Modeling uncertainty of activity duration in probabilistic time estimation of tunneling projects

Mohammad Mohammadi
Division of Soil and Rock Mechanics, KTH Royal Institute of Technology, Stockholm, Sweden

Johan Spross
Division of Soil and Rock Mechanics, KTH Royal Institute of Technology, Stockholm, Sweden

ABSTRACT: The PERT distribution may be a suitable distribution for modeling activities' duration in probabilistic time estimation of tunnel projects as it puts more emphasis on the mean value of the distribution. In this paper, we compared the outcome of time estimations for a tunnel, using the triangular and PERT distributions for modeling the uncertainty of activities' duration. The results indicate that the choice of the distribution affects the total estimated time considerably. In addition, the skewness of the distribution also affects the results of estimation meaning that realistic assessment of the parameters of the distributions is important.

Keywords: Time estimation, probabilistic approaches, tunneling, activity duration, production effort.

1 INTRODUCTION

Probabilistic approaches are used to account for the uncertainties when estimating time and cost of tunneling projects. In recent decades, researchers developed probabilistic time and cost estimation models for tunneling projects (Isaksson & Stille 2005; Min et al. 2008; Špačková et al. 2013; Mohammadi et al. 2022). However, most of these studies did not discuss the effect of the accuracy in the probabilistic models of the unit activity durations on the outcome of the estimation. According to Hajdu & Bokor (2016) the accuracy of three-point estimations is more important than choosing the type of the distribution of the unit activities' duration. However, we believe that the type of distribution is an important factor that affects the outcome of probabilistic time estimation in tunneling projects. In the literature of probabilistic time and cost estimation of tunneling projects, solely the three-point triangular distribution is used for modeling the uncertainty of activity durations, where experts assign the minimum, most likely (mode), and maximum times it takes to perform a unit activity. However, the PERT distribution, which also uses the same three-point parameters, can also be used to model this uncertainty. Both distributions can be skewed to either the left or right depending on the type of the unit activity. The PERT distribution puts more emphasis on the most likely value than the triangular distribution does, which according to our experience from tunneling projects is more realistic in terms of modeling the uncertainty of the unit activity durations. Thus, we investigated the effect of the type of distribution for modeling the uncertainty of unit activities' duration in time estimation of tunneling projects.

Isaksson & Stille (2005) developed a model for probabilistic time and cost estimation of tunneling projects ("the KTH model"), which was later updated by Mohammadi et al. (2022). The KTH model can be used in many types of geological conditions and for various construction methods. In order to investigate the importance of the choice of the distribution, we used in this paper the KTH model for a probabilistic time estimation of a simplified example, where we compared the use of the triangular and PERT distributions in the modelling of the uncertainty of unit activities' duration. In addition, we investigated the importance of the accuracy of the three-point estimations by looking into the effect of skewness of the distributions of the duration of the unit activities on the total time of tunneling, and discussed the practical implications of the skewness of these distributions.

2 TIME ESTIMATION USING THE KTH MODEL

2.1 Theoretical framework of the KTH model

In the KTH model (Isaksson & Stille 2005; Mohammadi et al. 2022), total tunneling time, T, is the sum of normal excavation time, T_N, and exceptional time, T_E ($T=T_N+T_E$). The T_N can be calculated using the concept of production effort, Q [h/m], i.e. the time spent for completing construction of a unit length of the tunnel. The Q for a tunnel section with unit length (l) is affected by geological conditions that are more or less unknown in the planning phase. The vector $\mathbf{x}$ is used in the model to describe these conditions. The elements of $\mathbf{x}$ are a chosen set of geotechnical characteristics such as rock quality, groundwater volume, and other properties describing the ground conditions at the site. This gives the conceptual key function which shows how the model accounts for geological conditions in estimation of construction time. The model assigns Q to be a stochastic variable which is estimated by the planning team. The production effort of all tunnel sections, $Q = [Q_1, Q_2, \dots, Q_L]$ are summed along the tunnel length L to obtain T_N:

$$T_{\mathrm{N}} = \int_L g(\mathbf{x})dl \approx \sum_{l=1}^{L} Q_l. \tag{1}$$

The Q_l are identical stochastic variables, meaning that they have the same mean μ_Q and standard deviation σ_Q. Thus T_N is also a stochastic variable. According to the central limit theorem, T_N tends to a normal distribution when it is the sum of a large number of production efforts. Thus, the mean value ($\mu_{T,\mathrm{N}}$) and standard deviation ($\sigma_{T,\mathrm{N}}$) of T_N can be obtained as:

$$\mu_{T,\mathrm{N}} = L\mu_Q, \tag{2}$$

$$\sigma_{T,\mathrm{N}} = L\sqrt{\frac{\delta}{\delta+L}}\sigma_Q, \tag{3}$$

where δ is the scale of fluctuation which accounts for spatial correlation along the tunnel length (see Vanmarcke, 1977, for more).

2.2 Practical application of the model

In practice it is not easy to describe the exact meaning of the function $Q=g[\mathbf{x}(l)]$. It is more convenient to assess the unit activities' duration for a set of classes than for a specific, exact combination of characteristics at a given location. Thus, the range of possible values of the geotechnical characteristics can be divided into intervals known as classes. For instance, for the characteristic rock mass quality, the range of possible RMR values can be divided into five classes.

The construction method is described in terms of its main production activities. For example, the main production activities for the Drill & Blast method of excavation can be pre-excavation grouting, excavation sequence, and permanent lining, each of which can be broken further down into their unit activities. For instance, the excavation sequence can be broken down into drilling, charging and blasting, ventilation, scaling, mucking, shotcrete, and rock bolt installation. Each production activity is usually affected by one or several geotechnical characteristics in **x** in such a way that the production effort of unit activities, $q_{a,j}$, are different in different classes of the relevant characteristics. The production effort of any production activity in the k^{th} class of the relevant geotechnical characteristic can therefore be obtained as (Mohammadi 2021):

$$Q_{a,\mathrm{k}} = \sum_{j=1}^{n_q^a} q_{j,a}, \tag{4}$$

where n_q^a is the number of unit activities of the a^{th} production activity. The $q_{j,a}$ are assigned by experts based on data from past similar projects, and the experts' subjective assessment, which is the main focus of this paper.

The distribution of production effort of the a^{th} production activity (Q_{a}) can be obtained as the weighted sum of the production efforts of the production activity in each class of the relevant geotechnical characteristics with respect to the probabilities of occurrence of each class:

$$Q_a = \sum_{k=1}^{n_k} p_{a,k} Q_{a,k}, \tag{5}$$

where n_k is the number of classes of the relevant geotechnical characteristic and $p_{a,k}$ are the proportion of the k^{th} class of the relevant geotechnical characteristic. The values of $p_{a,k}$ are also assigned by experts, the process of which however is not discussed in this paper. The production effort (Q) can then be obtained as:

$$Q = \sum_{a=1}^{n_a} Q_a, \tag{6}$$

where n_a is the number of main production activities that are required for the construction method. The variability in the values of $q_{a,j}$ can be modeled by using the triangular or PERT distributions. Thus, these distributions are described in section 2.3.

2.3 *The triangular and PERT distributions*

The triangular distribution is defined by three parameters: minimum (a), most likely (b), and maximum (c) values. The mean value ($\mu_1 = [a + b + c]/3$) and standard deviation ($\sigma_1 = [a^2 + b^2 + c^2 - ab - ac - bc]^{0.5}/18$) of this distribution can be obtained using its parameters (Back et al. 2000). The PERT distribution is defined by the same three parameters where the mean value ($\mu_2 = [a + 4b + c]/6$) and standard deviation ($\sigma_2 = [c - a]/6$) of this distribution can also be obtained using these parameters (Malcolm et al. 1959).

3 THE TUNNEL DESCRIPTION

We considered a simplified tunnel example for our calculations, where it is assumed that full-face excavation with the drill & blast method is used to construct a horse-shoe shaped tunnel with a width of 10 m, height of 6 m, and a length of 2000 m. The tunnel has a cross-sectional area of 50 m^2. The lithology along the whole tunnel is fine-grained sandstone which is classified as fair rock, meaning that the RMR ranges from 41 to 60, i.e. rock class III. The entire length of the tunnel is assumed to be completely dry and have an overburden varying between 100 and 150 m. The primary support system is systematic rock bolts with a spacing of 1.5*2 m and a length of 4 m, coupled with fibre-reinforced shotcrete with a thickness of 10 cm.

We limit the analysis to only one production activity, namely the excavation sequence, which in turn is broken down into its unit activities: drilling, charging and blasting, ventilation, scaling, mucking, surveying, shotcrete, and rock bolt installation. The input parameters required for calculating production effort are $p_{a,k}$ and $q_{j,a}$. The probability of occurrence of rock class III is equal to 1 while the probability of occurrence of the other rock classes are all equal to zero. A scale of fluctuation of 200 m is assumed. Moreover, it is assumed that the construction of the tunnel is completed without the occurrence of any disruptive event, i.e. $T_E = 0$.

Based on our experience as practicing engineers, we assigned the minimum, most likely and maximum times that it would take to perform each unit activity per unit length of the tunnel, as presented in Table 1. These values are the three parameters of the PERT and triangular distributions used for calculation of Q and T_N. A Monte Carlo simulation of 100000 samples was used to obtain the distribution of Q from equations 4-6 and the mean value and standard deviation of T_N from equations 2 and 3 respectively.

Usually the assessment of maximum values of all unit activities' duration is more challenging than the modes and minimum values. This is due to the fact that the maximum values do not occur regularly and thus, the experts are prone to make mistakes when assessing the maximum values (Hajdu & Bokor 2016). Consequently, to investigate the effect of the skewness of the PERT and triangular distributions on the results, in addition to the base case, the simulations are done once by increasing the maximum production effort (c in table 1) of all the unit activities by 1 h/m, and once again by reducing the maximum values in a way that all the unit activities have the symmetric distribution of production effort (the minimum and most likely values are kept unchanged).

Table 1. The assigned minimum, most likely and maximum production efforts of unit activities ($q_{j,a}$). Here a=1 as there is only one production activity, and j represents the 7 unit activities.

Unit activity, $q_{j,a}$	Drill (h/m)	Charge (h/m)	Ventilate (h/m)	Scale (h/m)	Muck (h/m)	Survey (h/m)	Support (h/m)
Min. (a)	0.67	0.39	0.07	0.14	0.54	0.10	0.56
Mode (b)	0.83	0.61	0.14	0.25	0.64	0.17	1.03
Max. (c)	1.57	1.17	0.33	0.67	1.58	0.50	1.86

4 RESULTS

The distributions of q for some unit activities are shown in Figure 1 for both the triangular and PERT distributions. The probability density in the mean value is higher when using the PERT distribution, indicating that it puts more emphasis on the mean value than the triangular distribution. The distribution of Q obtained by using the PERT and triangular distributions are shown in Figure 2. The mean value of Q when using the triangular distribution is higher than that of the PERT distribution. The mean values and standard deviations of the Q and T_N in both cases are shown in Table 2. By using the PERT distribution, the estimated mean value and standard deviation of the T_N are lower than that of the triangular distribution by 10% and 18% respectively. Thus, the choice of the type of the distribution to model the uncertainty of the unit activities' duration is important.

Table 2. The calculated mean value and standard deviation of Q and T_N, by using the triangular and PERT distributions for modeling the variability of unit activities' duration.

	Production effort, Q (h/m)		Normal time, T_N (h)	
	Mean value	Standard deviation	Mean value	Standard deviation
PERT distribution	4.14	0.38	8276	229
Triangular distribution	4.61	0.46	9216	280

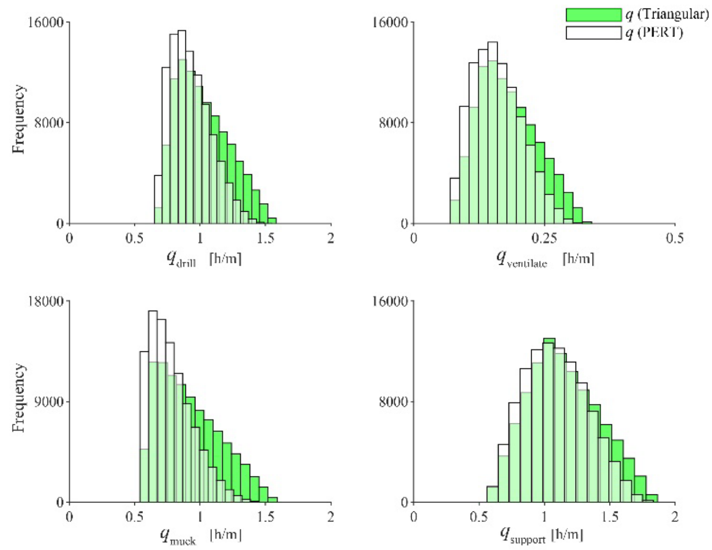

Figure 1. Modeling the uncertainty of the unit activities' duration by the triangular and PERT distributions.

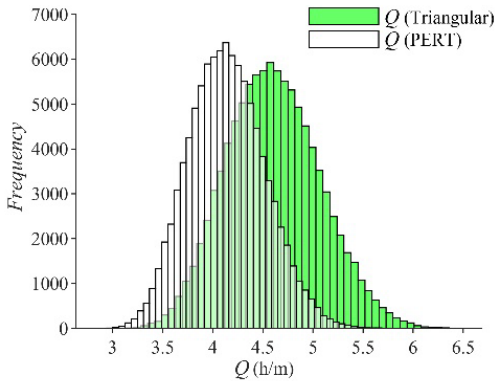

Figure 2. The distributions of Q obtained by using the PERT and triangular distributions for modeling the uncertainty of the unit activities' duration. Q is the sum of all unit activities' duration.

5 DISCUSSION

When the skewness of the production efforts of the unit activities is increased (meaning that the maximum values are increased), the estimated mean value and standard deviation of the T_N by using the PERT distribution are lower than that of the triangular distribution by 24% and 29% respectively. When the production effort of unit activities, $q_{j,a}$, have symmetrical distributions the mean values of Q and T_N are similar using both the triangular and PERT distributions while the use of triangular distributions results in slightly higher standard deviation.

The skewness of these distributions is related to the representativeness of three-point estimations. It means that if experts by mistake assign maximum values that are unrealistically large, the skewness will increase. Our conclusion that the representativeness of three-point estimations affects the outcome is in line with that of Hajdu & Bokor (2016), but contrary to their suggestion that the choice of the type of the distribution is not important, we found that the type of distribution to model the uncertainty of the unit activity durations affects the normal time, T_N, considerably.

The practical implication of the effect of skewness is that experts' ability to assign realistic maximum values (c) of the $q_{j,a}$ affects the outcome of the estimations considerably. Assigning the representative maximum values (c) of the $q_{j,a}$ are usually harder for experts than the minimum (a) and most likely (b) values (Hajdu & Bokor, 2016). Thus, it is of great importance to make sure that representative maximum values are assigned for all the unit activities. This can be achieved by using a group of experts that have experience from various similar past projects as well as accounting for the effect of the experts' bias.

6 CONCLUSIONS

The triangular distribution is often used to model the uncertainty of activity durations in probabilistic time and cost estimation models for tunneling projects. However, we believe that the PERT distribution can be a good alternative for this purpose as it puts more emphasis on the mean value than the triangular distribution. The results indicated that by using the PERT distribution the estimated mean value and standard deviation of the normal time (T_N) are lower than that of using the triangular distribution. In addition, when the skewness of the distributions of the unit activities' duration is increased, by using the PERT distribution the mean value and standard deviation of the normal time, T_N, are lower than that of using the triangular distributions. Thus, the first conclusion is that the choice of the distribution affects the total estimated time considerably. The second conclusion is that since the skewness has a noticeable effect on the outcome, the experts must pay serious attention for realistic assessment of the maximum values of production efforts of the unit activities

ACKNOWLEDGEMENTS

This research was funded by the Swedish Research Council for sustainable development (Formas, grant no. 2017-01218) and the Swedish Rock Engineering Research Foundation (BeFo, grant no. 400). The research was conducted without the involvement of funding sources.

REFERENCES

Back, W. E., Boles, W. W., & Fry, G. T. 2000. Defining triangular probability distributions from historical cost data. *Journal of Construction Engineering and Management, 126* (1), pp. 29-37.

Hajdu, M., & Bokor, O. 2016. Sensitivity analysis in PERT networks: Does activity duration distribution matter? *Automation in Construction, 65*, pp. 1-8.

Isaksson, T., & Stille, H. 2005. Model for estimation of time and cost for tunnel projects based on risk evaluation. *Rock Mechanics and Rock Engineering, 38* (5), pp. 373-398.

Malcolm, D. G., Roseboom, J. H., Clark, C. E., & Fazar, W. 1959. Application of a technique for research and development program evaluation. *Operations Research, 7* (5), pp. 646-669.

Min, S. Y., Kim, T. K., Lee, J. S., & Einstein, H. H. 2008. Design and construction of a road tunnel in Korea including application of the decision aids for tunneling–a case study. *Tunnelling and Underground Space Technology, 23* (2), pp. 91-102.

Mohammadi, M. 2021. *Probabilistic time estimation in tunnel projects.* Licentiate thesis, TRITA-ABE-DLT-2137, KTH Royal Institute of Technology, Stockholm.

Mohammadi, M., Spross, J. and Stille, H., 2022. Probabilistic Time Estimation of Tunneling Projects: The Uri Headrace Tunnel. *Rock Mechanics and Rock Engineering*, pp. 1-15. https://doi.org/10.1007/s00603-022-03022-3

Špačková, O., Šejnoha, J., & Straub, D. 2013. Probabilistic assessment of tunnel construction performance based on data. *Tunnelling and Underground Space Technology, 37*, pp. 62-78.

Vanmarcke, E.H., 1977. Probabilistic modeling of soil profiles. *Journal of the geotechnical engineering division*, *103* (11), pp. 1227-1246.