ORIGINAL PAPER

Mamdani fuzzy inference systems and artificial neural networks for landslide susceptibility mapping

Luísa Vieira Lucchese[1] · **Guilherme Garcia de Oliveira[2]** · **Olavo Correa Pedrollo[1]**

Abstract

Two Artificial Intelligence (AI) methods, Fuzzy Inference System (FIS) and Artificial Neural Network (ANN), are applied to Landslide Susceptibility Mapping (LSM), to compare complementary aspects of the potentials of the two methods and to extract physical relationships from data. An index is proposed in order to rank and filter the FIS rules, selecting a certain number of readable rules for further interpretation of the physical relationships among variables. The area of study is Rolante river basin, southern Brazil. Eleven attributes are generated from a Digital Elevation Model (DEM), and landslide scars from an extreme rainfall event are used. Average accuracy and area under Receiver Operating Characteristic curve (AUC) resulted, respectively, in 81.27% and 0.8886 for FIS, and 89.45% and 0.9409 for ANN. ANN provides a map with more amplitude of outputs and less area classified as high susceptibility. Among the 40 (10%) best-ranked FIS rules, 13 have high susceptibility output, while 27 have low; a cause is that low susceptibility areas are larger on the map. Slope is highly connected to susceptibility. Elevation, when high (plateau) or low (floodplain), inhibits high susceptibility. Six attributes show the same fuzzy set for the 18 best-ranked rules, meaning this fuzzy set is common on the map. Overall findings point out that ANN is best suited for LSM map generation, but, based on them, using FIS is important to help researchers understand more about AI models for LSM and about landslide phenomenon.

Keywords Rule set · Mass movement · Natural disasters · Map analysis · Map validation · Fuzzy rule interpretation

✉ Luísa Vieira Lucchese
luisa.lucchese@ufrgs.br

[1] Instituto de Pesquisas Hidráulicas, Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves, 9500, Porto Alegre, RS 91501-970, Brazil

[2] Departamento Interdisciplinar, Universidade Federal do Rio Grande do Sul, Rodovia RS 030, 11700, km 92. Emboaba, Tramandaí, RS 95590-000, Brazil

1 Introduction

Landslide is defined by Cruden (1991) as the movement of a mass of rock, earth or debris down a slope. Brazil has a history of these natural disasters causing death and property damage (Dourado et al. 2012; Guha-Sapir 2019).

Landslide susceptibility mapping (LSM) is the mapping of sites that are susceptible to landslides, which means to identify the areas subject to future landslides to prevent further loss of lives and properties. The objective of performing LSM differs from the one of other geomorphology research areas because, in this case, the analysis of observed distribution patterns, related to environmental conditions, is not of primary interest (Brenning 2005). Susceptibility maps show "where" landslides can be expected, they do not provide an estimate of "when" they will occur (Neuhäuser and Terhorst 2007).

Within the vast number of possibilities for LSM, the role of machine learning (ML) (Murphy 2012) and artificial intelligence (AI) (Minsky 1961) models are highlighted. AI can be defined as a branch of computer science focused on creating machines to perform tasks that typically require humans. It includes but is not limited to logistic regression (LR), random forest (RF), Naïve Bayes (NB), adaptive-network-based fuzzy inference system (ANFIS), fuzzy inference systems (FIS) and artificial neural networks (ANN). Previous studies on LSM area showed that AI methods usually outperform other methods (Xiao et al. 2019; Dou et al. 2019).

Both FIS and ANN are universal approximators (Hornik et al. 1989; Wang 1992), which means that, correctly set, they can approximate any existing measurable relation. Therefore, both have potential for an accurate representation of the landslide phenomenon. According to Kosko (1992), both ANN and FIS process inexact information in an inexact way, though ANNs do not use an explicit set of rules. Conversely, FIS are rule-based systems (Kosko 1992) that are applied to fuzzy sets (Zadeh 1965), within the theory of fuzzy logic. Instead of using traditional crisp logic, variables are classified in one or more fuzzy sets each, for which they have different grades of membership. Rules which are simple relations between fuzzy variables are characterized by conditional statements, e.g., "if X is small then Y is large" (Zadeh 1973). The rules can be either expert-based or train-based and Mamdani-type FIS (Mamdani 1977) returns numbers, and not functions, as outputs to the rules. When applying a FIS to a problem set in the crisp logic domain, fuzzification must be performed before the execution of the rules. After, defuzzification should be performed in order to return the results to crisp logic.

Even so, particularities of the phenomena or of the locations may cause one of the models to be more suited to the LSM task. Kanungo et al. (2006) compared maps generated by ANN and by FIS, although the main goal of their paper was to compare those to a conventional weighting procedure and to a combined (ANN and fuzzy) method. Kanungo et al. (2006) ANN presented an accuracy of 72.6% for the validation set used on cross-validation training. This metric was also used for the choice of the ANN architecture between 39 ANNs trained. It was the metric used for external verification as well. This is not the standard procedure, because generally, a separate verification set is generated to be used for test purposes only. The analysis of the area occupied by landslides for each output class showed that the model is acceptable, although Kanungo et al. (2006) did not present accuracy metrics for their FIS, which made a direct comparison between FIS and ANN more difficult. Pradhan and Pirasteh (2010) also compared ANN to FIS for LSM. Their ANN was trained without cross-validation and the training was set to stop when a goal value of Root Mean Squared Error (RMSE) of 0.01 was achieved or by 300 epochs. Fuzzy procedure

was performed for some fuzzy operators, from which gamma operator with $\gamma = 0.975$ was chosen to generate the map. Comparing the resulting susceptibility indexes to the actual landslides, Pradhan and Pirasteh (2010) concluded that the ANN outperformed the FIS.

Nevertheless, even if comparisons between FIS and ANN for LSM were shown in the two previous studies, some knowledge gaps are still to be investigated. The knowledge field can benefit from a comparison between Mamdani FIS and well-trained ANNs for LSM, and from a thorough exploration of the similarities and differences between the outputs of those two methods.

The existence of a human-friendly representation (visible rules) (Benitez et al. 1997) is one of the main differences between FIS and ANN and can be taken as an advantage of FIS. Interpretation of ANN behavior is not as apparent, so FIS can be used as a tool for understanding data and their relations. In Ercanoglu and Gokceoglu (2002), FIS rules were based on each variable separately, generating a total of 23 rules that were combined later. In this case, rules did not need interpretation as they had been provided by the authors, based on a prior factor analysis. Pourghasemi et al. (2012) FIS rules were also based on each separate variable. Their 40 rules were generated based on frequency ratio technique and were not further discussed. A total of 17 rules were generated by Zhu et al. (2014) on their expert-based FIS, but no thorough discussion is presented in the paper. Some authors focused their research on the susceptibility maps and did not present the FIS rules (Pradhan 2010; Peethambaran et al. 2019). The existence of explicit rules is a particularity of the FIS methodology that can be used for a better understanding of the method and of the phenomenon itself. However, in LSM field, studies on FIS rule interpretation cannot be found in the present literature. Therefore, the approach to rule interpretation presented in the present paper is needed to fill a knowledge gap and is one of the main contributions of this paper.

In this manuscript, we draw comparisons between Mamdani FIS and ANN for LSM, analyzing the similarities and differences between the outputs, as well as the rules generated by the FIS. We design and adapt the Mamdani FIS used for the presented research and perform an automated procedure to generate the rules. Using the designed resources of the FIS to our favor, the rules are filtered and interpreted, which is the main novelty of the present paper. The discussion here presented helps us to understand not only the landslide phenomenon, but also the methods employed to model it. The ANNs are trained by using cross-validation procedure with independent training, validation and verification sets, which provides the ANN a better adjustment to the problem and favors a fairer comparison to FIS than the ones currently found in the literature. The same metrics, accuracy and AUC, are calculated for both methods, by analogous procedures based on the verification sets, in order to enable a direct comparison between FIS and ANN.

2 Materials and methods

2.1 Area of study

The area of study is Rolante river basin, with an area of 828.26 km^2. It is a sub-basin of Sinos river basin, located in the state of Rio Grande do Sul, southern Brazil (Fig. 1). The altitude in the area used for this study varies from 15 to 997 m and the average slope angle is 13.1°. Serra Geral formation, from which the scarps of Rolante river basin are derived, was formed by basaltic and rhyolitic spills that happened during the Triassic period (White 1908), approximately 127–137 Ma ago (Turner et al. 1994). The lava spills that generated
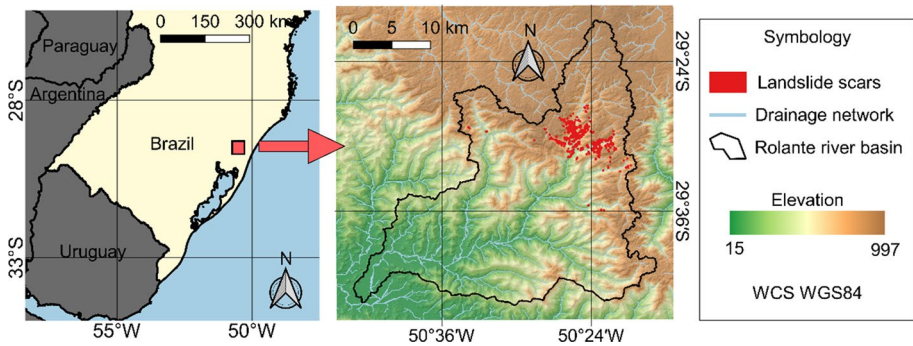
Fig. 1 Location of Area of Study within its regional context, and location of the Rolante river basin and of the landslide scars in the Area of Study

Serra Geral formation cover an area of approximately 917,000 km^2 (Frank et al. 2009). This includes about half of the territory of Rio Grande do Sul state, where many of its scarps are located. Rolante river basin and other on the scarps of Serra Geral have had landslide events in the last decades, showing that this geological setting may pose a risk of future landslides.

A series of landslides occurred in Rolante river basin on January 5, 2017. It did not cause any casualties, but properties were damaged, and some people lost their homes. The rainfall event that triggered Rolante river basin landslides accumulated 142.44 mm of rain in 57 h, according to the rainfall measurements of Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA) (Huffman et al. 2007), and the highest rainfall intensity was 21 mm/h. For reference, the mean annual precipitation in the basin is 1625 mm, according to Serviço Geológico do Brasil—CPRM (2011) (measurements between 1977 and 2006). TRMM estimates of convective rainfall are known to underestimate extreme rainfall in subtropical South America (Rasmussen et al. 2013). According to a Technical Report (SEMA and GPDEN, 2017) released shortly after the incident, farmers rainfall measurements varied from 90 to 272 mm in 24 h, showing that the spatial variability of the rain was high, making it harder for the satellites to capture it precisely.

2.2 Data sources

The scars from the landslides that occurred are mapped based on satellite imagery of the site and are revised. The images used for mapping were acquired from the satellite constellation of Maxar Technologies ©, available in the software Google Earth ©. The location of the scars is presented in Fig. 1. Because of the high spatial variability of the rainfall event and the localized nature of the scarps, the landslide scars mapped are clustered in a region of the map.

The methodology here presented follows a protocol; there are two sources of information for retrieving data to use in our models: one of them is the scar polygons, and the other one is a digital elevation model (DEM), which is a data source available for most of the globe. The DEM used is the Phased Array type L-band Synthetic Aperture Radar (PALSAR) by Advanced Land Observing Satellite-1 (ALOS) from Japan Aerospace Exploration Agency (ASF DAAC 2018), with 12.5 m spatial resolution. Lee et al. (2004) indicated that a spatial

resolution of about 10 m is good for mapping susceptibility. Arnone et al. (2016) analyzed the spatial resolution of the DEM especially for ANN landslide susceptibility applications and found out that the best resolutions were 10, 20 and 30 m, and that higher resolutions did not necessarily result in better maps. Susceptibility mapping is generally carried out in scales in the order of 1:50,000 to 1:100,000, that are compatible with the DEM used for this research.

2.3 Attribute generation

The default attribute is (a) elevation, which is the DEM itself. Other ten attributes are generated on QGis: (b) aspect: the slope aspect; (c) hillshade: a shaded image based on the DEM; (d) natural logarithm of flow accumulation: flow accumulation, based on flow direction, in log scale; (e) planar curvature: horizontal curvature of the terrain; (f) profile curvature: vertical curvature of terrain; (g) slope angle: declivity of the slope; (h) slope length and steepness factor (LS factor): a term used on Universal Soil Loss Equation; (i) topographic wetness index (TWI); (j) valley depth: vertical distance to the channel network base level; (k) vertical distance to channel network: vertical distance to the nearest draining channel. The attributes are shown in Fig. 2. The equation for the TWI is:

$$TWI = ln\left(\frac{upstream\,drainage\,area}{tan(Slope\,angle)} \right). \tag{1}$$

Some of the attributes present high intercorrelations with each other. However, in Lucchese et al. (2020), it was shown that even attributes that were intercorrelated by a coefficient higher than 0.7 brought useful information to the model.

2.4 Multilayer perceptron

The type of ANN used is a multilayer perceptron (MLP) with one hidden layer (Fig. 3), consisting of $nh = 30$ neurons. The number of neurons in the hidden layer is determined by using a novel approach (Lucchese et al., 2020). It consists in choosing the minimum number of neurons in the hidden layer for which the relationship between the input and the output variables is well represented. This analysis is performed by using the validation sample and all neural networks trained for this purpose are carefully observed not to be overfitted. The overall goal is to avoid producing a purposely oversized neural network. The neuron number choice is performed by analyzing the metrics attained using a range of neuron numbers in the hidden layer and plotting this data so that the tendencies can be carefully analyzed by the researcher.

The activation function used is unipolar sigmoid, and the inputs are represented by $\vec{p} = p_1, p_2, \ldots p_n$. The weights of the hidden layer are $\vec{w_h} = w_{h,1}, w_{h,2}, \ldots w_{h,k}$ (in which $k = nh \bullet n$), while the output weights by $\vec{w_o} = w_{o,1}, w_{o,2}, \ldots w_{o,nh}$, and b_h are the bias weights of the hidden layer and b_o is the bias of the output layer. The output is susceptibility, given in index values between 0 and 1.

The inputs, as they are derived from attributes that have multiple different orders of magnitude, are linearly scaled to [0,1] by

$$p_i = \frac{P_i - min_i}{max_i - min_i}, i = 1,2,\ldots 11, \tag{2}$$

where P_i is the original attribute, acquired by sampling on QGis, p_i is the normalized attribute, min_i is the global minimum value for the attribute and max_i is the global maximum.
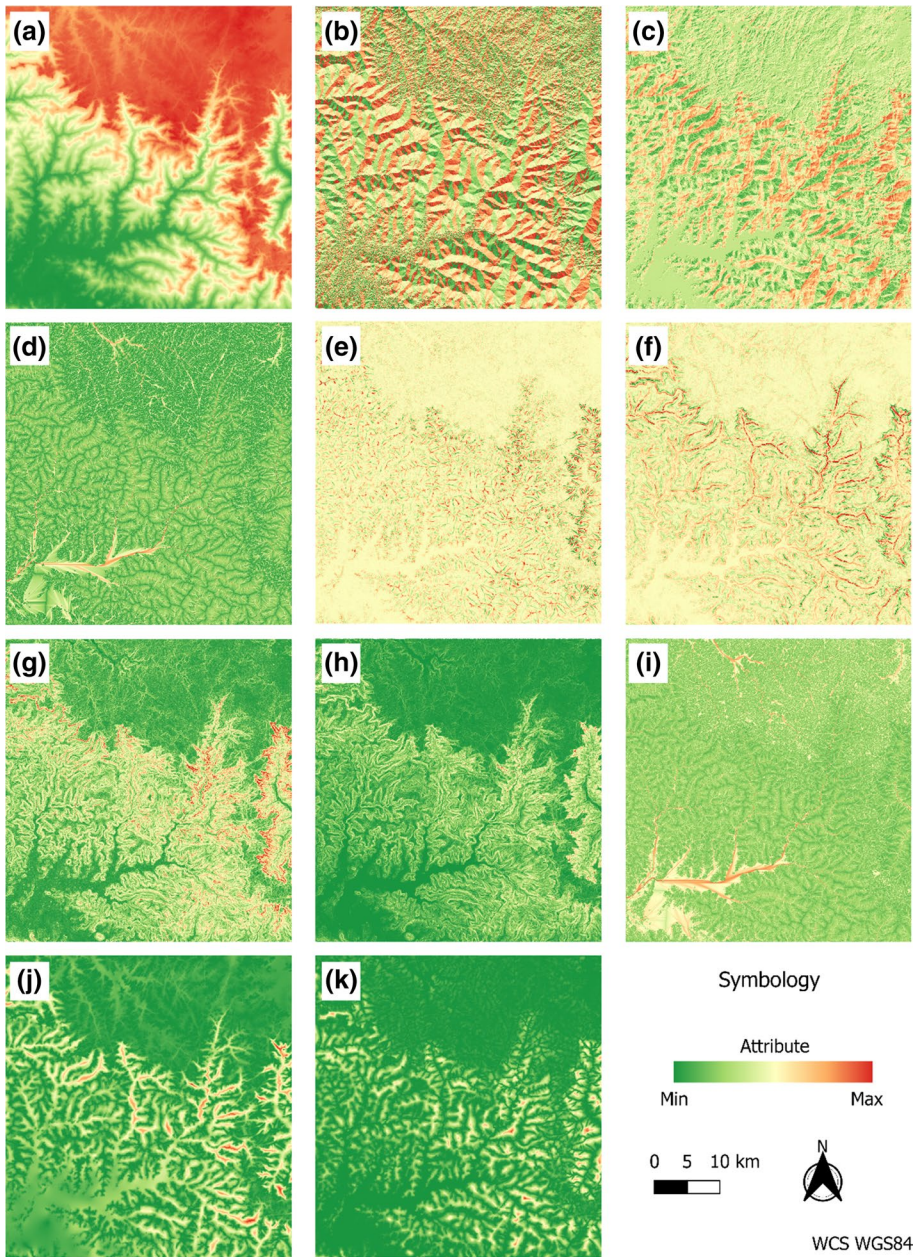
Fig. 2 Attributes on the Area of Study. **a** Elevation (min: 15, max: 997); **b** Aspect (min: 0, max: 360); **c** Hillshade (min: 0.00, max: 2.29); **d** Natural Logarithmic of Flow Accumulation (min: 5.05, max: 20.60); **e** Planar curvature (min: −0.5, max: 0.5; real minimum −1.36, real maximum 1.18); **f** Profile curvature (min: -0.5, max: 0.5; real minimum:−3.17, real maximum: 1.70); **g** Slope (min: 0, max: 50; real minimum: 0, real maximum: 79); **h** LS factor (min: 0, max: 30; real minimum: 0, real maximum: 123); **i** TWI (min:0.6, max: 25.2); **j** Valley Depth: (min:0, max: 401); **k** Vertical Distance to Channel Network (min: 0, max: 454)

Fig. 3 Scheme of the MLP employed
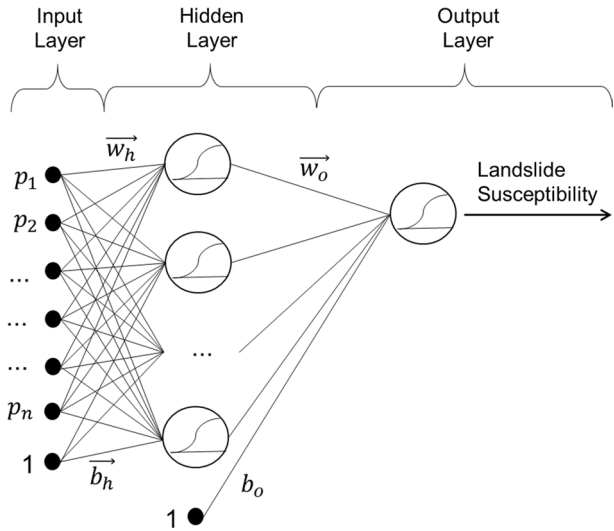
The MLP is trained by backpropagation (Rumelhart et al. 1986), on which the error obtained in the last layer is backpropagated to the previous layers progressively to the first one. It is based on the approximation given by the gradient descent of the errors, applied to progressively minimize the loss function value. The delta rule used is:

$$\overrightarrow{w_{k,t}} = \overrightarrow{w_{k,t-1}} + \tau\vec{\delta}\overrightarrow{p_k} + mo\left(\overrightarrow{w_{k,t-1}} - \overrightarrow{w_{k,t-2}}\right), \tag{3}$$

where delta is defined as $\vec{\delta} = \overrightarrow{e_k}\overrightarrow{s}'_k(\eta_k)$, the errors in k layer are $\overrightarrow{e_k}$, and $\overrightarrow{s}'_k(\eta_k)$ is the activation function derivative. In Eq. (3), the output weights of k layer are $\overrightarrow{w_k}$, τ is the learning rate, $\overrightarrow{p_k}$ are the inputs of layer k, and subscript t indicates the current epoch. A momentum term is used for faster training; mo is set to 0 if there was no improvement in training on the last epoch, and to 0.96 otherwise.

The learning rate used is heuristically varied (Vogl et al. 1988), in order for the MLP to achieve a near-optimum rate. The initial rate is set to $\tau = 0.00001$. If, in a given epoch, the quadratic error increases, the rate is reduced to $\tau = 0.5\tau$, otherwise, it is increased to $\tau = 1.1\tau$.

In the training phase, a procedure based on the cross-validation technique (Hecht-Nielsen 1990) is employed to avoid overfitting, using three sets, training (used for training the MLP), validation (used for cross-validation) and verification (used for testing and metrics calculation). Training is stopped when no improvement is made on validation set for over 10,000 epochs because a rise in the errors calculated based on the validation set, concomitantly with the errors of the training sample continuing to decrease as the training progresses, is a sign that overfitting is occurring. When this happens, training should be stopped. Training is also to be stopped if the overall training epochs exceed 300,000. To generate the three sample sets, the available samples are divided into 50% for training, 25% for validation and 25% for verification. This separation is random, but each set is ensured to consist in 50% occurrence (landslide) samples and 50% non-occurrence (non-landslide) samples.

Initial weights are randomly set. To constrain variability provided by the initial weights, five repetitions are performed, and the network that provides the lowest quadratic error on the validation set is chosen between the five.

Other source of variability is the separation between samples. To overcome it, five different separations are made, and an average of the maps is calculated. For each of these five separations, five sets of initial weights are tested (but only the best performing one is chosen to compose the final map), resulting in a total of 25 ANNs trained for this study. Both FIS and ANN models are custom programmed in Matlab® R2012b, which favors occasional improvements to the methods.

2.5 Mamdani fuzzy inference system

Mamdani fuzzy inference systems (FIS) are based on highly formalized insights about the structure of categories and articulated fuzzy "IF THEN" rules that can be based on expert knowledge. Fuzzy systems combine fuzzy sets with fuzzy rules to produce overall complex nonlinear behaviors (Kosko 1992). General steps are necessary for the application of a fuzzy model (see Fig. 4): (1) Input and output variables are fuzzified by considering convenient linguistic subsets; (2) Rules that relate input to output are constructed based on expert knowledge and/or extracted from data samples; (3) A reasoning mechanism is used to apply the rules to each input, resulting in a compound fuzzy set generated by the logical union of two or more fuzzy membership functions defined on the universe of discourse of the output variable; (4) A defuzzification procedure is performed to convert the fuzzy output to a crisp number.

In this paper, the number of FIS fuzzy sets employed for inputs is three (Low, Medium and High) and for the fuzzy sets employed for the output it is two (Low and High susceptibility). The prototype, which is the point where the fuzzy set achieves maximum membership, corresponds to the point where the other fuzzy sets have null membership. This results in every point belonging to a maximum of two fuzzy sets for each variable. The membership functions are initially sinusoidal for all fuzzy sets, but they can be altered later on the adaptive phase. All rules generated for this paper are of the type AND, which means they consider the effects of one variable classification combined to other variables classifications. Another possibility, not used, would be the OR-type rule, which considers the effects of one classification for one variable or another classification for other variable in the same rule.

The original dataset is divided in rule generation (50%), weight and function adaptation (5% + extreme values for each attribute) and verification (45%). Rule generation sets should be as large as possible so that more scenarios are presented, and more rules are generated, making the FIS model more generalizable. However, the ANN is using 50% of the samples, therefore 50% of the samples are used for FIS rule generation in order to keep
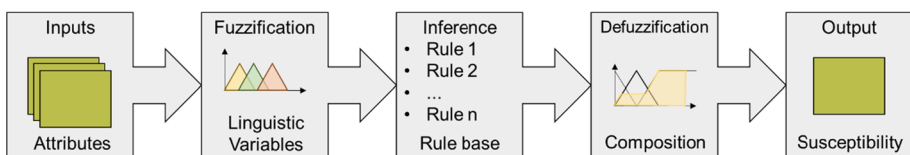

Fig. 4 Scheme of the general FIS structure. Based on Sen (2010)

the same proportion. But weight, location and function adaptation are different from cross-validation, because it is an adjustment, that needs to be performed to the rules that have been already formed. It needs a well-distributed set, therefore 5% plus the extreme values for each attribute.

FIS needs large datasets for rule generation because, in order to represent all possible combinations between variables, the FIS would need to have $3^{11} = 177147$ rules, which is much more than the total number of samples (13,480), but one must try to obtain sufficient rules to cover the situations that actually occur. To be consistent with ANN, five FIS based on different random separations of the dataset are generated and adapted, and the maps and metrics presented are an average of these five.

2.5.1 Self-organizing training

The rules are generated based on data. The method we employ for rule generation is based on Wang and Mendel (1992) and consists of fuzzification, assigning rules, assigning weights to the rules, and comparing similar rules generated by different registers. After this process, extra rules can be added by the expert.

Fuzzification of training data consists of transforming the data from crisp to fuzzy logic. Example: register 1, variable x_1 has 0.8 membership to fuzzy set Very low, 0.2 membership to fuzzy set Low; variable x_2 has 1.0 membership to fuzzy set high, and output has 1.0 membership to fuzzy set Medium. When assigning rules, the fuzzy set with the highest membership for each variable is saved, and a rule is created based on the information. Example: Very low, High → Medium. To generate the rule weight, the membership values of the variables that generated the rule are multiplied. Example: $0.8 \cdot 1.0 \cdot 1.0 = 0.8$. If another register generates a similar rule, but with a different output, the rule with the highest weight is saved and the other is discarded. Example: rule Very low, High → Low with 0.2 weight would be discarded.

2.5.2 Adaptive training

Adaptive training used for FIS was first developed by Driankov et al. (1996). Training performed is not exhaustive, as not all training options are explored, only variations in settings and parameters. The adaptations are about the membership function, the location of fuzzy set prototypes and the modification of the weights. For this research, we use a data subset for which the quadratic error should be lowered. The quadratic error is defined by:

$$E_q = \left(S_x - T_x\right)^2, \tag{4}$$

where S_x is the output response of the FIS, when executed, and T_x is the output (label) given on the dataset. The configuration that provides the lowest E_q is kept.

The possible modification of the membership functions is to trapezoidal, triangular, or inverted sinusoidal. A limitation is imposed that, for each variable, all membership functions should be the same. The next analysis is the possible modification of the location of the medium fuzzy set prototypes, on the input variables. Ten equally spaced displacement options between the neighboring prototypes (Low and Medium, Medium and High) are tested for each variable prototype. The weights of the fuzzy rules can also be altered at this point. For the weight of each rule, ten alternatives are tested varying between half and twice the original weight.

2.5.3 Defuzzification

Among many methods that have been proposed in the literature, we choose the center of area method, which is the most prevalent and physically appealing (Lee 1990a, b). It is given by the algebraic expression:

$$Z^* = \frac{\int \mu(z)zdz}{\int \mu(z)dz},$$ (5)

where Z^* is the resulting centroid, z is the set of numerical values of the output variable in its domain, and μ is the degree of membership of z to the composite output membership function. We use a numerical approximation for the integrals, performed along the domain of the output function. For this, we calculate integrals by the Riemann method for numerical integration, using domain discretization of $dz = 1 \bullet 10^{-4}$.

3 Theory and calculation

3.1 Dataset generation

The dataset is generated based on the attributes and scars location. First, every point of the rasters located within the scars is considered an occurrence sample and is acquired. There is a total of 6740 occurrence points available. Then, buffers of 5 km are generated around the scars and the same number of non-occurrence samples is acquired randomly in the area, resulting in 13,480 points, half of them occurrence, half of them non-occurrence samples. This sampling procedure was chosen based on the analysis of Lucchese et al. (2021).

3.2 Rule analysis

Since hundreds of rules are generated on each FIS, the rules have to be filtered to enable further analysis. Two factors are accounted for: the weight of the rule and the occurrence of it in the map. The following steps are proceeded:

– The attributes are scanned, and each time a pixel activates a rule, its counter is increased by one.
– The occurrence of each rule is normalized by the total number of pixels.
– The weight and the normalized occurrence are multiplied to calculate a factor.
– The rules are ordered based on the factor.
– The first 40 rules are shown.

3.3 Validation of methods

Accuracy is the rate of correct classifications made by the model and can be defined as $acc = (TP + TN)/(TP + TN + FP + FN)$, where TP is the number of True Positives, TN, True Negatives, FP, False Positives, and FN, False Negatives. To compute TP, TN, FP and FN, a threshold for classification is needed, since the models ANN and FIS outputs are continuous between zero and one. Our procedure is to calculate acc for every threshold and choose the threshold for the highest acc.

Another metric used to evaluate our models is the area under Receiver Operating Characteristic (ROC) curve, or AUC (DeLong et al. 1988). To plot a ROC curve, a range of thresholds is needed, as well. TP, TN, FP and FN calculated for each threshold are used to calculate True Positive Rate, or TPR, and False Positive Rate, or FPR. TPR is defined by

$$TPR = \frac{TP}{TP + FN} \tag{6}$$

and FPR is given by:

$$FPR = \frac{FP}{FP + TN}. \tag{7}$$

TPR is plotted on the abscissa and FPR on the ordinate, generating a graph that is expected to start on (0,0), as for threshold equal to one, both TP and FP will be null, and to end on (1,1) for a threshold equal to zero, because then TN and FN will be null. A flawless model would be located on (0,1) for the other thresholds because the FP and FN counts would be zero. This situation also maximizes the AUC to one. In opposition, a model with an AUC of 0.5 would not be able to discriminate between the output classes at all. AUCs of reliable models are usually over 0.7.

4 Results

4.1 Resulting fuzzy sets

Five FIS are generated and the membership functions and prototype locations for the Medium fuzzy set are adjusted. FIS 1 generates 396 rules, FIS 2, 390, FIS 3, 398, FIS 4, 405, FIS 5, 406 (average: 399).

The prototype of the fuzzy set relative to the medium class was initially placed on the middle point between each attribute minimum and maximum, and, in the adaptive phase, the membership function and the location of the medium prototype can be modified. In Fig. 5, fuzzy sets for FIS 1 are shown.

The resulting fuzzy sets presented in Fig. 5 show that for most attributes the trapezoidal membership function was chosen by the adaptive procedure. The only attribute for which sinusoidal function is kept is profile curvature. For elevation, planar curvature, profile curvature and slope, the Medium fuzzy set remains close to the center of the domain. Aspect middle fuzzy set migrates to the right, which can mean a separation between 300 and 350° benefits the model. The separation between Low and Medium fuzzy sets for TWI is adjusted to be lower.

For the five resulting FIS, prototype locations of the resulting medium fuzzy sets do not vary widely, and the functions chosen by the algorithm for each attribute resulted the same, making it possible to perform analysis based on FIS 1.

4.2 Metrics calculated

For FIS and ANN trained, metrics AUC and acc are calculated based on the verification sample of each model. These metrics are presented in Table 1, where an average for each model group is also shown. It is possible to observe that ANN provides higher metrics than FIS, although both present accuracies over 80%. Within both FIS and ANN groups, metrics
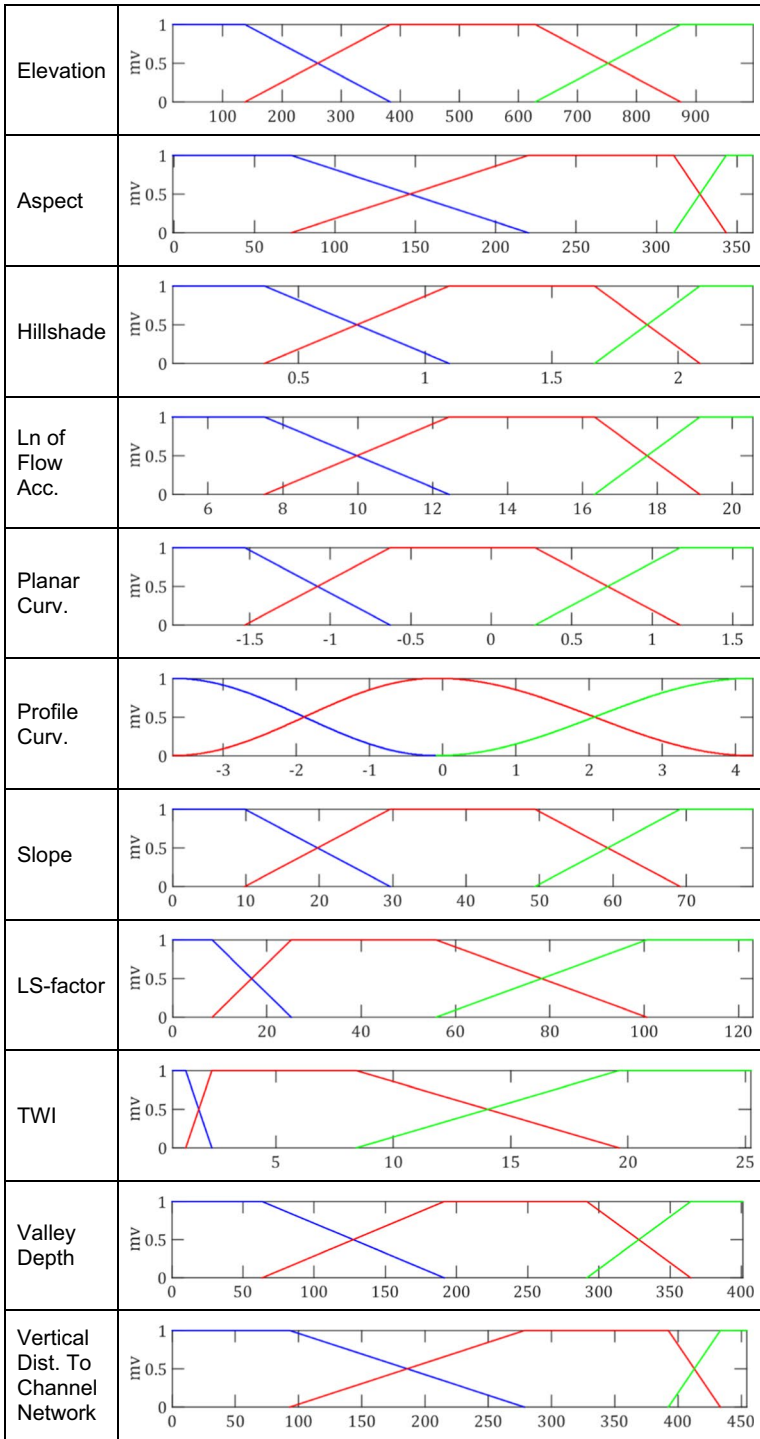
Fig. 5 Fuzzy sets of FIS 1 after adaptive phase. Blue: Low fuzzy set; red: Medium fuzzy set, green: High fuzzy set

Table 1 Metrics AUC and acc, calculated based on the verification samples

Model	AUC	Acc (%)	Model	AUC	Acc (%)
FIS 1	0.8876	81.63	ANN 1	0.9481	90.01
FIS 2	0.8874	80.49	ANN 2	0.9345	89.43
FIS 3	0.8901	81.43	ANN 3	0.9377	88.41
FIS 4	0.8869	80.99	ANN 4	0.9445	90.54
FIS 5	0.8909	81.81	ANN 5	0.9399	88.84
Average	0.8886	81.27	Average	0.9409	89.45

are very balanced between different models trained, showing the separation of samples is sufficiently random (Table 1).

4.3 Susceptibility maps

Resulting susceptibility maps are plotted, based on the output given by the models. This is done by executing each model individually to every pixel of the raster and saving the output as a raster as well. Later, the ensemble maps are composed by an average of the raster maps generated by each individual model. These are shown in Fig. 6, and an area is zoomed so that the landslide scars can be observed over the maps. The zoomed area is marked in yellow. The ANN resulting map shows more amplitude of outputs than the FIS
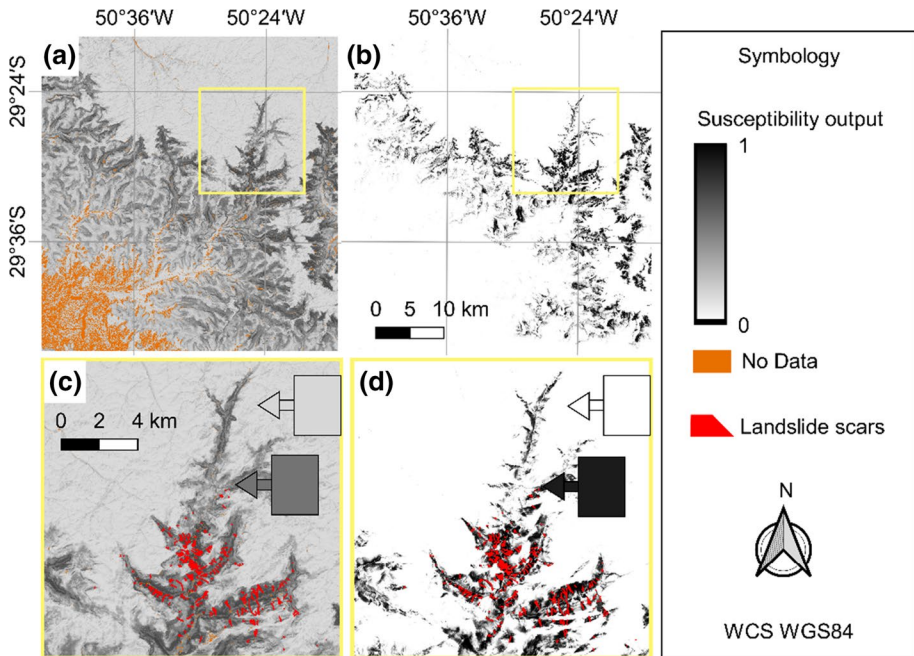
Fig. 6 Resulting susceptibility maps for **a** FIS and **b** ANN. Highlights on the scar area with the landslide scars plotted over it for **c** FIS and **d** ANN. In **c** and **d**, tags representing the susceptibility output in two locations were added to observe the amplitude of results for both models

map, as it can be observed in the color range in gray scale of the output map of each model, and by observing the susceptibility outputs in some locations, represented by the tags in Fig. 6c and d. Also, it shows to be more constrained regarding the areas classifiable as susceptible, as it is observable that for FIS a larger region has darker shades of gray, indicating outputs closer to 1. In the zoom (Fig. 6c and d), it is possible to observe that both maps can be considered acceptable, as the scars tend to be located in the areas marked as highly susceptible.

4.4 Comparisons between susceptibility maps

Based on the average thresholds, calculated as 0.49 for FIS and 0.53 for ANN models, one can generate maps that include the binary classification between low and high susceptibility, and then analyze in which areas do the resulting maps from FIS and ANN differ. First, the map is divided into four classes: both FIS and ANN resulting Low, both FIS and ANN resulting High, and the two cases in which the classifications differ between the models. The totals of each classification within the basin limits are presented in Table 2, a contingency table. It is possible to observe that, for most of the locations, the models agree on the classification, but, when they disagree, FIS providing high while ANN providing low outputs is more common than the opposite.

In order to further investigate the discrepancies between FIS and ANN, a map is generated with the differences calculated for the output values, between FIS and ANN outputs. Figure 7 allows for us to understand the numerical differences between the output values. Positive values mean the FIS provided a higher susceptibility for that area, and negative values mean that the ANN provided higher output, instead. An area in which the models mostly agree is the plateau on the north. Differences concentrate on the scarped slopes on mid- and lowlands. Three places with negative differences are plotted (Figs. 8, 9, 10), and then, three places with positive differences (Fig. 11), in order to analyze which factors contribute to these discrepancies between models. Along with the zooms in the highlighted areas, attributes considered useful for the analysis are plotted. The classifications on fuzzy sets of the attributes shown are relative to FIS 1.

As it can be observed in Figs. 8, 9 and 10, a combination of High aspect, Low hillshade and Medium slope can influence the output. Since the discrepancy is negative, it must mean that output susceptibility was higher for ANN than for FIS. Looking into the rules created by the FIS, those which include these variables fuzzy sets and result in low susceptibility have lower weights and are less common than rules that have these fuzzy sets but result in High susceptibility, indicating these could be particular cases. By the locations, it is not possible to infer which classification, Low or High, is right.

The opposite occurrence, which is when ANNs provide a lower output susceptibility than FIS, is also analyzed. This effect, seen in Fig. 11, is observed to be connected to the

		ANN	
		High	**Low**
FIS	**High**	292,652	536,669
	Low	100,416	4,086,217
No Data: 284,893			

Table 2 Contingency table of the output classifications provided by FIS and by ANN, inside the basin limits
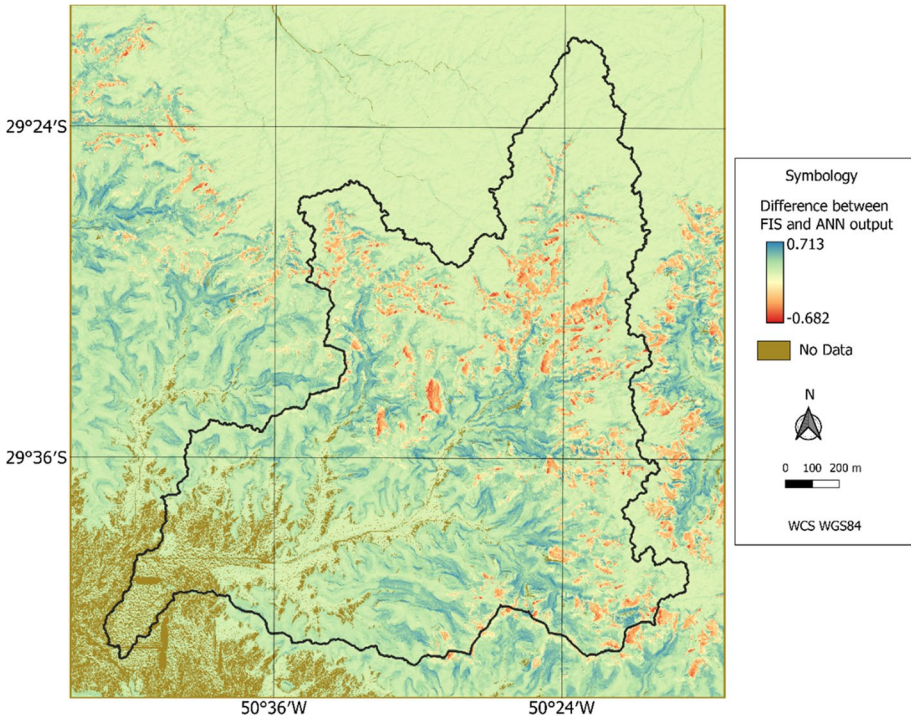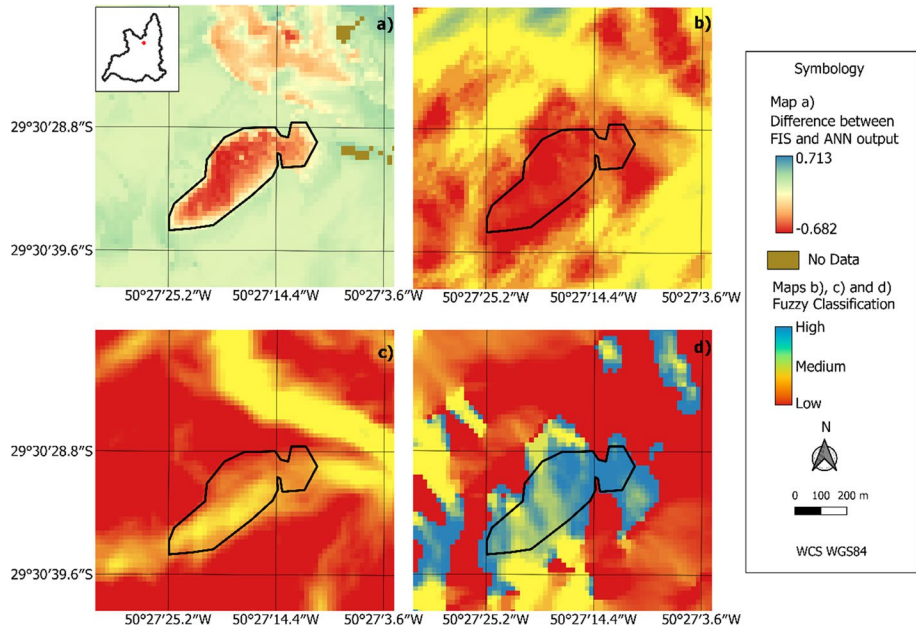
Fig. 7 Difference between FIS and ANN output for susceptibility. Rolante river basin is plotted for reference

attribute LS factor, as it transitions from Low to Medium fuzzy set in the areas with positive differences. This can either be an underestimation of the effect of LS factor on the ANN, or an overestimation of a rule connected to LS factor in FIS.

4.5 FIS rules

Because there are approximately 400 rules for each FIS model, only one FIS (FIS 1) is chosen to be shown. Even so, there are too many rules to interpret directly (396 for FIS 1). Some rules, even if they have high weights, are rarely used for map generation, while others, generated with low weights, are widely used instead. In order to balance these two effects, rule weight and rule use, a factor is calculated, and the rules are ordered and filtered based on it. FIS 1 rules are ordered based on this proposed factor which is the product between the rate of use of the rule on the map, and the rule weight. After ordering the rules, the first 40 (10% of total) are filtered and are shown in Table 3. The rule number is an indication of the order in which the rules are created by the original dataset (rule 1 was the first one created, then, rule 2, and so on). Therefore, observing the rule number, it is possible to see if the rule was one of the first or the last to be created. The rules related to Low susceptibility output are created before the ones related to High susceptibility, what is expected since the data were organized with all Low output first, and High output last, 50% each. This results in 228 rules (58%) with Low susceptibility, rules 1 to 228, and 168

Fig. 8 Four representations of the same area: **a** Difference between FIS and ANN output for susceptibility, **b** Hillshade, **c** Slope, **d** Aspect. First zoomed area, negative difference between FIS and ANN highlighted
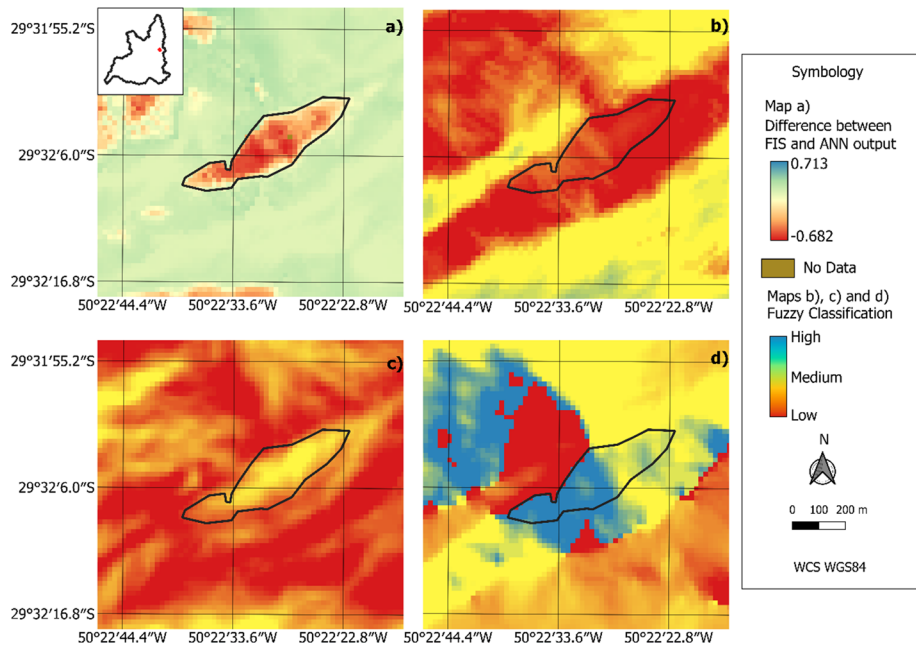
Fig. 9 Four representations of the same area: **a** Difference between FIS and ANN output for susceptibility, **b** Hillshade, **c** Slope, **d** Aspect. Second zoomed area, negative difference between FIS and ANN highlighted
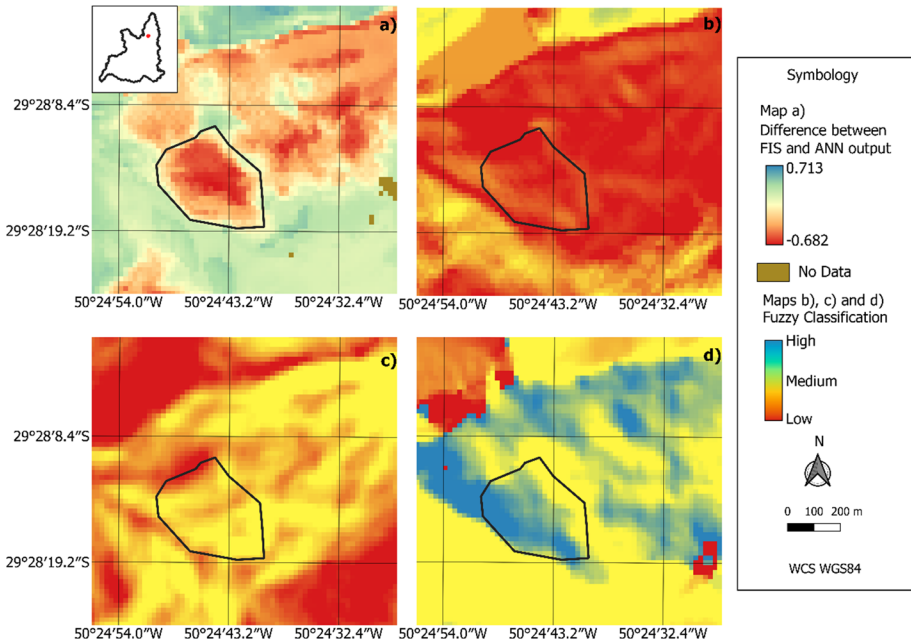
Fig. 10 Four representations of the same area: **a** Difference between FIS and ANN output for susceptibility, **b** Hillshade, **c** Slope, **d** Aspect. Third zoomed area, negative difference between FIS and ANN highlighted

rules with High susceptibility output (42%), rules 229 to 396. Rules with High susceptibility output appear less on the filtered list (13 out of 40), possibly because areas with high susceptibility are less abundant on the map and this is accounted for in the factor.

5 Discussion

In this manuscript, we provide a workaround for a FIS issue, i.e., the difficulty to interpret a large set of rules, making it possible to analyze the rules by implementing a factor and filtering some of them. FIS rule analysis had not been shown in previous LSM papers. This novel approach allows for a focused analysis of the rule set, in which the most important and most used rules gain a spotlight to be actively discussed.

ANN metrics are significantly higher than those achieved by FIS model. The amplitude of results portrayed by ANN is larger than the FIS one, showing it is a model that carries less uncertainty on the results. In a visual inspection, both susceptibility maps provided by ANN and FIS seem acceptable. In the binary analysis, FIS provided more areas with High susceptibility output than ANN.

FIS model is rule based, so the rules used to generate the maps can be observed directly. In opposition, because it is rule based, the system depends on specific rules to generate the output, which can be a weakness of FIS. Some or all rules may not be available for a given point during FIS execution, generating outputs that are either: a) nonexistent (No Data), or b) generated based on some of the possible rules, but not all. The case "a" happens when
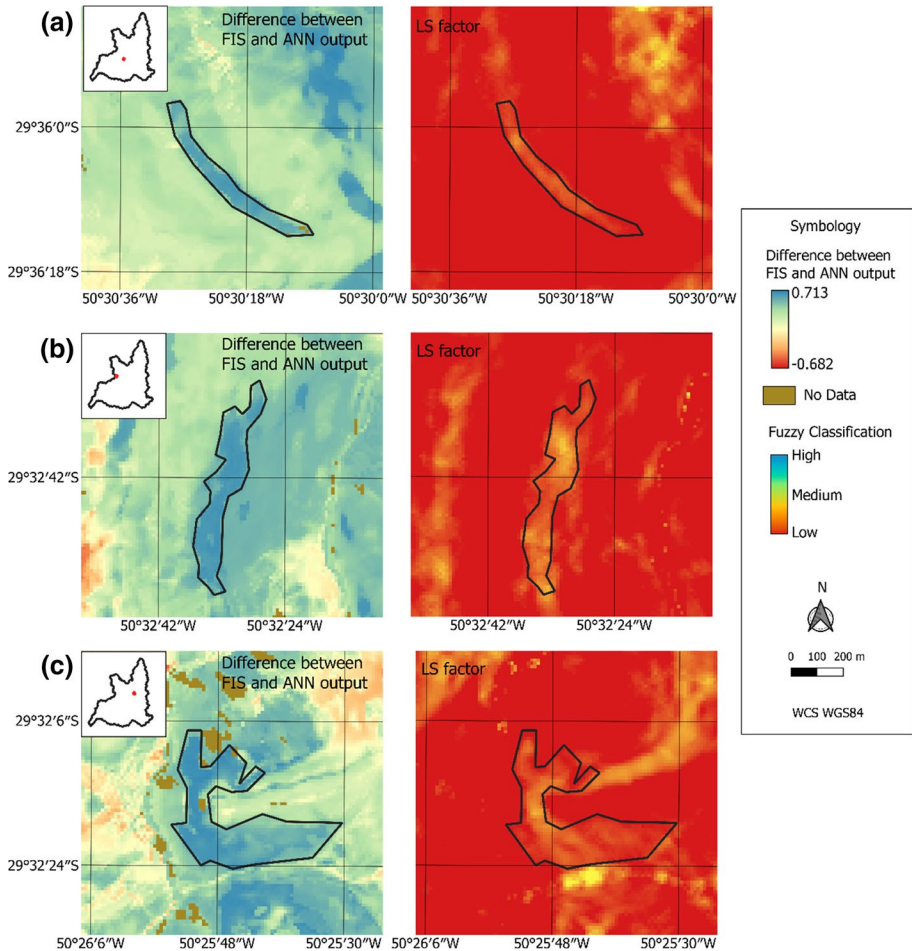
Fig. 11 Difference between FIS and ANN output for susceptibility. Three zoomed areas with positive differences highlighted

no rule could be found to suit any of the fuzzy classification combinations present on the said point. This can be directly observed in Figs. 6 or 7 (No Data points).

The case "b" happens when the output is generated based on the rules that exist on the rule set, ignoring rules that, though possible, do not exist. Case "b" is not directly observable in any map, because it results in non-null outputs. The existing rules are a small subset of the universe of possible rules (∼400/177,147). Each pixel in the domain can activate up to $2^{11} = 2048$ rules (2 fuzzy classifications for each attribute, maximum, and 11 attributes total), depending on its attribute fuzzy set memberships. Because only approximately 400 rules exist in the rule set, the output needs to be generated based on a subset of the generated rules, even if the map pixel is predominantly a member of the adjacent fuzzy set, i.e., a combination of fuzzy sets for which no rules exist. Based on that, particular cases existing in the sample that generated the rules can have large influence on the rule set. Additionally, rules do not merge or blend with each other at any point of model generation or adaptation. That means it is possible to have 20 rules, for which 19 that relate Medium slope to High

Table 3 Filtered FIS rules (*L* Low, *M* Medium, *H* High). Attributes: (a) Elevation; (b) Aspect; (c) Hillshade; (d) Natural logarithmic of flow accumulation; (e) Planar curvature; (f) Profile curvature; (g) Slope; (h) LS factor; (i) TWI; (j) Valley depth; (k) Vertical distance to channel network

Rule no.	Rule rank	Attributes (a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	Output	Rule weight	Rel. use	Factor
5	1	H	M	M	L	M	M	L	L	M	L	L	L	1.000	0.291	0.291
15	2	H	L	M	L	M	M	L	L	M	L	L	L	1.000	0.252	0.252
18	3	M	M	M	M	M	M	L	L	M	L	L	L	1.000	0.250	0.250
78	4	M	M	M	L	M	M	L	L	M	L	L	L	0.529	0.377	0.200
58	5	H	M	M	M	M	M	L	L	M	L	L	L	1.000	0.126	0.126
38	6	M	L	M	L	M	M	L	L	M	L	L	L	0.357	0.318	0.114
295	7	M	L	M	M	M	M	M	L	M	M	L	H	0.597	0.184	0.110
232	8	M	M	M	M	M	M	M	L	M	M	L	H	0.772	0.142	0.109
238	9	M	L	M	M	M	M	L	L	M	M	L	H	1.000	0.109	0.109
47	10	H	L	M	M	M	M	L	L	M	L	L	L	1.000	0.108	0.108
70	11	M	L	M	M	M	M	L	L	M	L	L	L	0.498	0.212	0.106
267	12	M	L	M	M	M	M	M	L	M	M	L	H	0.794	0.112	0.089
82	13	L	L	M	M	M	M	L	L	M	L	L	L	0.595	0.138	0.082
294	14	M	M	M	L	M	M	M	L	M	M	L	H	0.455	0.159	0.072
16	15	L	M	M	M	M	M	L	L	M	M	L	L	0.741	0.094	0.070
114	16	L	L	M	M	M	M	L	L	M	M	L	H	0.881	0.077	0.068
119	17	L	M	M	L	M	M	L	L	M	M	L	L	0.632	0.102	0.065
13	18	M	M	M	L	M	M	L	L	M	M	L	L	0.423	0.151	0.064
96	19	H	L	L	L	M	M	L	L	M	L	L	L	0.381	0.145	0.055
59	20	M	M	M	M	M	M	L	L	M	M	L	L	0.397	0.139	0.055
253	21	M	M	M	M	M	M	L	L	M	M	L	H	0.376	0.139	0.052
77	22	L	M	M	L	M	M	L	L	M	L	L	L	0.300	0.159	0.048
288	23	M	L	M	L	M	M	M	L	M	M	L	H	0.342	0.125	0.043
49	24	H	H	M	L	M	M	L	L	M	L	L	L	1.000	0.041	0.041
98	25	L	L	M	L	M	M	L	L	M	M	L	L	0.436	0.083	0.036

Table 3 (continued)

Rule no.	Rule rank	Attributes											Output	Rule weight	Rel. use	Factor
		(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)				
81	26	M	M	M	M	M	M	M	L	M	L	L	L	0.163	0.218	0.035
109	27	L	M	M	M	M	M	M	L	M	L	L	L	0.319	0.096	0.031
6	28	H	H	L	L	M	M	L	L	M	L	L	L	0.546	0.055	0.030
343	29	M	M	H	M	M	M	M	M	M	M	L	H	0.977	0.027	0.027
333	30	H	L	M	M	M	M	M	L	M	L	L	H	0.495	0.053	0.026
68	31	H	L	L	M	M	M	L	L	M	L	L	L	0.410	0.063	0.026
230	32	M	M	H	M	M	M	M	L	M	M	L	H	0.615	0.042	0.026
60	33	H	M	L	L	M	M	L	L	M	L	L	L	0.120	0.211	0.025
336	34	M	M	M	M	M	M	M	M	M	M	L	H	0.399	0.062	0.025
84	35	M	H	L	L	M	M	L	L	M	L	L	L	0.363	0.066	0.024
74	36	M	M	L	L	M	M	L	L	M	L	L	L	0.110	0.213	0.024
67	37	M	M	M	M	M	M	M	L	M	M	L	L	0.163	0.142	0.023
381	38	M	L	M	M	M	M	M	M	M	M	L	H	0.479	0.047	0.023
243	39	M	H	L	M	M	M	L	L	M	M	L	H	0.919	0.024	0.022
72	40	M	H	L	M	M	M	L	L	M	L	L	L	0.517	0.042	0.022

susceptibility, and 1 that relates Medium slope to Low susceptibility, but no action is taken to suppress the latter. Supposing a given point falls into that one rule, output provided is Low. The effects of this issue (case "b") can be observed in the negative differences between FIS and ANN outputs in the map. For the fuzzy sets present in these areas, many rules exist, most of them providing the opposite output (High), but by the combination of those attribute fuzzy classifications to others, the model provides output Low instead.

As the two (ANN and FIS) are models and there is no definitive answer for the susceptibility on the area, it is not possible to be sure which output is correct, because the other attributes that influenced the FIS response can be relevant, although, they can be a source of misguidance for the model, as well. ANNs do not deal with the discussed issue, since there are no apparent rules, and merging continuous function inputs and outputs on the layers is part of its architecture.

For the areas where the difference is positive (FIS output is higher), it is possible to observe that, for FIS, LS factor guards a strong influence on the resulting output. This means that, in these places, the outputs provided by the ANN disagree with those from the FIS, and this can be happening due to the ANN not connecting LS factor to susceptibility as strongly as FIS does. As discussed, it is not possible to know if this influence is justified because the ANN, on the other hand, can be erasing the strong connection LS factor (a factor used for erosion calculation) has with landslide triggering.

Overall, one of the advantages of FIS is the possibility to directly observe the rules. The rule that achieves the highest factor, activated by over 29% of the pixels in the domain, was the fifth one to be generated. Probably, it is very abundant in the set that generated the rules, as it is abundant in the map. Rule number 78 ranks 4th even if its weight is not as high as the previous rules, but it is the most activated rule (37.7% of the pixels). Pixels can activate many rules, so some of the rules that are activated serve only to fine adjustment purposes, which can be the case for rule 78.

Some fuzzy sets of the attributes repeat themselves in many rules. Every one of the first 18 rules shown is activated by Medium hillshade, Medium planar curvature, Medium profile curvature, Low LS factor, Medium TWI, and Low vertical distance to channel network. It does not mean all these factors are unimportant for landslide susceptibility, but that these classifications are the most abundant in the map. Most of these attributes are on Medium fuzzy set (4 out of 6), that can either appear alone or combined with Low and High fuzzy sets, so it tends to be more activated if the variable has an approximately centered distribution of occurrence in the map. Following this line of thought, LS factor and vertical distance to channel network are predominantly Low in the area (see Figs. 2 and 5), and that explains the occurrence of the Low fuzzy set.

Some attributes show patterns in their connection to the output. Slope Low fuzzy set appears in 27 out of 40 rules, being 24 resulting in Low susceptibility and 3 in High. Medium slope results in High susceptibility in 10 rules out of 13. It shows that slope, as expected, is deeply connected to the landslide phenomena. Researchers of the area can agree that slope angles ranged from 10 to 70 degrees (Medium slope) are more prone to landslides than slope angles ranged from 0 to 30 degrees (Low slope), as shown by the rules of the FIS model. Elevation also shows to be connected to the landslide susceptibility output, in this case. Among the first 29 rules ranked, all of them that present High output had Medium elevation. And as expected, all of these rules that have High elevation resulted in Low susceptibility. This is probably a way that the FIS made itself able to differentiate the plateau from the susceptible areas. Lowlands, which are typically also not susceptible areas, are differentiated by the model using the same principle. The susceptibility to landslides is not typically connected to terrain altitude. However, in the study area, there

is a preferential altitude for the occurrence of landslides, which can be associated with a specific pedological or geological formation. And it is possible that the elevation attribute indirectly brought information about these formations to the model. Also, a disclaimer for the findings about elevation attribute is that the fuzzy set classifications for elevation are particular of the Rolante river basin. Different geological settings may need different fuzzy set classifications and can generate different rules than those presented here.

The findings show that the FIS model can propose rules for LSM that are human-friendly and could as well be produced by rational thought. In addition to those rules that seem apparent to researchers, FIS provides a range of other rules that can be physically logical, however are not as easy to be interpreted. In 20 out of 22 rules (in Table 3) with Low valley depth, output was Low susceptibility, showing these variables are related, as expected. This is not a simple effect of elevation attribute, since elevation fuzzy sets vary widely between the 22 rules. Based on these findings, it is possible to observe that valid physical relationships between variables can be extracted from data alone on LSM modeling with FIS.

Rule 5 and rule 15 (the first two ranked) differ only in aspect fuzzy classification (one is Low and other is Medium) and, for both, output is Low. Rule number 2 (rank 316), not shown in Table 3, differs from rules 5 and 15 only by being based on High aspect, and returns Low susceptibility as well. It can mean that aspect alone is not a defining factor for susceptibility, as, in the presence of these other variable classifications, the aspect of the slope does not alter the susceptibility output. However, this finding is limited to this study area (Rolante river basin). With a latitude of approximately 29°30′S, in this region, landslides occurred in all directions, but this is not generalizable. In higher latitudes, the moisture of the soil is expected to vary with the slope aspect, because the insolation varies more between the scarps with different orientations (Isard 1986). Hillsides with lower solar incidence due to shading tend to have deeper and moister soils (Franzmeier et al. 1969; Wang et al. 2011), which can be a compounding factor for mass movements. Therefore, in other settings, especially in high latitude locations, it is expected that the rules generated would be different regarding the aspect attribute.

6 Conclusions

In this manuscript, we compare two models, FIS and ANN, for LSM. FIS has inherent advantages, based on the possibility to understand the paths that lead to the output. In FIS, the rules are explicit, and their occurrence and weight can be analyzed. However, it should be noted that the FIS (as well as the ANN) generated is applicable only to this study area, because different settings may ask for different fuzzy set classifications and rules.

It is observed that the FIS rules can compartmentalize the relations between cause and effect, as similar rules that differ in only one attribute fuzzy set, but result in different outputs can coexist. This way, ANN has a greater possibility of merging the effects of different attributes, and it has generated outputs with more amplitude of results, and higher metrics. The compartmentalization of FIS rules can cause another issue, which is that only a small part of the universe of possible rules is covered, due to the limitations of the data used for training, so that outputs may be generated by an incomplete set of rules, ignoring combinations of fuzzy set attributes for which rules could, but do not exist.

ANN provides better maps, with greater accuracy. Although, FIS rules help us understand the phenomena and formulate clear rules to represent what occurs. The use of

both models for the same area of study, Rolante river basin, integrated the possibility to present a susceptibility map with high accuracy with the possibility to analyze the maps and their differences thoroughly for a better understanding of both the phenomenon and the processes performed by the models.

Based on the findings presented, many areas could benefit from this type of analysis, using one model in which the rules can be analyzed, and another one, without apparent rules, but providing higher accuracy. By comparing the resulting maps from both models and analyzing the locations in which their outputs differ, and why they do, it is possible to develop a more reliable knowledge of each place is more certainly susceptible to landslides, within the study area.

For future studies, it is suggested that more fuzzy sets can be employed in each variable to analyze the changes in rules. Also, different factors can be developed to rank the most relevant rules for this analysis.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Arnone E, Francipane A, Scarbaci A et al (2016) Effect of raster resolution and polygon-conversion algorithm on landslide susceptibility mapping. Environ Model Softw 84:467–481
ASF DAAC (2018) ALOS PALSAR Radiometric Terrain Corrected High res
Benitez JM, Castro JL, Requena I (1997) Are Artificial Neural Networks Black Boxes? IEEE Trans Neural Netw 8:1156–1164
Brenning A (2005) Spatial prediction models for landslide hazards: review, comparison and evaluation. Nat Haz Earth Sys Sci 5(6):853–862
Cruden DM (1991) A simple definition of a landslide. Bull Eng Geol Environ 43:27–29
DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44:837–845
Dou J, Yunus AP, Tien Bui D et al (2019) Evaluating GIS-based multiple statistical models and data mining for earthquake and rainfall-induced landslide susceptibility using the LiDAR DEM. Remote Sens 11:638
Dourado F, Arraes TC, Fernandes M et al (2012) O Megadesastre da Região Serrana do Rio de Janeiro: as causas do evento, os mecanismos dos movimentos de massa e a distribuição espacial dos investimentos de reconstrução no pós-desastre. Anuário do Inst Geociências 35:43–54
Driankov D, Hellendoorn H, Reinfrank M (Michael) (1996) An introduction to fuzzy control. Springer.
Ercanoglu M, Gokceoglu C (2002) Assessment of landslide susceptibility for a landslide-prone area (north of Yenice, NW Turkey) by fuzzy approach. Environ Geol 41:720–730

Frank HT, Gomes MEB, Formoso MLL (2009) Review of the areal extent and the volume of the Serra Geral Formation, Paraná Basin, South America. Pesqui em Geociências 36:49–57

Franzmeier DP, Pedersen EJ, Longwell TJ et al (1969) Properties of some soils in the cumberland plateau as related to slope aspect and position. Soil Sci Soc Am J 33:755–761. https://doi.org/10.2136/sssaj1969.03615995003300050037x

Guha-Sapir D (2019) EM-DAT: the emergency events database. Univ Cathol Louvain Brussels, Belgium

Hecht-Nielsen R (1990) Neurocomputing. Addison-Wesley

Hornik K, Stinchcombe M, White H (1989) Multilayer feedforward networks are universal approximators. Neural networks 2:359–366

Huffman GJ, Bolvin DT, Nelkin EJ et al (2007) The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. J Hydrometeorol 8:38–55

Isard SA (1986) factors influencing soil moisture and plant community distribution on niwot ridge. Arct Alp Res 18:83–96. https://doi.org/10.1080/00040851.1986.12004065

Kanungo DP, Arora MK, Sarkar S, Gupta RP (2006) A comparative study of conventional, ANN black box, fuzzy and combined neural and fuzzy weighting procedures for landslide susceptibility zonation in Darjeeling Himalayas. Eng Geol 85:347–366

Kosko B (1992) Neural networks and fuzzy systems : a dynamical systems approach to machine intelligence. Prentice Hall

Lee CC (1990a) fuzzy logic in control systems: fuzzy logic controller—part I. IEEE Trans Syst Man Cybern 20:404–418. https://doi.org/10.1109/21.52551

Lee CC (1990b) Fuzzy logic in control systems: fuzzy logic controller, part II. IEEE Trans Syst Man Cybern 20:419–435. https://doi.org/10.1109/21.52552

Lee S, Choi J, Woo I (2004) The effect of spatial resolution on the accuracy of landslide susceptibility mapping: a case study in Boun. Korea Geosci J 8:51

Lucchese LV, de Oliveira GG, Pedrollo OC (2020) Attribute selection using correlations and principal components for artificial neural networks employment for landslide susceptibility assessment. Environ Monit Assess 192:129. https://doi.org/10.1007/s10661-019-7968-0

Lucchese LV, de Oliveira GG, Pedrollo OC (2021) Investigation of the influence of nonoccurrence sampling on Landslide Susceptibility Assessment using Artificial Neural Networks. Catena 198:105067. https://doi.org/10.1016/j.catena.2020.105067

Mamdani EH (1977) Application of fuzzy logic to approximate reasoning using linguistic synthesis. IEEE Trans Comput 26:1182–1191. https://doi.org/10.1109/TC.1977.1674779

Minsky M (1961) Steps Toward Artificial Intelligence. In: The Guest Editor (ed) Institute of Radio Engineers. IEEE

Murphy KP (2012) Machine Learning A Probabilistic Perspective. Massachusetts Institute of Technology

Neuhäuser B, Terhorst B (2007) Landslide susceptibility assessment using "weights-of-evidence" applied to a study area at the Jurassic escarpment (SW-Germany). Geomorphology 86:12–24. https://doi.org/10.1016/j.geomorph.2006.08.002

Peethambaran B, Anbalagan R, Shihabudheen KV (2019) Landslide susceptibility mapping in and around Mussoorie Township using fuzzy set procedure, MamLand and improved fuzzy expert system-a comparative study. Nat Hazards 96:121–147. https://doi.org/10.1007/s11069-018-3532-4

Pourghasemi HR, Pradhan B, Gokceoglu C (2012) Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed. Iran Nat Hazards 63:965–996. https://doi.org/10.1007/s11069-012-0217-2

Pradhan B (2010) Application of an advanced fuzzy logic model for landslide susceptibility analysis. Int J Comput Intell Syst 3:370–381. https://doi.org/10.1080/18756891.2010.9727707

Pradhan B, Pirasteh S (2010) Comparison between Prediction Capabilities of Neural Network and Fuzzy Logic Techniques for Landslide Susceptibility Mapping

Rasmussen KL, Choi SL, Zuluaga MD, Houze RA (2013) TRMM precipitation bias in extreme storms in South America. Geophys Res Lett 40:3457–3461. https://doi.org/10.1002/grl.50651

Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. Nature 323:533–536

Secretaria Estadual do Meio Ambiente, Grupo de Pesquisa em Desastres Naturais (2017) Diagnóstico preliminar

SEN Z, (2010) Fuzzy logic and hydrological modeling. CRC Press, Taylor and Francis Group

Serviço Geológico do Brasil - CPRM (2011) Atlas pluviométrico do Brasil: isoietas mensais, isoietas trimestrais, isoietas anuais, meses mais secos, meses mais chuvosos, trimestres mais secos,

trimestres mais chuvosos. http://www.cprm.gov.br/publique/Hidrologia/Mapas-e-Publicacoes/Atlas-Pluviometrico-do-Brasil-1351.html. Accessed 19 Oct 2020

Turner S, Regelous M, Kelley S et al (1994) Magmatism and continental break-up in the South Atlantic: high precision 40Ar-39Ar geochronology. Earth Planet Sci Lett 121:333–348

Vogl TP, Mangis JK, Rigler AK et al (1988) Accelerating the convergence of the back-propagation method. Biol Cybern 59:257–263

Wang L-X (1992) Fuzzy systems are universal approximators. In: [1992 Proceedings] IEEE International Conference on Fuzzy Systems. pp 1163–1170

Wang L-X, Mendel JM (1992) Generating fuzzy rules by learning from examples. IEEE Trans Syst Man Cybern 22:1414–1427

Wang L, Wei S, Horton R, Shao M (2011) Effects of vegetation and slope aspect on water budget in the hill and gully region of the Loess Plateau of China. CATENA 87:90–100. https://doi.org/10.1016/j.catena.2011.05.010

White IC (1908) Report on the coal measures and associated rocks of South Brazil. Comm Estud Minas Brazil Rio Janeiro

Xiao T, Yin K, Yao T, Liu S (2019) Spatial prediction of landslide susceptibility using GIS-based statistical and machine learning models in Wanzhou County, THREE Gorges Reservoir, China. Acta Geochim 38:654–669

Zadeh LA (1965) Fuzzy Sets. Inf. Control 8:338–353

Zadeh LA (1973) Outline of a new approach to the analysis of complex systems and decision processes. IEEE Trans Syst Man Cybern SMC 3:28–44. https://doi.org/10.1109/TSMC.1973.5408575

Zhu AX, Wang R, Qiao J et al (2014) An expert knowledge-based approach to landslide susceptibility mapping using GIS and fuzzy logic. Geomorphology 214:128–138. https://doi.org/10.1016/j.geomorph.2014.02.003