Applied Mathematics and Nonlinear Sciences

<https://www.sciendo.com>

Engineering project management based on multiple regression equation and building information modelling technology

Chunyi Yu^{1*}, Mohammed Abdulrazzqa²¹Chongqing Jianzhu College, Chongqing, 400072, China²Department of Business Administration, Faculty of Administration Sciences,
Applied Science University, Al Eker, Kingdom of Bahrain

Submission Info

Communicated by Juan Luis García Guirao

Received June 16th 2021

Accepted September 24th 2021

Available online December 27th 2021

Abstract

To realise the optimisation of Building Information Modelling (BIM) technical engineering project management, the paper conducts data investigation and analysis of the technical problem which deals with a specific power grid tunnel construction. This article uses BIM to sample considerable data information for the construction and management of power grid tunnels. According to the geological factors of rock strength in construction, the fuzzy information fusion method is used to carry out the feature fusion and adaptive scheduling of management information. We will extract the characteristics of each surrounding rock category's geological BIM information association rules and use the multiple regression analysis methods to carry out the BIM information fusion and adaptive scheduling of the tunnel construction project management. The study results confirm the conclusion that the predicted scores are in good agreement with the geological scores on site.

Keywords: multiple regression method, BIM technology construction, engineering project management, surrounding rock quality prediction.

AMS 2010 codes: 62J05

1 Introduction

At present, there is no recognised Building Information Modelling (BIM) technology for the tunnel surrounding rock classification methods under construction conditions at home and abroad. Therefore, it is inappropriate to apply with the widespread application of BIM technology. Therefore, it is particularly urgent to find a suitable method of surrounding rock classification for BIM technology construction of tunnels [1].

The regression analysis method is a widely used and theoretical quantitative prediction method. It is a statistical method for dealing with multivariate dependencies. The basic idea is to analyse the interrelationship

*Corresponding author.

Email address: chuitanzhanxnhot@163.com

between the predicted object and related factors, express it with an appropriate regression prediction model (i.e. regression equation), and then predict its future state based on the mathematical model.

The regression analysis method mainly has the following three advantages. (1) This method can study the relationship between the predicted object and related factors. Thus, it grasps the substantial reason for the change of the forecast object, and the forecast result is more credible. (2) It can give the confidence interval and the confidence level of the prediction result to make the prediction result more complete and objective. (3) The method considers the correlation. It can use related mathematical statistics methods to perform statistical tests on regression equations. Therefore, the regression analysis method can have a specific ability to discriminate the turning point of the predicted object change.

BIM technology for tunnelling is susceptible to geological conditions, and changes in the excavation rate are closely related to geological conditions, especially the quality of surrounding rock [2]. This paper uses multiple regression methods to analyse the BIM technology tunnelling sample analysis to obtain the regression prediction equation to explore the types of surrounding rock based on the BIM technology tunnelling parameters.

2 Introduction to the diversion tunnel of a hydropower station in Southwest China

A hydropower station is located on the mainstream of the Yalong River at the bend of the Jinping River on the Yalong River in Liangshan Prefecture, Sichuan Province. It is the backbone power station developed by cascade on the Yalong River. The average length of the diversion tunnel is about 16.67 km. Two full-face road headers using BIM technology are used for construction [3]. The section is circular, and the hole diameter is 13 m in length. The overlying rock mass of the tunnel is generally buried at a depth of 1500–2000 m. The maximum buried depth is about 2525 m. Thus, the tunnel has considerable buried depth, a long tunnel line and a large tunnel diameter. The strata crossed are mainly the Middle Triassic Yantang Formation and Baishan Formation. The limestone and lithology are mainly marble, argillaceous banded limestone, crystalline limestone, sandstone and slate.

The essential characteristics of the surrounding rock of the deep tunnel are as follows. (1) Most of the strata crossed are hard rock, and the saturated compressive strength of the rock is 50–120 MPa. (2) The rock mass structure is dominated by layered structure and block structure, and there are integral, fragmented mosaic structures and fragmented structures at the same time. (3) Increase with buried depth. The maximum principal stress increases, but the maximum principal stress does not have a linear relationship with the buried depth. The measured value of the maximum principal stress is 60 MPa. (4) The high ground stress failure types of excavated tunnel sections mainly include rockburst, stress type collapse, and tectonic stress type collapse.

3 Selection of tunnelling parameters

We analyse the BIM technology advancement speed method and the rock mass strength theory. We can find that it is feasible to estimate the quality index of the surrounding rock through the operational parameters of BIM technology. However, it isn't easy to directly use the Innaurato method and QBIM technology method to calculate the quality index of the surrounding rock. The feasibility of this method for the classification of surrounding rock is not high [4]. A lot of engineering practice shows a strong correlation between the cutter head speed, the driving speed, the propulsion pressure, the torque, and the quality indicators of the surrounding rock in the BIM technology tunnelling process. Therefore, we predict the quality i . Therefore, this method can most intuitively reflect this change in the extraction of the excavation parameters. The cutter head speed, propulsion pressure, torque, penetration and machine utilisation are the main tunnelling parameters reflecting the quality of the rock mass. At the same time, it helps us establish a BIM technology construction tunnel based on the tunnelling parameters of the surrounding rock quality prediction model. According to the current data collection situation, based on the 3# diversion tunnels under construction, the surrounding rock quality

prediction model of the BIM technology construction tunnel suitable for a hydropower station is established.

4 Establishment of BIM technology surrounding rock quality prediction model

4.1 Sample selection and correlation analysis

The selection of samples at the initial stage of predictive model establishment is more important. Therefore, we follow the requirements of the sampling conditions and consult relevant information. At the same time, it is reasonable to choose the average value of the tunnelling parameters in each round of BIM technology as the statistical value [5]. Following the above principles and combined with the actual situation of the data, we eliminate weak points, too large parameter values, and too long holes in the original data. The article selects 807 samples from the lithological section of the Yang Formation. The structure of the sample rock mass is mainly layered, of which the principal type III surrounding rock.

The best combination of sample variables is expressed as follows: (1) There is no correlation between variables, that is, non-collinearity. (2) The correlation between the variables entering the model and the dependent variable (geological score) is good. However, in practice, the natural geological conditions vary greatly, and the heterogeneity and anisotropy of rock and soil reduce the correlation between the collected parameters and the dependent variables. Therefore, we adopt two methods when establishing the prediction equation to improve the model's prediction accuracy: (1) consider the BIM technology tunnelling parameters. (2) Combination of tunnelling parameters and geological factors. Table 1 shows the correlation coefficients between the excavation parameter samples and the geological factors. It can be seen that the correlation coefficient between tunnelling parameters and geological factors is low. Therefore, the correlation is lacking, which is in line with the best combination of sample variables.

The construction speed of BIM technology is greatly affected by the surrounding rock geological conditions. The better the geological conditions of the surrounding rock, the faster the construction and excavation speed. On the contrary, the slower. Why is the correlation between the BIM technology excavation parameters and surrounding rock geological factors not apparent, as shown in Table 1? This is determined by the characteristics of the diversion tunnel of the hydropower station [6]. The water diversion tunnel has a considerable buried depth, and the effect of in-situ stress is pronounced, and the rockburst damage is severe. The better the geological conditions during the BIM technology excavation process, the easier it is for rock bursts. This resulted in the excavation speed not being too fast in the tunnel section with better geological conditions.

Table 1 Correlation coefficient table between tunnelling parameters and geological factors

	Rotating speed	Torque	Usage rate	Propulsion	Penetration	Rock strength
Rotating speed	1					
Torque	0.21	1				
Usage rate	0.389	0.726	1			
Propulsion	0.479	0.394	0.482	1		
Penetration	-0.585	0.214	0.111	-0.367	1	
Rock strength	0.053	0.091	0.086	-0.041	-0.005	1

4.2 Multiple regression analysis

We use multiple linear regression and multiple stepwise regression to perform regression analysis on the tunnelling parameters and the combined model of tunnelling parameters and rock strength.

Definition. Let $\{Z(t, x), t \geq 0\}$ be the second-moment process under condition $x = (x_1, x_2, \dots, x_q)^T$. For a given strictly monotonically increasing function $x_t = \varphi(t), y = Z[\varphi^{-1}(x_t), x]$, the following conditions are met

$$y = a + bx_t + \sum_{i=1}^q b_i x_i + \varepsilon(x_t) \quad (1)$$

$$\varepsilon(x_t) \sim N(0, d_0 + d_1 x_t + d_2 x_t^2) \quad (2)$$

$$\text{Cov}(\varepsilon(x_t), \varepsilon(x_t^*)) = d_0 + d_1 x_t + d_2 x_t^2, x_t \leq x_t^* \quad (3)$$

In the formula, $a, b, b = (b_1, b_2, \dots, b_q)^T, d_0, d_1, d_2$ is a parameter to be determined and $d_0 \geq 0, d_2 \geq 0$ is. Then $y = y(x_t | x_1, x_2, \dots, x_q) = Y(x_t | x)$ is called a multivariate linear process. x_t is the process independent variable. $x = (x_1, x_2, \dots, x_q)^T$ is the state argument. Let $y = y(x_t | x_1, x_2, \dots, x_q)$ be a multivariate linear process. Then it can be expressed as

$$y = a + bx_t + \sum_{i=1}^q b_i x_i + \varepsilon(x_t) \quad (4)$$

$$\varepsilon(x_t) \sim N(0, \sigma_1^2 + \Delta\sigma^2) \quad (5)$$

Where:

$$\Delta\sigma^2 = (x_t - x_{t1})I\left(\frac{x_t + x_{t1}}{2}\right)\sigma_0^2 \quad (6)$$

$$I(x_t) = 1 + \theta(x_t - \bar{x}_t) \quad (7)$$

Among them $a, b, b = (b_1, b_2, \dots, b_q)^T, \sigma_0, \sigma_1, \theta$ are undetermined parameters. $\bar{x}_t$ and x_{t1} are constants and $\sigma_0 \geq 0, \sigma_1 \geq 0, \theta \geq 0, \sigma_0$ and σ_1 are not zero at the same time.

Let $y_{kl} = y(x_{tkl} | x_{1k}, x_{2k}, \dots, x_{qk})$ be the l observation on the k sample function of the multivariate linear process y . x_{tkl} is the process independent variable on the k sample function. The i value $(x_{tk1} < x_{tk2} < \dots < x_{tkn_k})$: $x_{1k}, x_{2k}, \dots, x_{qk}$ of x_t is the value of the state independent variable $x_1, x_2, \dots, x_q$ on the k -th sample function. $k = 1, 2, \dots, m; l = 1, 2, 3, \dots, n_k$ and $m > q$ respectively [7]. When the first test time is the same, we can set $x_{t1} = x_{t11} = x_{t21} = \dots = x_{tm1}$ in Eq. (6). Then the likelihood function of the multivariate linear process y is

$$L = \prod_{k=1}^m \left\{ \frac{1}{\sqrt{2\pi}\sigma_1} \exp \left[-\frac{1}{2\sigma_1^2} \left(y_{k1} - a - bx_{t1} - \sum_{i=1}^q b_i x_{ik} \right)^2 \right] \times \prod_{l=2}^{n_k} \frac{1}{\sqrt{2\pi\Delta x_{tkl} I(\bar{x}_{tkl})}} \exp \left[-\frac{(\Delta y_{kl} - b\Delta x_{tkl})^2}{2\sigma_0^2 \Delta x_{tkl} I(\bar{x}_{tkl})} \right] \right\} \quad (8)$$

According to the maximum likelihood principle and unbiasedness of the estimator, the estimators of $a, b, b, \sigma_0^2, \sigma_1^2$ and θ can be obtained as

$$\hat{a} = \bar{y}_1 - \hat{b}x_{t1} - \sum_{i=1}^q \hat{b}_i \bar{x}_1 = \bar{y}_1 - \hat{b}x_{t1} - \hat{b}x_{t1} - \hat{b}^T \bar{x} \quad (9)$$

$$\hat{b} = \frac{l_{xy}}{l_{xx}} \quad (10)$$

$$\hat{\sigma}_0^2 = \frac{1}{v} \sum_{k=1}^m \sum_{l=2}^{n_k} \frac{(\Delta y_{kl} - \hat{b}\Delta x_{tkl})^2}{(\Delta x_{tkl} I(\bar{x}_{tkl}))} = \frac{1}{v} \left(l_{yy} - \frac{l_{xy}^2}{l_{xx}} \right) \quad (11)$$

$$\hat{b} = (\hat{b}_1, \hat{b}_2, \dots, \hat{b}_q)^T = L_{xx}^{-1} l_{xy} \quad (12)$$

$$\hat{\sigma}_1^2 = \frac{1}{m-q-1} \sum_{k=1}^m \left(y_{k1} - \hat{a} - \hat{b}x_{t1} - \sum_{i=1}^q \hat{b}_i x_{ik} \right)^2 = \frac{1}{m-q-1} (l_{yy} - l_{xy}^T L_{xx}^{-1} l_{xy}) \quad (13)$$

$$E(\theta) = 2 \frac{l_{txy}}{l_{txx}} l_{txy\theta} - \frac{l_{txy}^2}{l_{txx}} l_{txx\theta} - l_{tyy\theta} + \frac{n^* + v + m - n}{v} \left(l_{tyy} - \frac{l_{txy}^2}{l_{txx}} \right) = 0 \quad (14)$$

Where v is the degree of freedom of $\hat{\sigma}_0^2$. When $\theta \neq 0$ is $v = n - m - 2$ when $\theta = 0$ is $v = n - m - 1$, $n = \sum_{k=1}^m n_k$; $n^* = \sum_{k=1}^m \sum_{l=2}^{n_k} \frac{1}{l(\bar{x}_{tkl})}$ and

$$\bar{x}_t = \frac{1}{n} \sum_{k=1}^m \sum_{l=1}^{n_k} x_{tkl} \quad (15)$$

$$\bar{x}_i = \frac{1}{m} \sum_{k=1}^m x_{ik} \quad (16)$$

$$\bar{y}_1 = \frac{1}{m} \sum_{k=1}^m y_{k1} \quad (17)$$

$$\bar{x}_{tkl} = \frac{x_{tkl} + x_{tk(l-1)}}{2} \quad (18)$$

4.2.1 Multiple linear regression

We consider two situations. (1) The article selects 5 tunnelling parameters to enter the model. (2) Six parameters of tunnelling parameters and rock strength are entered into the model. The paper uses SPSS software to perform regression analysis on a large amount of data to obtain the summary table (Table 2) and model equations. That is, 2 equations are obtained. Equation (1) is a linear equation composed of BIM technology tunnelling parameters. Equation (2) is a linear equation composed of the BIM technology tunnelling parameters and the saturated uniaxial compressive strength of the rock. The establishment of the two equations is to allow all selected parameters to enter the equation [8]. The number of variables is not significant. Although there is collinearity between the parameters, it does not affect the use of the equation.

$$T_{score} = 55.071 + 1.161r + 0.321T + 0.0928S - 0.377F - 0.180P \quad (19)$$

$$T_{score} = -21.738 + 0.99r - 0.123T + 0.043S - 0.128F - 0.055P + 1.015R_b \quad (20)$$

Table 2 Model overview table

Model		Model 1	Model 2
Correlation coefficient		0.243	0.82
Fitting coefficient		0.059	0.673
Adjusted fitting coefficient		0.053	0.67
Standard error of the estimate		6.9209	4.0857
Corrected statistics	Fitting coefficient	0.059	0.673
	F	10.078	273.827
	Degree of freedom 1	5	6
	Degrees of freedom 2	801	800
	Significant level	0	0

Equation (2) is much better than that of Eq. (1). The introduction of rock strength significantly improves the linear relationship between model dependent variables and independent variables. Use 2 regression equations to

Fig. 1 Comparison of the predicted score trend of Eqs. (1) and (2) and the geological field score.

Table 3 Equations (1), (2) score span table

Score span	Equation (1)	Equation (2)
Minimum	50.17	38.92
Max	61.46	77.46

Table 4 Equations (1), (2) points value coincidence rate table

Surrounding rock category	Field length	Equation (1)	Equation (2)
IIs	148.2	0	148.2
II	80.3	0	77.4
IIIs	30.7	30.7	13.6
III	1002.4	1002.4	787.7
IV	14.7	0	14.7
Coincidence rate	-	80.94%	81.61%

verify the selected 807 sets of data [9]. The score span of the two equations in the original 807 sets of data and the coincidence length of the scores are shown in Figure 1, Tables 3 and 4.

The following points can be obtained from Tables 3, 4, and Figure 1. (1) From the perspective of the prediction score range, the prediction score range of Eq. (2) is more extensive (38–78), and the range of Eq. (1) is smaller (50–61). Thus, there is a big difference between the surrounding rock score and the original geological score. (2) From the perspective of the coincidence rate of geological scores, the coincidence rates of Eqs. (1) and (2) are 80.94% and 81.61%, respectively. Still, equation one can only reflect type III surrounding rock, and Eq. (2) covers three types of surrounding rock: II, III, and IV. (3) The type of surrounding rock above the red

area in Figure 2 is II or IIb, and the score of Eq. (1) does not reach this height [10]. The scores of Eq. (1) in the red area are all distributed between 50 and 60. Class IV surrounding rocks below the red zone cannot be effectively predicted. This makes the prediction of the equation meaningless. Although the predicted score of Eq. (2) does not partly match the geological field score, it does not exceed the category of the surrounding rock. From the graph, the trend of scattered points is the same. Therefore, it can be seen that the accuracy and practicability of Eq. (2) are better than that of Eq. (1).

4.2.2 Multiple stepwise regression

Because of the collinearity between variables in multiple linear regression. We test each introduced independent variable one by one. When the introduced variable becomes no longer significant due to the introduction of the latter variable, we remove it to ensure that only significant variables are included in the regression equation before each new variable is introduced. This process is repeated until no significant independent variables are selected into the regression equation, and all insignificant independent variables are eliminated from the regression equation. Based on this, the BIM technology tunnelling parameters and the combination of tunnelling parameters and rock strength are respectively subjected to multiple stepwise regression [11]. Stepwise regression's calculation and implementation process is still automatically completed on the computer using SPSS software to obtain model 3 and model 4. The regression equations are:

$$T_{\text{Points}} = 51.814 + 1.55r - 0.312F + 0.089S \quad (21)$$

$$T_{\text{Points}} = -22.616 + 1.018R_b + 1.190r - 0.094F \quad (22)$$

To compare visually, we use these two equations to verify and analyze the 807 sets of data of the sample. The comparison results are shown in Tables 5, 6, and Figure 2. After comparison, Eq. (4) is better than Eq. (3).

Table 5 Equations (3), (4) score span table

Score span	Equation (3)	Equation (4)
Minimum	51.4	34.3
Max	60.8	69.7

Table 6 Equations (3) and (4) points value coincidence rate table

Surrounding rock category	Field length	Equation (3)	Equation (4)
II	80.3	0	36.9
IIs	148.2	0	93.2
IIIs	30.7	30.7	30.7
III	1002.4	1002.4	977
IV	14.7	0	14.7
Coincidence rate	-	80.94%	90.30%

4.2.3 Comparison of multiple linear regression and multiple stepwise regression

We can see that the equations that rely solely on BIM technology tunnelling parameters for regression analysis have poor goodness of fit and low reliability among the two analysis methods. This is meaningless for the prediction of surrounding rock. The regression equations obtained by the two methods are pertinent after introducing the rock strength of geological conditions under the predictable idea [12]. The range of predicted scores covers multiple categories of the surrounding rock. We compare Eq. (2) and Eq. (4), which have a better regression effect. We mainly analyse the coincidence rate of geological scores, the span of scores, and the trend of equation scores and field geological scores. See Table 7 for details.

Fig. 2 Equations (3), (4) prediction score trend and on-site geological score comparison diagram.

Table 7 Comparison table of Eqs. (2), (4)

Surrounding rock category	Field length	Equation (2)	Equation (4)
II	148.2	148.2	93.2
I	80.3	77.4	36.9
III	30.7	13.6	30.7
II	1002.4	787.7	977
IV	14.7	14.7	14.7
Coincidence rate	-	81.61%	90.30%
Score span	37.3–73.0	38.9–77.5	34.3–69.7

It can be seen from Table 7 that the coincidence rate between Eq. (4) and the scene is higher than Eq. (2). This is embodied in the anastomotic length of type III and type IIIs surrounding rocks. This shows that the predicted value of Eq. (4) is closer to the sample value based on taking into account the consistency with the overall surrounding rock category. Equation (4) is closer to the actual surrounding rock score from the span of the score, while the score of Eq. (2) is higher [13]. Eq. (4) fluctuates around the field score, with a small amplitude. Equation (2) deviates far from the scatter points of the field scores, and most of the points have higher

scores than the field scores. So overall, the trend and fit of Eq. (4) are closer to reality.

Through the example mentioned above analysis and comparison, the multiple stepwise regression eliminates the dependence between variables in calculation and analysis. Furthermore, we have selected the collinearity variables in the multiple linear regression to reduce the influence between the parameter variables. At the same time, this method also reduces the influence of variables on dependent variables and simplifies the number of variables in the model. This improves the accuracy of the equation and makes the equation more applicable.

5 Conclusion

Through multiple linear regression and multiple stepwise regression analysis of BIM technology tunnelling parameters and the combination of tunnelling parameters and geological factors, a regression prediction equation with a 90% coincidence rate with the actual geological score was obtained. The applicability is mainly verified on the geological score of each surrounding rock category. The current research results have a particular guiding significance for identifying the surrounding rock types of tunnels constructed by BIM technology. However, due to the anisotropy of natural rock masses, the parameters and geological factors are not linear. This increases the difficulty of in-depth and precise research.

Acknowledgements: This project is supported in part by the project of science and technology research programme of Chongqing Education Commission of China. (No. KJQN201904303); and in part by Humanities-Society Scientific Research Programme of Chongqing Education Commission of China (No. 21SKGH377).

References

- [1] Son, A., Shin, J. Y., & Moon, H. S. BIM Platform Resource Management for BaaS (BIM as a Service) in Distributed Cloud Computing. *Journal of KIBIM*, 2020. 10(3): 43-53.
- [2] Alasfour, F. S., & Mirzal, A. Impact of Project Management on Project Success at Engineering Firms in Kuwait. *International Journal of Project Management and Productivity Assessment (IJPMPA)*, 2021. 9(2): 68-89.
- [3] Alasfour, F. S., & Mirzal, A. Impact of Project Management on Project Success at Engineering Firms in Kuwait. *International Journal of Project Management and Productivity Assessment (IJPMPA)*, 2021. 9(2): 68-89.
- [4] Mohsenijam, A., & Lu, M. Framework for developing labour-hour prediction models from project design features: case study in structural steel fabrication. *Canadian Journal of Civil Engineering*, 2019. 46(10): 871-880.
- [5] Badi, S., Ochieng, E., Nasaj, M., & Papadaki, M. Technological, organisational and environmental determinants of smart contracts adoption: UK construction sector viewpoint. *Construction Management and Economics*, 2021. 39(1): 36-54.
- [6] Tian, J., & He, G. The five-in-one teaching mode in the teaching of engineering courses. *Computer Applications in Engineering Education*, 2020. 28(6): 1683-1695.
- [7] Yarramsetty, S., Rohullah, M. S., & Sivakumar, M. V. N. An investigation on energy consumption in residential building with different orientation: A BIM approach. *Asian Journal of Civil Engineering*, 2020. 21(2): 253-266.
- [8] Elmousalami, H. H. Comparison of artificial intelligence techniques for project conceptual cost prediction: a case study and comparative analysis. *IEEE Transactions on Engineering Management*, 2020. 68(1): 183-196.
- [9] Iglesias Martínez, M., Antonino-Daviu, J., de Córdoba, P. & Conejero, J. Higher-Order Spectral Analysis of Stray Flux Signals for Faults Detection in Induction Motors. *Applied Mathematics and Nonlinear Sciences*, 2020. 5(2): 1-14.
- [10] Ruiz-Fernández, J., Benlloch, J., López, M. & Valverde-Gascuña, N. Influence of seasonal factors in the earned value of construction. *Applied Mathematics and Nonlinear Sciences*, 2019. 4(1): 21-34.
- [11] Skitmore, M., Xiong, B., Xia, B., Masrom, A., & Newton, S. Relationship between contractor satisfaction and project management performance. *Construction Economics and Building*, 2020. 20(4): 1-24.
- [12] Daslim, S., & Mursadin, A. A Control Model for Reinforced Concrete Work Materials on the Construction of Duta Mall 2. *Eng. Technol. Appl. Sci. Res.*, 2020. 10(2): 5483-5486.
- [13] Sanni-Anibire, M. O., Zin, R. M., & Olatunji, S. O. Developing a machine learning model to predict the construction duration of tall building projects. *Journal of Construction Engineering*, 2021. 4(1): 022-036.

This page is intentionally left blank