

Prediction of the mechanical properties of granites under tension using DM techniques

Francisco F. Martins*, Graça Vasconcelos^a and Tiago Miranda^b

ISISE, Department of Civil Engineering, University of Minho, Guimarães, Portugal

(Received June 22, 2017, Revised September 25, 2017, Accepted November 8, 2017)

Abstract. The estimation of the strength and other mechanical parameters characterizing the tensile behavior of granites can play an important role in civil engineering tasks such as design, construction, rehabilitation and repair of existing structures. The purpose of this paper is to apply data mining techniques, such as multiple regression (MR), artificial neural networks (ANN) and support vector machines (SVM) to estimate the mechanical properties of granites. In a first phase, the mechanical parameters defining the complete tensile behavior are estimated based on the tensile strength. In a second phase, the estimation of the mechanical properties is carried out from different combination of the physical properties (ultrasonic pulse velocity, porosity and density). It was observed that the estimation of the mechanical properties can be optimized by combining different physical properties. Besides, it was seen that artificial neural networks and support vector machines performed better than multiple regression model.

Keywords: granite; tensile behavior; mechanical properties; physical properties; data mining techniques

1. Introduction

The knowledge of the tensile behavior of rocks is important in the design, construction and rehabilitation of geotechnical works and ancient stone buildings. Underground openings, dam foundations and slope stability works should be highlighted in the geotechnical field. Regarding the stone buildings are noteworthy the historic buildings that require periodic interventions for its preservation.

The tensile strength of rock materials can be determined through laboratory tests or using empirical equations published in literature. The laboratory tests include the direct and the indirect tensile tests. The latter include, for example, Brazilian test, ring test, hoop test, bending test, etc.

The direct determination of rock behavior requires sophisticated apparatus and is time consuming. Therefore, the use of non-destructive tests, index tests and other simple tests for obtaining physical properties (generally, porosity and density) has been increasing to obtain parameters that are used in models where they are correlated with the main rock properties (Marques *et al.* 2012). The most commonly used non-destructive tests are the Schmidt hammer and the ultrasonic velocity tests. The index test that is generally used is the point load test. In the past, the correlations

between the main rock properties and the other parameters were obtained by simple or multiple regressions (MR) (Irfan and Dearman 1978, Christaras *et al.* 1994, Begonha and Sequeira Braga 2002, Sharma and Singh 2007, Kiliç and Teymen 2008). However, nowadays there are more sophisticated methods based on soft computing techniques such as data mining (neural networks, support vector machines, etc.) and genetic programming that can model more complex and nonlinear relationships. These techniques have been used to develop models to predict rock properties using mainly index parameters, physical parameters and petrographical properties. In published literature one can find models to predict UCS based on neural networks (Meulenkamp and Grima 1999, Martins *et al.* 2012, Yesiloglu-Gultekin *et al.* 2013), genetic programming (Karakus 2001, Alemdag *et al.* 2016) and fuzzy inference system (Mishra and Basu 2013, Yesiloglu-Gultekin *et al.* 2013). ANN and fuzzy inference systems (Gokceoglu *et al.* 2009, Dagdelenler *et al.* 2011) have also been used to develop weathering degree prediction models. Elasticity modulus prediction models based on genetic programming (Karakus 2001) and based on ANN and SVM (Martins *et al.* 2012) were also developed.

Some prediction models of tensile strength based on soft computing can also be found in the literature (Baykasoğlu *et al.* 2008, Ceryan *et al.* 2013, Çanakçı and Pala 2007, Çanakçı *et al.* 2009, Singh *et al.* 2017, Gurocak *et al.* 2012).

Çanakçı and Pala (2007) proposed a neural network based formula for the determination of tensile strength of basalt in terms of ultrasonic pulse velocity, dry density and water absorption parameters. They concluded that the NN-based formula is practical in predicting the tensile strength of basalt. Baykasoğlu *et al.* (2008) applied a set of genetic programming techniques which are known as multi expression programming, gene expression programming

*Corresponding author, Professor

E-mail: ffm@civil.uminho.pt

^aProfessor

E-mail: graca@civil.uminho.pt

^bProfessor

E-mail: tmiranda@civil.uminho.pt

and linear genetic programming to the uniaxial compressive strength and tensile strength prediction of chalky and clayey soft limestone. Ultrasonic pulse velocity, water absorption, dry density, saturated density and bulk density were used as input features. The results determined in their study indicate that genetic programming techniques are able to provide good prediction equations for strength prediction. Çanakci *et al.* (2009) used two soft computing approaches, which are known as artificial neural networks and Gene Expression Programming (GEP) in strength prediction of basalts. The parameters, ultrasound pulse velocity, water absorption, dry density, saturated density, and bulk density were used to predict uniaxial compressive strength and tensile strength of basalts. It was found out that neural networks give far better results than GEP and regression analysis. Gurocak *et al.* (2012) developed two regression models, and two ANN models, namely, radial basis function network (RBFN) and multi-layer perceptron network (MLPN), to correlate tensile strength with point load index, Schmidt rebound number and unit weight of rocks. They used samples of different rock types, namely igneous, metamorphic and sedimentary. They concluded that the MLPN model was the best model and suggested an equation on ANN model to estimate the tensile strength of rocks. Ceryan *et al.* (2013) examined the capability of two SVM algorithms (Least Square-SVM and SVM) for the prediction of tensile strength of carbonate rocks and compared its performance with ANN and linear regression models. Total porosity, sonic velocity, slake durability index and aggregate impact value were used as input parameters. They concluded that both SVM and ANN methods are successful tools for prediction of tensile strength and have better predictive capacity than linear regression model. Furthermore, LS-SVM makes the running time considerably faster with the higher accuracy. Singh *et al.* (2017) determined uniaxial compressive strength, tensile strength, point load index and Young's modulus of basaltic rocks from physical parameters like density, porosity and compressional wave velocity using multiple variable regression analysis and adaptive neuro-fuzzy inference system (ANFIS). They developed five ANFIS models and selected one of them as the best suitable model with high predictability.

All of these models are based on Brazilian tensile strength tests. This study was not based on Brazilian tensile tests but on direct tension tests.

The main aim of this paper consists of using data mining techniques for the prediction of the mechanical properties defining the mechanical behavior under tension through the tensile strength, which is a parameter more easily obtained. Besides, the prediction of the tensile strength and the other mechanical properties based on physical properties, like ultrasonic pulse velocity, porosity and density is also analyzed.

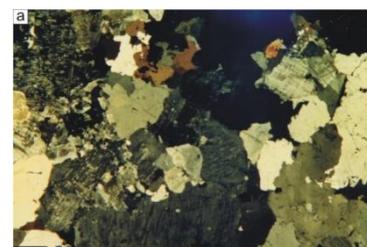
2. Materials

2.1 A brief description of granites

Granite is one of the most used stones in the

Table 1 Brief petrologic description of granites

| Granite designation | Description | Mean length (mm) | Grain size range (mm) | Loading directions |
|---------------------|--|------------------|-----------------------|--|
| BA | Fine to medium-grained porphyritic biotite granite | 0.5-0.6 | 0.2-6.5 | Parallel to the rift plane |
| GA, GA* | Fine to medium-grained, with porphyritic trend, two mica granite | 0.5-0.6 | 0.3-7.5 | Parallel to the rift plane |
| RM | Medium-grained biotite granite | 1.3-2.3 | 0.4-13.5 | Parallel to the rift plane |
| MC | Coarse-grained porphyritic biotite granite | 1.6-2.4 | 0.3-16.5 | Parallel to the rift plane |
| AF | Fine to medium-grained two-mica granite | 0.5-0.6 | 0.1-4.0 | Parallel (L) and perpendicular (P) to the foliation plane |
| MDB MDB* | Medium-grained two-mica granite | 0.7-0.9 | 0.3-14.5 | Parallel and perpendicular to the foliation plane |
| PTA PTA* | Fine to medium-grained two-mica granite | 0.7-0.8 | 0.3-12.0 | Parallel and perpendicular to the foliation and rift plane |
| PLA PLA* | Medium to coarse-grained porphyritic biotite granite | 0.5-1.1 | 0.2-14.0 | Parallel and perpendicular to the rift plane |



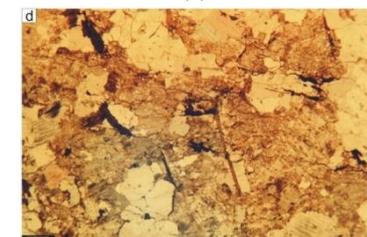
(a)



(b)



(c)



(d)

Fig. 1 Grain size and weathering features of studied granites observed with the stereomicroscope, (a) RM granite, (b) PTA granite, (c) GA*, besides grain size it could be observed weathering features in plagioclase sections and (d) MDB*, clouding of feldspar sections and iron compounds in fissures can be observed. White bar at the lower-left corner of each photograph represents 1 mm; total magnification: 5X; (a), (b) and (c), crossed polars, (d) polarized light

construction of vernacular and historical ancient buildings, ornamental elements and movable stone heritage artifacts (e.g., statues, altar pieces, benches, etc.) in Portugal, particularly in the northern regions, either in monumental or vernacular architecture. A wide range of granitic rocks can be seen in masonry buildings and artifacts, with different petrographic properties related to the grain size and internal texture and with distinct levels of weathering resulting from aging degradation processes through years. The granitic types considered in this study were mostly collected from the Northern region of Portugal, being selected according to mineralogical, textural and structural characteristics. In addition to these criteria, presence of planar anisotropies and weathering condition were also considered. In some of the studied granitic facies, samples of two different weathering degrees were collected in order to extend the range of values of the possible relations between properties. Weathering was assessed, as is frequently done in Engineering Geology studies, see the classic classifications of ISRM (1981) and GSL(1995), on the field (quarries), based on macroscopic evidences, namely discolouration (yellowing in the case of granites) and reduction of strength. The weathered types, identified by an (*) in Table 1, correspond to the initial stages of weathering and can be classified according to the classification of weathering of intact rock material proposed in GSL(1995). They are mostly between grades II and III (in all situations samples were extracted from blocks of rocks that could be used to extract building stone). Studies with the petrographic microscope reveal clouding of feldspar sections and the presence of iron compounds in fissures (responsible for discolouration), see Fig. 1. Based on visual assessments, tests specimens were considered homogeneous regarding weathering degree. With respect to planar anisotropies, two possibilities were considered to be relevant for further analysis of the variation of the mechanical properties (Vasconcelos *et al.* 2008a): (a) natural orientation planes of granitic rocks or preferred orientation of minerals (foliation); (b) three orthogonal planes identified as rock splitting planes (quarry planes), defined as planes of preferred rupture, being the rift plane the one corresponding to the easiest splitting in the quarry. To evaluate the grain size of the granitic types the mean length of sections intercepted by a single circle was measured following the principles of the Hilliard single-circle procedure described in ASTM E112-88 [28]. Four circles were studied for each granitic facies (the interval of values for the mean length obtained with the four circles is reported in Table 1). Individualized sections in the less weathered granitic types were considered. Grain size range, considering the maximum length of the smallest grain intercepted by the circular scanline and the maximum length of the biggest section present in the studied thin sections, is also reported in Table 1.

2.2 An overview of the tensile behavior of granitic rocks

Based on an extensive experimental characterization through direct tensile tests (Vasconcelos *et al.* 2008a) it was possible to observe that the shape of the complete stress-

displacement diagrams is essentially composed by four stages, similarly to what has been found for other quasi-brittle materials, such as concrete (Prado and Van Mier 2003), see Fig. 2: (1) the linear stretch of the stress-displacement diagram associated with the elastic behavior of the material. The linear behavior is characterized by the value of initial stiffness, k_0 ; (2) the stable microcracking process reflected by a nonlinear stretch before the peak stress is reached. This phase is characterized by the nonlinear pre-peak displacement, δ_p , calculated as the subtraction of the linear displacement at peak stress to the total displacement at peak stress, δ_{π} ; (3) the macrocracking propagation observed through the increasing on the crack opening, which is visible with naked eye, being also associated with the steep negative stretch in the softening branch with a slope that depends on the type of granite; (4) the stress-transfer mechanism, due to the bridging effect, which is responsible for the long tail of the softening branch of the complete stress-displacement diagram.

According to Tang *et al.* (2007), in heterogeneous rocks, the failure is not abrupt and there exist a transition in which the tensile strength decreases at much slower rate until it reaches its residual strength. The post-peak behavior (softening stress-displacement branch) is characterized by the tensile strength, f_t , fracture energy, G_f , and the critical crack opening, w_c , defined as the intersection of the crack opening axis with a linear fitting to the softening branch below 0.15 ft. In addition to latter parameter, the ductility index parameter, d_u , is also considered to characterize the post-peak behavior being defined as the ratio G_f/f_t (Chiaia *et al.* (1998). This parameter represents the fracture energy normalized by the tensile strength and allows the characterization of the brittleness of the granites.

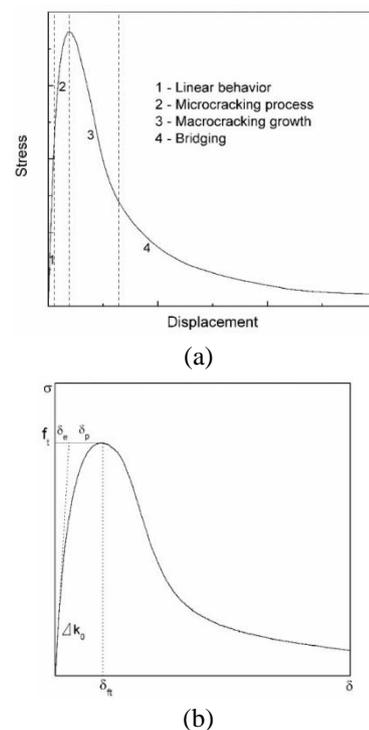


Fig. 2 Tensile behavior of granites, (a) global behavior and (b) characteristic mechanical properties

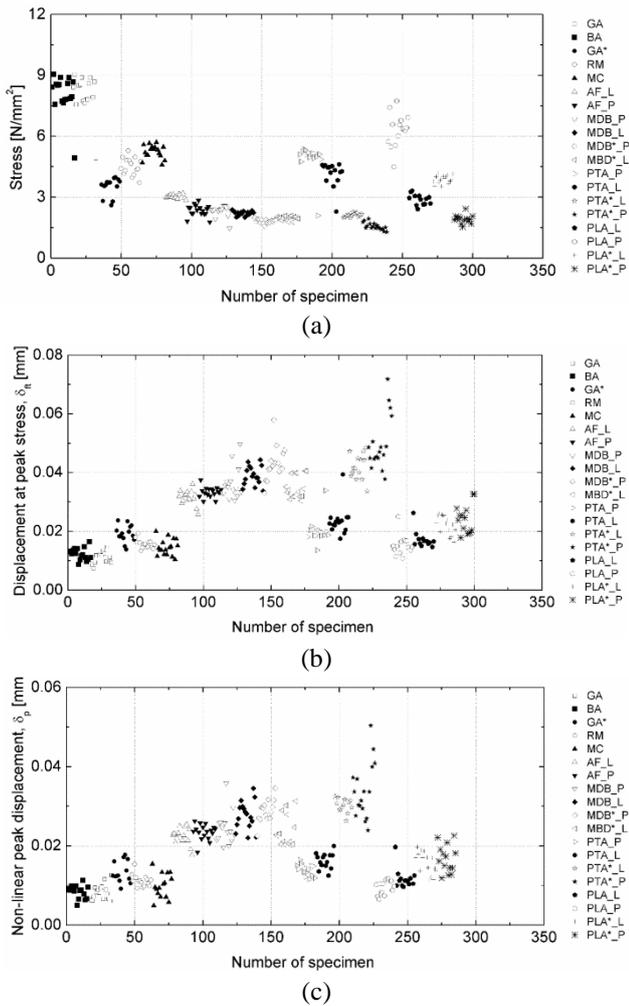


Fig. 3 Experimental data, (a) tensile strength, (b) displacement at peak stress and (c) nonlinear displacement at peak stress

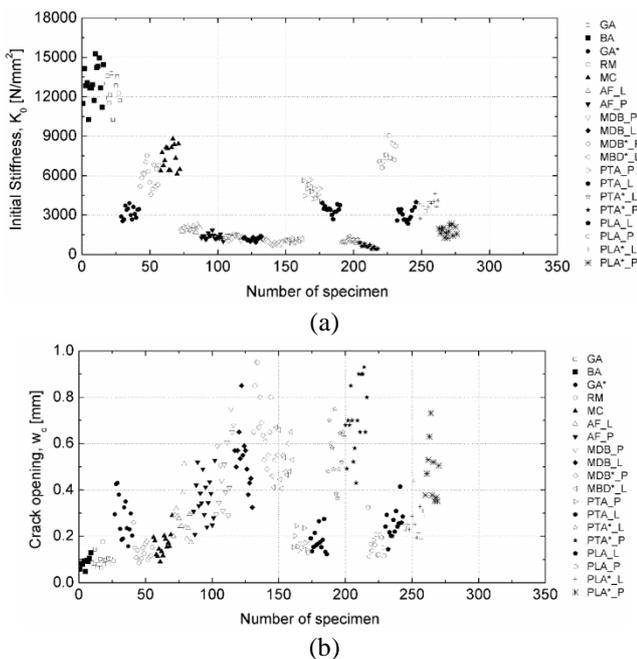


Fig. 4 Experimental data, (a) initial stiffness, (b) critical crack opening and (c) ductility index

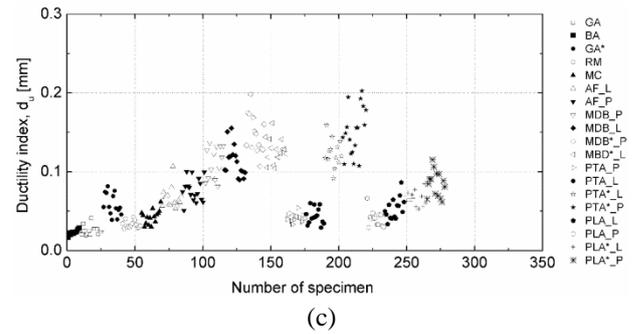


Fig. 4 Continued

The data of the key tensile properties are presented graphically in Fig. 3 and Fig. 4. It is seen that granites exhibit a large range of values for the stiffness and fracture parameters as a result of different characteristic shapes of the stress-displacement diagrams, which can be explained essentially by microstructural aspects, weathering and planar anisotropy, as discussed in detail by Vasconcelos *et al.* (2008a).

2.3 Data mining techniques

Since the 1960s that the data mining techniques have come to constitute a branch of applied artificial intelligence (Liao *et al.* 2012). Data mining consists in the application of methods and techniques in large databases to search trends or patterns in order to discover valuable information. The data mining techniques used in this study are the multiple regression (MR), artificial neural networks (ANN) and support vector machines (SVM). These techniques have their advantages and disadvantages. The main advantage of MR is their simplicity. They are quite easily developed and provide a model that is readable and understandable. However, they perform poorly in highly non-linear problems or when data is noisy. Both ANN and SVM have shown high learning capabilities even when working with complex data and are particularly useful for problems that do not have an analytical formulation or where explicit knowledge does not exist. SVM are a very specific class of algorithms, which are characterized by the use of kernels, absence of local minima during the learning phase, sparseness of the solution and capacity control obtained by acting on the margin, or on the number of support vectors. When compared with other types of base learners, such as the well-known multilayer perceptron (also known as backpropagation neural networks) used by ANN, SVM represents a significant enhancement in functionality. The supremacy of SVM lies in their use of nonlinear kernel functions that implicitly map inputs into high dimensional feature spaces. ANN are robust when dealing with noisy data. However, the initial values of the weights may affect the results accuracy.

These techniques were previously described by the authors (Martins *et al.* 2012, Martins and Miranda 2012). A brief description of these techniques will be presented below. Further details can be found in many publications. Aleksander and Morton (1990), Ilonen *et al.* (2003), Downing (2015) and Souza and Soares (2016) for ANN;

Vapnik (1998), Cristianini and Shawe-Taylor (2000), Dibiki *et al.* (2001), Ben-Hur and Weston (2010) and Liang *et al.* (2011) for SVM.

The traditional MR uses an analysis similar to the simple regression, but rather than describing the relationship between two variables, describes the relationship between several independent variables and the dependent variable.

The ANN technique is based on the architecture of the human brain where artificial neurons are linked each other according to a given architecture. The communication among neurons is performed by signals through links. A weight, $w_{i,j}$ (i and j are neurons or nodes) is associated to each link and an activation function that introduces a non-linear component is associated to each neuron. This study used a logistic activation function f given by $1/(1+e^{-x})$ and the following general equation (Hastie *et al.* 2001)

$$\hat{y} = w_{o,0} + \sum_{j=1}^{o-1} f\left(\sum_{i=1}^I x_i w_{j,i} + w_{j,0}\right) w_{o,i} \quad (1)$$

where x_i are the input parameters or nodes, I is the number of input parameters and o is the output parameter.

The multilayer perceptron (feed forward network) architecture (Haykin 1999) with one hidden layer of HN hidden nodes was adopted in this study. The grid search of the number of hidden nodes HN was $\{0,2,4,6,8,10,12,14,16,18,20\}$.

The SVM technique was initially developed to classification problems by Cortes and Vapnik (1995). This method uses a non-linear mapping to transform the input data into a multidimensional feature space. After this transformation the SVM finds the best hyperplane inside the feature space. The non-linear mapping depends on a kernel function $k(x,x')$. This work uses the following kernel function

$$k(x, x') = e^{\left(-\gamma \cdot \|x-x'\|^2\right)}, \quad \gamma > 0 \quad (2)$$

The performance of the regression is affected by the kernel parameter, γ , a penalty parameter, C, and the width of the ϵ -insensitive zone. In order to limit searching space, C and ϵ were set using heuristics proposed by Cherkassy and Ma (2004): $C=3$ and $\epsilon = \hat{\sigma} / \sqrt{N}$, where $\hat{\sigma} = 1.5 \times \sum_{i=1}^N (y_i - \hat{y}_i)$, \hat{y}_i is the value predicted by a 3-nearest neighbor algorithm and N the number of examples. Under this setup, the search space was limited to the input values of γ .

Therefore, the search space was limited to the input values of γ which in this study were $\{2^{-15}, 2^{-13}, 2^{-11}, 2^{-9}, 2^{-7}, 2^{-6}, 2^{-5}, 2^{-4}, 2^{-3}, 2^{-2}, 2^{-1}, 2^0, 2^1, 2^2, 2^3\}$.

The forecasting capacity of the data mining techniques was tested using only part of the whole dataset. This part, corresponding to two-thirds of the whole dataset, was used in an evaluation scheme using a 10-fold cross validation. On this scheme, nine subsets were used to fit the model and the remaining subset was used to test the model. This process was repeated until all the subsets have been tested. Ten runs of this process were carried on this study. After

this process, the model was fitted with all the training dataset. Finally, the fitted model was tested using the testing dataset corresponding to one-third of the whole dataset.

The performance of the models was evaluated along the training process using three metrics: coefficient of determination R^2 , root mean square error (RMSE) (Eq. (3)) and mean absolute deviation (MAD) (Eq. (4)).

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (3)$$

$$MAD = \frac{1}{N} \times \sum_{i=1}^N |y_i - \hat{y}_i| \quad (4)$$

where y_i is the measured value, \hat{y}_i is the predicted value and N is the number of samples. R^2 equal 1 corresponds to an excellent performance whereas R^2 equal 0 corresponds to a very bad performance. The lower RMSE and MAD the better is the performance of the model.

The computing process was performed in the R environment (R Development Core Team) using the RMiner library developed by Cortez (2010) that facilitates the use of DM algorithms.

In the next chapters the DM techniques will be applied to evaluate the parameters defining the tensile-displacement diagrams and to predict the mechanical properties in function of the physical properties. It must be stressed that all the figures shown in the next chapters and corresponding to DM techniques were based on the testing dataset.

3. Evaluation of the parameters defining the tensile-displacement diagrams

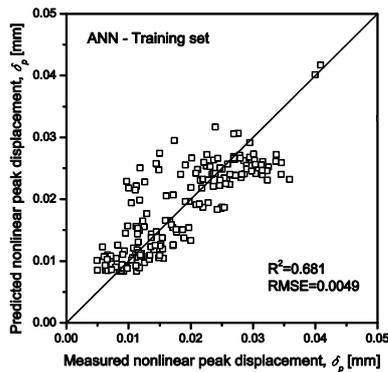
In this section an analysis is carried on the possibility of obtaining the parameters characterizing the complete tensile behavior of granites through the knowledge of the tensile resistance by considering distinct data mining techniques, namely the MR, ANN and SVM. This analysis represents a step forwarding in relation to the simple correlations presented by Vasconcelos *et al.* (2008a), given that besides the MR models, also ANN and SVM are used. For this, four relationships intend to analyze the feasibility of using the tensile strength for prediction of the non-linear peak displacement, δ_p , the initial stiffness, k_0 , the ultimate crack opening, w_c , and the normalized tensile fracture energy, d_u . Additionally, the prediction of the nonlinear peak displacement, δ_p , through the displacement at peak stress, δ_{fp} , is also assessed. Notice that these relations encompasses the principle that the tensile strength is a parameter easier to be obtained, when compared to the other mechanical properties defining the complete tensile behavior. The number of tensile tests considered for cases 2 to 4 is 240, whereas for cases 1 and 5 about 285 experimental results are considered. A general statistical overview of the analysed cases is presented in Table 2, which also presents the mean values and the coefficient of variation of the input and output parameters. The coefficients of variation of the tensile resistance and the other mechanical properties is high, which is justified by the great variation of the petrographic characteristics, weathering and planar anisotropy features. As discussed in detail in previous

Table 2 Range of variation of the input and output parameters and the corresponding average values

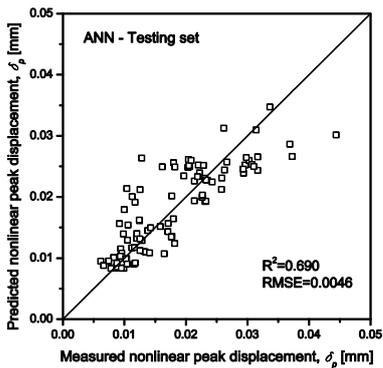
| Case | Input | Min. | Mean | Max. | CV (%) | Out. | Min. | Mean | Max. | CV (%) |
|------|------------|-------|-------|-------|--------|------------|-------|--------|---------|--------|
| 1 | f_i | 1.279 | 3.652 | 9.040 | 51.1 | δ_p | 0.005 | 0.018 | 0.050 | 46.5 |
| 2 | f_i | 1.279 | 3.375 | 9.040 | 47.8 | k_0 | 401.5 | 3413.0 | 15246.0 | 93.3 |
| 3 | f_i | 1.279 | 3.375 | 9.040 | 47.8 | d_u | 0.020 | 0.079 | 0.198 | 52.8 |
| 4 | f_i | 1.279 | 3.375 | 9.040 | 47.8 | w_c | 0.048 | 0.355 | 0.950 | 59.2 |
| 5 | δ_p | 0.009 | 0.026 | 0.072 | 46.1 | δ_p | 0.005 | 0.019 | 0.050 | 46.6 |

Table 3 Coefficients of correlation and root mean square errors obtained for all the cases according to different models

| Case | Coefficient of correlation | | | Root mean square error | | |
|------|----------------------------|-------|-------|------------------------|---------|---------|
| | ANN | SVM | MR | ANN | SVM | MR |
| 1 | 0.802 | 0.799 | 0.756 | 0.0052 | 0.0053 | 0.0569 |
| 2 | 0.916 | 0.940 | 0.922 | 1261.80 | 1117.50 | 1209.69 |
| 3 | 0.846 | 0.845 | 0.773 | 0.0216 | 0.0217 | 0.0257 |
| 4 | 0.777 | 0.820 | 0.741 | 0.1415 | 0.1166 | 0.1366 |
| 5 | 0.979 | 0.980 | 0.980 | 0.0017 | 0.0017 | 0.0017 |



(a) Training set

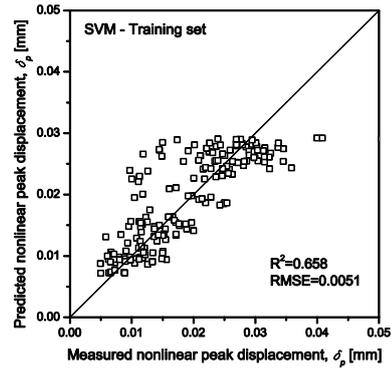


(b) Testing set

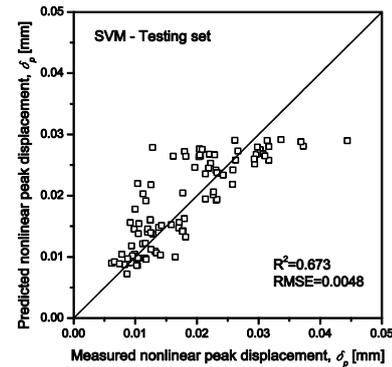
Fig. 5 Prediction of the nonlinear peak displacement, δ_p , from the input parameter, f_i using the ANN model

publications (Vasconcelos *et al.* 2008a), the different mineralogical, textural and structural characteristics of the granites control their mechanical properties.

Table 3 shows the coefficient of correlation of Pearson (R) and the root mean square error obtained in the training phase for the three algorithms used: ANN, SVM and MR.



(a) Training set



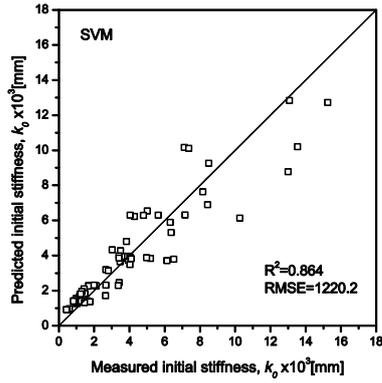
(b) Testing set

Fig. 6 Prediction of the nonlinear peak displacement, δ_p , from the input parameter, f_i using the SVM model

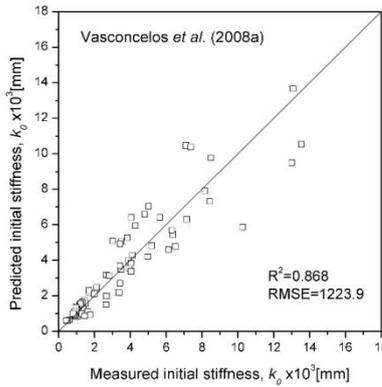
From the analysis of results, it is possible to conclude that, for cases 1 and 3, the ANN and SVM models have similar performances, even if ANN model seems to have a slightly better predictive capacity. SVM model presents the best predictive performance for cases 2 and 4. The results obtained for case 5 indicate similar performances for all the models.

The efficiency of the models to predict the mechanical properties of granites under tension can be also analysed through Fig. 5 to Fig. 9, where the predicted values versus the measured values for a certain variable are displayed. Figs. 5 and 6 present the relationships between predicted and measured nonlinear peak displacement, δ_p , using the ANN and SVM models with the training and testing dataset. It is seen in Figs. 5 and 6 that the results obtained with the training set are better than those obtained with the testing set. As in all the examples presented in this paper the results obtained with the training data were better than the results obtained with the testing data, from Fig. 6 only the results obtained with the test data will be presented. The presentation of all the figures with results of training data for all situations would make the article too long and tedious.

The representation of the predicted variables takes into account only one third of the whole data, which represent the test data. It must be emphasized that this data was not used in the training process. It is possible to see that there is an important dispersion of the values and that ANN model confirms a better performance, which is confirmed by the higher values of R^2 and a lower RMSE.

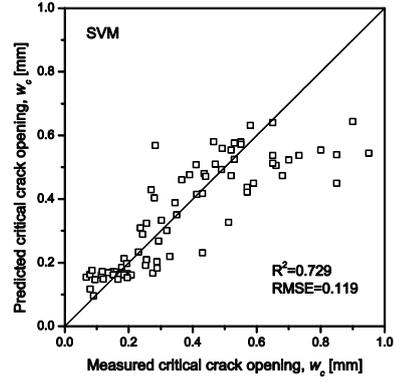


(a) SVM model

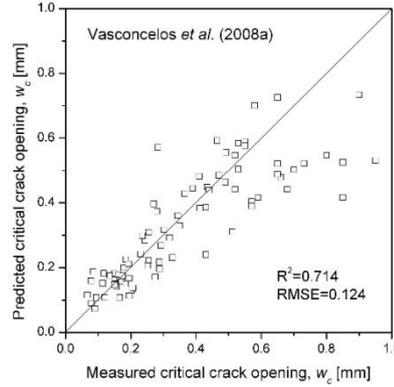


(b) Single regression model (Vasconcelos *et al.* 2008a)

Fig. 7 Prediction of the initial stiffness, k_0 , from the input parameter, f_i

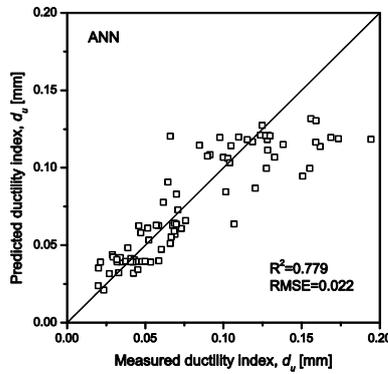


(a) SVM model

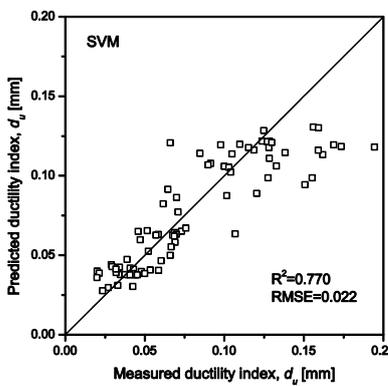


(b) Single regression model (Vasconcelos *et al.* 2008a)

Fig. 9 Prediction of the ultimate crack opening, w_c , from the input parameter f_i

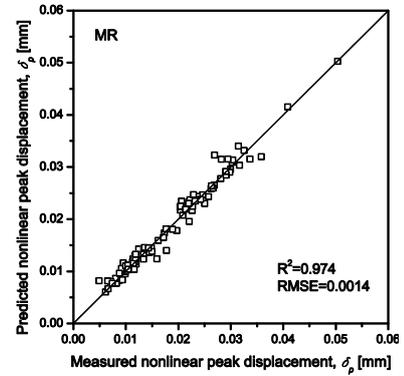


(a) ANN model

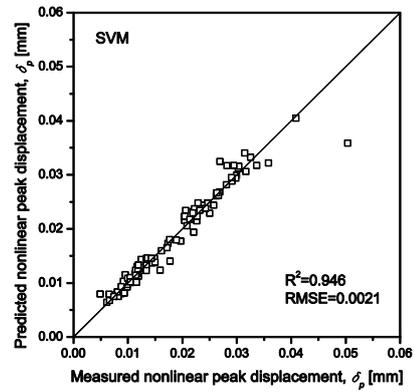


(b) SVM model

Fig. 8 Prediction of the ductility index, d_u from the input parameter, f_i



(a) MR model



(b) SVM model

Fig. 10 Prediction of the nonlinear peak displacement, δ_p , based on the displacement at peak stress, δ_f

The prediction of the initial stiffness based on the tensile strength by using the testing data and the SVM model is presented in Fig. 7. It can be seen a dispersion of values similar to the one evidenced by the equation given by Vasconcelos *et al.* (2008) obtained with simple regression analysis and using all data.

Fig. 8 shows the prediction of the ductility index as a function of the tensile strength obtained with the ANN and SVM models. It is observed that both models present some dispersion but the ANN model appears to give slightly better prediction. The prediction of the ultimate crack opening, w_c , presents also important scatter as can be seen in Fig. 9, where the performance of the SVM model is evaluated. The prediction found by Vasconcelos *et al.* (2008a) (Fig. 9(b)) presents also scatter but taken into account that all data was used for this, it can be concluded that the data mining techniques present better prediction performance.

Fig. 10 shows the very good performance of the data mining models about the prediction of the nonlinear peak displacement based on the peak displacement, as almost all points are aligned along the 45 degree line. In this case, the best prediction model is the MR model, which shows that the relationship between the variables is predominantly linear.

The MR model corresponds to Eq. (5).

$$\delta_p = 0.6984\delta_{ft} + 0.0001 \quad (5)$$

4. Prediction of the mechanical properties through the physical properties

This section aims at evaluating the possibility of using the physical properties to predict the mechanical parameters characterizing the mechanical behavior of granites under tensile loading. In fact, as the physical properties are easier to obtain, it can be beneficial to obtain the mechanical properties through physical properties. For this, it was decided to use the data mining models to evaluate the possibility and the performance of using different combinations of the physical properties to predict the mechanical properties, namely ultrasonic pulse velocity (UPV_{dry}) porosity (η) and density (ρ). The data regarding to the physical properties found in all of the granites under studied is presented in Fig. 11. It is observed that a great variety of physical properties are also found as the result of the variation of the petrographic features of the granites under analysis. The models that use a single input variable are denominated by M_1 . M_2 and M_3 are the designations attributed to models that use two and three input variables, respectively.

4.1 Prediction of the tensile strength

For the prediction of the tensile strength through the physical properties, 240 groups of data were used. Table 4 shows the general statistical overview of the rock properties used in database. As discussed previously, the coefficient of variation is relatively high, except in the case of ρ . This is related to the wide range of granitic rocks used in this study.

Table 4 General statistical overview of the rock properties used in database

| Symbol | Minimum | Mean | Maximum | Std. deviation | Coef. variation |
|-----------------------------|---------|---------|---------|----------------|-----------------|
| UPV_{dry} (m/s) | 1578.01 | 2851.91 | 4480.59 | 797.70 | 27.97 |
| η (%) | 0.61 | 2.96 | 7.40 | 2.11 | 71.27 |
| ρ (kg/m ³) | 2520.57 | 2611.78 | 2704.69 | 45.75 | 1.75 |
| f_t (N/mm ²) | 1.28 | 3.65 | 9.04 | 1.61 | 47.81 |

Table 5 Correlations between f_t with other parameters and the corresponding coefficients of correlation (R)

| Correlations | R |
|--|-------|
| $f_t = 0.6756 \times e^{0.0005 \cdot UPV_{dry}}$ | 0.931 |
| $f_t = 2 \times 10^{-5} \times UPV_{dry}^{1.5343}$ | 0.931 |
| $f_t = 4.3935 \times \eta^{-0.461}$ | 0.804 |
| $f_t = 0.0245 \times \rho - 60.723$ | 0.696 |

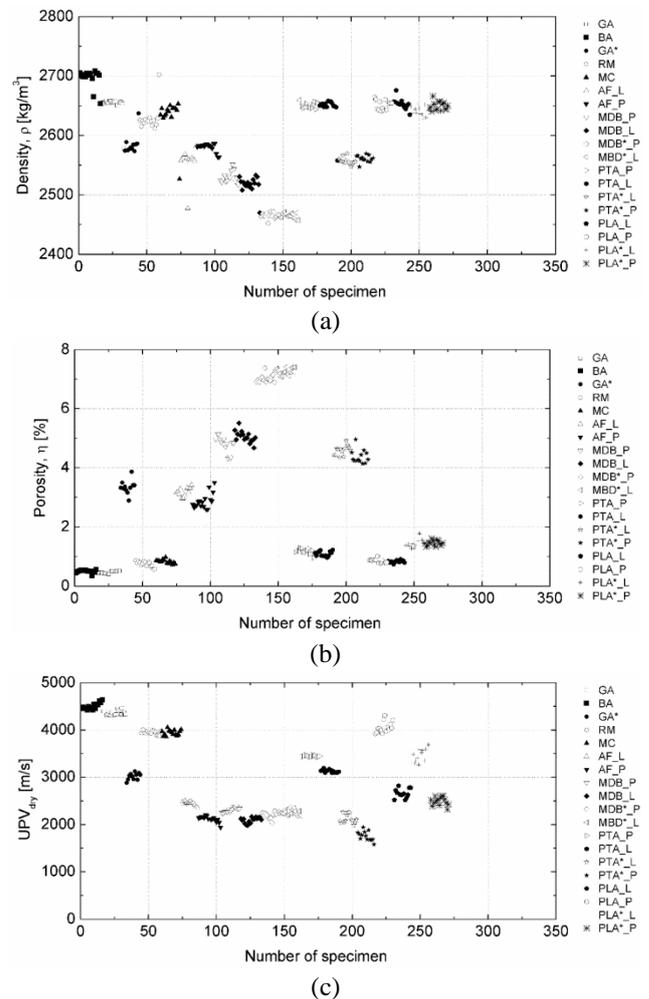


Fig. 11 Experimental data, (a) density, (b) porosity and (c) ultrasonic pulse velocity, UPV

The different mineralogical, textural and structural characteristics control the variables employed in this study. When compared to the other variables, the ρ is less affected.

In a first phase, simple regression analysis of the tensile strength versus UPV_{dry} and η and ρ was carried out to assess

Table 6 Mean values of RMSE obtained in the cross-validation scheme for different combination of input parameters

| Symbol | M ₁ | | | M ₂ | | M ₃ | |
|--------|----------------|--------|--------|------------------|-------------|------------------|-----------------------|
| | UPV_{dry} | η | ρ | $UPV_{dry}&\eta$ | $\eta&\rho$ | $UPV_{dry}&\rho$ | $UPV_{dry}&\eta&\rho$ |
| ANN | 0.5400 | 0.8050 | 1.0065 | 0.4965 | 0.8850 | 0.4294 | 0.4111 |
| SVM | 0.5782 | 0.9222 | 0.9961 | 0.5096 | 0.8457 | 0.4295 | 0.3943 |
| MR | 0.6321 | 1.1684 | 1.1827 | 0.6370 | 1.1598 | 0.6235 | 0.6215 |

Table 7 Mean values of R obtained in the cross-validation scheme for different combination of input parameters

| Symbol | M ₁ | | | M ₂ | | M ₃ | |
|--------|----------------|--------|--------|------------------|-------------|------------------|-----------------------|
| | UPV_{dry} | η | ρ | $UPV_{dry}&\eta$ | $\eta&\rho$ | $UPV_{dry}&\rho$ | $UPV_{dry}&\eta&\rho$ |
| ANN | 0.942 | 0.868 | 0.782 | 0.951 | 0.839 | 0.964 | 0.967 |
| SVM | 0.933 | 0.821 | 0.786 | 0.949 | 0.854 | 0.964 | 0.970 |
| MR | 0.920 | 0.687 | 0.678 | 0.918 | 0.693 | 0.922 | 0.922 |

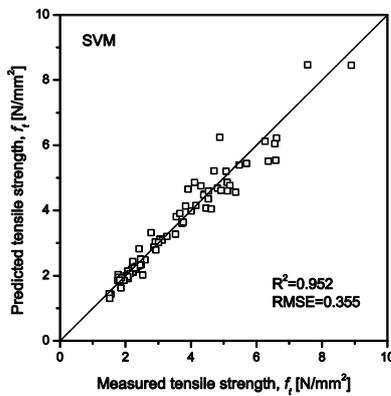


Fig. 12 Performance of SVM model in the prediction of f_t using the combination with all the input parameters

the statistical performance of the correlations. As can be seen through the equations shown in Table 5, the correlations found between the tensile strength and the physical properties presented are statistically meaningful. In this context, the relation between the tensile strength and the density is the poorest one. This result is in line with the ones presented by Vasconcelos *et al.* (2008a, b).

For the construction of the data mining models, the data was split into two sets. The training set with two thirds of the groups of the data (160 cases) and the testing set with one third of the groups of data (80 cases). The mean values of RMSE and R obtained in the training process are presented in Tables 6 and 7. According to Johnson (1984), values of R higher than ± 0.8 are considered statistically significant at 95% confidence. It can be seen that most of the values presented in Table 7 are greater than 0.8, which confirms the good predictive capacity of the majority of the models. Despite the correlation presented in the Table 5 have been obtained with the whole dataset, the values of R obtained from the ANN and SVM approaches with the training data were higher (Table 7). From the results, it is also seen that the MR model gives the poorest results for all the combinations of input parameters, even if they remain clearly statistically meaningful. For models M₁ and M₂, better prediction results were obtained when the UPV_{dry} ,

was used. On the other hand, when two input parameters were used, the best prediction was obtained with ANN model that also include the ρ . The better prediction obtained when considering the combination of the three physical properties is achieved with the SVM model.

The SVM model was fitted with all the testing set and the result is graphically presented in Fig. 12. It can be seen that the model has a great accuracy in predicting the measured results.

Regarding the predictive capacity of the models in predicting the minimum and maximum values of tensile strength, among all models presented in this study, the best results were obtained with the ANN model using the three input data. The minimum and maximum tensile strength values used in the test set were 1.498 and 8.895. The ANN model predicted the values 1.490 and 8.634 while the SVM model with three input parameters predicted the values 1.445 and 8.454. This means that the ANN model can better predict the extreme values of tensile strength, especially the maximum value.

4.2 Prediction of the ductility index

The prediction of the ductility index (d_u) used the same database with 240 experimental results, being the range of variation already presented in Table 4 for the physical properties. The ductility index ranges from 0.02 to 0.20 with a mean value equal to 0.08. The methodology followed for the training and testing processes was the same explained in the previous section.

A simple regression analysis of the d_u versus UPV_{dry} , η and ρ was also performed (see Table 8). It must be underlined the statistically meaningful relation between the d_u and UPV_{dry} , and between d_u and η . Again, the relation between the d_u and ρ is the poorest one.

The mean values of RMSE and R obtained for the training process are presented in Tables 8 and 9. Also in this case, despite the correlation presented in Table 8 has been obtained with the whole dataset, the values of correlation coefficients obtained from M₁ models based on ANN and SVM approaches were higher for the ρ and η . In case of the UPV_{dry} , as the independent variable, the R is slightly lower (Table 10).

Table 8 Correlations between d_u with other parameters and the corresponding coefficients of correlation (R)

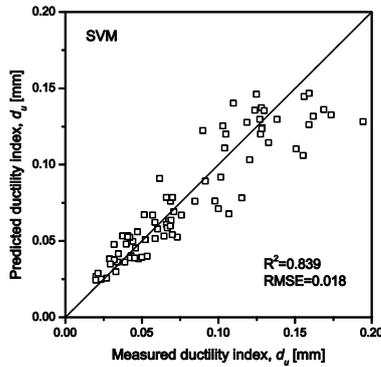
| Correlations | R |
|--|-------|
| $d_u=82601 \times UPV_{dry}^{-1.769}$ | 0.863 |
| $d_u=0.0421 \times \eta^{0.6073}$ | 0.851 |
| $d_u=3 \times 10^8 \times e^{-0.009 \cdot \rho}$ | 0.695 |

Table 9 Mean values of RMSE obtained in the cross-validation scheme for different combination of input parameters

| Symbol | M ₁ | | | M ₂ | | M ₃ | |
|--------|----------------|--------|--------|------------------|-------------|------------------|-----------------------|
| | UPV_{dry} | η | ρ | $UPV_{dry}&\eta$ | $\eta&\rho$ | $UPV_{dry}&\rho$ | $UPV_{dry}&\eta&\rho$ |
| ANN | 0.0223 | 0.0194 | 0.0263 | 0.0185 | 0.0194 | 0.0195 | 0.0192 |
| SVM | 0.0225 | 0.0194 | 0.0257 | 0.0181 | 0.0189 | 0.0201 | 0.0176 |
| MR | 0.0250 | 0.0230 | 0.0301 | 0.0204 | 0.0221 | 0.0241 | 0.0189 |

Table 10 Mean values of R obtained in the cross-validation scheme for different combination of input parameters

| Symbol | M ₁ | | | M ₂ | | M ₃ | |
|--------|----------------|--------|--------|------------------|-------------|------------------|-----------------------|
| | UPV_{dry} | η | ρ | $UPV_{dry}&\eta$ | $\eta&\rho$ | $UPV_{dry}&\rho$ | $UPV_{dry}&\eta&\rho$ |
| ANN | 0.836 | 0.878 | 0.763 | 0.890 | 0.880 | 0.878 | 0.884 |
| SVM | 0.832 | 0.879 | 0.773 | 0.895 | 0.885 | 0.870 | 0.901 |
| MR | 0.787 | 0.824 | 0.667 | 0.864 | 0.838 | 0.803 | 0.884 |

Fig. 13 Performance of SVM model in the prediction of d_u using the combination with the three input parameters

It is seen that when only one independent variable is used (Models M₁), the best predictive performance for the ductility index is achieved using the porosity. In this case, the performances of ANN and SVM models are similar and better than that obtained with the MR model. When combining two independent input variables, the best results are obtained with the SVM model considering the variables the UPV_{dry} and η . Among all the combinations and models, the best performance was obtained with the SVM model using all the input variables. This model was also applied to a database containing the testing set and the result is graphically presented in Fig. 13. It can be observed that the model has a good performance to predict the measured values of the ductility index up to 0.07. For higher values the results are more scattered around the line of 45 degrees. However, it should be mentioned that the coefficient of correlation is higher than 0.8.

4.3 Prediction of the displacement at peak stress

In the prediction of the displacement at peak stress (δ_{ft}), 251 experimental results were used. A general overview of the rock properties used in database is presented in Table 11. In this case, the data was also split into two sets: (1) the training set with two thirds of the groups of the data (167 results); (2) the testing set with one third of the data (84 results).

The best single correlations between the displacement at peak stress and the three input parameters are given in Table 12. It can be seen one more time that the UPV_{dry} and the η are considerably more relevant for the description of the δ_{ft} , as it is confirmed by the coefficient of correlation found for model M₁. The density is the independent variable that leads to a worse prediction. The values found for the coefficient of correlation for models M₁ (Table 14) confirm

Table 11 General statistical overview of the rock properties used in database

| Symbol | Minimum | Mean | Maximum | Std. deviation | Coef. variation |
|-----------------------------|---------|---------|---------|----------------|-----------------|
| UPV_{dry} (m/s) | 1578.01 | 2973.32 | 4577.14 | 851.98 | 28.65 |
| η (%) | 0.41 | 2.70 | 7.42 | 2.18 | 80.74 |
| ρ (kg/m ³) | 2452.48 | 2596.11 | 2705.41 | 66.01 | 2.54 |
| δ_p (mm) | 0.008 | 0.026 | 0.062 | 0.012 | 44.37 |

Table 12 Correlations between δ_{ft} with other parameters and the corresponding coefficients of correlation (R)

| Correlations | R |
|--|-------|
| $\delta_{ft}=3397.3 \times UPV_{dry}^{-1.494}$ | 0.911 |
| $\delta_{ft}=0.0174 \times \eta^{0.4708}$ | 0.896 |
| $\delta_{ft}=-0.0001 \times \rho + 0.3702$ | 0.761 |

Table 13 Mean values of RMSE obtained in the cross-validation scheme for different combination of input parameters

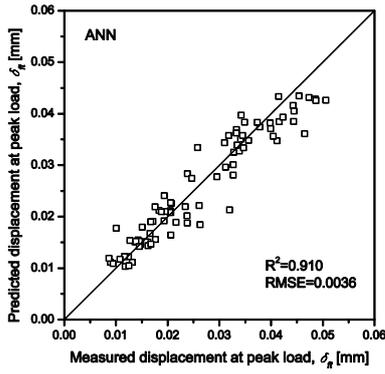
| Symbol | M ₁ | | | M ₂ | | M ₃ | |
|--------|----------------|--------|--------|------------------|-------------|------------------|-----------------------|
| | UPV_{dry} | η | ρ | $UPV_{dry}&\eta$ | $\eta&\rho$ | $UPV_{dry}&\rho$ | $UPV_{dry}&\eta&\rho$ |
| ANN | 0.0041 | 0.0047 | 0.0066 | 0.0041 | 0.0059 | 0.0052 | 0.0043 |
| SVM | 0.0041 | 0.0048 | 0.0068 | 0.0041 | 0.0048 | 0.0044 | 0.0042 |
| MR | 0.0054 | 0.0067 | 0.0075 | 0.0049 | 0.0068 | 0.0050 | 0.0049 |

Table 14 Mean values of R obtained in the cross-validation scheme for different combination of input parameters

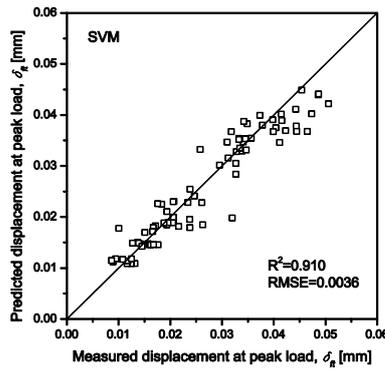
| Symbol | M ₁ | | | M ₂ | | M ₃ | |
|--------|----------------|--------|--------|------------------|-------------|------------------|-----------------------|
| | UPV_{dry} | η | ρ | $UPV_{dry}&\eta$ | $\eta&\rho$ | $UPV_{dry}&\rho$ | $UPV_{dry}&\eta&\rho$ |
| ANN | 0.931 | 0.911 | 0.817 | 0.933 | 0.862 | 0.898 | 0.926 |
| SVM | 0.931 | 0.908 | 0.802 | 0.932 | 0.909 | 0.922 | 0.931 |
| MR | 0.880 | 0.808 | 0.752 | 0.904 | 0.805 | 0.897 | 0.904 |

also that the data mining techniques are more reliable than the single regression analysis. By comparing the performance of the different data mining methods, it is observed that MR model gives the poorest results for all the combinations of input parameters (Tables 12 and 13). This confirms the nonlinear relationships between δ_{ft} with other parameters. The best performance for M₁ models is obtained with UPV_{dry} . In this case, the ANN and SVM models have similar accuracy to predict δ_{ft} . When two input parameters are considered, it is observed that the combination that leads to better prediction is composed by the ultrasonic pulse velocity and the porosity. For the other two combinations, it is interesting to notice that the prediction performance decreases, and both of them perform worse than model M₁ when the independent variable is the porosity.

Using three input parameters, the model with the best predictive capacity is the SVM and it should be noticed that the performance between models M₁ and M₃ is the same as the coefficient of correlation is equal and the RMSE is also practically equal. As already mentioned, the best prediction of the displacement at peak load is achieved when two independent variables are adopted (UPV_{dry} and η), which



(a) ANN model



(b) SVM model

Fig. 14 Prediction of the displacement at peak load with two independent variables UPV_{dry} and η

also confirmed when the prediction is obtained with the testing data through ANN and SVM models (see Fig. 14). It is seen that both models have a great accuracy to predict the peak displacement at peak load (coefficient of correlation of about 0.9).

4.4 Prediction of the nonlinear peak displacement

For the prediction of the nonlinear peak displacement (δ_p), 271 experimental results were considered. Table 15 shows a general statistical overview of the rock properties used in database. Similarly to the physical properties under study, also the scatter found for the δ_p is high, given the wide range of granites analysed in this study. In this case 181 results were considered as training data and 90 test results were used for the test. Table 16 presents the best simple regression analysis of the δ_p versus UPV_{dry} and η and ρ . It must be stressed the statistically meaningful relation between δ_p and the pair UPV_{dry} and η . Similarly to the other mechanical parameters, also in the prediction of the nonlinear peak displacement the density leads to weaker single correlation. It should be mentioned that these single correlations were obtained with the whole dataset.

The mean values of the RMSE and the R obtained during the training process are presented in Tables 17-18. Similarly to what was seen for the other mechanical properties, in spite of the data mining models (M_1) uses only part of the dataset, the coefficient of correlation obtained with the ANN and SVM models are very close to or even higher than the ones obtained for the single

Table 15 General statistical overview of the rock properties used in database

| Symbol | Minimum | Mean | Maximum | Std. Deviation | Coef. Variation |
|-----------------------------|---------|---------|---------|----------------|-----------------|
| UPV_{dry} (m/s) | 1578.01 | 2992.44 | 4632.69 | 874.50 | 29.22 |
| η (%) | 0.34 | 2.69 | 7.42 | 2.19 | 81.41 |
| ρ (kg/m ³) | 2452.47 | 2597.06 | 2708.31 | 67.77 | 2.61 |
| δ_p (mm) | 0.005 | 0.018 | 0.044 | 0.008 | 45.69 |

Table 16 Correlations between δ_p with other parameters and the corresponding coefficients of correlation (R)

| Correlations | R |
|---|-------|
| $\delta_p=2034.6 \times UPV_{dry}^{-1.474}$ | 0.884 |
| $\delta_p=0.0122 \times \eta^{0.4642}$ | 0.868 |
| $\delta_p=18298 \times e^{-0.005 \cdot \rho}$ | 0.748 |

Table 17 Mean values of RMSE obtained in the cross-validation scheme for different combination of input parameters

| Symbol | M ₁ | | M ₂ | | M ₃ | |
|--------|----------------|--------|----------------|---------------------|----------------|---------------------|
| | UPV_{dry} | η | ρ | $UPV_{dry} \& \eta$ | $\eta \& \rho$ | $UPV_{dry} \& \rho$ |
| ANN | 0.0039 | 0.0037 | 0.0049 | 0.0036 | 0.0194 | 0.0039 |
| SVM | 0.0037 | 0.0037 | 0.0050 | 0.0036 | 0.0189 | 0.0037 |
| MR | 0.0043 | 0.0052 | 0.0057 | 0.0040 | 0.0221 | 0.0041 |

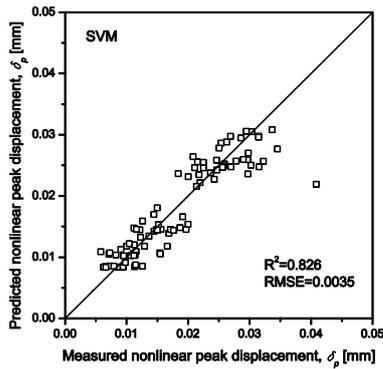
Table 18 Mean values of R obtained in the cross-validation scheme for different combination of input parameters

| Symbol | M ₁ | | M ₂ | | M ₃ | |
|--------|----------------|--------|----------------|---------------------|----------------|---------------------|
| | UPV_{dry} | η | ρ | $UPV_{dry} \& \eta$ | $\eta \& \rho$ | $UPV_{dry} \& \rho$ |
| ANN | 0.878 | 0.893 | 0.802 | 0.901 | 0.880 | 0.882 |
| SVM | 0.896 | 0.893 | 0.795 | 0.902 | 0.885 | 0.892 |
| MR | 0.857 | 0.780 | 0.720 | 0.877 | 0.838 | 0.870 |

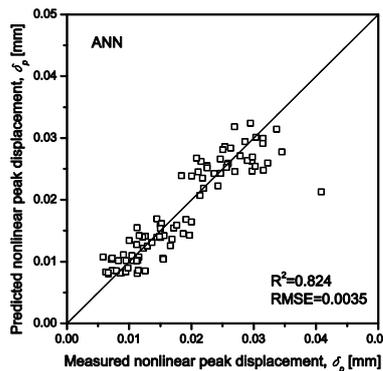
correlations when the independent variables UPV_{dry} , η and ρ are used. MR model have a lower predictive capacity but have also a good performance.

It is observed that the ANN and SVM models have good and similar performances in predicting the nonlinear peak displacement. Using only one input parameter, the best result was obtained with the SVM model using UPV_{dry} . When using two input parameters to obtain the prediction of the nonlinear peak displacement, the SVM model with UPV_{dry} and η is the one that performs better. It should be stressed that the simultaneous use of all physical properties does not improve the prediction, if it is compared with the case where two variables (UPV_{dry} and η) were used with the ANN and SVM models. On the other hand, in case of SVM model, the performance is even slightly worse when three input variables are used when compared to the case when only the UPV_{dry} is considered. This means that there are no gains in considering the three variables. This prediction is influenced negatively by the consideration of the density, as already seen it represents the poorest correlation with the mechanical properties.

The performance of the model M_2 when the porosity and



(a) SVM model



(b) ANN model

Fig. 15 Prediction of nonlinear peak displacement, δ_p , with two independent variables UPV_{dry} and η

ultrasonic pulse velocity are combined for the test dataset, can be evaluated based on the diagrams presented in Fig. 15.

It can be seen that both models have a good performance, with a coefficient of correlation higher than 0.8. Only one value is too far from the 45-degree line. The SVM model confirms the slightly better performance than ANN model.

5. Conclusions

This paper provides and discusses the results of data mining techniques for the prediction of the parameters characterizing the tensile behavior of granites.

In a first phase, a prediction of the parameters characterizing the complete shape of the stress-displacement diagrams based on the tensile strength, which is considered the property of easier obtainment, is carried out. It is seen that, in general, the tensile strength can be used very satisfactorily in the prediction of the mechanical parameters characterizing global tensile behavior of granites. It is clear that, in general, the prediction given by artificial neural networks (ANN) or supporting vector machines (SVM) is more meaningful than the prediction provided by multiple or single regression models. This is more evident in case of comparison of ANN and SVM applied in the test data in relation to single regression models applied to the whole data.

It was possible also to observe that there is a strong

relation between the mechanical properties characterizing the tensile behavior of granites and the physical properties of granites, namely ultrasonic pulse velocity, porosity and density. However, among these variables, the density provides poorest predictions results. This should be related to the lower range of variation when compared to the wide variation of the mechanical properties of granites, much more connected to variation on the porosity. The simultaneous combination of the three input variables almost always conducted to the best results, even if not so much difference was recorded to the combination of input variables ultrasonic pulse velocity and the porosity of the data mining techniques, particularly related to ANN and SVM models. On the other hand, it is stressed that as already pointed out in previous works, the ultrasonic pulse velocity plays a central role on the estimation of the mechanical properties of granites, which is also confirmed in this work by the application of the data mining techniques with very important correlation coefficients found. Finally, it is important to stress, that among the MR, ANN and SVM models, the ANN and SVM were the ones that led to the best prediction results, being the performance very close.

Acknowledgements

This work was partly financed by FEDER funds through the Competitivity Factors Operational Programme-COMPETE and by national funds through FCT-Foundation for Science and Technology within the scope of the project POCI-01-0145-FEDER-007633.

References

- Aleksander, I. and Morton, H. (1990), *An Introduction to Neural Computing*, Chapman & Hall, London, U.K.
- Alemdag, S., Gurocak, Z., Cevik, A., Cabalar, A.F. and Gokceoglu, C. (2016), "Modeling deformation modulus of a stratified sedimentary rock mass using neural network, fuzzy inference and genetic programming", *Eng. Geol.*, **203**, 70-82.
- Baykasoğlu, A., Güllü, A., Çanakçı, H. and Özbakır, L. (2008), "Prediction of compressive and tensile strength of limestone via genetic programming", *Expert Syst. Appl.*, **35**(1-2), 111-123.
- Begonha, A. and Sequeira Braga, M.A. (2002), "Weathering of the Oporto granite: geotechnical and physical properties", *Catena*, **49**(1-2), 57-76.
- Ben-Hur, A., Weston, J. (2010), *A User's Guide to Support Vector Machines*, in *Data Mining Techniques for the Life Sciences*, Humana Press, New York, U.S.A., 223-239.
- Çanakçı, H. and Pala, M. (2007), "Tensile strength of basalt from a neural network", *Eng. Geol.*, **94**(1-2), 10-18.
- Çanakçı, H., Baykasoğlu, A. and Güllü, H. (2009), "Prediction of compressive and tensile strength of Gaziantep basalts via neural networks and gene expression programming", *Neural Comput. Appl.*, **18**(8), 1031-1041.
- Ceryan, N., Okkan, U., Samui, P. and Ceryan, S. (2013), "Modeling of tensile strength of rocks materials based on support vector machines approaches", *J. Numer. Anal. Meth. Geomech.*, **37**(16), 2655-2670.
- Cherkassy, V. and Ma, Y. (2004), "Practical selection of SVM parameters and noise estimation for SVM regression", *Neural*

- Networks*, **17**(1), 113-126.
- Chiaia, B., Van Mier, J.G.M. and Vervuurt, A. (1998), "Crack growth mechanisms in four different concretes: Microscopic observations and fractal analysis", *Cement Concrete Res.*, **28**(1), 103-114.
- Chararas, B., Auger, F. and Mosse, E. (1994), "Determination of the moduli of elasticity of rocks. Comparison of the ultrasonic velocity and mechanical resonance frequency methods with direct static methods", *Mater. Struct.*, **27**(4), 222-228.
- Cortes, C. and Vapnik, V. (1995), "Support vector networks", *Mach. Learn.*, **20**(3), 273-297.
- Cortez, P. (2010), "Data mining with neural networks and support vector machines using the R/rminer tool", *Proceedings of the 10th Industrial Conference on Data Mining*, Berlin, Germany, July.
- Cristianini, N. and Shawe-Taylor, J. (2000), *An Introduction to Support Vector Machine*, Cambridge University Press, Cambridge, U.K.
- Dagdelenler, G., Sezer, E.A. and Gokceoglu, C. (2011), "Some non-linear models to predict the weathering degrees of a granitic rock from physical and mechanical parameters", *Expert Syst. Appl.*, **38**(6), 7476-7485.
- Dibike, Y.B., Velickov, S., Solomatine, D.P. and Abbott, M.B. (2001), "Model introduction with support vector machines; Introduction and applications", *J. Comput. Civ. Eng.*, **15**(3), 208-216.
- Downing, K.L. (2015), *Intelligence Emerging: Adaptive and Search in Evolving Neural System*, MIT Press, U.S.A.
- Gokceoglu, C., Zorlu, K., Ceryan, S. and Nefeslioglu, H. A. (2009), "A comparative study on indirect determination of degree of weathering of granites from some physical and strength parameters by two soft computing techniques", *Mater. Charact.*, **60**(11), 1317-1327.
- GSL (Geological Society of London) (1995), "The description and classification of weathered rocks for engineering purposes, geological society engineering group working party report", *Quart. J. Eng. Geol.*, **28**, 207-242.
- Gurocak, Z., Solanki, P., Alemdag, S. and Zaman, M.M. (2012), "New considerations for empirical estimation of tensile strength of rocks", *Eng. Geol.*, **145**, 1-8.
- Hastie, T., Tibshirani, R. and Friedman, J. (2001), *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, New York, USA.
- Haykin, S. (1999), *Neural Networks-A Comprehensive Foundation*, Prentice-Hall, Upper Saddle River, New Jersey, U.S.A.
- Ilonen, J., Kamarainen, J.K. and Lampinen, J. (2003), "Differential evolution training algorithm for feed-forward neural network", *Neural Process. Lett.*, **17**(1), 93-105.
- Irfan, T.Y. and Dearman, W.R. (1978), "The engineering petrography of weathered granite in Cornwall, England", *Quart. J. Eng. Geol.*, **11**(3), 233-244.
- Johnson, R. (1984), *Elementary Statistics*, Duxbury Press, Boston, U.S.A., 86-106.
- Karakus, M. (2001), "Function identification for the intrinsic strength and elastic properties of granitic rocks via genetic programming (GP)", *Comput. Geosci.*, **37**(9), 1318-1323.
- Kiliç, A. and Teymen, A. (2008), "Determination of mechanical properties of rocks using simple methods", *Bull. Eng. Geol. Environ.*, **67**(2), 237-244.
- Liang, Y., Xu, Q.S., Li, H.D. and Cao, D.S. (2011), *Support Vector Machines and their Application in Chemistry and Biotechnology*, CRC Press, Boca Raton, Florida, U.S.A.
- Liao, S.H., Chu, P.H. and Hsiao, P.Y. (2012), "Data mining techniques and applications-A decade review from 2000 to 2011", *Expert Syst. Appl.*, **39**(12), 11303-11311.
- Marques, M.L., Chastre, C.M. and Vasconcelos, G. (2012), "Modelling the compressive mechanical behavior of building stones", *Construct. Build. Mater.*, **28**(1), 372-381.
- Martins, F.F., Begonha, A. and Sequeira Braga, M.A. (2012), "Prediction of the mechanical behavior of the Oporto granite using data mining techniques", *Expert Syst. Appl.*, **39**(10), 8778-8783.
- Martins, F.F. and Miranda, T.F.S. (2012), "Estimation of the rock deformation modulus and RMR based on data mining techniques", *Geotech. Geol. Eng.*, **30**(4), 787-801.
- Meulenkamp, F. and Grima, M.A. (1999), "Application of neural networks for the prediction of the unconfined compressive strength (UCS) from Equotip hardness", *J. Rock Mech. Min. Sci. Geomech.*, **36**(1), 29-39.
- Mishra, D.A. and Basu, A. (2013), "Estimation of uniaxial compressive strength of rock materials by index tests using regression analysis and fuzzy inference system", *Eng. Geol.*, **160**, 54-68.
- Prado, E.P. and Van Mier, J.G.M. (2003), "Effect of particle structure on mode I fracture process of concrete", *Eng. Fract. Mech.*, **70**(14), 1793-1807.
- Sharma, P.K. and Singh, T.N. (2007), "A correlation between P-wave velocity, impact strength index, slake durability index and uniaxial compressive strength", *Bull. Eng. Geol. Environ.*, **67**(1), 17-22.
- Singh, R., Umrao, R.K., Ahmad, M., Ansari, M.K., Sharma, L.K. and Singh, T.N. (2017), "Prediction of geomechanical parameters using soft computing and multiple regression approach", *Measurement*, **99**, 108-119.
- Smola, A.J. and Schölkopf, B. (2004), "A tutorial on support vector regression", *Stat. Comput.*, **14**(3), 199-222.
- Souza, A.M.F. and Soares, F.M. (2016), *Neural Network Programming with Java*, Packt Publishing Ltd, Birmingham, U.K.
- Tang, C.A., Tham, L.G., Wang, S.H., Liu, H. and Li, W.H. (2007), "A numerical study of the influence of heterogeneity on the strength characterization of rock under uniaxial tension", *Mech. Mater.*, **39**(4), 326-339.
- Vapnik, V.N. (1998), *Statistical Learning Theory*, Wiley, New York, U.S.A.
- Vasconcelos, G., Lourenço, P.B., Alves, C.A.S. and Pamplona, J. (2008a), "Experimental characterization of the tensile behaviour of granites", *J. Rock Mech. Min. Sci.*, **45**(2), 268-277.
- Vasconcelos, G., Lourenço, P.B., Alves, C.A. and Pamplona, J. (2008b), "Ultrasonic evaluation of the physical and mechanical properties of granites", *Ultrasonics*, **48**(5), 453-466.
- Yesiloglu-Gultekin, N., Gokceoglu, C. and Sezer, E.A. (2013), "Prediction of uniaxial compressive strength of granitic rocks by various nonlinear tools and comparison of their performances", *J. Rock Mech. Min. Sci.*, **62**, 113-122.