Determination of earth pressure balance tunnel-related maximum surface settlement: a multivariate adaptive regression splines approach

Goh, Anthony Teck Chee; Zhang, Wengang; Zhang, Yanmei; Xiao, Yang; Xiang, Yuzhou

2016

Goh, A. T. C., Zhang, W., Zhang, Y., Xiao, Y., & Xiang, Y. (2016). Determination of earth pressure balance tunnel-related maximum surface settlement: a multivariate adaptive regression splines approach. *Bulletin of Engineering Geology and the Environment*, in press.

<https://hdl.handle.net/10356/84586>

<https://doi.org/10.1007/s10064-016-0937-8>

© 2016 Springer-Verlag Berlin Heidelberg. This is the author created version of a work that has been peer reviewed and accepted for publication by *Bulletin of Engineering Geology and the Environment*, Springer Berlin Heidelberg. It incorporates referee's comments but changes resulting from the publishing process, such as copyediting, structural formatting, may not be reflected in this document. The published version is available at: [<http://dx.doi.org/10.1007/s10064-016-0937-8>].

Downloaded on 26 Apr 2023 18:50:31 SGT

[Click here to view linked References](#)

Determination of EPB tunnel-related maximum surface settlement: A multivariate adaptive regression splines approach

Goh ATC³, Zhang Wengang^{123*}, Zhang Yanmei⁴, Xiao Yang¹², Xiang Yuzhou²

¹Key Laboratory of New Technology for Construction of Cities in Mountain Area, Chongqing University, Ministry of Education, Chongqing 400045, China

²School of Civil Engineering, Chongqing University, Chongqing 400045, China

³School of Civil and Environmental Engineering, Nanyang Technological University, 639798, Singapore

⁴School of Mechanical and Aerospace Engineering, Nanyang Technological University, 639798, Singapore

Abstract: A major consideration in urban tunnel design is to estimate the ground movements and surface settlements associated with the tunnelling operations. Excessive ground movements may result in damage to adjacent buildings and utilities. Numerous empirical and analytical solutions have been proposed to relate the shield tunnel characteristics and surface/subsurface deformation. Also, numerical analyses, either 2-D or 3-D, have been used for such tunnelling problems. However, substantially fewer approaches have been developed for earth pressure balance (EPB) tunnelling. Based on instrumented data on ground deformation and shield operation from three separate EPB tunnelling projects in Singapore, this paper utilizes a multivariate adaptive regression splines (MARS) approach to establish relationships between the maximum surface settlement and the major influencing factors including the operation parameters, the cover depth and the ground conditions. Since the method has the ability to map input to output patterns, MARS enables one to map all influencing parameters to surface settlements. The main advantages of MARS over other soft computing techniques such as ANN, RVM, SVM and GP are its capacity to produce simple, easy-to-interpret model, its ability to estimate the contributions of the input variables, and its computational efficiency.

Keywords: EPB shield; surface settlement; multivariate adaptive regression splines; shield tunnelling; operational parameters

*Corresponding author, Professor, E-mail: cheungwg@126.com;

Introduction

Due to population growth and urbanization, there is an increasing demand for the construction of tunnels for mass rapid transportation services. Mechanized excavations using earth pressure balance (EPB) have been successfully applied around the world, especially in urban environments where there is less surface space available. Ground movements and surface settlements associated with shield tunnelling operations are a major concern in the design of tunnels in urban areas as excessive movements can damage nearby buildings and utilities.

Generally, there are three major classes of ground settlement prediction models: empirical and analytical methods, numerical models and the soft computing approaches driven by the measured data from case histories. These soft computing techniques include the artificial neural networks (ANNs), the support vector machines (SVMs), the relevance vector machines (RVMs), the Gaussian processes (GP), the regression method, the decision tree method, the adaptive neuro-fuzzy inference system (ANFIS), and the gene expression programming (GEP).

In the early stage of design, the designer bases his estimations on previous experience and may use simple empirical equations, such as those described by Peck (1969), Attewell and Farmer (1974), Atkinson and Potts (1977), Yoshikoshi et al. (1978), O'Reilly and New (1982), Attewell et al. (1986), Mair et al. (1993), or simple equations based on the theory of elasticity (Uriel and Sagaseta 1989). Recently, several works have extended Peck's formula to model the surface settlement of twin- and quadruple-tube tunnels (Chen et al. 2012, Gui and Chen 2013). Analytical solutions were developed by Sagaseta (1987), Verruijt and Booker (1996), Loganathan and Poulos (1998), Bobet (2001), Chou and Bobet (2002), Park (2005). However, the final design generally requires more accurate stress and deformation analyses and usually the finite-difference or finite-element methods may therefore be used, as discussed by Rowe and Lee (1992). Moreover, there are doubts as to the accuracy of some of these methods since they fail to take into account all the relevant factors, which include many shield operational parameters that all concurrently influence ground settlement (Kim et al., 2001).

With the rapid development of computer technology and numerical algorithms, numerical models have been widely used in tunnelling projects (Rowe and Lee 1992, Swoboda and Abu-Krishna 1999, Addenbrooke and Potts 2001, Mroueh and Shahrour 2002, Kasper and

Meschke2004, Callari 2004, Ng and Lee 2005, Ocak 2009, Chakeri et al. 2010, Ercelebi et al. 2011, Lambrughi et al. 2012, Gong et al. 2014, Huang et al. 2015). Recently, Fahimifara and Zareifard (2013) proposed an analytical-numerical approach to simulate the response of tunnels under different hydro-mechanical conditions. Chen et al. (2014) presented a full 3D seismic analysis on a water conveyance tunnel. Talebinejad et al. (2014) investigated the surface and subsurface displacements due to multiple tunnels excavation in the Tehran region using a full 3D finite difference analysis with special attention paid to the effect of subsequent tunnelling on the support system. Nevertheless, the implementation of a numerical model is relatively complex, particularly when mechanized processes of shield excavation are considered. In addition, the predictive performance of numerical models significantly depends on the model describing the soil behaviour (Karakus and Fowell 2003). Detailed information on soil properties, which is required for simulation, is lacking or unavailable in many cases, and thus building a practical constitutive soil model for tunnelling-induced settlement prediction is rather difficult (Karakus and Fowell 2005). Moreover, the formulation of conventional finite elements has difficulties in capturing the onset of strain localization and its propagation from the tunnel up to the ground surface, and considering the coupled relationship between the soil response and fluid flow is of great importance (Callari 2004; Callari et al. 2010). Even with powerful numerical tools, significant effort and computational time are needed to correlate ground movements with these various parameters.

During the past decade, artificial neural networks (ANNs) have also used as an alternative method for solving this complex nonlinear problem. Most of the ANN-based analyses were implemented by extracting the relationship between the influencing parameters, such as the shielding operation factors, tunnel depth and diameter, soil properties, and the induced settlement. ANN has been successfully adopted for predicting the maximum surface settlement in a number of tunnelling projects (Shi et al. 1998; Kim et al. 2001; Suwansawat and Einstein 2006; Santos and Celestino 2008; Goh and Hefney 2010; Xu and Xu 2011; Pourtaghi and Lotfollahi-Yaghin 2012; Ocak and Seker 2013; Mohammadi et al. 2015). In all these projects, the neural networks were trained and tested using instrumented data and data from the tunnel operational parameters and the geological parameters. However, when using ANNs, it is difficult to determine the optimal network architecture, and ANNs often suffer the problem of poor generalization and model interpretation.

Support Vector Machines (SVMs), which are based on statistical learning, have also been successfully applied in predicting ground settlement due to underground excavations (Neaupane and Adhikari 2006; Samui 2008; Yao et al. 2010; Jiang et al. 2011; Wang et al. 2012; Ocak and Seker 2013). SVMs also have drawbacks such as the determination of model parameters (e.g., the penalty weight C and the insensitivity parameter ε), relative high model complexity and kernel function restrictions.

Other computing methods and their ground settlement estimation applications are listed in Table 1.

Table 1. Summary of the computing methods and their use in ground settlement estimation

The purpose of this paper is to propose an alternative approach to determine the EPB tunnelling induced maximum surface settlement, using multivariate adaptive regression splines (MARS), based on three separate mass rapid transit projects in Singapore. Section 2 gives the ground conditions of the EPB projects and section 3 provides the data sets collected and the influential parameters. In section 4, the MARS methodology and its associated procedures are explained in detail. Section 5 focuses on the developed MARS model and comprehensive performance comparison between MARS and ANN. Parametric sensitivity analyses are performed in section 6, demonstrating the effect of each influencing parameter on the shielded ground settlement. Finally, a brief summary of the work and several conclusions are provided in Section 7.

Ground conditions

The major works for the construction of mass rapid transit tunnels in Singapore have been described in a number of publications (Hulme and Burchell 1999; Izumi et al. 2000; Shirlaw et al. 2003). The four major soil types encountered in this study were:

- (a) Kallang formation. This is comprised of near-normally consolidated marine clay as well as loose fluvial sands and moderately stiff fluvial clay.
- (b) Old Alluvium. This is comprised of mainly dense alluvial silty sands and clays.
- (c) Fort Canning Boulder Bed. This is comprised of a colluvial deposit of strong to very strong quartzite boulders in a hard clayey silt matrix.

(d) Jurong formation. This is comprised of primarily residual soils of clayey silt and sandy clay of medium plasticity and clayey to silty sand.

The engineering properties of these formations are described in Sharma et al. (1999) and Hulme and Burchell (1999).

The tunnel settlement data used in this paper were obtained from three separate mass rapid transit projects of Circle Line in Singapore. The locations of the projects are shown in Fig.1. Fig.2 shows the soil profiles of the three projects. Twin-bored tunnel drives using EPB machines were used (Table 2). Precast concrete tunnel linings of 5.8 m internal diameter were used throughout. Surface settlement points were installed at approximately 25 m intervals along the tunnel alignment. A total of 148 settlement patterns were obtained from the three projects.

Fig. 1. Map showing the project locations

Fig. 2. Longitudinal soil profiles of projects

Table 2. Summary of tunnel project details.

Data sets and the influential parameters

A total of 148 instrumented sections of settlement data (patterns) were obtained from the three projects. The denoted symbols, descriptions and ranges of the various parameters are shown in Table 3.

Table 3. Statistical summary of settlement data sets

There are actually numerous factors that influence the surface settlement. They can be subdivided into three major categories (Suwansawat and Einstein 2006):

(1) tunnel geometry,

(2) geological conditions, and

(3) EPB operation factors.

Tunnel geometry

The two important geometrical factors that influence the surface settlement are the tunnel diameter and the tunnel depth H (measured from the tunnel crown to the ground surface). However, since the internal diameter of the tunnels for all three projects were 5.8 m, this parameter was omitted from the MARS and neural network analysis. Fig.3 shows the tunnel cover depth H versus the measured maximum surface settlement S_t .

Fig. 3. Relationship between H and surface settlement S_t .

Geological conditions

The geological factors used as inputs for the analyses were S_1 the mean SPT N value (standard penetration test) of the soil layers above crown level up to ground surface, S_2 the average of the SPT N values at the crown, middle and invert levels, MC the average moisture content of the soil layer driven through by the tunnel machine, and E the average modulus of elasticity of the soil layer driven through by the tunnel machine. As the depth of the groundwater table did not vary significantly in these three projects, it was omitted as one of the input parameters. Figs 4-7 plot the relationship between the measured maximum surface settlement S_t and S_1 , S_2 , MC and E , respectively.

Fig. 4 Relationship between S_1 and surface settlement S_t .

Fig. 5 Relationship between S_2 and surface settlement S_t .

Fig. 6 Relationship between MC and surface settlement S_t .

Fig. 7 Relationship between E and surface settlement S_t .

EPB operation factors

The EPB operational factors used as inputs were AR the tunnel advance rate, EP the EPB earth (face) pressure, and GP the grout pressure used for injecting grout into the tail void. Fig.8a shows the tunnel advance rate AR versus the measured maximum surface settlement as a function of S_1 . Fig.8b shows the plot of the tunnel advance rate AR versus the measured maximum surface settlement as a function of S_2 . In general while it can be seen that the surface settlements are less than 20 mm for S_1 and S_2 with N values greater than 30, it is difficult to establish any well-defined relationship between the surface settlement, the advance rate, S_1 and S_2 . Fig.9a and b shows the plot of EP versus S_t as a function of S_1 and S_2 , respectively. Fig.10a

and b shows the plot of GP versus S_t as a function of S1 and S2, respectively. Generally, these figures show considerable scatter in the data. Hence a multivariate adaptive regression splines algorithm, as a statistical, nonlinear and nonparametric regression method for fitting the relationship between a set of input variables and dependent variables, was used to determine the relationship between the surface settlement, and the geological and EPB operational factors. As will be shown in the following section, the MARS approach is capable of learning from the training examples and can capture the nonlinear and complex interactions among the various inputs and the surface settlement.

Fig. 8. Relationship between S1, S2, tunnel advance rate AR and surface settlement.

Fig. 9. Relationship between S1, S2, Earth pressure EP and surface settlement.

Fig. 10. Relationship between S1, S2, Grout pressure GP and surface settlement.

Methodology of MARS

Friedman (1991) introduced MARS as a statistical method for fitting the relationship between a set of input variables and dependent variables. It is a nonlinear and nonparametric regression method based on a divide and conquer strategy in which the training data sets are partitioned into separate piecewise linear segments (splines) of differing gradients (slope). No specific assumption about the underlying functional relationship between the input variables and the output is required. The end points of the segments are called knots. A knot marks the end of one region of data and the beginning of another. MARS automatically selects which variables to use (some variables are important, others not) and the knot locations for each chosen variable. The resulting piecewise curves (known as basis functions), give greater flexibility to the model, allowing for bends, thresholds, and other departures from linear functions through the use of these knots.

MARS generates basis functions by searching in a stepwise manner. An adaptive regression algorithm is used for selecting the knot locations. MARS models are constructed in a two-phase procedure. The forward phase adds functions and finds potential knots to improve the performance, resulting in an overfit model. The backward phase involves pruning the least effective terms. An open source code on MARS from Jekabsons (2010) is used in carrying out the analyses presented in this paper.

Let y be the target output and $\mathbf{X} = (X_1, \dots, X_P)$ be a matrix of P input variables. Then it is assumed that the data are generated from an unknown “true” model. In case of a continuous response this would be

$$y = f(X_1, \dots, X_P) + e = f(\mathbf{X}) + e \quad (1)$$

in which e is the distribution of the error. MARS approximates the function f by applying basis functions (BFs). BFs are splines (smooth polynomials), including piecewise linear and piecewise cubic functions. For simplicity, only the piecewise linear function is expressed. Piecewise linear functions are of the form $\max(0, x - t)$ with a knot occurring at value t . The equation $\max(\cdot)$ means that only the positive part of $(\cdot)$ is used otherwise it is given a zero value. Formally,

$$\max(0, x - t) = \begin{cases} x - t, & \text{if } x \geq t \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

The MARS model $f(\mathbf{X})$ is constructed as a linear combination of BFs and their interactions, and is expressed as

$$f(X) = \beta_0 + \sum_{m=1}^M \beta_m \lambda_m(X) \quad (3)$$

where each $\lambda_m(X)$ is a basis function. It can be a spline function, or the product of two or more spline functions already contained in the model (higher orders can be used when the data warrants it; for simplicity, at most second-order is assumed in this paper). The coefficients β are constants, estimated using the least-squares method.

Fig. 11 shows a simple example of how MARS would use piecewise linear spline functions to attempt to fit data. The MARS mathematical equation is expressed as

$$y = -44.08 + 4.24 \times BF1 - 3.67 \times BF2 + 6.31 \times BF3 - 2.50 \times BF4 \quad (4)$$

where $BF1 = \max(0, 16 - x)$, $BF2 = \max(0, x - 10)$, $BF3 = \max(0, x - 5.5)$ and $BF4 = \max(0, 5.5 - x)$. The knots are located at $x = 5.5, 10$ and 16 . They delimit four intervals where different linear relationships are identified. It is obvious that MARS approach is good at dealing with

problems in which there are great scatters in both the explanatory independent variables and the target responses.

Fig. 11 Knots and linear splines for a simple MARS example

The MARS modelling is a data-driven process. To fit the model in Eq. (3), first a forward selection procedure is performed on the training data. A model is constructed with only the intercept, β_0 , and the basis pair that produces the largest decrease in the training error is added. Considering a current model with M basis functions, the next pair is added to the model in the form

$$\hat{\beta}_{M+1} \lambda_m(X) \max(0, X_j - t) + \hat{\beta}_{M+2} \lambda_m(X) \max(0, t - X_j) \quad (5)$$

with each β being estimated by the method of least squares. As a basis function is added to the model space, interactions between BFs that are already in the model are also considered. BFs are added until the model reaches some maximum specified number of terms leading to a purposely overfit model.

To reduce the number of terms, a backward deletion sequence follows. The aim of the backward deletion procedure is to find a close to optimal model by removing extraneous variables. The backward pass prunes the model by removing the BFs with the lowest contribution to the model until it finds the best sub-model. Thus, the BFs maintained in the final optimal model are selected from the set of all candidate BFs, used in the forward selection step. Model subsets are compared using the less computationally expensive method of Generalized Cross-Validation (GCV). The GCV equation is a goodness of fit test that penalizes large numbers of BFs and serves to reduce the chance of overfitting. For the training data with N observations, GCV for a model is calculated as follows (Hastie et al. 2009)

$$GCV = \frac{\frac{1}{N} \sum_{i=1}^N [y_i - f(x_i)]^2}{[1 - \frac{M + d \times (M - 1) / 2}{N}]^2} \quad (6)$$

in which M is the number of BFs, d is the penalizing parameter, N is the number of observations, and $f(x_i)$ denotes the predicted values of the MARS model. The numerator is the mean squared error of the evaluated model in the training data, penalized by the denominator. The denominator accounts for the increasing variance in the case of increasing model complexity. Note that $(M - 1)/2$ is the number of hinge function knots. The GCV penalizes not only the number of the model's basis functions but also the number of knots. A default value of 3 is assigned to penalizing parameter d (Friedman 1991). At each deletion step a basis function is removed to minimize Eq. (3), until an adequately fitted model is found. MARS is an adaptive procedure because the selection of BFs and the variable knot locations are data-based and specific to the problem at hand.

After the optimal MARS model is determined, by grouping together all the BFs that involve one variable and another grouping of BFs that involve pairwise interactions (and even higher level interactions when applicable), the procedure known as analysis of variance (ANOVA) decomposition (Friedman 1991) can be used to assess the contributions from the input variables and the BFs through comparing (testing) variables for statistical significance. Previous applications of MARS algorithm in civil engineering can be found in Attoh-Okine et al. (2009), Zarnani et al. (2011), Samui and Karup (2011), Lashkari (2012), Zhang and Goh (2013), Adoko et al. (2013), Goh and Zhang (2014), Khoshnevisan et al. (2015), Zhang et al. (2015).

The developed MARS model

The MARS model used to model the tunnel settlement consisted of eight inputs that represented the tunnel geometry, geological conditions and the EPB operation factors as listed in Table 3. From the 148 datasets, a total of 115 sets of settlement data samples (patterns) were randomly selected as the training data and the remaining 33 data samples were used for testing the validity of the developed MARS model. In this paper, a data sample refers to a set of input data and associated measured settlement corresponding to an instrumented section. Since for most practical cases, the serviceability limit is less than 30 mm, the testing samples that were less than 30 mm were randomly selected, while ensuring that the training and testing samples are statistically consistent. The full database is summarized and listed in Appendix A.

For the MARS model, the logarithmic values of parameters EP , E , GP and S_t were used as it was found that this substantially improves the MARS's training process. The tunnel settlement analysis using MARS adopted 16 BFs of linear spline functions with second-order interaction. The plots of the MARS predicted versus the measured settlement values are shown in Fig.12. For comparison, the same sets of training and testing patterns are analyzed using the ANN method. Based on trial and error, the optimal ANN model consist of five hidden neurons. The results indicated a fairly high coefficient of determination R^2 between the actual and predicted settlement values of 0.906 and 0.721 for the training and testing samples, respectively, compared with the fitting results of 0.873 and 0.689 using ANN at the same time. The mean average error MAE was 3.24 mm for the training samples and 3.58 mm for the testing samples. These values are smaller than the ones of 4.19 mm for the training patterns and 3.64 mm for the testing patterns obtained through ANN.

Fig. 12. Comparison of measured and MARS predicted S_t .

Table 4 displays the ANOVA decomposition of the developed MARS model. The first column in Table 4 lists the ANOVA function number. The second column gives an indication of the importance of the corresponding ANOVA function, by listing the GCV score for a model with all BFs corresponding to that particular ANOVA function removed. This GCV score can be used to evaluate whether the ANOVA function is making an important contribution to the model, or whether it just marginally improves the global GCV score. The third column provides the standard deviation of this function. The fourth column gives the number of BFs comprising the ANOVA function. The last column gives the particular input variables associated with the ANOVA function. Fig. 13 gives the plots of the relative importance of the input variables for the developed MARS model, which is evaluated by the increase in the GCV value caused by removing the considered variables from the developed MARS model. It can be observed that the earth pressure EP is the most important parameter, followed by mean moisture content MC , the advance rate AR and the grout pressure GP .

Table 4 ANOVA decomposition of MARS model.

Fig. 13. Relative importance of the input variables selected in the MARS model

Table 5 lists the BF_s and their corresponding equations. It is noted from Table 5 that of the 16 BF_s, 13 BF_s with interaction terms are integrated in this model (except BF₁, BF₅ and BF₁₀), indicating that the model is not simply additive and that interaction terms play a significantly important role. The mathematical expression for this interpretable MARS model is given by

$$\begin{aligned} \log(S_t) = & 1.4764 - 4.0327 \times BF_1 - 13.466 \times BF_2 + 3.8852 \times BF_3 + 0.7561 \times BF_4 \\ & - 0.0619 \times BF_5 - 0.3573 \times BF_6 - 12.328 \times BF_7 - 2.9412 \times BF_8 + 0.082 \times BF_9 \\ & + 0.4218 \times BF_{10} - 1.7684 \times BF_{11} - 3.8127 \times BF_{12} - 88.602 \times BF_{13} + 0.4682 \times BF_{14} \\ & - 0.0036 \times BF_{15} + 0.0256 \times BF_{16} \end{aligned} \quad (7)$$

Table 5 Basis functions and corresponding equations of MARS model for settlement prediction.

It should be noted that the model uncertainty of the developed model for EPB tunnel-related maximum surface settlement can be further characterized using the novel methods developed by Wang et al. (2012) and Juang et al. (2013). For example, Wang et al. (2012) developed an original method using the Bayesian theory to estimate the model uncertainty for the excavation induced maximum ground surface settlement. Juang et al. (2013) developed a new maximum likelihood based algorithm for characterizing the model uncertainty for the liquefaction induced ground surface settlement. This part will be further extended in future research.

Parametric sensitivity analyses

To validate the MARS EPB tunnel related settlement model, a parametric analysis was performed, with the aim to study the effect of TBP shield operation factors on the induced maximum ground settlement. This parametric sensitivity analysis investigates the response of $\text{Log}(S_t)$ predicted by the MARS model to a set of hypothetical input data generated over the ranges of the minimum and maximum data sets. One input variable was changed each time within its range while the others were kept at the average values of their entire data sets. As suggested by Alavi et al. (2011), a set of synthetic data for the single varying parameter was generated by increasing the value of this in increments. These values were presented to the MARS prediction model and $\text{Log}(S_t)$ was calculated. This procedure was repeated using another variable until the responses of the models were tested for all of the predictor variables (Alavi et al. 2011). Fig. 14 *a-c* presents the effects of the $\text{Log}(S_t)$ predictions to the variations of *AR*, *EP* and *GP*, respectively. Fig. 14a confirms the findings of Santos and Celestino (2008) that the

maximum surface settlement value increases with increase of AR and reaches a plateau. Subsequently, it then decreases as the AR continuously rises. It is obvious from Fig. 14b and c that the maximum surface settlement decreases as the earth pressure and the grout pressure become larger.

Fig. 14. Parametric analysis of the EPB tunnel related settlement MARS model

Summary and conclusions

This paper proposed a new estimation model for the maximum ground settlement, using a non-parametric, multivariate adaptive regression spline algorithm. Major findings obtained in this research include:

(1) MARS is capable of capturing the nonlinear relationships involving a multitude of variables with interaction among each other (e.g., E and GP in BF2, AR and EP in BFs 3 and 4) without making any specific assumption about the underlying functional relationship between the input variables and the response.

(2) The MARS technique is able to provide the relative importance of the input variables.

(3) The developed MARS model can be used for determination of EPB tunnel related settlement in similar ground. The interaction terms integrated in this model (Eq. 7 and Table 5) indicates that the interactions between the geological conditions and EPB operation factors play a significant role in determining the maximum settlement. Also, geological conditions essentially influences set-up of the EPB operational factors.

However, it should be noted that there are still some limitations with regard to use of the built MARS model:

(1) Since the built MARS model makes predictions based on the knot values and the basis functions, thus interpolations between the knots of design input variables are more accurate and reliable than extrapolations. Consequently, it is not recommended that the model be applied for values of input parameters beyond the specific ranges in this study.

(2) The diameter of the shield tunnel and the other operational factors such as the segment thickness and length, cutter head rotational speed and the maximum torque should also be taken into account.

(3) It should also be pointed out that the groundwater level change is excluded in this analysis since the depth of groundwater table during construction did not vary significantly in these three projects. In general, minimal lowering of the groundwater is permitted for the construction of mass rapid tunnels in Singapore.

Acknowledgements

A portion of the work was completed while the corresponding author was on sabbatical leave at Nanyang Technological University, Singapore. The corresponding author is grateful to the support by the 111 Project (No. B13024), the Program for Changjiang Scholars and Innovative Research Team in University (No. IRT_15R17) and the National Natural Science Foundation of China (No. 51608071).

References

- Addenbrooke TI, Potts DM. (2001). Finite element analysis of St. James Park greenfield reference site. In: Burland JB, Standing JR, Jardine FM (eds) Building response to tunnelling, Vol 1. Thomas Telford, London, pp 177–194.
- Adoko AC, Jiao YY, Wu L, Wang H, Wang ZH. (2013). Predicting tunnel convergence using Multivariate Adaptive Regression Spline and Artificial Neural Network. *Tunn Undergr Sp Tech* 38(3): 368-376.
- Ahangari K, Moeinossadat SR, Behnia D. (2015). Estimation of tunnelling-induced settlement by modern intelligent methods. *Soils Found*, 55(4): 737-748.
- Alavi AH, Ameri M, Gandomi AH, Mirzahosseini MR. (2011). Formulation of flow number of asphalt mixes using a hybrid computational method. *Constr Build Mater* 25, 1338-1355.
- Atkinson JH, Potts DM. (1977). Subsidence above shallow tunnels in soft ground. In: *Proc. ASCE Geotechnical Eng Div*, 59-64.
- Attewell PB, Farmer IW. (1974). Ground deformations resulting from shield tunnelling in London Clay. *Can Geotech J*, 11, 380–395.
- Attewell PB, Yeates J, Selby AR. (1986). *Soil Movements Induced by Tunneling*. Chapman & Hall, New York.

- Attoh-Okine NO, Cooger K, Mensah S. (2009). Multivariate adaptive regression spline (MARS) and hinged hyper planes (HHP) for doweled pavement performance modelling. *Constr Build Mater* 23, 3020-3023.
- Bobet A. (2001). Analytical solutions for shallow tunnels in saturated ground. *J. Eng. Mech.* 12: 1258-1266.
- Bouayad D, Emeriault F, Maza M. (2015). Assessment of ground surface displacements induced by an earth pressure balance shield tunneling using partial least squares regression. *Environ Earth Sci* 73:7603–7616.
- Callari C. (2004). Coupled numerical analysis of strain localization induced by shallow tunnels in saturated soils. *Comput Geotech*, 31, 193–207.
- Callari C, Armero F, Abati A. (2010). Strong discontinuities in partially saturated poroplastic solids. *Comput. methods Appl. Mech. Engrg*, 199, 1513–1535.
- Chakeri H, Hasanpour R, Hindistan MA, Unver B. (2010). Analysis of interaction between tunnels in soft ground by 3D numerical modeling bulletin of engineering geology and the environment. Springer, Berlin.
- Chen SL, Gui MW, Yang MC. (2012). Applicability of the principle of superposition in estimating ground surface settlement of twin- and quadruple-tube tunnels. *Tunn Undergr Space Technol*, 28, 135–149.
- Chen Z, Yu H, Yuan Y. (2014). Full 3D seismic analysis of a long distance water conveyance tunnel. *Struct Infrastruct E*, 10, 128–140.
- Chou WL, Bobet A. (2002). Predictions of ground deformations in shallow tunnels in clay. *Tunn Undergr Space Technol*, 17: 3-19.
- Dindarloo SR, Siami-Irdemoosa E. (2015). Maximum surface settlement based classification of shallow tunnels in soft ground. *Tunn Undergr Space Technol*, 49: 320-327.
- Ercelebi SG, Copur H, Ocak I. (2011). Surface settlement predictions for Istanbul Metro tunnels excavated by EPB-TBM. *Earth Sci Rev*, 62: 357-365.
- Fahimifara A, Zareifard MR. (2013). A new elasto-plastic solution for analysis of underwater tunnels considering strain-dependent permeability. *Struct Infrastruct E*, 10, 1432–1450.
- Friedman JH. (1991). Multivariate adaptive regression splines. *Ann Stat* 19: 1-141.
- Goh ATC, Hefney AM. (2010). Reliability assessment of EPB tunnel-related settlement. *Geomech Eng*, 2(1), 57-69.
- Goh ATC, Zhang WG. (2014). An improvement to MLR model for predicting liquefaction-induced lateral spread using Multivariate Adaptive Regression Splines. *Eng Geol* 170, 1-10.

- Gong W, Luo Z, Juang CH, Huang H, Zhang J, Wang L. (2014). Optimization of site exploration program for improved prediction of tunneling induced ground settlement in clays. *Comput Geotech*, 56, 69-79.
- Gui MW, Chen SL. (2013). Estimation of transverse ground surface settlement induced by DOT shield tunneling. *Tunn Undergr Space Technol*, 33, 119–130.
- Hastie T, Tibshirani R, Friedman J. (2009). *The elements of statistical learning: Data mining, inference and prediction*, 2nd edition, Springer.
- Hou J, Zhang MX, Tu M. (2009). Prediction of surface settlements induced by shield tunneling: an ANFIS model. In: Huang, Liu (Eds.), *Geotechnical Aspects of Underground Construction in Soft Ground* - Ng. Taylor & Francis Group, London, 551-554.
- Huang H, Gong W, Khoshnevisan S, Juang CH, Zhang D, Wang L. (2015). Simplified procedure for finite element analysis of the longitudinal performance of shield tunnels considering spatial soil variability in longitudinal direction. *Computers and Geotechnics*, 64, 132-145.
- Hulme TW, Burchell AJ. (1999). Tunneling projects in Singapore: an overview. *Tunn Undergr Space Technol* 14(4), 409-418.
- Izumi C, Khatri NN, Norrish A, Davies R. (2000). Stability and settlement due to bored tunneling for LTA, NEL. *Proc. Int. Conf. on tunnels and underground structures*, Singapore, 555-560.
- Jekabsons G. (2010). *VariReg: A software tool for regression modeling using various modeling methods*. Riga Technical University. <<http://www.cs.rtu.lv/jekabsons/>>
- Jiang AN, Wang SY, Tang SL. (2011). Feedback analysis of tunnel construction using a hybrid arithmetic based on support vector machine and particle swarm optimisation. *Automat Constr*, 20, 482–489.
- Juang CH, Ching J, Wang L, Khoshnevisan S, Ku CS. (2013). Simplified procedure for estimation of liquefaction-induced settlement and site-specific probabilistic settlement exceedance curve using cone penetration test (CPT). *Can Geotech J*, 50, 1055–1066.
- Karakus M, Fowell RJ. (2003). Effects of different tunnel face advance excavation on the settlement by FEM. *Tunn Undergr Space Technol*, 18, 513–523.
- Karakus M, Fowell RJ. (2005). Back analysis for tunnelling induced ground movements and stress redistribution. *Tunn Undergr Space Technol*, 20, 514–524.
- Kasper T, Meschke G. (2006). A numerical study of the effect of soils and grout material properties and cover depth in shield tunneling. *Comput Geotech* 33(4–5):234–247.
- Khoshnevisan S, Juang H, Zhou YG, Gong W. (2015). Probabilistic assessment of liquefaction-induced lateral spreads using CPT-Focusing on the 2010–2011 Canterbury earthquake sequence. *Eng Geol* 192:113-128

- Kim CY, Bae GJ, Hong SW, Park CH, Moon HK, Shin HS. (2001). Neural network based prediction of ground settlements due to tunneling. *Comput Geotech* 28(6-7), 517-547.
- Lambrugh A, Medina Rodriguez L, Castellanza R. (2012). Development and validation of a 3D numerical model for TBM-EPB mechanised excavations. *Comput Geotech*, 40: 97-113.
- Lashkari A. (2012). Prediction of the shaft resistance of non-displacement piles in sand. *Int J Numer Anal Met* 37, 904-931.
- Loganathan N, Poulos HG. (1998). Analytical prediction for tunneling-induced ground movements in clay. *J Geotech Geoenviron*, 124(9): 846-856.
- Mair RJ, Taylor RN, Bracegirdle A. (1993). Subsurface settlement profiles above tunnels in clays. *Geotechnique* 43(2), 315-320.
- Mohammadi SD, Naseri F, Alipour S. (2015). Development of artificial neural networks and multiple regression models for the NATM tunnelling-induced settlement in Niayesh subway tunnel, Tehran. *Bull Eng Geol Environ*, 74:827–843.
- Mroueh H, Shahrour I. (2002). Three-dimensional finite element analysis of the interaction between tunneling and pile foundations. *Int J Numer Anal Met*, 26: 217-230.
- Neaupane KM, Adhikari NR. (2006). Prediction of tunneling induced ground movement with the multi-layer perceptron. *Tunn Undergr Space Technol*, 21, 151–159.
- Ng CW, Lee GTK. (2005). Three-dimensional ground settlements and stress-transfer mechanisms due to open-face tunneling. *Can Geotech J*, 42: 1015-1029.
- O'Reilly MP, New BM. (1982). Settlement above tunnels in the United Kingdom, their magnitude and prediction. In *Proceedings of Tunnelling82*, Brighton London: 173–181.
- Ocak I. (2009). Environmental effects of tunnel excavation in soft and shallow ground with EPBM: the case of Istanbul. *Environ Earth Sci* 59(2):347–352.
- Ocak I, Seker SE. (2013). Calculation of surface settlements caused by EPBM tunneling using artificial neural network, SVM, and Gaussian processes. *Environ Earth Sci* 70:1263–1276.
- Park KH. (2005). Analytical solution for tunneling-induced ground movement in clays. *Tunn Undergr Space Technol*, 20(3): 249-261.
- Peck RB. (1969). Deep excavations and tunneling in soft ground. 7th ICSMFE, state of the art report, Mexico, 225-290.
- Pourtaghi A, Lotfollahi-Yaghin MA. (2012). Wavenet ability assessment in comparison to ANN for predicting the maximum surface settlement caused by tunnelling. *Tunn Undergr Space Technol*, 28, 257–271.
- Rowe RK, Lee KM. (1992). An evaluation of simplified techniques for estimating three dimensional undrained ground movements due to tunneling in soft soils. *Can Geotech J* 29, 39-52.

- Samui P. (2008). Support vector machine applied to settlement of shallow foundations on cohesionless soils. *Comput Geotech*, 35, 419– 427.
- Samui P, Karup P. (2011). Multivariate adaptive regression spline and least square support vector machine for prediction of undrained shear strength of clay. *IJAMC* 3(2), 33-42.
- Santos OJ.Jr., Celestino TB. (2008). Artificial neural networks of Sao Paulo subway tunnel settlement data. *Tunn Undergr Space Technol* 23, 481-491.
- Segaseta C. (1987). Analysis of undrained soil deformation due to ground loss. *Geotechnique* 37: 301-320.
- Sharma JS, Chu J, Zhao J. (1999). Geological and geotechnical features of Singapore: an overview. *Tunn Undergr Space Technol* 14(4), 419-431.
- Shi J, Ortigao JAR, Bai J. (1998). Modular neural networks for predicting settlements during tunneling. *Geotech Geoenviron Eng* 124(5), 389-395.
- Shirlaw JN, Ong JCW, Rosser HB, Tan CG, Osborne NH, Heslop PE. (2003). Local settlements and sinkholes due to EPB tunneling. *Geotechnical Engineering, ICE* 156(GE4), 193-211.
- Suwansawat S, Einstein HH. (2006). Artificial neural networks for predicting the maximum surface settlement caused by EPB shield tunneling. *Tunn Undergr Space Technol* 21(2), 133-150.
- Swoboda G, Abu-Krish A. (1999). Three dimensional numerical modeling for TBM tunneling in consolidated clay. *Tunn Undergr Space Technol*, 14(3): 327-333.
- Talebinejad A, Chakeri H, Moosavi M, Özçelik Y, Ünver, B, Hindistan, M. (2014). Investigation of surface and subsurface displacements due to multiple tunnels excavation in urban area. *Arab J Geosci*, 7:3913–3923.
- Uriel AO, Sagaseta C. (1989). Selection of design parameters for underground construction. *Proc. 12th ICSMFE, Rio de Janeiro, Brazil, Balkema* 4, 2521-2551.
- Verruijt A, Booker JR. (1996). Surface settlement due to deformation of a tunnel in an elastic half place. *Geotechnique* 46(4): 753-756.
- Wang DD, Qiu GQ, Xie WB, Wang Y. (2012). Deformation prediction model of surrounding rock based on GA-LSSVM-markov. *Nat Sci* 4(2):85–90.
- Wang F, Gou BC, Qin YW. (2013). Modeling tunneling-induced ground surface settlement development using a wavelet smooth relevance vector machine. *Comput Geotech* 54: 125-132.
- Wang L, Ravichandran N, Juang CH. (2012). Bayesian updating of KJHH model for prediction of maximum ground settlement in braced excavations using centrifuge data. *Comput Geotech*, 44, 1–8.
- Xu J, Xu Y. (2011) Grey correlation-hierarchical analysis for metro-caused settlement. *Environ Earth Sci* 64(5):1246–1256.

Yao BZ, Yang CY, Yao JB, Sun J. (2010). Tunnel surrounding rock displacement prediction using support vector machine. *Int J Comput Int Sys* 3(6):843–852.

Yoshikoshi W, Osamu W, Takagaki N. (1978). Prediction of ground settlements associated with shield tunnelling. *Soils Found*, 18, 47–59.

Zarnani S, El-Emam M, Bathurst RJ.(2011).Comparison of numerical and analytical solutions for reinforced soil wall shaking table tests. *Geomech Eng* 3(4), 291-321.

Zhang WG, Goh ATC.(2013). Multivariate adaptive regression splines for analysis of geotechnical engineering systems. *Comput Geotech*48, 82-95.

Zhang WG, Goh ATC, Zhang YM, Chen YM, Xiao Y. (2015). Assessment of soil liquefaction based on capacity energy concept and multivariate adaptive regression splines. *Eng Geol* 188: 29-37.

Fig. 1 Map showing the project locations (Google map)

(a)

(b)

(c)

Fig. 2 Longitudinal soil profiles of projects

Fig. 3 Relationship between H and surface settlement S_t .

Fig. 4 Relationship between S1 and surface settlement S_t .

Fig. 5 Relationship between S2 and surface settlement S_t .

Fig. 6 Relationship between MC and surface settlement S_t .

Fig. 7 Relationship between E and surface settlement S_t .

(a)

(b)

Fig. 8 Relationship between S1, S2, tunnel advance rate AR and surface settlement.

(a)

(b)

Fig. 9 Relationship between S1, S2, Earth pressure EP and surface settlement.

(a)

(b)

Fig. 10 Relationship between S1, S2, Grout pressure GP and surface settlement.

Fig. 11 Knots and linear splines for a simple MARS example

Fig. 12 Comparison of measured and MARS predicted S_t .

Fig. 13 Relative importance of the input variables selected in the MARS model

(a) $\text{Log}(S_t)$ versus AR

(b) $\text{Log}(S_t)$ versus EP

(c) $\text{Log}(S_t)$ versus GP

Fig. 14 Parametric analysis of the EPB tunnel related settlement MARS model

Table 1. Summary of the computing methods and their use in ground settlement estimation

Computing methods	References
Relevance vector machines (RVMs)	Wang et al. 2013
Gaussian Processes (GP)	Ocak and Seker 2013
Partial least squares regression method	Bouayad et al. 2015
Decision tree method	Dindarloo and Siامي-Irdemoosa 2015
Adaptive Neuro-Fuzzy Inference System (ANFIS)	Hou et al. 2009
Gene expression programming (GEP)	Ahangari et al. 2015

Table 2. Summary of tunnel project details.

Contract	C705	C823	C825
TBM manufacture	Hitachi-Zosen	Hitachi-Zosen	Herrenknecht
Drive length (km)	1.3	3.3	1.5
Tunnel drive (No.)	2	2	2
Outside diameter (m)	6.44	6.63	6.58
Internal diameter (m)	5.8	5.8	5.8
Stations	Boon Keng&PotongPasir	Mountbatten, Dakota and PayaLebar	DhobyGhaut, Bras Basah, Esplanade and Promenade
Geology (General description)	Mostly Old Alluvium with marine clay	Fill overlying Kallang formation and Old Alluvium	Soft marine clay, Old Alluvium, Fort Canning Boulder bed, Jurong formation

Table 3. Statistical summary of settlement data sets

Parameter and description	Min	max	ave	Stdev
Cover H (m)	8.5	30	17.5	4.3
Advance rate AR (mm/min)	9.5	52.1	30.8	10.9
Earth pressure EP (kPa)	11	370	193.6	81.5
Mean SPT above crown level S1 (blows/300 mm)	0.66	80.33	27.9	28.2
Mean tunnel SPT S2 (blows/300 mm)	0	100	57.0	41.8
Mean moisture content MC (%)	5.95	66.48	27.1	18.7
Mean soil elastic modulus E (MPa)	5	120	72.9	50.8
Grout pressure GP (kPa)	27.7	700	258.6	154.9
Surface settlement (mm)	0.2	98.5	13.6	17.0

Table 4 ANOVA decomposition of MARS model.

Function No.	GCV	STD	#basis	Variable(s)
1	42.591	3.966	1	<i>EP</i>
2	56.969	4.653	2	<i>MC</i>
3	36.700	2.790	2	<i>H&SI</i>
4	7.600	1.107	1	<i>H&E</i>
5	20.900	2.227	1	<i>H&GP</i>
6	92.566	5.905	2	<i>AR&EP</i>
7	31.720	2.562	1	<i>EP&MC</i>
8	4.610	0.919	1	<i>EP&E</i>
9	47.800	4.231	2	<i>EP&GP</i>
10	12.800	1.752	1	<i>MC&E</i>
11	11.260	1.643	2	<i>E&GP</i>

Table 5 Basis functions and corresponding equations of MARS model for settlement prediction.

Basis function	Equation
BF1	$\max(0, \log(EP) - 2.0492)$
BF2	$\max(0, \log(E) - 1.8921) \times \max(0, \log(GP) - 2.1761)$
BF3	$BF1 \times \max(0, AR - 21.2)$
BF4	$BF1 \times \max(0, 21.2 - AR)$
BF5	$\max(0, MC - 35.7)$
BF6	$\max(0, \log(E) - 1.8921) \times \max(0, H - 23)$
BF7	$\max(0, 2.0492 - \log(EP)) \times \max(0, \log(GP) - 2.5065)$
BF8	$\max(0, 2.0492 - \log(EP)) \times \max(0, 2.5065 - \log(GP))$
BF9	$\max(0, 35.7 - MC) \times \max(0, 2.0622 - \log(E))$
BF10	$\max(0, MC - 17.5)$
BF11	$\max(0, 1.8921 - \log(E)) \times \max(0, \log(GP) - 1.6628)$
BF12	$BF10 \times \max(0, \log(EP) - 1.9469)$
BF13	$\max(0, 1.8921 - \log(E)) \times \max(0, 2.0362 - \log(EP))$
BF14	$\max(0, 22 - H) \times \max(0, \log(GP) - 2.4771)$
BF15	$\max(0, 22 - H) \times \max(0, 25.67 - SI)$
BF16	$\max(0, 22 - H) \times \max(0, 3.25 - SI)$