UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA - CURSO DE ENGENHARIA MECÂNICA
TRABALHO DE CONCLUSÃO DE CURSO

PREDICTING RHEOLOGICAL PARAMETERS OF NATURAL VISCOPLASTIC FLUIDS THROUGH ARTIFICIAL NEURAL NETWORKS

por

Henrique Rosset Ferreira

Monografia apresentada ao Departamento de Engenharia Mecânica da Escola de Engenharia da Universidade Federal do Rio Grande do Sul, como parte dos requisitos para obtenção do diploma de Engenheiro Mecânico.

Porto Alegre, maio de 2021

DADOS INTERNACIONAIS DE CATALOGAÇÃO

Rosset Ferreira, Henrique
Predicting Rheological Parameters of Natural Viscoplastic Fluids through Artificial Neural Networks / Henrique Rosset Ferreira. -- 2021.
25 f.
Orientador: Guilherme Henrique Fiorot.

Trabalho de conclusão de curso (Graduação) -- Universidade Federal do Rio Grande do Sul, Escola de Engenharia, Curso de Engenharia Mecânica, Porto Alegre, BR-RS, 2021.

1. Corridas de Detritos. 2. Reologia. 3. Redes Neurais Artificiais. I. Fiorot, Guilherme Henrique, orient. II. Título.

Henrique Rosset Ferreira

PREDICTING RHEOLOGICAL PARAMETERS OF NATURAL VISCOPLASTIC FLUIDS THROUGH ARTIFICIAL NEURAL NETWORKS

ESTA MONOGRAFIA FOI JULGADA ADEQUADA COMO PARTE DOS REQUISITOS PARA A OBTENÇÃO DO TÍTULO DE
ENGENHEIRO MECÂNICO
APROVADA EM SUA FORMA FINAL PELA BANCA EXAMINADORA DO CURSO DE ENGENHARIA MECÂNICA

Prof. Mário Roland Sobczyk Sobrinho
Coordenador do Curso de Engenharia Mecânica

Área de concentração: Energia e Fenômenos de Transporte

Orientador: Prof. Guilherme Henrique Fiorot

Comissão de Avaliação:

Prof./Profª. Alexandre Vagtinski de Paula

Prof./Profª. Diogo Elias da Vinha Andrade

Prof./Profª. Paulo Smith Schneider

Porto Alegre, maio de 2021

AGRADECIMENTOS

A minha família, sem a qual essa jornada não seria possível.

NOMENCLATURA

Símbolos

f	Activation function	
h_j	Hidden Unit	
n	Flow index	
N	Number of observation	
w_{ij}	Weight of the j-th hidden unit to the i-th input neuron	
w_{jk}	Weight of the k-th output neuron to the j-th hidden unit	
x_i	Input Neuron	
y_k	Output Neuron	
z	Observated value	
$\hat{z}$	Predicted value	

Símbolos gregos

α_k	Bias of the k-th output neuron	
α_j	Bias of the j-th hidden unit	
λ	Regularization Parameter	
κ	Consistency	[Pa.s^n]
μ_B	Plastic Viscosity	[Pa.s]
τ	Shear Stress	[Pa]
τ_0	Yield Stress	[Pa]
$\dot{\gamma}$	Sheat Rate	[s^{-1}]

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA - CURSO DE ENGENHARIA MECÂNICA
TRABALHO DE CONCLUSÃO DE CURSO – 2021

PREDICTING RHEOLOGICAL PARAMETERS OF NATURAL VISCOPLASTIC FLUIDS THROUGH ARTIFICIAL NEURAL NETWORKS

Ferreira, Henrique Rosset
henriquerosset@hotmail.com

Abstract. *Debris flows are geological phenomena that can represent serious threats to life and property. They are fast mass movements of water-sediment mixtures, saturated with sediments that scale from smaller than 2 µm to up to 2 cm. Debris flow has heavy bulk densities and can travel at speeds up to 30 kph. They have high erosive capacity and can carry large items such as rocks (up to 1 m), trees, and cars. To assess the occurrence susceptibility and hazard risk of these phenomena, the evaluation of constitutive laws for the debris flow materials is required. Phenomenological models that describe these materials' behavior by fixed rheometric parameters, such as yield stress and consistency, had been attested as an effective alternative. Thus, the evaluation of these material rheological parameters plays a key role in this rheology framework. The objective of this study is to develop a predictive model using artificial neural networks to identify the mixture's rheological properties (yield stress, consistency, and flow index), based on the mixture solid concentration and the sediment grain-size distribution. 178 data points obtained from previous studies were used to build the model. Three single-hidden layer neural networks were built to predict each rheological property. The models showed that the prediction of these rheological properties based on mixture composition is feasible but needs further efforts to improve the model's accuracy and precision.*

Keywords: *Debris flow, Neural Networks, Bingham fluid, Herschel–Bulkley fluid, Predictive Modeling.*

Resumo. *As corridas de detritos são fenômenos geológicos com grande potencial de causar prejuízos e danos a vida humana. Esses fenômenos são rápidos escoamentos de um fluido composto por uma mistura de água saturada com sedimentos, de alta densidade e que pode alcançar velocidades de mais de 30km/h. As corridas de detritos têm alto poder erosivo e são capazes de arrastar grandes objetos como árvores e carros. Para avaliar as áreas sujeitas a esses fenômenos e estimar seus impactos, é necessário conhecer as equações constitutivas que governam os fluidos desses fenômenos. Uma eficiente alternativa é o uso de modelos que descrevem o comportamento desses fluidos com base em parâmetros reométricos fixos, como a tensão limite de escoamento e a consistência, o que torna crítica a mensuração destes parâmetros reológicos. O objetivo deste trabalho é desenvolver um modelo utilizando o método de redes neurais para prever propriedades reológicas desses fluidos com base na sua composição sedimentar. 178 observações coletadas de estudos já publicados foram empregadas na construção e validação do modelo. Três redes neurais, uma para cada propriedade reológica, foram construídas. Os resultados mostraram que é possível usar dessa tecnologia para prever as propriedades reológicas com base na composição sedimentar. Mais estudos, porém, são necessários para melhorar a precisão e acurácia desses modelos.*

Palavras-chave: *Corridas de detritos, Redes neurais, Fluido de Bingham, Fluido de Herschel–Bulkley, Modelagem preditiva.*

1. INTRODUCTION

Debris flows are a type of mass movement in which a fast-moving saturated mixture of water, sediments, and debris flows down slopes attracted by gravity forces. These geological phenomena happen in regions with steep inclinations and occasional rainfalls. Most debris flows start from static, nearly rigid masses of sediments, usually located on unstable slopes. When water from rainfalls infiltrates these masses, it increases the soil weight and decreases the soil strength. At some point, this sediment-water mixture transforms in a liquid-like state, which triggers the mass movement (IVERSON, 1997; JAKOB; HUNGR, 2005).

Debris flows distinguish themselves from other related events (e.g., rock avalanches or mudflows) because both solid and fluid forces influence these mass movement's dynamics. Rock avalanches, for example, are dominated only by solid grain forces, whereas in mudflows the fluid forces are the main driver of the event's physics. These combination of solid and fluid forces gives debris flow events a unique destructive power. Debris flows can travel long distances, have bulk densities comparable to rock avalanches, and flow with velocities peaking up to 10 m/s. These velocities are high enough to make bulk inertial forces relevant, given these mass movements enough energy to exert huge loads on objects in the flow path (IVERSON, 1997).

The prediction and prevention of debris-flow events are in the interest of the engineers, given the threat these mass movements represent to life and property. To perform this hazard assessments, it is necessary to understand the dynamics of debris flow. Debris flow phenomena are traditionally modeled as a homogeneous fluid and constitutive models are applied to predict somehow its viscoplastic behavior (COUSSOT; PIAU, 1994; PHILLIPS; DAVIES, 1991). Some of these models are Herschel-Bulkey, Bingham, among others (IVERSON, 1997).

In this frame, the bulk debris flow is assumed to be controlled by the rheologic properties of the fluid matrix (i.e., the debris flow mixture of water and sediments in suspension) (COSTA, 1984), where the particle size distribution and solid volumetric concentration could be used to characterize the fluid behavior (ANCEY, 2001; COUSSOT; PROUST; ANCEY, 1996).

The literature reports experimental studies showing the effects that water content and grain size distribution have in rheological parameters (KAITNA; RICKENMANN; SCHATZMANN, 2007; MALET *et al.*, 2003; PELLEGRINO; SCOTTO DI SANTOLO; SCHIPPA, 2015; SCOTTO DI SANTOLO; PELLEGRINO; EVANGELISTA, 2010; ZEGERS *et al.*, 2020). However, the experimental measure of rheological parameters for natural debris flow is a challenging task where similarity criteria is hard to achieve (TURNBULL; BOWMAN; MCELWAINE, 2015).

In this context, the present study aims to explore the use of predictive models, in this case artificial neural networks (ANN), to predict debris flow rheological parameters. One of the major advantages of ANN methods is their ability to approximate any existing nonlinear relationship between input and output parameters (WASZCZYSZYN, 1999). This characteristic suit the debris flow rheological parameter prediction, given the problem complexity. More specifically, the present work aims to:

- identify a collection of data presenting similar input and output variables;
- train an existing ANN model;
- understand how database could be better equipped to enhance the model’s response;
- explore the possible advantages and drawbacks of the approach.

2. FUNDAMENTALS

2.1. Debris flow rheology

Typical debris flows are hazardous mass movements characterized by being dense, poorly sorted solid-fluid mixtures with constituent particles that range widely in size, from smaller than 2 micrometers clay particles to gravel particles, that have up to 2 cm diameter sizes (MAJOR; PIERSON, 1992). The understanding of the flow dynamics of these phenomena is key to the risk assessment of these events. The evaluation of constitutive laws to model these solid-fluid mixtures is crucial to allow the use of the continuous mechanics' framework (i.e., Cauchy momentum equation) to describe the dynamics or the kinematics of these events.

Debris flow phenomena are traditionally modeled as a homogeneous fluid. That is, the poorly sorted mixture is regarded as one medium when moving, and its flow behavior is presumed to be controlled by the rheologic properties of the 'matrix'. This matrix is a mixture of fine sediment and water in which coarse particles are dispersed (COSTA, 1984; SCOTTO DI SANTOLO; PELLEGRINO; EVANGELISTA, 2010).

Due to the high variability of matrix compositions, no generally applicable model is yet able to cover the full range of possible flow behaviors.

2.1.1. Rheologic Models

Coussot et al. (1996), proposed a simplified classification criterion to separate the models applied to describe the 'rheologic' properties of the debris flow matrix based on the interaction of solid and fluid forces. Beyond a critical solid volumetric concentration the mixture typically behaves like a non-Newtonian, viscoplastic fluid (PARSONS; WHIPPLE; SIMONI, 2001; PELLEGRINO; SCHIPPA, 2013). A viscoplastic fluid has a yield stress under which it will not deform. That is, it is necessary to apply a shear stress that overcomes the material yield stress to induce the flow.

The mentioned approaches do not provide a unique rheological formula for the solid-fluid composite. Phenomenological models that employ a fixed rheology of debris flows material mixtures, such as the Bingham or the generalized Herschel-Bulkley models have been used to effective describe these mixtures' rheological behavior (ANCEY, 2007; PELLEGRINO; SCHIPPA, 2013). In this frame, the flow is characterized by relating shear stress to shear rate and by material parameters using viscosity and yield stress.

2.1.1.1. Bingham Model

The simplest viscoplastic model is the Bingham model. The shear stress τ by the Bingham law is given in the Equation 1:

$$\tau = \tau_0 + \mu_B \dot{\gamma} \tag{1}$$

where the shear stress is a linear function of the shear rate $\dot{\gamma}$ and the constant material parameters yield stress τ_0 and plastic viscosity μ_B.

For many natural debris flows, the Bingham model provides a rough approximation of the flow behavior, but with good representativity over large shear rate values. For some range of mixture parameters, such as grain size distribution, the linear relationship between shear rate and shear stress fails to describe the mixture rheological properties (KAITNA; RICKENMANN; SCHATZMANN, 2007), specially at low shear rate.

2.1.1.2. Herschel Bulkley

The Herschel Bulkley is a generalized Bingham model. Here, the shear stress τ law is given by the Equation 2:

$$\tau = \tau_0 + \kappa \dot{\gamma}^n \quad (2)$$

where the shear stress is a non-linear function of the shear rate $\dot{\gamma}$, with the addition of the flow index n exponent, and the constant material parameters yield stress τ_0 and consistency κ.

The flow behavior is still viscoplastic, which means that a certain level of stress is necessary to induce the shearing into the flow. The difference remains in the flow index exponent. Experiments showed that for most mudflows and debris flow matrix n is lower than the unit, given a shear thinning behavior for the flow (COUSSOT; PROUST; ANCEY, 1996; KAITNA; RICKENMANN; SCHATZMANN, 2007).

2.1.2. Debris flow rheological parameters measurement

The experimental measure of rheological parameters for natural debris flow is a challenging task. The main difficulty of these measurements remains in replicating the field scale of these events. Analogue experiments of large-scale phenomena need to satisfy at least a relevant amount of similarity criteria to be trustworthy. However, this is hard to achieve for debris flow (TURNBULL; BOWMAN; MCELWAINE, 2015).

Different experimental approaches that achieve similarity for a limit set of parameters were develop through the years to deal with this scale challenge. The evaluation of the material parameters yield stress and consistency (or plastic viscosity) is key to rheological models such as Bingham and Herschel Buckley. The determination of these rheological parameters has been attempt with a variety of equipment's such as penetrometers, vane testers, viscosimeters, conventional rheometers and large-scale rheometers (ANCEY, 2001; PELLEGRINO; SCHIPPA, 2013; PHILLIPS; DAVIES, 1991).

These methods are not suitable for the complete flow and are mostly limited to specific operating conditions, often governed by the granular material size (PHILLIPS; DAVIES, 1991). Any generally applicable method would need to be able to accommodate reasonable large ranges of grain-sizes to produce conditions similar to an actual debris flow material.

The rheological parameters, yield stress and consistency (or plastic viscosity), are strictly related to grain characteristics and can be assessed from the particle size distribution and solid volumetric concentration (BAGNOLD; A, 1954; COUSSOT; PROUST; ANCEY, 1996; PELLEGRINO; SCHIPPA, 2013). Experimental studies showing the effects that water content and grain size distribution have in rheological parameters have been reported for specific ranges of grain size composition (KAITNA; RICKENMANN; SCHATZMANN, 2007; MALET *et al.*, 2003; PELLEGRINO; SCOTTO DI SANTOLO; SCHIPPA, 2015; SCOTTO DI SANTOLO; PELLEGRINO; EVANGELISTA, 2010; ZEGERS *et al.*, 2020).

2.2. Predictive Modeling

Predictive modeling is the process of developing a model to try to generate accurate predictions (KUHN; JOHNSON, 2013). A simple example of a prediction model is the ordinary linear regression, which provides a numeric response for a linear combination of its predictors.

In general terms, the approach to build a predictive model consists of the following steps: obtain representative data, pick a modeling technique, plug in data, and generate a prediction. However, a good understanding of the data and the scope of the aimed problem is necessary to

develop a trustworthy model to give accurate predictions for new samples (KUHN; JOHNSON, 2013).

The problem scope will point between a classification model and a regression model. To predict categorical outcomes, classification methods are necessary. To predict numerical continuous outcomes, regression methods are the best fit. The understanding of the data characteristics gives insights to choose between a linear and a non-linear model, for example. It also gives hints about the necessary data pre-processing treatments.

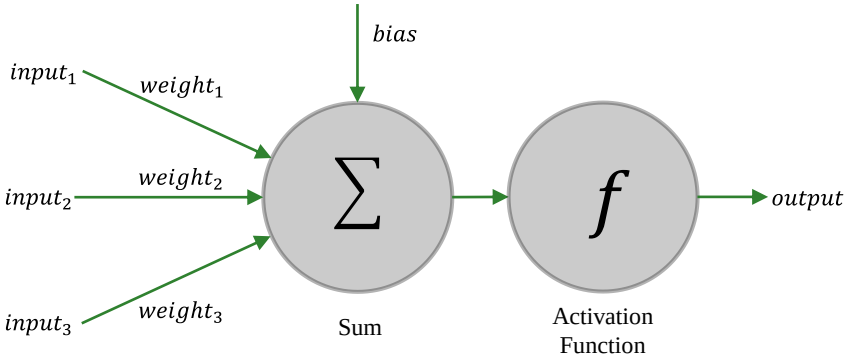
2.2.1. Artificial neural networks

Artificial Neural Networks are powerful nonlinear regression techniques. The roots of these methods rely in a scientific effort to replicate the human's decision-making approach. The decision-making processes of humans are related to the recognition of patterns which led to an interchange between engineers and psychologists. This interchange gave birth to powerful mathematical methods: the perceptron in the late 1950s, and neural networks in the 1980s. Even though both studies quickly departed from their biological roots to reach the potential of mathematical techniques, the psychological approach heavily influenced the name of these methods (RIPLEY, 1996).

2.2.1.1. Artificial neuron (AN) model

The concept borrowed from the nervous system is how a neuron receives electrical signals and reacts to these stimulations. A neuron will transmit information (i.e., an electrical signal) only if receives a charge superior to a specific threshold (WASZCZYSZYN, 1999). The AN model, shown in Fig. 1, mimics the biological neuron's ability to receive signals from several neurons, process these signals, and decide whether to send a signal forward or not.

Figure 1 – Generic artificial neuron model receiving information from 3 input neurons.



To the AN model, the numerical values inputs act as electrical signals from other neurons. The weights and the bias are variables to allow to change how the artificial neuron will interpret different neuron signals (i.e., numerical values) from previous AN. The AN model activation function f, which can work as a switch, is a way to evaluate whether a signal must go forward or not. One classical example of the activation function is the binary step, or Heaviside function.

Mathematically, the AN model sums the product of the input neurons and its relative weights to the bias associated with the neuron. Then, it applies the activation function to this sum. For the Heaviside function, for example, the neuron will transmit information (i.e, assume the value 1) only if the sum is greater than 1.

2.2.1.2. Feed-forward neural network

A neural network is an assembly of connected neurons aggregated in at least three layers. The simplest but most common architecture of neural network is an input layer, one hidden layer and an output layer. Feed-forward means that the signals are transmitted in one direction only (i.e., from inputs to outputs). The Figure 2 shows the architecture of a generic single-hidden layer neural network.

The mathematical formulation for one output neuron y_k in a single hidden-layered neural network is given in the Equation 3:

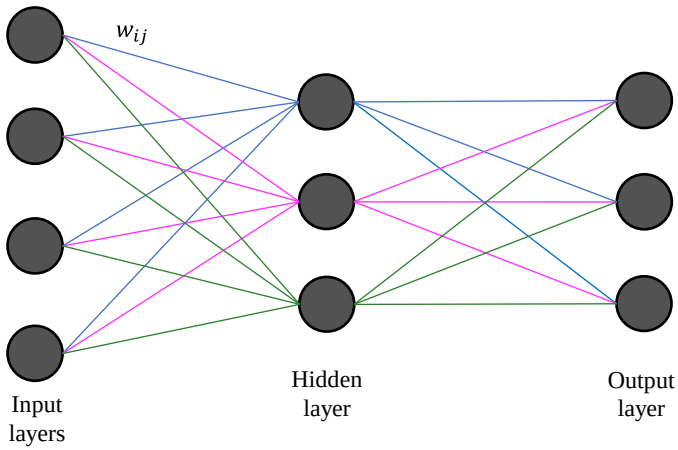
$$y_k = f_k\left(\alpha_k + \sum_{j \to k} w_{jk} h_j\right) \tag{3}$$

where the hidden neuron h_j is given by the Equation 4:

$$h_j = f_{jk}\left(\alpha_j + \sum_{j \to k} w_{ik} x_i\right) \tag{4}$$

and i are the input neurons indexes, j are the hidden units indexes, k the output neurons indexes, w_{ik} and w_{jk} are the neurons weight, and α is the neuron bias.

Figure 2 – Generic single-hidden layer neural network with 4 input neurons, 3 hidden units and 3 output layers.



From the Eq. (3), is easy to understand that the sum is relative to the interaction between the input neurons (x_i) and the hidden units' weights (w_{ij}) associated to each input neuron. The sum of the Eq. (4) is the product of the hidden units h_j and the associated k-th output neuron weights w_{jk}. The output neuron, then, is the value returned by the activation function applied to this sum plus the output neuron bias α_k.

Neural networks are capable of massive parallel processing of information. This characteristic implies that these methods have low sensitivity to errors cause by noisy

information. Another feature of ANN is plasticity (i.e., adaptivity to new information). ANN methods are not programmed, but instead they are trained. This means that the weights and bias for an ANN are oriented by examples (patterns) taken from experimental labeled data (i.e., data with the expected predict values) (WASZCZYSZYN, 1999).

2.2.1.3 Supervised learning

The supervised learning problem is, for a given approximation function (i.e., regression) called model, to find the optimal parameters that minimize the distance between a known observed-value (training data) and the prediction (i.e., approximation) made by this model (BONETTO; LATZKO, 2020). The collection of algorithms design to solve this problem are usually known as machine learning or artificial intelligence techniques.

For neural networks, one of these algorithms is the back-propagation algorithm. It is a highly efficient method that works with calculus concepts (i.e., derivatives) to search the optimal parameters for a neural network (KUHN; JOHNSON, 2013).

Due to its large number of variable parameters, ANNs have the tendency to over-fit the model parameters to a given set of training data. To avoid this issue, a penalization parameter (λ), known as regularization parameter or weight decay, is introduced to these models. For large regularization values, the fitted model becomes more smooth and less likely to over-fitting. This regularization parameter is a given parameter for a neural network architecture, same as the number of hidden layers and units (KUHN; JOHNSON, 2013). The introduction of the regularization parameter in the networks demand that the predictor (i.e., input data) is on the same scale.

2.2.2. Data pre-processing

Data pre-processing are a group of manipulation techniques that transform a given set of data. This stage of the modeling is crucial to the model final predictive ability. One straightforward and common data transformation is to center scale the all the predictor variables (KUHN; JOHNSON, 2013). The ‘centering’ part of the method subtracts from each predictor value the average of the predictor, resulting in a zero mean. The ‘scaling’ part divide each value for the predictor standard deviation, leading to a common standard deviation of one.

2.2.3. Performance metrics

The metrics chosen to measure the model’s predictive performance play an important role in the predictive modeling process. For regression models, accuracy measures are usually employed to evaluate the effectiveness of the models. There are many methods to measure accuracy, which should be chosen accordingly to the application (KUHN; JOHNSON, 2013). The choice of the appropriate method to measure predictive performance is critical to understand the quality of the model regarding its purpose.

2.2.3.1. Root mean squared error (RMSE)

The RMSE is a common method to evaluate a model’s predictive capabilities. This metric can be interpreted as how far, on average, is the distance between the observed values and the model predictions. The $RMSE$ is given by the Equation 5:

$$RMSE = \sqrt{\frac{\sum|z_i - \hat{z}_i|^2}{N}} \quad (5)$$

where z is i-th observed value, $\hat{z}$ is i-th predicted value and N is the number of measurement predictions.

The RMSE is an adequate prediction performance metric when the model is predicting values within the same scale.

2.2.3.2. Mean absolute percent error (MAPE)

The MAPE is an alternative metric to evaluate predictive capabilities when the model needs to predict values that are in different scales. The $MAPE$ is given by the Equation 6:

$$MAPE = \frac{1}{N}\sum\left|\frac{z_i - \hat{z}_i}{z_i}\right| \quad (6)$$

where z is i-th observed value, $\hat{z}$ is i-th predicted value and N is the number of measurement predictions.

The MAPE needs to be interpreted carefully hence it troubles to deal with near-zero values. Also, this performance metric cannot be used for datasets with zero values.

3. METHODOLOGY

Artificial neural networks can model complex nonlinear relationships such as the one between the debris flows mixture's composition and its rheological properties. ANNs can be built by learning from a set of input data which is the best value for its internal weights and bias, method known as supervised learning.

The reliability of ANNs models built through supervised learning methods rely heavily on the data (i.e., training data) they learned from. The architecture parameters of the networks (i.e., number of hidden layers and hidden units) are another key performance component. The metrics used to evaluate the prediction performance (i.e., MAPE, RMSE, R-squared) of the models also play a role in obtaining reliable models from a supervised learning process.

This section shows the study's methodology used to gather the training and test set process. It also presents the iterative process used to identify the best network architecture for a chosen set of data points. In the last subsection, a methodology to select which data points will participate in the training and testing data sets are also presented.

3.1. The data set

The dataset was built through a review of previous works in the literature with experimental studies on the relation of geomechanical characteristics (i.e., grain size distribution) and rheological characteristics (i.e., yield stress) of the natural mixtures. Six studies were chosen to compose the dataset. The common characteristics in these works were the presence of a geomechanical characterization of the soil, rheometrical measures for different soil volumetric concentrations in the mixture, and the indication of the flow index used to model the mixture's flow behavior.

Six datasets were chosen, and a brief description of each work is presented:

A) Major *et al.* (1992) performed rheology measurements in sediments from a natural debris flow deposit sampled from the North Fork Toutle River located in Washington, United States.

The material rheological properties were evaluated using a concentric cylinder viscometer designs and calibrated by the authors.

B) Coussot *et al.* (1994) studied industrial clay, natural clay from Grenoble in France and seven different debris flows samples using a stress-controlled and a strain-rate-controlled rheometer in the experiments.

C) Malet *et al.* (2003) tested deposits from seven mass movement events and three weathered soils, all originated from the French Alps. The article shows the rheological characterization and rheometric tests of natural mixtures and artificial mixtures (i.e., mixtures of soil made by the authors). The tests were carried out in rotation rheometers with different geometries such as parallel-plates and coaxial cylinders.

D) Kaitna *et al.* (2007) studied materials taken from the Scalära torrent in Eastern Switzerland. The studied used a ball measuring system implemented in a Paar Physica MCR 300 rheometer to measure torque and rotational speed which later were translated as rheological parameters.

E) Santolo *et al.* (2009) performed rheometrical measurements on three debris flows that originated from the Campania region in southern Italy. A rotational rheometer was used to test the mixture of soils sampled from the debris flow events.

F) Del Galdio *et al.* (2018) studied the rheological properties of clay-silt and sand-silt artificial mixtures. The rheological characterization was carried out with a Anton Paar MCR301 rheometer with a vane rotor geometry.

A total of 176 observations were collected. The collection of data from all datasets is shown in Figures 3 and 4. Fig. 3 presents the yield stress observations for the five predictors (i.e., solid volumetric percent concentration and the percentage of clay, sand, silt, and gravel). Fig. 4 presents the consistency (or plastic viscosity) observations for the five same predictors.

Figure 3 – Yield Stress training set with its five predictors: solid volumetric concentration and soil percent content of clay, gravel sand, and silt.

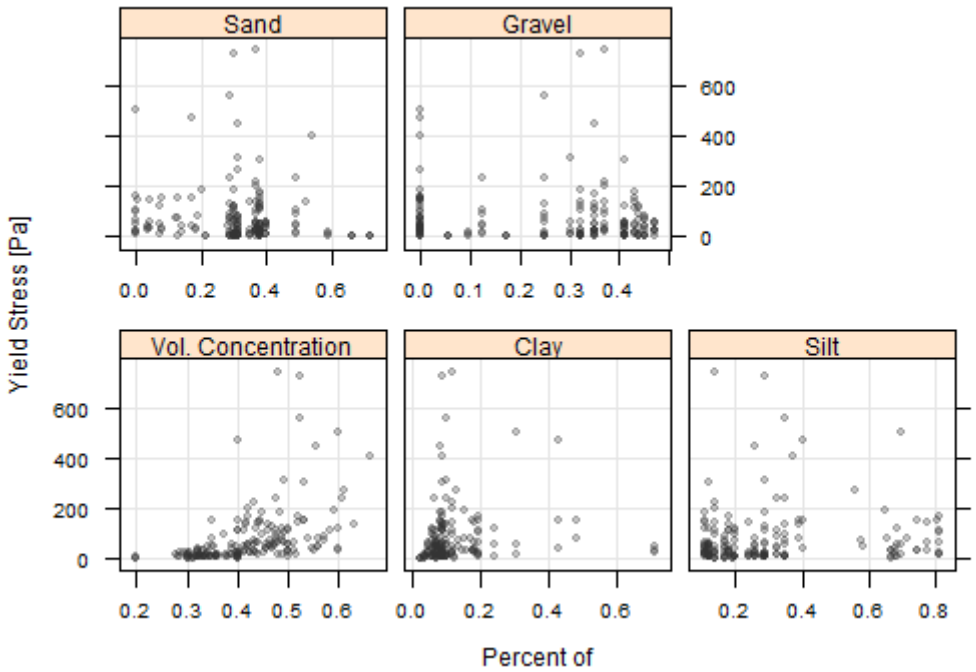
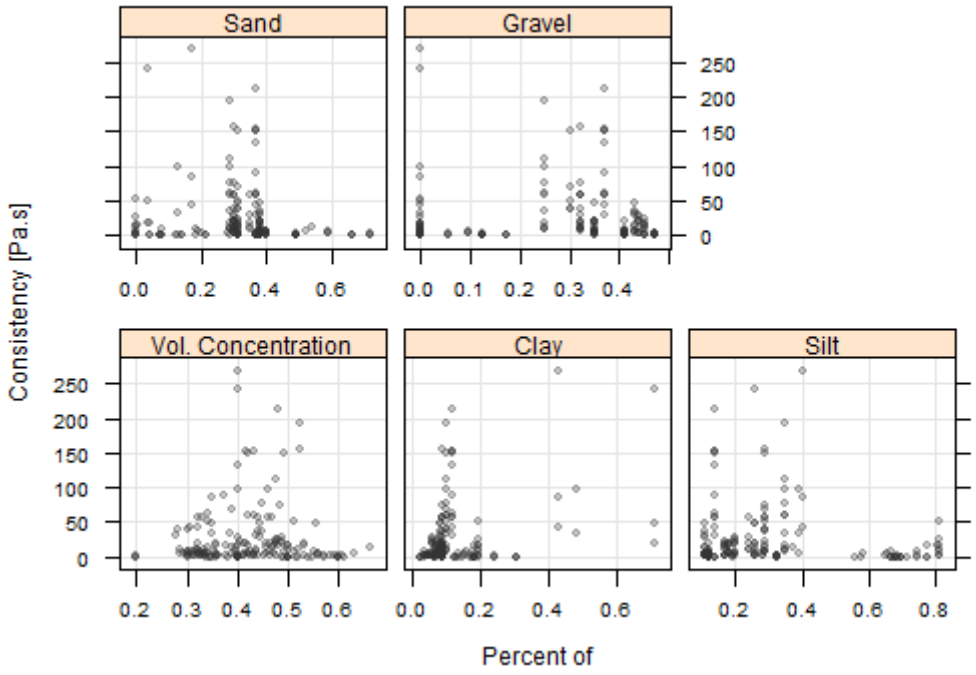


Figure 4 – Consistency (plastic viscosity) training set with its five predictors: solid volumetric concentration and soil percent content of clay, gravel sand, and silt.



The data set obtained from previous articles shown in Fig. 3 and Fig. 4 shows a very heterogeneous collection of data. It contains several spaces with sparse amounts of representative data points which can be misleading for neural network models. The percent of clay particles in the soil, for example, contains almost 95% of the observations in the range of concentrations between 0 and 0.20, with the last 5% of the observations between the concentrations range of 0.20 and 0.70.

All the possible combinations between the studies were tested as candidates of training and test data to address the heterogeneous nature of the dataset. For the yield stress and consistency (or plastic viscosity) neural networks, the study by Malet *et al*. (2003) had to be present in every combination because it was the work with the most data points (117) between the studies. For the flow index, the study by Major *et al*. was the one included since it was the work with most flow index data points (23).

3.2. The neural network models

Three averaged neural network models were attempted to be built in this work. Each model was built to predict one specific rheological parameter, which were the yield stress, the consistency (or plastic viscosity), and the flow index of debris flow fluids using the Herschel-Bulkley rheological model.

This study chose to build single hidden-layered architecture networks which is the simples but most common form of neural networks (VENABLES; RIPLEY, 2002). All the networks have five neurons in the input layer, with each neuron allocating the values of one of the five predictors. The networks have only one neuron in the output layer, which gives the rheological propriety value (i.e., yield stress) for the combination of the five predictors. The number of units in the single hidden layer was one of the architecture parameters chosen to be modified for prediction performance improvement. The networks were built following the algorithm proposed by Venables *et al.* (2002).

A single architecture of a neural network has an optimized value for its weights and bias that better fits the given training data. The best fit for a single architecture may, or may not, produce the best prediction performance for a given data set.

Two network parameters were modified to identify the best architecture for a given data set: the number of units in the hidden layer and regularization parameter γ. The number of units in the hidden layer was tested between 1 and 15 for each network. The regularization parameter (i.e., penalization parameter to prevent model overfitting) was tested with values between 0.1, 0.01, and 0.001.

For each rheological parameter, 45 neural networks, one per combination of these two parameters, was build and evaluated for a prediction performance metric. The neural network with the best performance measure was chosen as the final model.

As data pre-processing, in each model built, both the predictors and outcomes were centered and scaled, leading to the dataset having a mean equal to zero and standard deviation equal to one.

The mean average percent error metric was chosen to evaluate the prediction performance of the yield point and consistency (or plastic viscosity) models. Since the yield point and the consistency (or plastic viscosity) observations have an exponential behavior, this metric is more adequate to evaluate the performance of these models. For the flow index, the root mean square error metric was chosen, as the flow indexes in the dataset are in the same unit.

4. RESULTS

4.1. Data set

A total of 26 combinations of studies for each rheological parameter was experimented. For each combination, 45 neural networks, one for ANN architecture, were fitted and had their prediction performance measured. Tables A.1, A.2, and A.3, in the appendix, show the complete results of the tests.

For the yield stress, the MAPE metric was within 17.73% and 123.96%. For the consistency parameter, the MAPE metric was between 18.68% and 704.52%. For the flow index, the lowest RMSE was 0.233, and the highest 0.724.

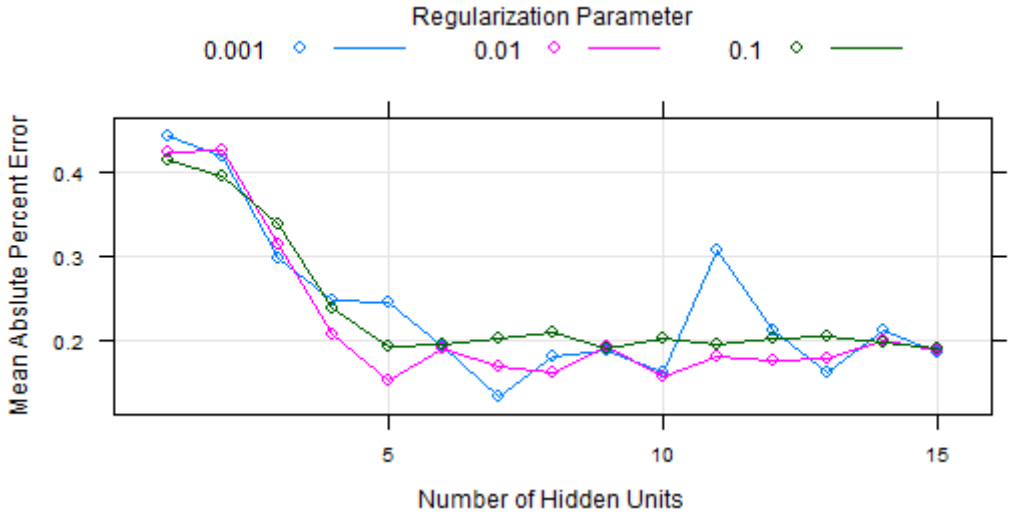
The final choice of studies for the yield stress and consistency (or plastic viscosity) was the ones performed by Malet *et al.* (2003), Kaitna *et al.* (2007), and Santolo *et al.* (2010). For the flow index, the lowest RMSE was reached with Major *et al.* (1992), Malet *et al.* (2003) and Kaitna *et al.* (2007).

The diversity of rheometers used by different authors could explain the performance contrast between studies data set combinations. Malet *et al.* (2003) show that the relative error between different methods for yield stress determination fluctuates around a mean value close to 15%.

4.2. Yield stress (τ_0) neural network model

The single-hidden-layer neural network model with the best performance (i.e., the model with the smallest mean percentage error) was obtained with 7 units in the hidden layer with a regularization parameter of 0.001. For this set of parameters, the performance metrics for the training set was a mean average percent error of 13,29% with a standard deviation of 24.23% and a determination coefficient of 0.980. Figure 5 shows the MAPE metric for all the parameter combinations experimented during the model training process.

Figure 5 – MAPE profiles for the yield stress neural network model.
The best model used λ = 0.001 and 7 hidden units.

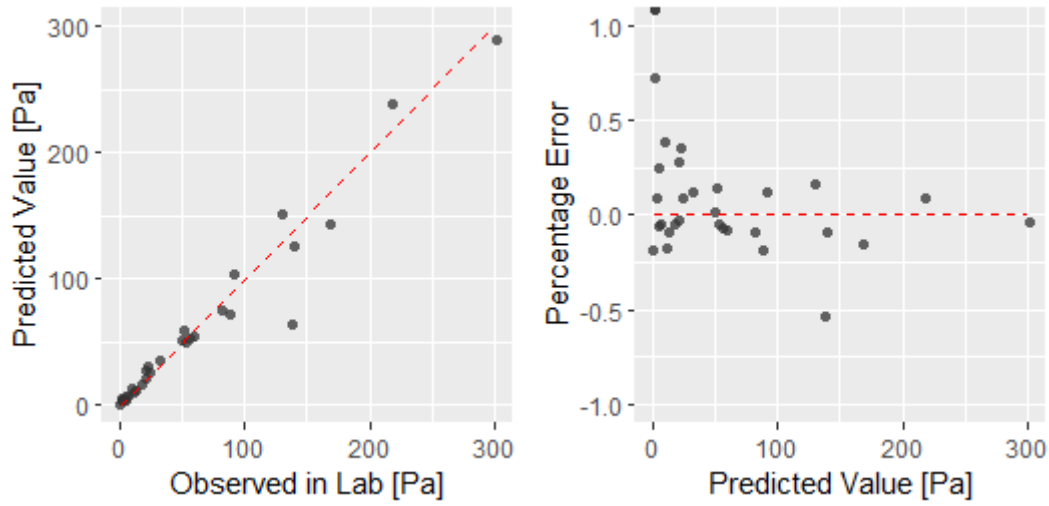


The model's performance improvement by increasing the units in the hidden layer was ceased, or only marginal, after the seven units. The effects of the regularization parameter on the performance were similar for values of 0.01 and 0.1.

By the proposed methodology, the best parameters were 7 hidden units with a regularization parameter of 0.001. Yet, the results indicate that a single-hidden-layer neural network with seven hidden units and regularization parameter of 0.01 and 0.001 were also good choices for ANN regression models fitted on this dataset.

The performance of the model measured by out-of-bag predictions showed a mean average percent error of 22.40% with a standard deviation of 27.85%. Figure 6 shows the out-of-bag predictions (i.e., model predictions for data points not included in the training process) and the percentage error for each prediction made by the model.

Figure 6 – Diagnostic plots of the test set results for the yield stress model.

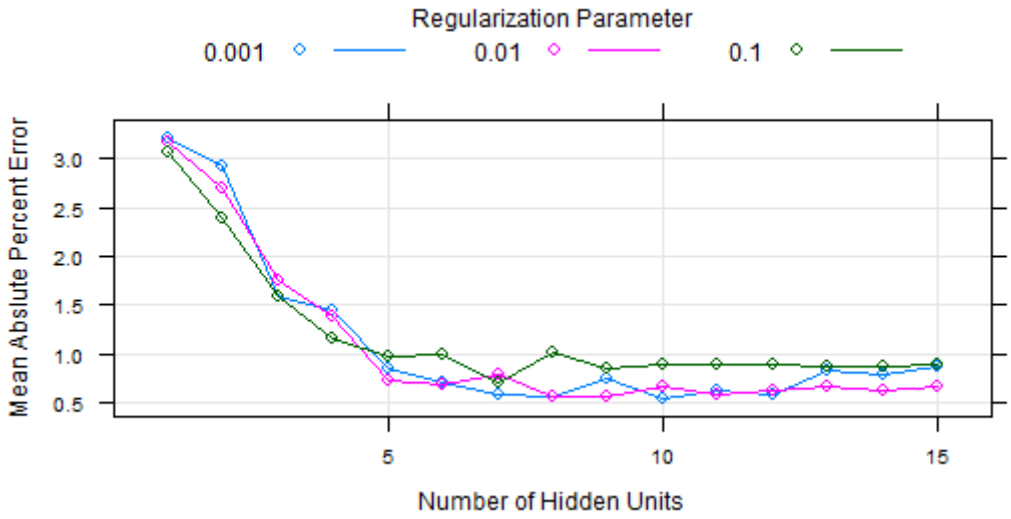


The out-of-bag predictions of the yield stress ANN, in general, showed reasonably good accuracy with low precision. The predictions for higher values of yield points associated with high percentage errors are indicatives of bad precision for these ranges.

4.3. **Consistency** (or plastic viscosity) **neural network model**

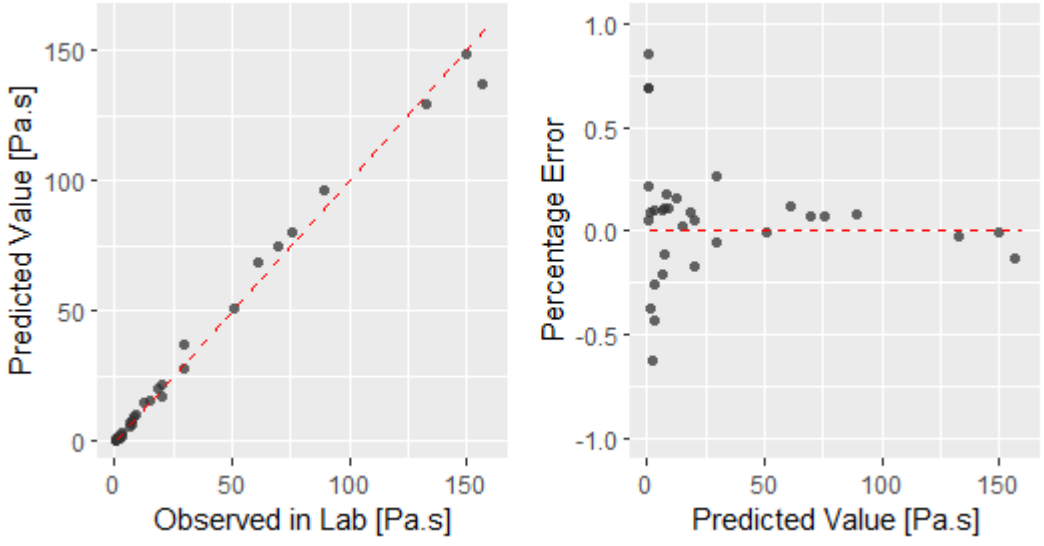
The best single-hidden-layer model for the consistency (or plastic viscosity) property was constituted by 10 units in the hidden layer with a regularization parameter of 0.001. The performance with the training data set for these parameters showed a mean percent error of 54.41% with a standard deviation of 138.80% and a determination coefficient of 0.953. Figure 7 shows the MAPE metric for all the parameter combinations experimented during the model training process.

Figure 7 – MAPE profiles for the consistency (or plastic viscosity) neural network model. The best model used λ = 0.001 and 10 hidden units.



The improvement in the model's performance by adding units in the hidden layer did not increase or was marginal after nine units. The variation in the number of hidden units showed similar behavior to the yield stress ANN.

Figure 8 – Diagnostic plots of the test set results for the consistency (or plastic viscosity) model.



The model's performance for out-of-bag predictions was a mean average percent error of 20.38% with a standard deviation of 22.09% and a coefficient determination of 0.991. Figure 8 shows the predictions the percentage error for each prediction for the test dataset and.

The single-hidden-layer for the consistency (or plastic viscosity) parameter showed general good accuracy and precision results. MAPE metrics are known to show distortions for close to zero prediction values, which can mislead the interpretation about the model precision. The model predict consistency (or plastic viscosity) values in a range between 0 and 150 Pa.s with a mean absolute error of 2.46 Pa.s.

4.2. Flow index (n) neural network model

The lack of data points for flow index in the database led to poor ANNs for this rheological parameter, and it was chosen to not include the flow index ANNs in the present work. The flow index is an exponential parameter of the fluid's constitutive model. Even with the lowest RMSE network, the error is still large to provide useful predictions for this parameter. The lowest RMSE achieved during the training process was 0.233 and the flow indexes in the collected data set have values between 0.255 and 1, with a mean of 0.634.

There are documented works that successfully predict fluid flow index with single-hidden-layer neural networks for specific fluids. Gowida *et al.* (2019) predict flow index for calcium chloride brine-based drilling fluids with a MAPE of 2.4%. One of the reasons for the lack of prediction performance of the flow index ANN model developed can be the wide scope of debris flow materials used in the present work.

5. CONCLUSION

The use of neural networks to parametrize general non-linear functions is well known, with several authors showing that is possible to approximate any continuous functions by these methods (CYBENKO, 1989; FUNAHASHI, 1989; VENABLES; RIPLEY, 2002). Neural networks also were successfully applied to predict the rheology parameters of specific, well behaved, mixtures (ALSABAA *et al.*, 2021; GOWIDA *et al.*, 2019; ROOKI, 2015; WANG; HUANG; CUI, 2003).

In this study, three single-hidden-layered neural network models were built to describe the yield stress, the consistency (or plastic viscosity), and the flow index of mixtures based on the soil granulometric distribution and volumetric concentration in the debris flow mixture. The models obtained by the study showed that the prediction of these rheological properties from the five proposed predictors (i.e., clay, silt, sand, and gravel percent content and the water-soil ratio) with artificial neural networks is feasible but needs further studies.

The lack of data from a single methodology of experimental tests, the great diversity of soil samples, and the low number of data points made it difficult to identify which was the main cause for the lack of prediction performance for the models trained in the collected dataset. The addition of different predictors could also improve the model's performance. One example is the temperature which is known to influence these rheological parameter values.

Further efforts to build a large data set, with soil samples from a wider range of granulometric distributions, and with an established experimental methodology (i.e, fixed vane rotor geometry or temperature) would provide an ideal framework to develop these technologies. A dataset from such controlled experiment would allow to distinguish between the issues mentioned above which one is limiting the models' predictions power.

REFERENCES

ALSABAA, Ahmed *et al*. New correlations for better monitoring the all-oil mud rheology by employing artificial neural networks. **Flow Measurement and Instrumentation**, [*s. l.*], v. 78, n. July 2020, p. 101914, 2021. DOI: https://doi.org/10.1016/j.flowmeasinst.2021.101914

ANCEY, Christophe. Plasticity and geophysical flows: A review. **Journal of Non-Newtonian Fluid Mechanics**, [*s. l.*], v. 142, n. 1–3, p. 4–35, 2007. DOI: https://doi.org/10.1016/j.jnnfm.2006.05.005

ANCEY, Christophe. Role of lubricated contacts in concentrated polydisperse suspensions. **Journal of Rheology**, [*s. l.*], v. 45, n. 6, p. 1421–1439, 2001. DOI: https://doi.org/10.1122/1.1413504

BAGNOLD, R A; A, Proc R Soc Lond. Experiments on a gravity-free dispersion of large solid spheres in a Newtonian fluid under shear. **Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences**, [*s. l.*], v. 225, n. 1160, p. 49–63, 1954. DOI: https://doi.org/10.1098/rspa.1954.0186

BONETTO, Riccardo; LATZKO, Vincent. Machine learning. *In*: COMPUTING IN COMMUNICATION NETWORKS. [*S. l.*]: Elsevier, 2020. p. 135–167. DOI: https://doi.org/10.1016/B978-0-12-820488-7.00021-9

COSTA, J. E. Physical geomorphology of debris flows. **Developments and applications of geomorphology**, [*s. l.*], p. 268–317, 1984. DOI: https://doi.org/10.1007/978-3-642-69759-3_9

COUSSOT, P.; PIAU, J. M. On the behavior of fine mud suspensions. **Rheologica Acta**, [*s. l.*], v. 33, n. 3, p. 175–184, 1994. DOI: https://doi.org/10.1007/BF00437302

COUSSOT, Philippe; PROUST, Sébastien; ANCEY, Christophe. Rheological interpretation of deposits of yield stress fluids. **Journal of Non-Newtonian Fluid Mechanics**, [*s. l.*], v. 66, n. 1, p. 55–70, 1996. DOI: https://doi.org/10.1016/0377-0257(96)01474-7

CYBENKO, G. Approximation by superpositions of a sigmoidal function. **Mathematics of Control, Signals, and Systems**, [*s. l.*], v. 2, n. 4, p. 303–314, 1989. DOI: https://doi.org/10.1007/BF02551274

DEL GAUDIO, Pierdomenico; VENTURA, Guido. Flow Behavior of Clay-Silt to Sand-Silt Water-Rich Suspensions at Low to High Shear Rates: Implications for Slurries, Transitional Flows, and Submarine Debris-Flows. **Acta Geologica Sinica**, [*s. l.*], v. 92, n. 6, p. 2395–2404, 2018. DOI: https://doi.org/10.1111/1755-6724.13735

FUNAHASHI, Ken Ichi. On the approximate realization of continuous mappings by neural networks. **Neural Networks**, [*s. l.*], v. 2, n. 3, p. 183–192, 1989. DOI: https://doi.org/10.1016/0893-6080(89)90003-8

GOWIDA, Ahmed *et al.* Data-Driven Framework to Predict the Rheological Properties of CaCl2 Brine-Based Drill-in Fluid Using Artificial Neural Network. **Energies**, [*s. l.*], v. 12, n. 10, 2019. DOI: https://doi.org/10.3390/en12101880

IVERSON, Richard M. The physics of debris flows. **Reviews of Geophysics**, [*s. l.*], v. 35, n. 3, p. 245–296, 1997. DOI: https://doi.org/10.1029/97RG00426

JAKOB, Matthias; HUNGR, Oldrich. **Debris-flow Hazards and Related Phenomena**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005. (Springer Praxis Books). DOI: https://doi.org/10.1007/b138657

KAITNA, R.; RICKENMANN, D.; SCHATZMANN, M. Experimental study on rheologic behaviour of debris flow material. **Acta Geotechnica**, [*s. l.*], v. 2, n. 2, p. 71–85, 2007. DOI: https://doi.org/10.1007/s11440-007-0026-z

KUHN, Max; JOHNSON, Kjell. **Applied Predictive Modeling with Applications in R**. [*S. l.: s. n.*], 2013. *E-book*.

MAJOR, Jon J.; PIERSON, Thomas C. Debris flow rheology: Experimental analysis of fine-grained slurries. **Water Resources Research**, [*s. l.*], v. 28, n. 3, p. 841–857, 1992. DOI:

https://doi.org/10.1029/91WR02834
MALET, J. P. *et al.* Flow susceptibility of heterogeneous marly formations: Implications for torrent hazard control in the Barcelonnette Basin (Alpes-de-Haute-Provence, France). **International Conference on Debris-Flow Hazards Mitigation: Mechanics, Prediction, and Assessment, Proceedings**, [*s. l.*], v. 1, n. 1996, p. 351–362, 2003.
PARSONS, J. D.; WHIPPLE, K. X.; SIMONI, A. Experimental study of the grain flow, fluid-mud transition in Debris flows. **Journal of Geology**, [*s. l.*], v. 109, n. 4, p. 427–447, 2001. DOI: https://doi.org/10.1086/320798
PELLEGRINO, Anna Maria; SCHIPPA, Leonardo. Macro viscous regime of natural dense granular mixtures. **International Journal of GEOMATE**, [*s. l.*], v. 4, n. 1, p. 482–489, 2013. DOI: https://doi.org/10.21660/2013.7.2123
PELLEGRINO, Anna Maria; SCOTTO DI SANTOLO, Anna; SCHIPPA, Leonardo. An integrated procedure to evaluate rheological parameters to model debris flows. **Engineering Geology**, [*s. l.*], v. 196, p. 88–98, 2015. DOI: https://doi.org/10.1016/j.enggeo.2015.07.002
PHILLIPS, Christopher J.; DAVIES, Timothy R.H. Determining rheological parameters of debris flow material. **Geomorphology**, [*s. l.*], v. 4, n. 2, p. 101–110, 1991. DOI: https://doi.org/10.1016/0169-555X(91)90022-3
RIPLEY, Brian D. **Pattern Recognition and Neural Networks**. [*S. l.*]: Cambridge University Press, 1996. DOI: https://doi.org/10.1017/CBO9780511812651
ROOKI, Reza. Estimation of Pressure Loss of Herschel–Bulkley Drilling Fluids During Horizontal Annulus Using Artificial Neural Network. **Journal of Dispersion Science and Technology**, [*s. l.*], v. 36, n. 2, p. 161–169, 2015. DOI: https://doi.org/10.1080/01932691.2014.904793
SCOTTO DI SANTOLO, A.; PELLEGRINO, A. M.; EVANGELISTA, A. Experimental study on the rheological behaviour of debris flow. **Natural Hazards and Earth System Science**, [*s. l.*], v. 10, n. 12, p. 2507–2514, 2010. DOI: https://doi.org/10.5194/nhess-10-2507-2010
TURNBULL, Barbara; BOWMAN, Elisabeth T.; MCELWAINE, Jim N. Debris flows: Experiments and modelling. **Comptes Rendus Physique**, [*s. l.*], v. 16, n. 1, p. 86–96, 2015. DOI: https://doi.org/10.1016/j.crhy.2014.11.006
VENABLES, W. N.; RIPLEY, B. D. **Modern Applied Statistics with S**. New York, NY: Springer New York, 2002. (Statistics and Computing). DOI: https://doi.org/10.1007/978-0-387-21706-2
WANG, Xie-kang; HUANG, Er; CUI, Peng. Simulation and prediction of debris flow using artificial neural network. **Chinese Geographical Science**, [*s. l.*], v. 13, n. 3, p. 262–266, 2003. DOI: https://doi.org/10.1007/s11769-003-0028-1
WASZCZYSZYN, Zenon. **Neural Networks in the Analysis and Design of Structures**. Vienna: Springer Vienna, 1999. ISSN 1098-6596.(CISM International Centre for Mechanical Sciences).v. 404 DOI: https://doi.org/10.1007/978-3-7091-2484-0
ZEGERS, Gerardo *et al.* Sensitivity and identifiability of rheological parameters in debris flow modeling. **Natural Hazards and Earth System Sciences**, [*s. l.*], v. 20, n. 7, p. 1919–1930, 2020. DOI: https://doi.org/10.5194/nhess-20-1919-2020

APPENDIX

Table A.1 – Yield Stress neural network performance metrics and architecture parameters for different dataset combinations. The indexes A, B, C, D, E, F stands for the studies of Malet *et al.* (2003), Kaitna *et al.* (2007), Del Gaudio *et al.* (2018), Coussot *et al.* (1994), Major *et al.* (1992) and Santolo *et al.* (2010) respectively.

Study Index	*MAPE*	*SDPE*	Hidden Units	Regularization Parameter
ABDE	0.177	0.219	10	0.001
ABCDF	0.221	0.190	12	0.001
ABF	0.224	0.278	**7**	**0.001**
ADEF	0.226	0.256	10	0.001
ADE	0.236	0.287	9	0.001
ABCD	0.243	0.264	15	0.01
ACF	0.246	0.233	8	0.001
ABE	0.248	0.264	9	0.001
ABCF	0.249	0.184	8	0.01
ABD	0.254	0.423	8	0.001
ABEF	0.255	0.305	9	0.001
AEF	0.266	0.252	12	0.001
ABCE	0.267	0.392	9	0.001
ABC	0.271	0.285	15	0.01
ACE	0.288	0.257	9	0.001
ACD	0.327	0.688	8	0.001
ACDF	0.342	0.458	11	0.01
ACEF	0.351	0.441	10	0.001
ABDF	0.370	0.503	11	0.001
ABCDE	0.391	0.853	10	0.001
ABDEF	0.405	0.675	8	0.001
ADF	0.479	0.992	8	0.001
ACDE	0.485	1.046	9	0.001
ACDEF	0.492	1.111	10	0.001
ABCEF	0.667	1.280	8	0.001
ABCDEF	1.240	5.175	12	0.001

Table A.2 – Consistency (or plastic viscosity) neural network performance metrics and architecture parameters for different dataset combinations. The indexes A, B, C, D, E, F stands for the studies of Malet *et al.* (2003), Kaitna *et al.* (2007), Del Gaudio *et al.* (2018), Coussot *et al.* (1994), Major *et al.* (1992) and Santolo *et al.* (2010) respectively.

Study Index	*MAPE*	*SDPE*	Hidden Units	Regularization Parameter
ABC	0.187	0.219	10	0.001
ADF	0.200	0.199	11	0.001
ABF	**0.204**	**0.221**	**10**	**0.001**
ABDF	0.217	0.204	9	0.001
ABD	0.240	0.223	15	0.001
ACF	0.256	0.483	10	0.001
ABCD	0.558	1.591	11	0.001
ABE	0.628	1.471	13	0.001
ABCDF	0.826	3.405	15	0.001
ACDF	1.078	4.888	14	0.001
ACDE	1.200	2.418	9	0.001
ABCE	1.543	3.125	5	0.1
ABDE	1.567	4.971	9	0.001
ACDEF	1.680	4.683	15	0.01
AEF	1.900	7.590	8	0.01
ACD	1.949	9.999	13	0.001
ABCF	2.416	12.296	15	0.001
ADE	2.733	7.497	8	0.01
ABEF	4.876	26.417	11	0.01
ABCEF	7.691	20.942	15	0.01
ABDEF	14.271	76.997	11	0.001
ABCDE	27.903	106.238	15	0.001
ADEF	34.934	150.720	13	0.001
ACE	39.464	217.166	12	0.01
ACEF	70.452	280.846	13	0.01

Table A.3 – Flow index neural network performance metrics and architecture parameters for different dataset combinations. The indexes A, B, C, D, E, F stands for the studies of Malet *et al.* (2003), Kaitna *et al.* (2007), Del Gaudio *et al.* (2018), Coussot *et al.* (1994), Major *et al.* (1992) and Santolo *et al.* (2010) respectively.

Study Index	*RMSE*	R^2	Hidden Units	Regularization Parameter
EAB	0.233	0.588	1	0.1
EBCD	0.291	0.793	5	0.01
EACDF	0.335	0.534	14	0.1
EDF	0.375	0.249	1	0.1
EABC	0.408	0.664	7	0.01
EABDF	0.456	0.324	1	0.1
EAC	0.478	0.337	6	0.01
EBCDF	0.494	0.330	4	0.1
EABCF	0.501	0.263	7	0.01
EBCF	0.502	0.208	14	0.1
EACD	0.516	0.211	13	0.1
ECDF	0.523	0.280	15	0.1
EABCD	0.537	0.277	15	0.01
EBD	0.567	0.164	1	0.1
EBDF	0.595	0.169	1	0.1
EACF	0.617	0.296	1	0.1
EABD	0.626	0.140	1	0.1
EADF	0.626	0.001	10	0.1
EABF	0.627	0.077	1	0.1
EAF	0.640	0.183	1	0.1
ABCDEF	0.656	0.204	7	0.01
EAD	0.658	0.224	1	0.1
ECF	0.675	0.165	10	0.1
ECD	0.707	0.093	14	0.1
EBF	0.708	0.106	5	0.01