

# Artificial neural networks for soil embankments stability condition identification

## Réseaux de neurones artificiels pour l'identification des conditions de stabilité des remblais de sol

J. Tinoco

*ISISE – Institute for Sustainability and Innovation in Structural Engineering/ALGORITMI Research  
Center, University of Minho, Guimarães, Portugal*

A. Gomes Correia

*ISISE – Institute for Sustainability and Innovation in Structural Engineering, University of Minho,  
Guimarães, Portugal*

P. Cortez

*ALGORITMI Research Center, University of Minho, Guimarães, Portugal*

D. Toll

*School of Engineering and Computing Sciences, University of Durham, Durham, UK*

**ABSTRACT:** Today challenge concerning to transportation infrastructure networks is how to keep them operational under all conditions. Budgetary constraints and network dimension are between the main factors that make the management of a transportation network such a challenging task. Accordingly, aiming to support transportation network management tasks, a data-driven model for stability condition prediction of soil embankment slopes is proposed based on the well known Artificial Neural Networks (ANN) algorithm. For that, the ANN was feed with more than fifty visual features that usually are collected during routine inspections. The proposed model was addressed following two different strategies: nominal classification and regression. Moreover, to overcome the problem of imbalanced data, three training sampling approaches were explored: no resampling, SMOTE (Synthetic Minority Over-sampling TEchnique)and Oversampling. The main results are presented and discussed, comparing ANN predictive performance under the two strategies implemented. Also the effect of the three training sampling approaches is discussed. Moreover, aiming a better understanding of the proposed data-driven models, a detailed sensitivity analysis was applied, allowing to quantify the relative importance of each model input.

**RÉSUMÉ:** Le défi aujourd'hui relatif aux réseaux d'infrastructure de transport est comment le garder opérationnel dans toutes les conditions. Les contraintes budgétaires et la dimension de réseau sont entre les principaux facteurs qui rendent la gestion d'un service de transport réseau telle un véritable enjeu. En conséquence, visant à soutenir les tâches de gestion de réseau de transport, un modèle guidé sur les données pour la prevision de condition de stabilité des talus de remblais en terre est proposé basé sur l'algorithme bien connu des réseaux de neurones artificiels (RNA). Pour cela, le RNA a été construit avec plus de cinquante caractéristiques visuelles qui sont habituellement enregistrées au cours des inspections de routine. Le modèle proposé a suivi deux stratégies d'apprentissage différentes: nominale de classification et régression. Par ailleurs, pour surmonter le problème des données déséquilibrées, trois approches de formation d'échantillonnage ont été

explorées : aucun rééchantillonnage, technique de sur-échantillonnage « SMOTE » et suréchantillonnage. Les principaux résultats sont présentés et discutés, en comparant la performance prédictive RNA sur les deux stratégies mises en œuvre. Par ailleurs, l'effet des trois approches d'échantillonnage sont discutées. En outre, visant une meilleure compréhension des modèles guidés par les données, une analyse de sensibilité détaillée a été appliquée, permettant de quantifier l'importance relative de chaque entrée du modèle.

**Keywords:** Slopes Stability, Soil Embankments, Soft Computing, Neural Networks, Sensitivity Analysis

## 1 INTRODUCTION

In the framework of a transportation networks, one of the biggest challenges today is to keep it operational under all conditions. Their extension and the increased budget limitations for maintenance and repair purposes are two of the main factors that make it such a challenging task. Indeed, this is one of the main concerns of every developed or countries undergoing development that have invested and keep investing to build a safe and functional transportation network. Thus, taken into account the strategic importance of the transportation system in modern societies, it is fundamental to develop new tools able to help on its management.

In the framework of transportations networks, in particular for a railway, slopes are perhaps the element for which their failure can have the strongest impact at several levels. Therefore, it is important to develop ways to identify potential problems before they result in failures. Over time, several efforts have been made toward the development of a system to detect slope failures. However, most of the systems were developed for natural slopes, presenting some constraints when applied to engineered (humanmade) slopes. In addition, they have limited applicability as most of them were developed based on particular case studies or using small databases. Moreover, their applicability can also be conditioned by the type of information required to feed them, such as data taken from complex tests or from expensive monitoring systems. Pourkhosravi and Kalantari (2011) summarize some of the current methods for slope failure detection,

which were grouped into Limit Equilibrium (LE) methods, Numerical Analysis methods, Artificial Neural Networks and Limit Analysis methods. There are also approaches based on finite elements methods (Suchomel et al. 2010), reliability analysis (Malkawi,et al., 2000), as well as some methods making use of data mining (DM) algorithms (Cheng and Hoang, 2014). More recently, a new system was proposed by Pinheiro et al., (2015), based on the assessment of different factors that usually affect slope stability. By weighting the different factors, a final indicator of the slope stability condition is calculated.

It is well known that the assessment of the stability condition of a given slope is a multi-variable problem characterized by a high dimensionality. Keeping this in mind, the high learning capabilities of Artificial Neural Networks (ANNs) (Tinoco et al., 2014; Tinoco et al., 2018a) were applied in this work in the development of a data-driven model for stability condition prediction of soil embankment slopes based on a pre-defined classification scale comprising four levels (classes). One of the underlying premises of this work is to identify the real stability condition of a given slope based on information that can be in a somewhat easily obtained during visual routine inspections. For that, more than fifty visual features collected during routine inspections as well as geometric, geological and geographic data were used to feed the models. At the end, the proposed approach is intended to support railway network management companies to allocate the available funds in the priority assets according to its stability condition.

## 2 METHODOLOGY

### 2.1 Data Characterization

To fit the proposed models for stability condition identification, from this point referred to as EHC (Earthwork Hazard Category (Power et al., 2016)), of soil embankments slopes, a database was compiled containing information collected during routine inspections and complemented with geometric, geological and geographic data of each slope. To be precise, a set of 53 variables were selected as models attributes, some of which are here enumerated: height, slope angle, tree cover, animal activity, construction activity toe, adjacent land drainage, sub drainage, catchment geology, composition crest, etc.

The data were gathered by Network Rail workers and are concerned with the railway network of the UK. For each slope a class of the EHC system was defined by the Network Rail Engineers based on their experience/algorithm (Power et al., 2016), which will be assumed as a proxy for the real stability condition of the slope for year 2015. The EHC system comprises 4 classes (“A”, “B”, “C” and “D”) where “A” represents a good stability condition and “D” a bad stability condition. In other words, the expected probability of failure is higher for class “D” and lower for class “A”.

The compiled database contains an impressive number of 25673 records. Figure 1 plots the distribution of EHC classes, from which it is possible to observe a high asymmetric distribution (imbalanced data) of the records. Indeed, more than 63% of the embankments are classified as “A” and only 2.5% belongs to class “D”. Although this type of asymmetric distribution, where most of the slopes present a low probability of failure (class “A”), is normal and desirable from the safety point of view and slope network management, it can represent an important challenge for DM models learning, as detailed in next section.

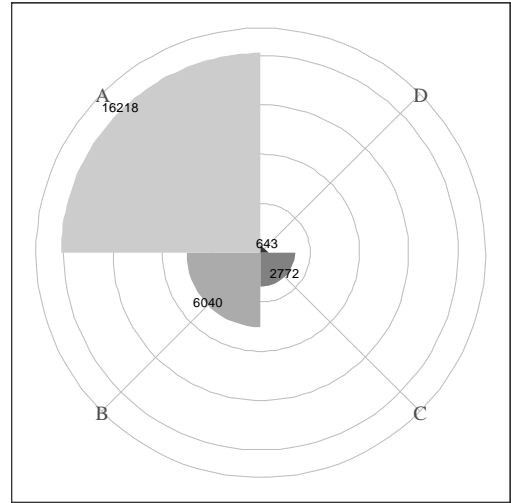


Figure 1. Soil embankment data distribution by EHC classes.

### 2.2 Modelling

For modelling purposes, the well known ANNs algorithm was applied to fit EHC prediction of soil embankment slopes. This algorithm, although not new, is supported in a strong background. Indeed, it has been applied in the past with high success in different knowledge domains including in civil engineering (Chou et al., 2016). Also some applications in slope stability analysis can be found in the literature. (Cheng et al., 2012).

ANN are learning machines that were initially inspired in functioning of the human brain (Kenig et al., 2001). The information is processed using iteration among several neurons. This technique is capable of modelling complex non-linear mappings and is robust in exploration of data with noise. In this study was adopt the multilayer perceptron that contains only feedforward connections, with one hidden layer containing  $H$  processing units. Because the network's performance is sensitive to  $H$  (a trade-off between fitting accuracy and generalisation capability), a grid search of  $\{0; 2; 4; 6; 8\}$  under an internal (i.e. applied over training data) three fold cross validation was adopt during the learning phase to find the best  $H$  value. Under this

grid search, the  $H$  value that produced the lowest MAE (Mean Absolute Error) was selected, and then the ANN was retrained with all of the training data. The neural function of the hidden nodes was set to the popular logistic function  $1/(1 + e^{-x})$ .

As a first attempt, EHC prediction of soil embankments was approached following a nominal classification strategy. Then, the problem was also addressed following a regression strategy, adopting a regression scale where  $A = 1, B = 2, C = 4, D = 10$ , which was that leading to the best performance.

Moreover, in order to overcome the problem of imbalanced data (since typically good conditions are much common than bad ones, see Figure 1), three training sampling approaches were explored: Normal (no resampling), OVERed (Oversampling (Ling and Li, 1998)) and SMOTEd (SMOTE – Synthetic Minority Oversampling Technique (Chawla et al., 2002)). In the case of regression, only two sampling approaches were applied: Normal (no resampling) and SMOTEd (SMOTE for regression). When approaching imbalanced classification tasks, where there is at least one target class label with a smaller number of training samples when compared with other target class labels, the simple use of a soft computing training algorithm will lead to data-driven models with better prediction accuracies for the majority classes and worst classification accuracies for the minority classes. Thus, techniques that adjust the training data in order to balance the output class labels, such as Oversampling and SMOTE, are commonly used with imbalanced datasets. In particular, Oversampling is a simple technique that randomly adds samples (with repetition) of the minority classes to the training data, such that the final training set is balanced. SMOTE is a more sophisticated technique that creates “new data” by looking at nearest neighbours to establish a neighbourhood and then sampling from within that neighbourhood. It operates on the assumptions that the original data is similar

because of proximity. More recently, Torgo et al. (2015) adapted the SMOTE method for regression tasks. We note that the different sampling approaches were applied only to training data, used to fit the data-driven models, and the test data (as provided by the 5-fold procedure) was kept without any change.

The R statistical environment (R Team, 2009) and the rminer package (Cortez, 2010), were used to conduct all experiments.

### 2.3 Model Assessment

For models comparison and accuracy measurement, three classification metrics were calculated: recall, precision and  $F_{1-score}$  (Hastie et al., 2009). The recall measures the ratio of how many cases of a certain class were properly captured by the model. In other words, the recall of a certain class is given by  $TruePositives = (TruePositives + FalseNegatives)$ . On the other hand, the precision measures the correctness of the model when it predicts a certain class. More specifically, the precision of a certain class is given by  $TruePositives = (TruePositives + FalsePositives)$ . The  $F_{1-score}$  was also calculated, which represent a trade-off between the recall and precision of a class. The  $F_{1-score}$  correspond to the harmonic mean of precision and recall, according to the following expression:

$$F_{1-score} = 2 \cdot \frac{precision \cdot recall}{precision + recall} \quad (1)$$

For all three metrics, the higher the value, the better are the predictions, ranging from 0% to 100%.

The generalization capacity of the models was accessed through a 5-fold cross-validation approach under 20 runs (Hastie et al., 2009). Also, the three prediction metrics are always computed on test unseen data (as provided by the 5-fold validation procedure).

In addition to models performance, understanding what was learned by them is also a

key point in any data driven project due to its mathematical complexity. With this in mind, Cortez and Embrechts (2013) proposed a novel visualization approach based on a sensitivity analysis (SA), which is used in this work. SA is a simple method that is applied after the training phase and measures the model responses when a given input is changed, allowing the quantification of the relative importance of each attribute. In particular, it was applied the Global Sensitivity Analysis (GSA) method (Cortez and Embrechts, 2013), which is able to detect interactions among input variables. This is achieved by performing a simultaneous variation of  $F$  inputs (in this study was performed a one dimensional (1-D) SA ( $F = 1$ )). Each input is varied through its range with  $L$  levels and the remaining inputs fixed to a given  $b$  baseline value. In this work, it was adopted the average input variable value as a baseline and set  $L = 7$  (number of levels), which allows an interesting detail level under a reasonable amount of computational effort.

With the sensitivity response of the GSA, the input importance barplot can be plotted, which shows the relative influence ( $R_a$ ) of each input variable in the model (from 0% to 100%). The rationale of GSA is that the higher the changes produced in the output, the more important is the input. To measure this effect, first the gradient metric ( $g_a$ ) for all inputs was calculated. After that, the relative influence was computed (Cortez and Embrechts, 2013).

### 3 MAIN RESULTS

Figure 2 compares recall, precision and  $F_{1-score}$  of ANN fitted models for EHC prediction of soil embankments, according to a nominal classification and regression strategies, as well as using SMOTE and Oversampling resampling approaches. As shown, the proposed models are robust on identifying very accurately soil embankments of class “A”, observing a slightly decreasing on its performance for the other three

classes. Taken  $F_{1-score}$  as reference, for class “A” a value higher than 92% was achieved. Concerning to class “D” also a very promising performance is observed with an  $F_{1-score}$  around 58% according to a regression strategy.

Analysing the effect of the training sampling approaches (oversampling e SMOTE), it is observed some effectiveness for class “D” (minority class) following a nominal classification strategy. For the other classes, the application of a sampling approach seems to be ineffective. Indeed, and considering  $F_{1-score}$  as reference, better results are achieved with no resampling. These results show that applying a training sampling approach allows improving models performance in the identification of the minority classes but decreasing its response for the other classes. In fact, taking in account that these training sampling approaches are tailed to address learning problems related with the minority classes in imbalanced datasets, it is acceptable and expected to observe a slightly decrease in the majority classes performance. Comparing oversampling and SMOTE approaches, the first one seems to be more effective. According to a regression strategy, the application of a resampling approach, i.e., SMOTE sampling, has a residual effect on models performance, even for minority classes.

Comparing both nominal classification and regression strategies, one can conclude that approaching the problem as a nominal classification is slightly more effective than following a regression strategy.

Figure 3 show the relation between observed and predicted EHC values according to the best fits, following a nominal classification and regression strategies respectively. From its analysis, it is observe that soil embankments of class “A” are very well identified in both cases, with a slightly superior performance following a regression strategy with no resampling. On the other hand, soil embankments of class “D” are better identified when a resampling approach is applied following a nominal classification strategy.

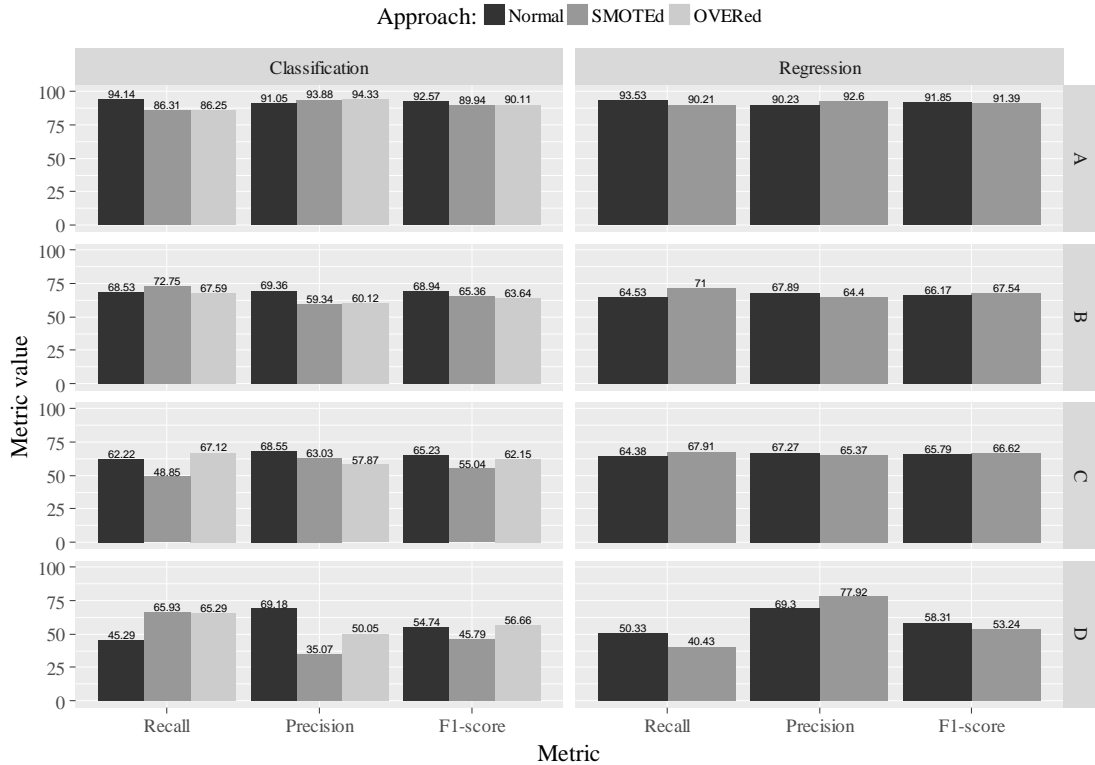


Figure 2. Models comparison based on recall, precision and F1-score, according to a nominal classification strategy in EHC prediction of soil embankments.

In overall, and keeping in mind the importance of identify correctly those slopes with a higher probability of failure (class “D”) one can conclude that the best model for stability condition identification of soil embankments is those based on a nominal classification strategy following an OVERed resampling approach. Moreover, and according to this model, it is also interesting to observe that when the stability condition of a given slope is not correctly identified, such slope is classified as belonging to the nearest class. For example, almost all soil embankments of class “D” not identified as it are classified as belong to class “C”.

As importance as model performance is its interpretability, in particular when are involved data-driven models, namely those based on ANNs algorithms. Accordingly, a GSA methodology (Cortez and Embrechts 2013) was

applied aiming to identify the key parameters (input importance bar plot, 1-D SA) in EHC prediction of soil embankments. Figure 4. shows the relative importance of the ten more relevant variables according to the ANN model with oversampling and following a nominal classification strategy, which was identified as the best model for stability condition identification of soil embankments. Thus, and according to this model, three of the most relevant variables in EHC prediction of soil embankments are related with the height of the slope, summing more than 20% of the total influence. Moreover, “Embankment Opposite Side Condition” as well as “Validate Track Movement” also play an important role in EHC prediction of soil embankments.

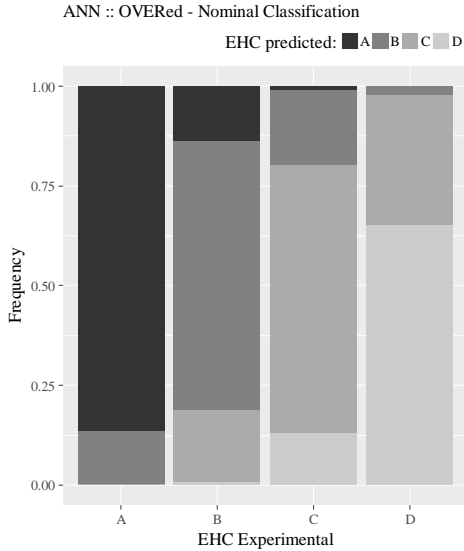


Figure 3. Models performance comparison according to a nominal classification strategy in EHC prediction of soil embankments, following an OVERed approach (Tinoco et al., 2018b).

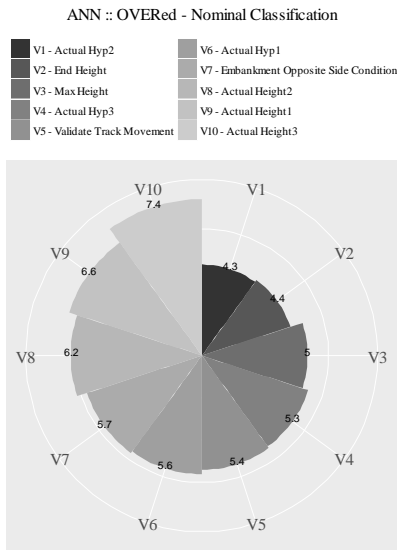


Figure 4. Relative importance bar plot for each variable according to ANN model with oversampling and following a nominal classification strategy.

#### 4 FINAL REMARKS

This study is the first known attempt to predict stability condition (EHC – Earthwork Hazard Category) of soil embankment, assessed by four class (“A”, “B”, “C” and “D”), through the application of soft computing techniques and considering as model attributes information usually collected during routine inspections (visual information). A very promising predictive performance was observed, with  $F_{1-score}$  values higher than 92% for class “A”, around 66% for classes “B” and “C” and close to 57% for class “D”. Moreover, the application of a resampling approach (aiming to overcome the problem of imbalanced data), namely the Oversampling method, allows to improve the predictive performance, particularly for the minority class “D”. Finally, based on a detailed sensitivity analysis method, it was observed that three of the most relevant parameters in EHC prediction of soil embankments (accounting for 20% of the influence) are related with the height of the slope, which was expected from a geotechnical point of view.

#### 5 ACKNOWLEDGEMENTS

This work was supported by FCT – “Fundação para a Ciência e a Tecnologia“, within ISISE, project UID/ECI/04029/2013 as well Project Scope: UID/CEC/00319/2013 and through the post-doctoral Grant fellowship with reference SFRH/BPD/94792/2013. This work was also partly financed by FEDER funds through the Competitivity Factors Operational Programme - COMPETE and by national funds through FCT within the scope of the project POCI-01-0145-FEDER-007633. This work has been also supported by COMPETE: POCI-01-0145-FEDER-007043. A special thanks goes to Network Rail that kindly made available the data (basic earthworks examination data and the Earthworks Hazard Condition scores) used in this work.

## 6 REFERENCES

- Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P.: Smote: synthetic minority oversampling technique. *Journal of artificial intelligence research* 32, 321–357 (2002).
- Cheng, M.Y., Hoang, N.D.: Slope collapse prediction using bayesian framework with k-nearest neighbor density estimation: Case study in taiwan. *Journal of Computing in Civil Engineering* 30(1), 04014116 (2014).
- Cheng, M.Y., Roy, A.F., Chen, K.L.: Evolutionary risk preference inference model using fuzzy support vector machine for road slope collapse prediction. *Expert Systems with Applications* 39(2), 1737–1746 (2012).
- Chou, J.S., Yang, K.H., Lin, J.Y.: Peak shear strength of discrete fiber-reinforced soils computed by machine learning and metaensemble methods. *Journal of Computing in Civil Engineering*, 04016036 (2016).
- Cortez, P. (2010). Data mining with neural networks and support vector machines using the r/miner tool. In: 10<sup>th</sup> Industrial Conference on Data Mining, Berlin, Germany, pp. 572–583. LNAI 6171, Springer.
- Cortez, P., Embrechts, M.: Using sensitivity analysis and visualization techniques to open black box data mining models. *Information Sciences* 225, 1–17 (2013).
- Hastie, T., Tibshirani, R., Friedman, J.: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Second Edition ed.). Springer-Verlag New York (2009).
- Malkawi, A., Hassan, W., Abdulla, F.: Uncertainty and reliability analysis applied to slope stability. *Structural Safety* 22(2), 161–187 (2000).
- Kenig, S., Ben-David, A., Omer, M., Sadeh, A.: Control of properties in injection molding by neural networks. *Engineering Applications of Artificial Intelligence* 14(6), 819–823 (2001).
- Ling, C.X., Li, C.: Data mining for direct marketing: Problems and solutions. In *KDD*, Volume 98, pp. 73–79 (1998).
- Pinheiro, M., Sanches, S., Miranda, T., Neves, A., Tinoco, J., Ferreira, A., Gomes Correia, A.: A new empirical system for rock slope stability analysis in exploitation stage. *International Journal of Rock Mechanics and Mining Sciences* 76, 182–191 (2015).
- Pourkhosravani, A., Kalantari, B.: A review of current methods for slope stability evaluation. *Electronic Journal of Geotechnical Engineering* 16 (2011).
- Power, C., Mian, J., Spink, T., Abbott, S., Edwards, M.: Development of an evidence-based geotechnical asset management policy for network rail, great britain. *Procedia Engineering* 143, 726–733 (2016).
- Suchomel, R. et al.: Comparison of different probabilistic methods for predicting stability of a slope in spatially variable c- $\phi$  soil. *Computers and Geotechnics* 37(1), 132–140 (2010).
- Team, R.: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Viena, Austria. Web site (accessed on October 22, 2018): <http://www.r-project.org/>.
- Tinoco, J., Gomes Correia, A., Cortez, P.: Support vector machines applied to uniaxial compressive strength prediction of jet grouting columns. *Computers and Geotechnics* 55, 132–140 (2014).
- Tinoco, J., Gomes Correia, A., Cortez, P.: Jet grouting column diameter prediction based on a datadriven approach. *European Journal of Environmental and Civil Engineering* 22(3), 338–358 (2018a).
- Tinoco, J., Gomes Correia, A., Cortez, P., Toll, D.: Data-driven model for stability condition prediction of soil embankments based on visual data features. *Journal of Computing in Civil Engineering*, 32(4), 04018027 (2018b).
- Torgo, L., Branco, P., Ribeiro, R., Pfahringer, B.: Resampling strategies for regression. *Expert Systems* 32(3), 465–476 (2015).