Applicability of Artificial Neural Networks (ANN) for equilibrium state prediction in tunnel excavation

Alec Tristani, Lina-María Guayacán-Carrillo, Jean Sulem, Sebastián Ariel Donzis
Laboratoire Navier/CERMES, Ecole des Ponts ParisTech, Université Gustave Eiffel, CNRS, Marne la Vallée, France

ABSTRACT: Simplified 2D analysis of ground support interaction using the convergence-confinement method, although useful in preliminary stages of tunnel design, may be inadequate for complex situations which necessitate 3D numerical simulations with high computational efforts and costs. An alternative approach based on machine learning is proposed here in order to evaluate the stresses and displacements at equilibrium in the lining of a supported tunnel. Based on data previously obtained by numerical simulations, the Bagging Method applied to Artificial Neural Networks (ANNs) is used. We consider a circular tunnel excavated in a Mohr-Coulomb elastoplastic medium. The analysis accounts for a large range of ground conditions, support characteristics, lay distances and tunnel radius. The results show that ANNs models perform well with the small dataset used here and can be considered as a useful alternative to complex 3D numerical simulations.

Keywords: Deep tunnels, Machine learning, Numerical simulations, Bagging, Surrogate models.

1 INTRODUCTION

1.1 Context

Preliminary stage of tunnel design is generally performed using a plane-strain approach based on the convergence confinement (cv-cf) method. This method however relies on strong geotechnical and geometrical assumptions whose relevance is to be checked for each application (Panet & Sulem 2022). Therefore, the limits for its applicability have been highlighted (e.g. Eisenstein & Branco 1991, De la Fuente et al. 2019).

In particular, when the ground exhibits large deformation and/or when the support is very stiff and installed close to the tunnel face, the classical cv-cf method appears to be inaccurate. In that case, because of the arching effect in the longitudinal direction, the ground pressure may be underestimated (Cantieni & Anagnostou, 2009). Therefore, the original method has been enhanced using the so-called implicit methods (Nguyen-Minh & Guo. 1996, Bernaud et al. 1994). In two recent papers (De la Fuente et al. 2019, 2021) the limits of these approaches have been highlighted and the domain of application of the cv-cf methods was evaluated (see Figure 8 in De la Fuente et al. 2021).

On the other hand, three-dimensional simulations may be performed but at rather high computational efforts and costs. Such simulations are more often performed for a posteriori interpretation of field data to validate constitutive models, (e.g. Liu Y. et al. 2021, Tran-Manh et al. 2015, Bonini et al. 2009).

1.2 Machine learning tools in tunneling

As an alternative to three dimensional numerical modeling, emergent artificial intelligence techniques begin to be more and more used in tunneling and underground construction fields (Jong et al. 2021). Many machine learning techniques have been tested to predict convergences at the wall of a tunnel. Three types of approaches can be distinguished: deterministic (e.g. Adoko et al. 2013, Mahdevari et al. 2013, Satici et al. 2020), Bayesian (Feng et al. 2019) or hybrid (Fei et al. 2020, Chang et al. 2022). A common observation is that machine learning tools are able to provide fast and rather accurate predictions.

However, a crucial point in the application of machine learning is the capacity of a model to perform well on previously unobserved inputs (Goodfellow 2015). Therefore, the dataset on which the tool is being trained is of utmost importance: the amount of available data but also its nature and quality will define how well the model will generalize. In other words, the developed tool should only be tested on a similar dataset as the one used for training to avoid bad performances. As every underground structure is unique, it is thus difficult to reuse a model trained on a specific tunnel to another one.

As a consequence, the use of surrogate models based on artificial intelligence tools that are trained on synthetic datasets (generated by numerical simulations) may be relevant. The main idea is to bring a complement to advanced numerical modeling in order to create a powerful link between the traditional paradigm of numerical modeling and the new paradigm of machine learning (Furtney et al. 2022). Synthetic datasets are convenient because they do not contain any noise. Moreover, they can take into account a large range of parameters (ground conditions, structure characteristics…) which enhance generalization. Still, the quantity of synthetic data is problem-dependent.

When dealing with scarce datasets, which is usually the case for tunneling applications (Liu L. et al. 2021), one has to be careful with overfitting. Overfitting occurs when a model adjusts excessively to the training data, seeing patterns that do not exist and consequently performing poorly in predicting new data (Bishop 2006). One way to limit overfitting is to avoid complex models with too much hyperparameters or to use regularization techniques. Another way is to combine several models together, called weak learners, and then to make a global prediction according to every single one. In most cases, this method leads to better results than using a unique strong model. These techniques are called ensemble methods: they contribute to reduce the variance/bias of final predictions and therefore improve generalization.

1.3 Scope of the study

In this paper, we explore the applicability and the accuracy of artificial neural networks (ANNs) models for analyzing the ground-support interaction and evaluate the ground displacement and support stress at equilibrium state. Using a synthetic data set obtained by previous three-dimensional numerical simulations (De La Fuente et al. 2019), different ANN models (weak learners) are trained according to the bagging ensemble method and a hyper-parameter analysis is performed. Motivated by the above considerations, this paper aims to answer the question whether a machine learning tool based on an ensemble method can accurately and reliably predict displacements and stresses in the lining of a supported tunnel excavated in an elastoplastic ground based on a synthetic dataset.

To that end, a Mohr–Coulomb elastic perfectly plastic model is used to describe the constitutive behavior of the ground and a linear elastic model is assumed for the support. This analysis takes into account a large range of ground conditions, support characteristics, lay distances (distance of support/lining installation from the tunnel face) and tunnel radius. Moreover, the calculations are performed for some representative values of the stability number N. This notion is related to the extension of a plastic zone near the tunnel face (Panet & Sulem 2022). The bagging technique (explained in section 2.2) combined with neural networks is applied. Results are discussed in section 3 and conclusions are proposed in section 4.

2 MACHINE LEARNING MODEL

2.1 Dataset

The dataset has been created by performing in total 720 three-dimensional simulations of excavations. Each simulation corresponds to one configuration given a specific selection of the parameters defined in Equation (1) and presented in Table 1 (De La Fuente et al. 2019).

Because of the geometry of the problem, a simple axisymmetric study has been carried out. The aim is to monitor the maximum hoop stresses σ^*_{max} and displacements u^*_{max} that occur in the lining at the equilibrium state far from the tunnel face for each configuration.

$$R^* = \frac{R}{e}; \quad d^* = \frac{d}{2R}; \quad E^* = \frac{E}{E_s}; \quad N = \frac{2\sigma_0}{\sigma_c}; \quad u^*_{max} = \frac{u(\infty)2G}{\sigma_0 R}; \quad \sigma^*_{max} = \frac{\sigma}{\sigma_0} \tag{1}$$

where E and E_s are respectively the Young moduli of the ground and of the lining and where d and R are respectively the lay distance and the tunnel radius, e is the thickness of the lining, σ_0 is the initial stress state in the ground and σ_c represents the compression resistance.

The stability number N was first introduced by Broms & Bennermark (1967) and is related to the extension of a plastic zone near the tunnel face (Panet & Sulem 2022).

Table 1. Range of values for the parameters used in numerical simulations.

Parameter	ν	ν_s	d^*	R^*	E^*	ϕ	ψ	N
Value	0.25	0.2	1	10, 12.5, 15	0.05, 0.025, 0.5, 0.75, 1	20°, 25°, 30°, 35°	0°, ϕ/3, ϕ	1, 2, 5, 10

As parameters ν, ν_s and d^* remain constant, they are not taken into consideration in the dataset used to train the model. Therefore, the total dataset is made of 5 input features (R^*, E^*, ϕ, ψ, N) and 2 output features (σ^*_{max} and u^*_{max}).

2.2 Bagging model

Bagging (that stands for bootstrap aggregating) is a type of ensemble machine learning algorithm that was first introduced by Breiman (1996). It consists in combining the predictions from multiple machine learning models together to make better predictions than the individual ones. It is mostly used to reduce the variance of an individual model. Bagging is the application of the bootstrapping procedure (in statistics, bootstrapping means resampling with replacement) to a machine learning model.

In this paper, bagging is applied to neural networks: multiple ANNs predictors (called weak learners or members) are trained on different random subsets (called bootstrap samples) of the training set. As every bootstrap sample can contain several times the same instance, each individual predictor is biased and therefore its final prediction error will be higher than the one of a unique model trained on the original training set. However, both bias and variance are reduced when we aggregate all the members together. Aggregation is simply taking the mean of the predictions of all predictors (see equation (2)).

In this study, we consider 15 individual predictors. All of these members are ANNs with identical architecture which was obtained after conducting a hyperparameter study (using GridSearch method from Scikit-Learn library and given a validation set).

The final prediction error is calculated according to the following equation:

$$MAE = \frac{1}{N}\sum\nolimits_{y_i \in test} |y_i - \hat{y_i}| \; ; \qquad \hat{y_i} = \frac{1}{N_b}\sum\nolimits_{y_i \in test} \widehat{y_{ib}} \tag{2}$$

where MAE stands for Mean Averaged Error, calculated over each example present in the test set, N is the total number of examples in the test set, N_b is the number of members (15) and $\widehat{y_{ib}}$ is a single prediction made by a unique weak learner.

To summarize, the steps of the algorithm used to apply bagging are listed below:

1. Decompose the total dataset in a training set and a test set (80-20%).
2. For $i = 1$ to $i = N_b = 15$:
   - Create a subset of the training set (same size) using the bootstrapping procedure;
   - Train an individual member on this bootstrap sample;
3. Calculate an average prediction from each weak learner given the test set using equation (2).

The architecture found for the neural network is constituted of 3 hidden layers each one composed of 40 neurons with the LeakyReLU activation function. Ensemble predictions are compared to the ones obtained when using only one neural network trained on the whole training set (called "Unique model"). Finally, each neural network is trained over 150 epochs.

3 RESULTS AND DISCUSSION

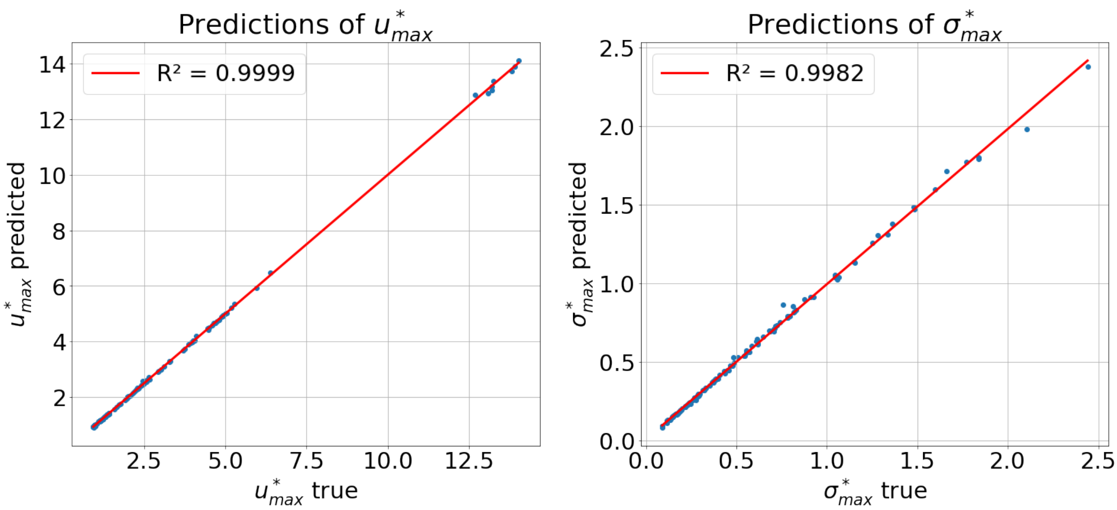

Figure 1. Predictions using the bagging technique.

Figure 1 shows the predictions calculated by the ensemble model in regards of the "true" value obtained by the numerical simulation. Figure 2 shows the evolution of the predictions according the number of members taken into account in the aggregation (orange continuous curve), every individual weak learner prediction (blue points) and the unique model prediction (red dotted line).

Several observations can be made: (i) The model succeeds in obtaining accurate predictions of non-linear elastoplastic response with only few data points; (ii) Bias (final error) is diminished when using bagging compared to the predictions made by the unique model. It can be noticed that the prediction errors of the weak learners are higher than the one of the unique model which was expected since every bootstrap sample is biased; (iii) Model predictions are more stable with bagging: about 7 members are sufficient to give accurate and reliable predictions compared to every weak learner. Variance is therefore reduced; (iv) Training the model took a few minutes and predictions took a few seconds. It may be convenient in practice and serve as a rapid estimation tool;

However, some limits and drawbacks appear: (i) There is a zone that presents a lack of data for the displacement: this requires to extend the dataset; (ii) As mentioned in the introduction, the amount of data needed to obtain good accuracy is problem dependent: if a more complex model is used for the ground, it is expected that the size of the dataset should be larger to achieve the same accuracy (Furtney et al. 2022); (iii) This model works only for interpolation: testing parameters outside the range of values used for training may result in high error outliers. (iv) Since it is based on a synthetic dataset, the model can only be as good as the numerical model.

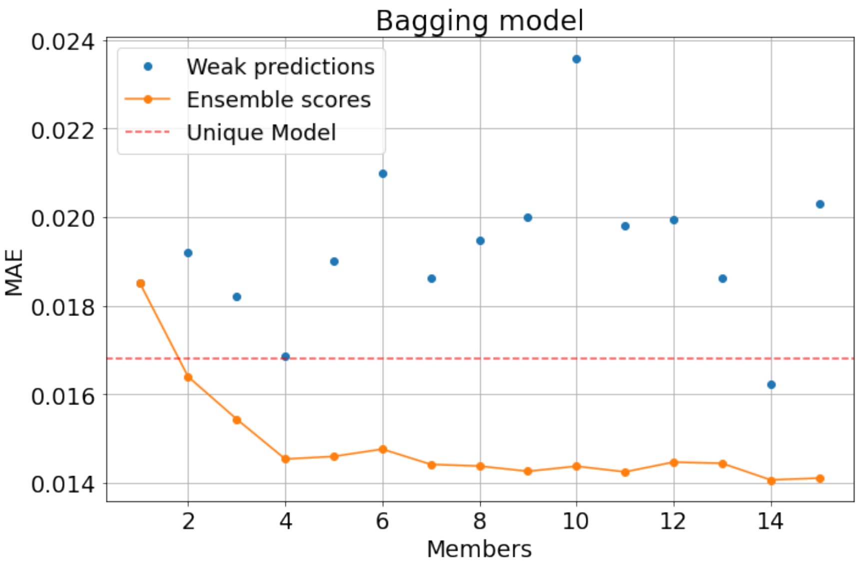

Figure 2. Ensemble predictions in regards of the number of members versus each weak learner prediction.

4 CONCLUSION

An ensemble machine learning approach for training artificial neural networks is presented in order to estimate maximum stresses/displacements at the wall a supported tunnel excavated in an elastoplastic ground. The results show that the model performs well with the small dataset used in this study and can capture the complex behavior of the ground. Based on a synthetic dataset, the ensemble method is able to make accurate and reliable predictions of both the maximum hoop stress and displacement that occur in the tunnel. Both variance and bias are reduced compared to a unique model. Thus, combining multiple ANNs could be considered as a useful alternative to three-dimensional modeling as it can be used as a quick and reliable estimation tool, which is of prime interest for applications in the engineering field.

However, high prediction errors could be obtained if the input parameters do not belong to the range used during training. Besides, one must be aware that the amount of data needed for training depends strongly on the complexity of the problem to be solved. Still, as they do not request high computational cost, neural networks could serve as surrogate models and help engineers as a complement to numerical simulations. Furthermore, this analysis can be extended by taking into account the time-dependent behavior of the ground.

REFERENCES

Adoko, A.C., Jiao, Y.Y., Wu L., Wang, H., Wang, Z.H. 2013. Predicting tunnel convergence using MARS and ANN. *Tunn and Undergr Space Technol* 38, pp. 368–76. DOI: 10.1016/j.tust.2013.07.023

Bernaud, D., Benamar, I. & Rousset, G. 1994. La nouvelle méthode implicite pour le calcul des tunnels dans les milieux élasto-plastiques et viscoplastiques. *Rev Franç Géotech* 68, pp. 3–19. DOI:10.1051/geotech/1994068003

Bishop, C.M. 2006. Pattern Recognition and Machine Learning. *New York: Springer*.

Bonini, M., Debernardi, D., Barla, M. & Barla, G. 2009. The mechanical behaviour of clay shales and implications on the design of tunnels. *Rock Mech. Rock Eng* 42 (2), pp. 361–388. DOI: 10.1007/s00603-007-0147-6

Breiman, L. 1996. Bagging predictors. *Mach Learn* 24 (2), pp. 123–40. DOI: 10.1007/BF00058655

Broms B.B. & Bennermark, H. 1967. Stability of clay at vertical openings. *J Soil Mech Found Div 93*, (1), pp. 71–94. DOI: 10.1061/JSFEAQ.0000946

Cantieni, L. & Anagnostou, G. 2009. The Effect of the Stress Path on Squeezing Behavior in Tunneling. *Rock Mech. Rock Eng*.42 (2), pp. 289–318. DOI: 10.1007/s00603-008-0018-9

Chang, X., Wang, H., Zhang, Y., Wang, F. & Li Z. 2022. Bayesian prediction of tunnel convergence combining empirical model and RVM. *Measurement* 188, 110621. DOI: 10.1016/j.measurement.2021.110621

De La Fuente, M., Taherzadeh, R., Sulem, J., Nguyen, X.S. & Subrin, D. 2019. Applicability of the convergence-confinement method to full-face excavation of circular tunnels with stiff support system. *Rock Mech and Rock Eng* 52 (7) pp 2361–2376. DOI: 10.1007/s00603-018-1694-8

De La Fuente, M., Berro, F., Khodr, S., Taherzadeh, R. & Sulem, J. 2021. Effect of the lay distance of a stiff support on the applicability of the Convergence-Confinement method. In: *IOP Conference Series: Earth and Environmental Science* 833 (1), 012168. DOI: 10.1088/1755-1315/833/1/012168

Eisenstein, Z. & Branco, P. 1991. Convergence-confinement method in shallow tunnels. *Tunn Undergr Space Technol* 6 (3) pp 343–346. DOI: 10.1016/0886-7798(91)90146-U

Fei, J., Wu, Z., Sun, X., Su, D. & Bao, X. 2021. Research on tunnel engineering monitoring technology based on BPNN neural network and MARS ML regression algorithm. *Neur Comput & Applic* 33 (1), pp. 239–55. DOI: 10.1007/s00521-020-04988-3

Feng, X., Jimenez, R., Zeng, P. & Senent, S. 2019. Prediction of time-dependent tunnel convergences using a Bayesian updating approach. *Tunnelling and Underground Space Technology* 94, 103118. DOI: 10.1016/j.tust.2019.103118

Furtney, J.K., Thielsen, C., Fu, W. & Le Goc, R. 2022. Surrogate Models in Rock and Soil Mechanics: Integrating Numerical Modeling and Machine Learning. *Rock Mech Rock Eng*. 2022 55 (5), pp. 2845–59. DOI: 10.1007/s00603-021-02720-8

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D. & Ozair, S., Courville, A. & Bengio, Y. 2014. Generative Adversarial Nets. *Advances in Neural Information Processing Systems*. DOI: 10.48550/arXiv.1406.2661

Jong, S.C., Ong D.E.L. & Oh, E. 2021. State-of-the-art review of geotechnical-driven artificial intelligence techniques in underground soil-structure interaction. *Tunn and Undergr Space Technol* 113, pp. 103946. DOI: 10.1016/j.tust.2021.103946

Liu, L., Zhou, W. & Gutierrez, M. 2022. Effectiveness of predicting tunneling-induced ground settlements using machine learning methods with small datasets. *Journal of Rock Mech and Geotech Eng* 14 (4), pp. 1028–41. DOI: 10.1016/j.jrmge.2021.08.018

Liu, Y., Sulem, J., Subrin, D., Tran-Manh, H. & Humbert, E. 2021. Time-Dependent Behavior of Saint-Martin-La-Porte Exploratory Galleries: Field Data Processing and Numerical Modeling of Excavation in Squeezing Rock Conditions. *Int J Geomech* 21 (12), 04021239. DOI: 10.1061/(ASCE)GM.1943-5622.0002200

Mahdevari, S. & Torabi, S.R. 2012. Prediction of tunnel convergence using Artificial Neural Networks. *Tunn and Undergr Space Technol* 28, pp. 218–28. DOI: 10.1016/j.tust.2011.11.002

Nguyen-Minh, D. & Guo, C. 1996. Recent progress in convergence-confinement method. In: *ISRM International Symposium – EUROCK 96*, Turin, Italy, September, 1996, pp 855–860.

Panet, M. & Sulem, J. 2022. Convergence-Confinement Method for Tunnel Design. *Cham: Springer International Publishing (Springer Tracts in Civil Engineering).* DOI: 10.1007/978-3-030-93193-3

Satici, Ö. & Topal, T. 2020. Prediction of tunnel wall convergences for NATM tunnels which are excavated in weak-to-fair-quality rock masses using DT tech and rock mass strength params. *SN Appl Sci* 2, 546, DOI: 10.1007/s42452-020-2311-5

Tran Manh,H., Sulem, J., Subrin, D. & Billaux, D. 2015. Anisotropic Time-Dependent Modeling of Tunnel Excavation in Squeezing Ground. *Rock Mech Rock Eng* 48 (6), pp.2301–17. DOI: 10.1007/s00603-015-0717-y