of using NN are: (i) its proven ability to model complex and nonlinear relationships between input and output variables of a hydrological process using only sets of observed data and still produce satisfactory solutions; (ii) the low computational execution time that enables its real-time use as an operational tool to support decision makers; (iii) possibility of rapid recalibration when new data are available, regardless of whether they come from different sources. The last item (iii) is very relevant for the optimization of the tradeoff between advance and accuracy of the warnings because integrating data from multiple sources into the design of the NN, such as data from weather radar, weather satellite image, numerical weather forecast model, in general, makes the NN forecasts more accurate even for longer forecast horizons as, for example, in Filho & Santos (2006) and Chaipimonplin et al. (2011).

The use of NN in early warning systems is also widespread. In Kanbua & Khetchaturat (2007) NN was designed to forecast whether precipitation in a mountainous region of Thailand will reach a threshold above which landslides and floods may occur. The NN was able to forecast this threshold 24 hours in advance with reasonable accuracy and it is used in a decision support system that can be fed with data by users. Windarto (2010) applied NN to forecast the level of the Kali Garang River in the western of Semarang City in Indonesia, which is a densely populated region, as part of a flood warning system integrated with information technology (SMS and Web) that allow access flood early warning anywhere. A Mean Square Error (MSE) of 0.0046 was achieved by NN. In Roy et al. (2012), wireless sensor clusters, installed at various points along the bed of the Damodar River in India, supply inputs to an NN designed to forecast the water level. The NN output tracks the observed values quite well, with a R^2 coefficient above 0.95, and provides flood early warning as well as flood situation for disaster management and preparedness to combat aftermath. Elsafi (2014), used NN models to forecast flooding along the Nile River in a study area located in Dongola town down-stream of the junction of the main tributaries to the Nile including the White Nile, Blue Nile, and Atbara River. This study would provide baseline information toward the establishment of a flood warning system for certain sections of the Nile River and R^2 coefficient above 0.9 and low Root Mean Square Error were obtained by the NN in forecasting flows. In all of the above cases, the NN used was a Multilayer Perceptron (MLP). Banihabib (2016) compares the performance of a NN with a conceptual model for the determination of flood warning lead-time (FWLT) in Tajrish watershed that is a steep urbanized watershed located in the north of metropolitan city, Tehran, Iran, and the main flash flooder watershed in north of Tehran. Dynamics artificial NN (DANN) with time delay units by recurrent connections was used. FWLT was estimated longer by DANN than by conceptual model. Silva et al. (2016) presents an approach based on NN to forecast the flow of the Claro River in Caraguatatuba City in Brazil. The study area was chosen due to the high occurrence of mass movements and floods, mainly during the rainy season from December to April. The chosen NN was an MLP that achieved good agreement to the observed flow data (Nash index of 0.77) and good ability for providing early warnings (efficiency index of 0.91). In Sankaranarayanan et al. (2020), deep NN (DNN) has been employed for forecasting the occurrence of flood in some districts selected from the states of Bihar and Orissa in India. In addition, the deep neural model was compared with other machine learning models (Support Vector Machine (SVM), K-Nearest Neighbors (KNN) and Naïve Bayes) in terms of accuracy and error. The results indicated that DNN performed better than the other methods. However, the use of NN in Brazil as operational tool in warning systems maintained by public or private institutions to preventing and reducing the negative impacts of natural hazards is still little explored. An example of use in Brazil is the Caí River watershed, in northeast of the Rio Grande do Sul State, which is monitored by Extreme Hydrological Events Warning System of Brazilian Geological Service (CPRM) (Pickbrenner et al., 2017).

In this article, a case study is presented in which the rainfall-runoff process of the small Capivari river watershed, located in the city of Campos do Jordão in Brazil, is modeled using NN. In fact, concentrated or distributed hydrological models could be used to model the process of transforming rainfall into flow. But in cases where the available hydrometeorological database is still reduced, simplified modeling techniques are often used. Due the irregular relief and the form of the land use and occupation in the studied area, the time to peak of the watershed is very short. The time to peak was estimated by the Cemaden operators themselves, as approximately 30 minutes, based on the observation of all cases in which the river overflow level was reached. In addition, during the summer (November to March), heavy, rapid and localized convective rains are common. These combined factors favor the occurrence of flash floods in the urban perimeter of the Campos do Jordão and the central area of the city is one of the risk areas monitored by Cemaden.

The work in the situation rooms is carried out largely only based on knowledge and experience of the operators who are sometimes overwhelmed with the enormous amount of information from various risk areas to be analyzed. To bridges this gap, a line of research focused on the use of machine learning techniques to develop data-driven models applicable in the natural hazards prevention phase, with an emphasis on semi-automation and improvement of flash flood early warnings, is under consolidation at Cemaden. The study reported in this article is part of this research effort and, in this sense, the contributions of the article are in two respects: (i) for the chosen study area (and for other watersheds in Brazil with similar physical characteristics), there are few similar studies (using NN to forecast water level); (ii) NN's potential to improve the quality of flash flood warnings (by optimizing the tradeoff between advance and accuracy) will be assessed for the first time at Cemaden.

In addition to this introduction, the article contains the following sections: the description of the study area and the description of the database are given in sections 2 and 3, respectively;

Section 4 is devoted to NN architecture and training algorithm; results and discussions are presented in the section 5, and finally, Section 6 addresses the concluding remarks.

STUDY AREA

Campos do Jordão city is situated in the Mantiqueira Mountains and its topography is very uneven with 85% of the municipality composed by undulating regions, 10% by mountain slopes and 5% by escarpments areas. The city is located within a valley, the extension of its flat part not exceeds 500 meters wide and some points reach more than 2000 meters of altitude (Instituto Geológico do Estado de São Paulo, 2014).

According to Zucherato et al. (2016), the pattern of urbanization from the drainage areas have conditioned and concentrated the occurrence of flash floods in the most densely populated and longest-occupied parts of the municipality, especially around its central area, as can be observed by the high number of flash floods in the Capivari neighborhood. Local urbanization began with the settlement of the surroundings of the Capivari River (Instituto Geológico do Estado de São Paulo, 2014) that gave rise to the homonymous neighborhood. So the neighborhood with the highest number of flash flood records is also the oldest neighborhood in the city with more complex urbanization processes and more heterogeneous form of occupation.

During periods of heavy rain, the river overflows causing flash floods on the city's avenues and streets as well as material loss in commerce and homes.

The Capivari River, inserted in the Serra da Mantiqueira watershed, is formed by the Serraria Stream (located in the Santa Cruz neighborhood) and Piracuama Stream (located in the Vila Albertina neighborhood). The river runs through the Abernéssia and Jaguaribe neighborhoods and flows into the Sapucaí-Guaçu River in Vila Capivari, as shown in Figure 1.
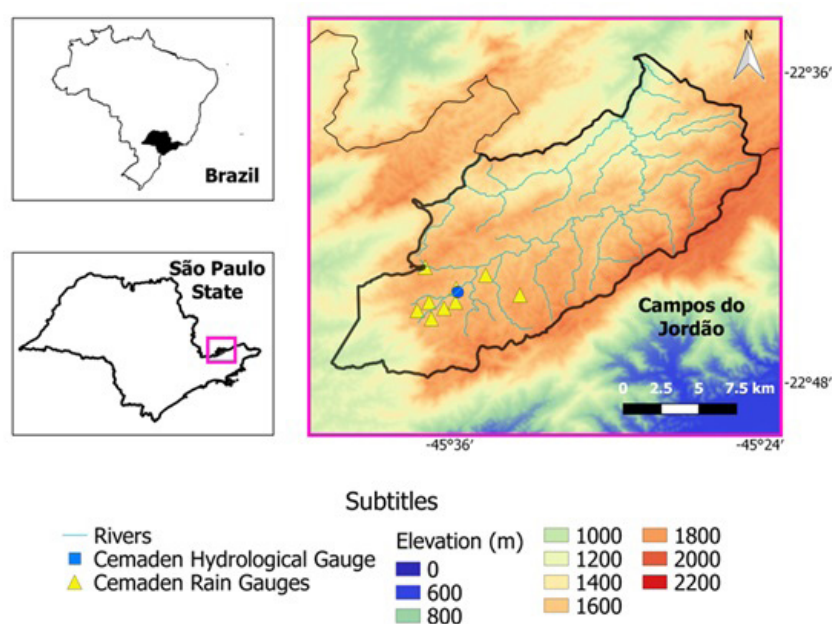

Figure 1. Map showing the area of the Capivari River watershed in Campos do Jordão city and the Cemaden's hydrometeorological stations installed there.

With an area of approximately 20 Km², Capivari River watershed is classified as a small watershed and its time to peak is approximately 30 minutes.

Campos do Jordão's climate is classified as Cwb - high altitude tropical climate - by the Koeppen Climate Classification (Köppen, 1918). The rainy season occurs usually during the summer (November to March) when the water surplus (the excess of water in the soil) is of the order of 1.130mm, making this period the most critical for the occurrence of landslides and flash floods (Departamento de Águas e Energia Elétrica do Estado de São Paulo, 2014; Instituto Geológico do Estado de São Paulo, 2014; Bosco, 2018).

DATABASE

The database used in this study to train and test the NN-based hydrological model was obtained from 11 Cemaden's hydrometeorological stations installed in Campos do Jordão. On the map in Figure 1, the representation (yellow dots) of one of the hydrometeorological stations is not distinguishable due to the proximity to another station. Among the 11 stations, 10 are pluviometric and 1 is pluviometric and hydrological. Most pluviometric stations are distributed upstream of the hydrological station but two of them are downstream. In the next stages of this study, a sensitivity analysis will be performed to identify which pluviometric stations are most important to forecast the water level at the point where the hydrological station is installed. Names, types, codes and coordinates of the hydrometeorological stations are shown in Table 1.

Rainfall and water level data from these stations are transmitted with 10 minutes temporal resolution when it is raining and with 1 hour temporal resolution during periods when there is no rain. To monitor measurements taken at the hydrological station, 3 quotas of water level, nominated as attention quota, alert quota, and overflow quota, are considered and their values were set in 1.61 m, 2.15 m and 2.69 m, respectively. The objective of this study is to design a forecasting model using this database and apply it to separate the data into two classes considering the attention quota of 1.61 m as the separation threshold. The database covers the period from September 2015 to January 2019 and has 15382 entries of which 15298 represent occurrences where the attention quota of the Capivari River was not exceeded and 84 are cases where the attention quota was exceeded.

THE NN

There is a vast literature available on NN (Bishop, 1995; Haykin, 1998; Goodfellow et al., 2016). NN is widely used to model the dynamics of real-world environmental systems (such as watersheds) in which the relationships between input and output variables are complex, nonlinear and non-deterministic, because NN is able to perform this task well enough, using only input-output pairs of observed values of the variables of interest.

NN ARCHITECTURE

The NN architecture used in this article to forecast an hour in advance if the Capivari River's water level will exceed its attention quota is derived from the standard MLP architecture and is shown in Figure 2.

NN architecture has an input layer with 23 nodes, 3 hidden layers with 23 neurons and an output layer with 2 neurons. Each layer is fully connected to the next. The input layer was configured to receive tuples with 23 input features: a water level (measured one hour before the time for which the forecast is made) and the average and maximum rainfall values observed at each of the 11 hydrometeorological stations in the previous one hour period. Each tuple also has an output feature that corresponds to the water level measured at the moment for which the forecast is made. For the sake of clarity, Figure 3 exemplifies a tuple used to train the NN to make a forecast of the water level for 2 pm.

The option to compose tuples with average and maximum values instead of using all six precipitation values measured in each station in the previous 1 hour period, is justified for two reasons: (i) using averages and maximums values works as a pre-processing that increases data separability making classification easier and, as shown in Figure 4, this hypothesis is plausible; (ii) in future stages of this study, when rainfall data from historical series will be replaced by radar data or numerical weather forecast models, the use of average and maximum values will be more advantageous due to their lower uncertainty.

Another type of tuple with a different combination of variables was also used to train and test the NN. In this new input tuple, in addition to the attributes illustrated in Figure 3,

Table 1. Cemaden hydrometeorological stations in the study area.

|	Name	Type	Code	Latitude (°)	Longitude (°)
1	Alto da Boa Vista	Pluviometric	350970006A	-22.73	-45.62
2	Bela Vista	Pluviometric	350970004A	-22.74	-45.60
3	Brancas Nuvens	Pluviometric	350970012A	-22.75	-45.62
4	Capivari	Pluviometric	350970013A	-22.74	-45.55
5	Jaguaribe	Pluviometric	350970005A	-22.73	-45.58
6	Monte Carlo	Pluviometric	350970009A	-22.75	-45.58
7	Nova Suíça	Pluviometric	350970008A	-22.75	-45.61
8	Ribeirão Capivari	Hydrological/Pluviometric	350970004H	-22.74	-45.59
9	Vila Abernéssia	Pluviometric	350970001A	-22.74	-45.60
10	Vila Cristina	Pluviometric	350970003A	-22.76	-45.61
11	Vila Matilde	Pluviometric	350970011A	-22.75	-45.60
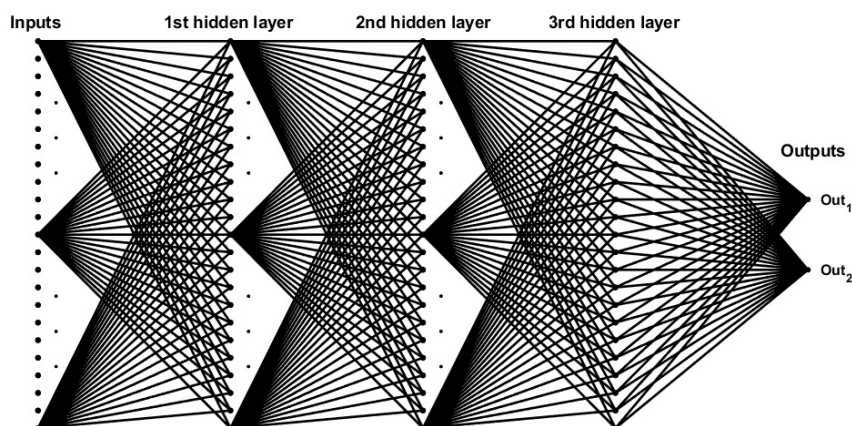
Figure 2. NN architecture used to forecast an hour in advance the water level of Capivari River.
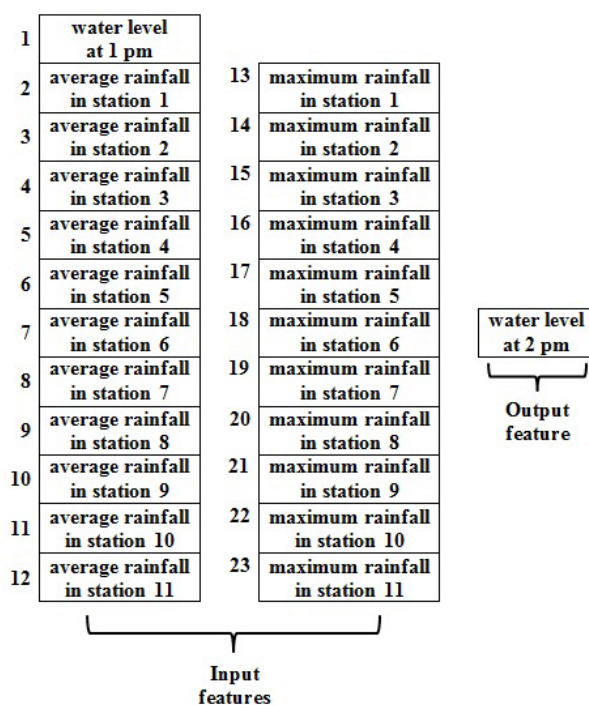

Figure 3. Example of a tuple of type 1 inputs used in the NN design to make a forecast of the water level for 2 pm.
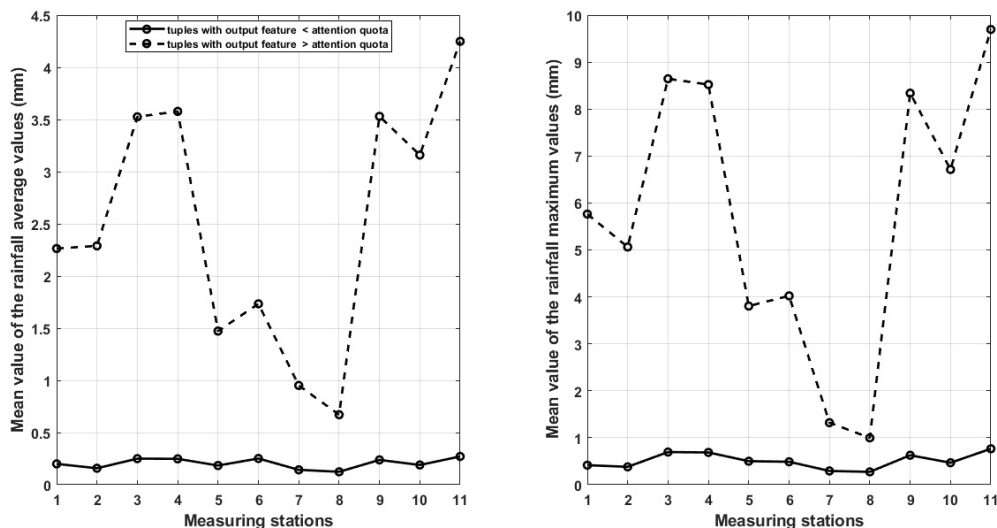

Figure 4. Mean values of the average and maximum rainfall input features of the type 1 inputs for the two classes in the database.

the following were added: the accumulated rainfall in each hydrometeorological station during the second hour prior to the time for which the forecast is made and also the water level two hours before the time for which the forecast is made. In order to process these inputs, the NN architecture was also changed and the size, both the input layer and the 3 hidden layers, was increased to 35 neurons. Hereafter, the two types of tuples are referred to as 'type 1 inputs' (23 features) and 'type 2 inputs' (35 features).

NN TRAINING

The database with 15382 tuples was split into a training set and a test set. The training set covers the period from September 2015 to December 2017 and has 10795 tuples of which 10754 have an output feature (1 hour ahead observed water level) that does not exceed the attention quota (1.61 m) and 41 have an output feature that exceed the attention quota. The test set covers the period from January 2018 to January 2019 and has 4587 tuples of which 4544 has output feature bellow the attention quota and 43 has output feature above this quota. Hereafter, these two classes will be referred to as C_{ne_at} (tuples whose output feature is below the attention quota) and C_{e_at} (tuples whose output feature is above the attention quota).

The 10754 C_{ne_at} training tuples were divided into 60 subsets. This sampling was done in such a way that the entire value range of the output feature was well represented in each subset. Each of these subsets has 84 C_{ne_at} training tuples. Thus, 60 different neural classifiers were trained using 60 different sets of 84 C_{ne_at} tuples and always the same set of 41 C_{e_at} tuples. The trainings were done using the standard Backpropagation algorithm (Rumelhart et al., 1986) and the Stochastic Gradient Descent (SGD) (Robbins & Monro, 1951; Bottou et al., 2018) applied on mini-batches of tuples with 1 C_{ne_at} tuple and 1 C_{e_at} tuple. These procedures were applied separately to type 1 and type 2 inputs.

A training iteration (or training epoch) is completed when the 84 C_{ne_at} tuples are presented to NN. Therefore, during iteration

each C_{e_at} tuple is presented to the NN 2 (~84/41) times while each C_{ne_at} tuple is presented to the NN only once. The sampling described above helps to minimize the harmful effect that class unbalance (C_{ne_at} has 10754 tuples and C_{e_at} has just 41 tuples) would have on NN performance. Training ends when a previously chosen maximum number of training epochs is reached.

Exponential Linear Unit (ELU) activation functions were used in the 3 hidden layers. ELU overcome vanishing gradient problem so that deep learning-NN architecture (Figure 2) could be designed. The hidden activation functions are expressed by Equation 1.

$$y(x) = \begin{cases} x & \text{if } x \geq 0 \\ \beta * (\exp(x) - 1) & \text{if } x < 0 \end{cases} \tag{1}$$

Where $y(x)$ is the activation of the hidden neuron, x is the net signal at the neuron input and β is the threshold at which the exponential function converges when x tends to minus infinity.

In the output neurons, this same type of activation function was used, however with some modifications. The activation functions of the outputs neurons out_1 and out_2 are given, respectively, by Equations 2 and 3.

$$Out_1(x) = \begin{cases} 1.61 - x & \text{if } x < 1.61 \\ \beta * (\exp(1.61 - x) - 1) & \text{if } x \geq 1.61 \end{cases} \tag{2}$$

$$Out_2(x) = \begin{cases} \beta * (\exp(x - 1.61) - 1) & \text{if } x < 1.61 \\ x - 1.61 & \text{if } x \geq 1.61 \end{cases} \tag{3}$$

Note that the displacement applied to the NN output activation functions is equal to 1.61m which is the same value of the attention quota for the Capivari River's water level. Figure 5 show the activation functions of the NN output neurons.
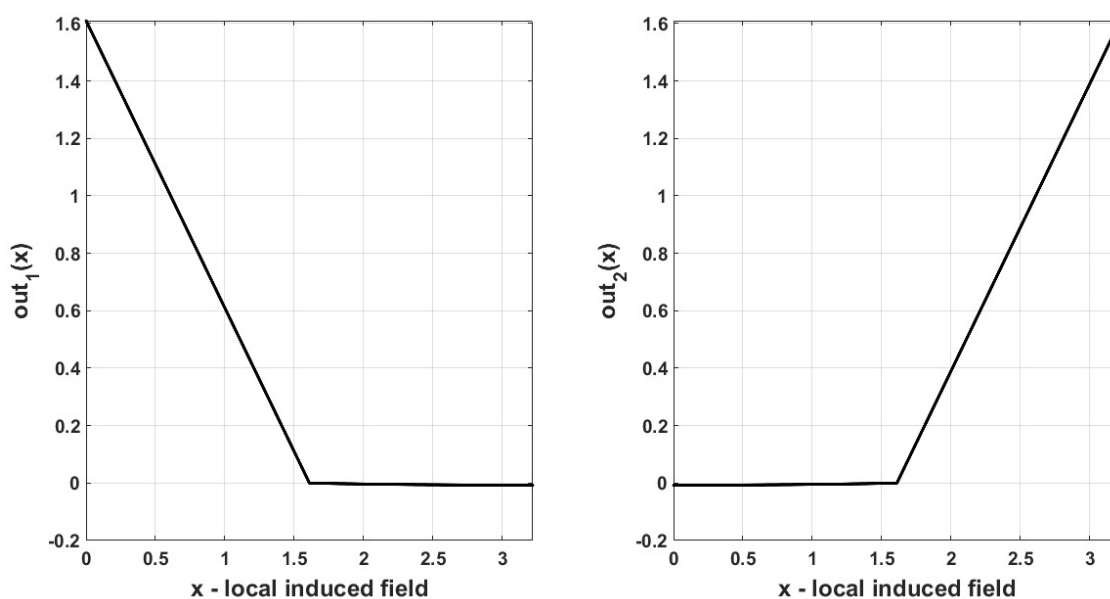

Figure 5. Activation functions used in the NN outputs.

Based on Equations 2 and 3, the errors for the two NN outputs are computed as follows:

Out$_1$

For tuples of the class C$_{ne\_at}$:

$$\text{error} = (1.61 - v_{ob}) - out_1 \qquad (4)$$

where v_{ob} is the observed value that is, the tuple's output feature.

For tuples of the class C$_{e\_at}$:

$$\text{error} = \beta * \left(\exp\left(1.61 - v_{ob}\right) - 1\right) - out_1 \qquad (5)$$

Out$_2$

For tuples of the class C$_{ne\_at}$:

$$\text{error} = \beta * \left(\exp\left(v_{ob} - 1.61\right) - 1\right) - out_2 \qquad (6)$$

For tuples of the class C$_{e\_at}$:

$$\text{error} = (v_{ob} - 1.61) - out_2 \qquad (7)$$

Thus, the class assigned by the NN to a tuple presented to its input layer is identified by comparing the two outputs of the NN:

- if out$_1$ is greater than out$_2$, the tuple is classified by the NN as belonging to the C$_{ne\_at}$ class;

- if out$_2$ is greater than out$_1$, the tuple is classified by the NN as belonging to the C$_{e\_at}$ class.

The results of the NN classification are presented in the form of confusion matrices and were evaluated by the following performance indices: accuracy, kappa and precision.

Accuracy is given by the following formula:

$$\text{Accuracy} = \frac{\sum\limits_{i=1}^{C} D_i}{T} \qquad (8)$$

where T is the total number of elements in the confusion matrix, C is the number of class considered in the classification task and D$_i$ is the value in the i-th diagonal position.

Kappa index (Cohen, 1960), is defined, according to Congalton & Green (2009),

$$\text{kappa} = \frac{\left(T \times \sum\limits_{i=1}^{C} D_i\right) - \left(\sum\limits_{i=1}^{C} TC_i \times \sum\limits_{i=1}^{C} TR_i\right)}{\left(T^2 \times \left(\sum\limits_{i=1}^{C} TC_i \times \sum\limits_{i=1}^{C} TR_i\right)\right)} \qquad (9)$$

where TC$_i$ is the sum of the i-th column and TR$_i$ is the sum of the i-th row.

Precision is given by:

$$\text{Precision} = \frac{D_2}{TC_2} \qquad (10)$$

In the training algorithm executions that generated the results presented in the next section, (i) the NN weights were initialized with small random values, (ii) the maximum number of epochs was chosen as training stopping criterion and was set at 4000, (iii) the learning rate decayed exponentially during the 4000 epochs from 0.1 to 0.01 and (iv) the parameter (β) of the hidden activation functions was set at 0.1 and for the output activation functions was set at 0.4

RESULTS AND DISCUSSIONS

Results

The NN was trained to separate the tuples of the database into two classes:

- C$_{ne\_at}$ - tuples whose data indicates that the water level (1 hour ahead) will not exceed 1.61m;

- C$_{e\_at}$ - tuples whose data indicate that the water level (1 hour ahead) will exceed 1.61m.

Figure 6 shows, for the two NN outputs, Mean Squared Error (MSE) typical curves obtained during the training of the 60 classifiers using type 1 inputs.
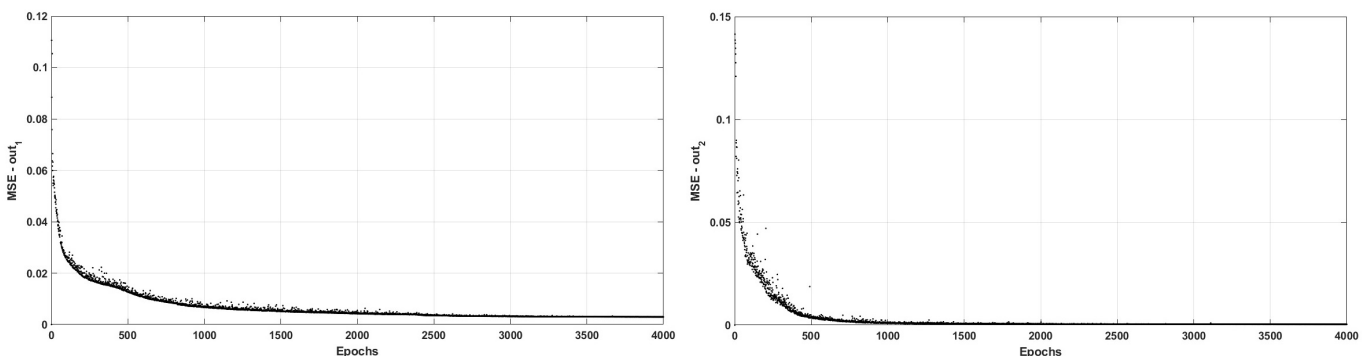

Figure 6. Typical MSE curves obtained in the training of one of the classifiers with type 1 inputs.

Figure 7 illustrates the outputs of one of the 60 classifiers trained with type 1 inputs. These outputs are relative to the subset formed by 84 C_{ne_at} inputs plus 41 C_{e_at} inputs that was sampled to train the classifier. Note, in the Figure 7, that for C_{ne_at} tuples, out_1 values are greater than the out_2 values and for C_{e_at} tuples the opposite is true.

Table 2 presents the average confusion matrices of the classification results produced by the 60 classifiers trained with type 1 inputs and type 2 inputs.

In the confusion matrix, rows indicate tuples of a given class, while the columns indicate the predicted class for these tuples.

Table 3 presents the confusing matrices of the classification results of the test set and the entire training set obtained by voting with the 60 classifiers trained with type 1 inputs.

Table 4 presents the confusing matrices of the classification results of the test set and the entire training set obtained by voting with the 60 classifiers trained with type 2 inputs.

DISCUSSIONS

The NN architecture (Figure 2) with three hidden layers was the one that provided the best relationship between performance and the number of epochs needed to train the NN.

Comparing the confusion matrices and respective performance indices, it can be seen that NNs trained with type 2 inputs have better classification performance than those trained with type 1 inputs.
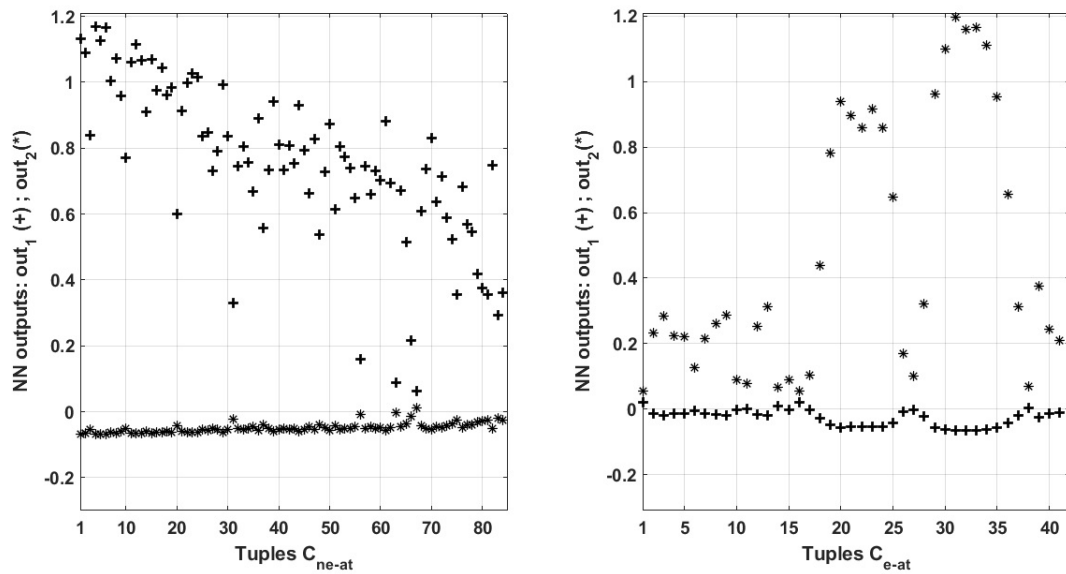

Figure 7. Classification for one of 60 training sets made by NN; C_{ne_at} tuples generate out_1 greater than out_2 while C_{e_at} tuples generate out_2 greater than out_1.

Table 2. Average confusion matrices of the NN classification results obtained with 60 different samples of training set and performance indices.

	Using type 1 training inputs			Using type 2 training inputs	
	C_{ne_at}	C_{e_at}		C_{ne_at}	C_{e_at}
C_{ne_at}	79	5	C_{ne_at}	84	0
C_{e_at}	3	38	C_{e_at}	0	41
	Accuracy: 0.9360			Accuracy: 1.000	
	Kappa: 0.8566			Kappa: 1.000	
	Precision: 0.8837			Precision: 1.000	

Table 3. Confusion matrices and performance indices of the NN classification results for training and test set obtained by voting with the 60 classifiers trained with type 1 training inputs.

	Training set			Test set	
	C_{ne_at}	C_{e_at}		C_{ne_at}	C_{e_at}
C_{ne_at}	10709	45	C_{ne_at}	4427	117
C_{e_at}	1	40	C_{e_at}	2	41
	Accuracy: 0.9957			Accuracy: 0.9741	
	Kappa: 0.6330			Kappa: 0.3991	
	Precision: 0.4706			Precision: 0.2595	

As shown in Table 4, NN trained with type 2 inputs is able to classify correctly almost all tuples of the C_{ne_at} class and all the tuples of the C_{e_at} class resulting in accuracy greater than 0.99. However, the confusion matrices (in Table 4) also show a considerable number of false positives (tuples of class C_{ne_at} classified as class C_{e_at}). There are 16 cases in the training set and 71 cases in the test set. Thus, the indices that best evaluate the performance of the classifiers (since they also take into account the existence of false positives) are the kappa and, mainly, the precision that for the training set was greater than 0.7 (reasonable value) and for the set of test was less than 0.4 (low value). Table 5 shows the sum of accumulated, averages and maximum rainfall of the 16 training tuples in Table 4 that were classified as false positives.

From the analysis of Table 5, it can be hypothesized that the tuples of the training set incorrectly classified are points in the input space that when mapped to the output space fall very close to the decision boundary of the two classes. This can be caused not only by output values very close to the decision boundary of the classes (1.61m), but also by sums of accumulated, averages and maximum rainfall of considerable magnitude. Thus, tuples containing such values represent cases whose classification is usually more difficult. A similar analysis applies to the false positives of the test set in Table 4.

The occurrence of false positives in the classification results demonstrates that the NN design requires improvements. This includes improving the training algorithm and NN architecture, testing new combinations of inputs, and, perhaps, applying more robust data qualification pre-processing techniques to eliminate noise and inconsistencies that can also generate false positives or even false negatives which may be an even more serious misclassification. Considering the above analysis on the false positives in Table 4, improvements can be directly targeted to correct this problem.

On the other hand, since the NN trained with type 2 inputs proved to be efficient in detecting the true positives even for data in the test set (this is very important because the test set contains tuples not presented to the NN during its training and serves to assess the NN generalization ability), the present study is a proof of concept of the potential that has NN as an auxiliary decision-making tool in an early warning system for flash floods and that should, therefore, be tested operationally. In practical terms, the role of an auxiliary tool in an early warnings system is make forecasts that will serve as a pre-warning for flash flood events that the system operators (decision-makers) may or not confirm based on information from other analysis tools and also based on their expertise. In this way, NN semi-automates the decision-making process. However, as previously mentioned, for the purpose of making the NN operational, historical series of observed rainfall data used in this study should be replaced by rainfall estimates given by radar, numerical weather forecast model and other sources. This will be done in the next stages of this study.

CONCLUDING REMARKS

This article reports the development of a NN-based hydrological forecast model to be integrated to the Cemaden

Table 4. Confusion matrices and performance indices of the NN classification results for training and test set obtained by voting with the 60 classifiers trained with type 2 training inputs.

	Training set			Test set	
	C_{ne_at}	C_{e_at}		C_{ne_at}	C_{e_at}
C_{ne_at}	10738	16	C_{ne_at}	4473	71
C_{e_at}	0	41	C_{e_at}	0	43
	Accuracy: 0.9985			Accuracy: 0.9983	
	Kappa: 0.8360			Kappa: 0.8192	
	Precision: 0.7193			Precision: 0.3772	

Table 5. Features of the tuples that generates false positives in the confusion matrices of type 2 training set

Sum of accumulated rain in the second previous hour (mm)	Sum of rainfall average in the first previous hour (mm)	Sum of rainfall maximum in the first previous hour (mm)	Output feature (m)
1.700	13.5766	23.5400	1.5530
55.0700	12.7433	36.0600	1.2900
109.8800	19.4650	84.8900	1.0430
43.5400	11.9666	27.9900	0.7170
0.2000	18.2766	42.7700	0.7963
13.1400	17.3133	41.9700	0.7963
0.6000	14.8983	29.7900	1.0568
25.5600	18.6733	49.3200	1.1580
50.7600	15.4717	48.6400	1.0140
82.8000	12.1650	43.9700	0.9635
45.4100	7.9042	26.3700	0.7480
0.8000	14.7100	27.0000	1.2540
18.3100	14.6417	25.8000	1.1540
47.7400	12.5233	25.0000	0.9970
69.6000	7.9517	19.0500	0.8430
33.0000	14.4883	66.0500	0.7300

flash flood early warning system as a decision-make auxiliary tool. The small Capivari River watershed located in a mountainous region in the Campos do Jordão city, with short time to peak and susceptible to intense convective rains and flash floods, was chosen as study case. This watershed is one of the natural hazards risk areas monitored by Cemaden. Historical series of rainfall and water level data, covering a period from 2015 to 2019, were used to train and test the NN. Two different combinations of input variables were used to train and test the NN. One containing the average and maximum precipitation values and the other containing these two attributes and also the accumulated rainfall. Exponential Linear Unit activation functions, which do not suffer from vanishing gradient problem, were used and so a deep-neural network with 3 hidden layers was designed. This number of hidden layers provided the best relationship between training iterations and performance. The task performed by the NN consisted of separating the tuple in database into two classes: (i) tuples whose data indicate that the water level, 1 hour ahead, will not exceed the attention quota (of 1.61m) and tuples whose data indicate that the water level, 1 hour ahead, will exceed the attention quota. This task was accomplished by voting with 60 different classifiers trained with subsets of inputs that were sampled from the total training set.

The training with inputs containing the accumulated precipitation produced better results. The results, presented in the form of confusion matrices and performance indicators (accuracy, kappa and precision) derived from these matrices, show an amount of false positive, which indicates that the NN design still needs to be improved. However, what stands out the most in the results is the 100% correctness in the classification of the true positives both in the training set and the test set. This indicates that the NN has the potential to be integrated into Cemaden's flash flood early warning system as auxiliary tool for decision-makers. However, the operationalization of this tool requires that the rainfall data obtained from historical series be replaced by rainfall forecasts made for the period of one hour ahead.

REFERENCES

Abrahart, R. J., Anctil, F., Coulibaly, P., Dawson, C. W., Mount, N. J., See, L. M., Shamseldin, A. Y., Solomatine, D. P., Toth, E., & Wilby, R. L. (2012). Two decades of anarchy? Emerging themes and outstanding challenges for neural network modelling of surface hydrology. *Progress in Physical Geography*, *36*(4), 480-513. http://dx.doi.org/10.1177/0309133312444943.

Banco Mundial. (2012). *Avaliação de perdas e danos: inundações e deslizamentos na Região Serrana do Rio de Janeiro - Janeiro de 2011*. Brasília. Relatório.

Banihabib, M. E. (2016). Performance of conceptual and black-box models in flood warning systems. *Cogent Engineering*, *3*(1), 1127798. http://dx.doi.org/10.1080/23311916.2015.1127798.

Bishop, C. M. (1995). *Neural network for pattern recognition*. New York: Oxford University Press.

Bosco, R. B. (2018). *Vulnerabilidade socioambiental na sub-região 2 da Região Metropolitana do Vale do Paraíba Paulista: consideração de indicadores precipitação-deslizamento* (Dissertação de mestrado). Universidade Federal do ABC, Santo André.

Bottou, L., Curtis, F. E., & Nocedal, J. (2018). Optimization methods for large-scale machine learning. *SIAM Review*, *60*(2), 223-311. http://dx.doi.org/10.1137/16M1080173.

Chaipimonplin, T., See, L., & Kneale, P. (2011, December 12–16). Improving neural network for flood forecasting using radar data on the Upper Ping River. In F. Chan, D. Marinova & R.S. Anderssen (Eds) *19th International Congress on Modelling and Simulation* (pp. 1070-1076). Perth, Australia. ISBN: 978-0-9872143-1-7. https://www.mssanz.org.au/modsim2011/C1/chaipimonplin.pdf.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37-46. http://dx.doi.org/10.1177/001316446002000104.

Congalton, R. G., & Green, K. (2009). *Assessing the accuracy of remotely sensed data: principles and practices*. Boca Raton: CRC Press.

Dawson, C. W., & Wilby, R. L. (1998). An artificial neural network approach to rainfall-runoff modelling. *Hydrological Sciences Journal*, *43*(1), 47-66. http://dx.doi.org/10.1080/02626669809492102.

Dawson, C. W., & Wilby, R. L. (1999). A comparison of artificial neural networks used for river flow forecasting. *Hydrological Sciences Journal*, *3*(4), 529-540.

Dawson, C. W., & Wilby, R. L. (2001). Hydrological modelling using artificial neural networks. *Progress in Physical Geography*, *25*(1), 80-108. http://dx.doi.org/10.1177/030913330102500104.

Departamento de Águas e Energia Elétrica do Estado de São Paulo – DAEE-SP. (2014). *Banco de dados pluviométricos do Estado de São Paulo*. Retrieved in 2020, October 20, from http://www.daee.sp.gov.br

Elsafi, S. H. (2014). Artificial Neural Networks (ANNs) for flood forecasting at Dongola Station in the River Nile, Sudan. *Alexandria Engineering Journal*, *53*(3), 655-662. http://dx.doi.org/10.1016/j.aej.2014.06.010.

Filho, A. J. P., & Santos, C. C. (2006). Modeling a densely urbanized watershed with an artificial neural network, weather radar and telemetric data. *Journal of Hydrology (Amsterdam)*, *317*(1-2), 31-48. http://dx.doi.org/10.1016/j.jhydrol.2005.05.007.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. Cambridge: The MIT Press.

Haddad, E. A., & Teixeira, E. (2015). Economic impacts of natural disasters in megacities: the case of floods in São Paulo, Brazil. *Habitat International*, *45*(2), 106-113. http://dx.doi.org/10.1016/j.habitatint.2014.06.023.

Haykin, S. S. (1998). *Neural networks: a comprehensive foundation.* Upper Saddle River: Prentice Hall.

Hsu, K.-L., Gupta, H. V., & Sorooshian, S. (1995). Artificial neural network modeling of the rainfall-runoff process. *Water Resources Research*, *31*(10), 2517-2530. http://dx.doi.org/10.1029/95WR01955.

Instituto Geológico do Estado de São Paulo – IG-SP. (2014). *Mapeamento de riscos associados a escorregamentos, inundações, erosão e solapamento de margens e drenagens, (II) - Introdução, Metodologia, caracterização do município, Resultado do Risco regional* (Technical Report, 108 p.). São Paulo.

International Network for Multi-Hazard Early Warning Systems – INMHEWS. (2017). Multi-hazard early warning systems: a checklist. In *1st Multi-hazard Early Warning Conference*, Cancún, México. Retrieved in 2020, October 20, from https://library.wmo.int/doc_num.php?explnum_id=4463

International Strategy for Disaster Reduction – ISDR. (2006). *Platform for the promotion of early warning*. Retrieved in 2020, October 20, from https://www.unisdr.org/2006/ppew/whats-ew/basics-ew.htm

Kanbua, W., & Khetchaturat, C. (2007). Decision support system for flash flood warning management using artificial neural network. In *The 11th Annual National Symposium on Computational Science and Engineering (ANSCSE11)* (pp. 28-30). Thailand: Faculty of Engineering, Prince of Songkla University.

Kobiyama, M., & Goerl, R. F. (2007). Quantitative method to distinguish flood and flash flood as disasters. *Hydrological Research Letters*, *1*, 11-14. http://dx.doi.org/10.3178/suisui.1.11.

Köppen, W. (1918). *Climatologia: con un estudio de los climas de la tierra*. Mexico-DF: Fondo de Cultura Economica.

Londhe, S., & Charhate, S. (2010). Comparison of data-driven modelling techniques for river flow forecasting. *Hydrological Sciences Journal*, *55*(7), 1163-1174. http://dx.doi.org/10.1080/02626667.2010.512867.

Mosavi, A., Ozturk, P., & Chau, K.-W. (2018). Flood prediction using machine learning models-literature review. *Water*, *10*(11), 1536-1575. http://dx.doi.org/10.3390/w10111536.

Oyebode, O., & Stretch, D. (2019). Neural network modeling of hydrological systems: a review of implementation techniques. *Natural Resource Modeling*, *32*(1), e12189. http://dx.doi.org/10.1111/nrm.12189.

Pickbrenner, K., Duarte, E., Weschenfelder, A. B., Germano, A. O., & Pinto, E. J. A. (2017). Cheia na bacia do Rio Caí (RS) em outubro de 2016. In *Anais do XXII Simpósio Brasileiro de Recursos Hídricos*. Florianópolis: ABRH.

Robbins, H., & Monro, S. (1951). A stochastic approximation method. *Annals of Mathematical Statistics*, *22*(3), 400-407. http://dx.doi.org/10.1214/aoms/1177729586.

Roy, J. K., Gupta, D., & Goswami, S. (2012). An improved flood warning system using WSN and Artificial Neural Network. In *Annual IEEE India Conference (INDICON)* (pp. 770-774). New York: IEEE. http://dx.doi.org/10.1109/INDCON.2012.6420720.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533-536. http://dx.doi.org/10.1038/323533a0.

Sankaranarayanan, S., Prabhakar, M., Satish, S., Jain, P., Ramprasad, A., & Krishnan, A. (2020). Flood prediction based on weather parameters using deep learning. *Journal of Water and Climate Change*, *11*(4), 1766-1783. http://dx.doi.org/10.2166/wcc.2019.321.

Silva, M. R., Santos, L. B. L., Scofield, G. B., & Cortivo, F. D. (2016). Using an artificial neural network for hidrological warnings on the Claro River in Caraguatatuba, São Paulo State. *Anuário do Instituto de Geociências*, *39*(1), 23-31. http://dx.doi.org/10.11137/2016_1_23_31.

Varoonchotikul, P. (2003). *Flood forecasting using Artificial neural networks*. Boca Raton: CRC Press.

Windarto, J. (2010). Flood early warning system develop at garang river semarang using information technology base on SMS and web. *International Journal of Geomatics and Geosciences*, *1*(1), 14-28.

Zucherato, B., Cunha, L., & Freitas, M. I. C. (2016). Análise das áreas de Inundação de Campos do Jordão – SP: caracterização por meio de bancos de dados. In A. Nunes, C. O. Moreira, I. R. Paiva & L. S. Cunha (Eds.), *Territórios de água* (pp. 162-170). Coimbra: CEGOT.

Authors contributions

Glauston Roberto Teixeira de Lima: Preprocessing of raw data (outlier detection, identification of missing data, filtering, etc); database formatting (search for the best format of data tuples); choosing the best neural network architecture and training algorithm; neural network training and testing; literature review; preparation, submission and review of the article.

Graziela Balda Scofield: Supply of raw data; literature review; preparation of the article.