Tunnel construction time and cost estimates: from deterministic to probabilistic approaches

O. Špačková¹² & J. Šejnoha² & D. Straub¹

¹Engineering Risk Analysis Group, Technische Universität München, Germany
²Dep. of Mechanics, Fac. of Civil Engineering, Czech Technical University in Prague, Czech Republic

ABSTRACT: This paper provides an overview of existing approaches to construction time and cost estimates. While the deterministic estimates are commonly based on analysis of data (average performance of the construction in past projects), the variability of the construction performance has not been studied systematically. The inputs for the probabilistic models of time and costs presented in the past have been typically assessed based on expert judgement and the models therefore do not provide a full estimate of the uncertainty in tunnel construction. A statistical study of the construction performance data has been thus carried out. The paper shows some results of this study estimating unit time (i.e. time for construction of a tunnel segment with given unit length), the failure rate (i.e. expected number of construction failures such as cave-in collapses etc.) and the delay caused by a failure. The probabilistic estimate of tunnel construction time using data from past tunnels is then demonstrated on a case study of a 610 m long road tunnel.

1 INTRODUCTION

Estimates of tunnel construction time and costs are a fundamental part of the tunnel project planning. The construction time and costs are commonly estimated on a deterministic basis; the estimator is asked to provide one number representing the best estimate. However, the deterministically assessed construction costs and time are often underestimated: Flyvbjerg et al. (2002) show that final construction costs of tunnel and bridge construction projects are, on average, 34% above original estimates made at the time of decision to build. The study further shows that there has been no improvement over the past seventy years. The reasons for this systematic underestimation are discussed in Flyvbjerg (2006): First, people tend to “judge future events in a more positive light than is warranted by actual experience”; this psychological phenomena is called optimism bias. Second, the system of administering and financing of transport projects often motivates the people interested in realization of the project to purposely underestimate the construction costs and time, because it improves the chance of the project to be financed; this political and organizational phenomenon is denoted as strategic misrepresentation. Another detailed study of cost overruns in infrastructure projects is available in Cantarelli (2011). Probably the only way of eliminating the biases and misinterpretations is a consistent learning from the experience from past projects. However, even if the biases and misinterpretations were eliminated, construction time and cost estimates would still be subject to significant uncertainty. This uncertainty results from the uncertain geological conditions, from natural variability of the construction performance and from unpredictable influence of common factors such as quality of planning and construction management or other external factors (political, economical). This uncertainty should be quantified and communicated with the stakeholders and with the public. The need of probabilistic prediction of construction time and costs and their communication with the stakeholders has been recognized in the tunnelling community in recent years (Lombardi, 2001; Reilly, 2005; Grasso et al., 2006; Edgerton, 2008) and the demand for applicable probabilistic models is apparent.

This paper summarizes the state of the art of time and cost estimates for tunnel construction, both on the deterministic and probabilistic basis (Section 2). Because the overview of the probabilistic approaches reveals a lack of reliable inputs for probabilistic models, statistical analysis of data from constructed tunnel, which should be a basis for learning the parameters of the probabilistic models, is presented in Section 3. The effect of utilization of such inputs on tunnel construction time estimate is demonstrated in Section 4.
2 ESTIMATION OF CONSTRUCTION TIME AND COST – STATE OF THE ART

Time and cost of tunnel construction depend primarily on the following factors:
- Geological conditions (e.g. mechanical properties of the ground, frequency and orientation of discontinuities)
- Hydrological conditions
- Frequency of changes of the geological and hydrological conditions
- Cross-section area of the tunnel
- Length of the tunnel
- Inclination of the tunnel
- Depth of the tunnel/height of overburden
- Affected structures and systems (requirements on maximal deformations, protection of water systems and environment, operational constraints)
- Quality of planning and design
- Construction management and control, quality of construction works

The effect of these factors is often not exactly predictable prior to construction and it thus introduces significant uncertainties into the time and cost prediction. Two types of uncertainties can be distinguished (Isaksson and Stille, 2005): the common variability of the construction process (here including the normal performance and small disturbances of the process) and extraordinary events (construction failure), i.e. events that causes significant delay and/or financial loss.

In the following, we provide an overview of the methods for estimate and presentation of construction time/cost. The different types of estimates are illustrated in Figure 1.

The deterministic approaches (Section 2.1) only assess one value and neglect the uncertainty of the estimate. It is often not clear, which value is actually estimated, but it can be assumed that the value is close to the mode (i.e. the most frequent value in a sample), which represents some kind of ideal case. The interval and percentile estimates (Section 2.2) take into account the uncertainty of the estimates: they represent the uncertainty using a confidence interval or using a percentile (i.e. a value which will not be crossed with a certain probability). A full probabilistic analysis (Section 2.3) provides a probability distribution of the construction time/costs.

In practice, the deterministic models are still the most common ones. They are sometimes accompanied by a qualitative or semi-quantitative analysis of the risks – these methods are not discussed here, for an overview see for example (Špačková, 2012).

2.1 Deterministic analyses

The methods of time and costs estimates vary in different project phases. In the early design phase, only little information is available on the geotechnical conditions, tunnel design and construction technology. Therefore, only rough estimates based on experience and/or data from tunnels constructed in the past can be provided. Analyses of tunnel construction time and costs are available in the literature: Burbaum et al. (2005) provides a detailed guidance for tunnel construction cost estimates in early design phases, which is based on the experiences in Hessen, Germany. Kim and Bruland (2009) study the dependence of tunnel construction time on the geotechnical conditions classified using the Q-system and on the cross-section area of the tunnel. Zare and Bruland (2007) analyse the construction time and costs in conventional tunnels. These two studies are based on Norwegian experiences and a detailed simulation of the construction process. Farrokh et al. (2012) present a number of models for the prediction of the TBM penetration rate and compares the estimates of these models with data from tunnels constructed in the past. Rostami et al. (2013) study the correlation between unit cost of tunnel construction and tunnel diameter on a sample of c. 270 tunnels constructed with different technologies. Another approach for estimating tunnel construction time as a function of geology, excavation technology, support system and other factors is presented in Singh and Goel (1999).

In later planning phases, geotechnical surveys are carried out and a detailed tunnel design is prepared. Based on this information, more precise time and costs estimates are made, taking into account the particular activities of the construction process and the associated utilization of resources such as material, labour and machinery. The estimates are done by experience cost estimators and projects planners. The estimators can additionally use simulation tools for modelling of the construction processes, an overview of the tools is given for example in Jimenez (1999) and Wu et al. (2010). Applications of simulation models for tunnelling are presented for example in Zhou et al. (2008). It should be, however, noted that even the most detailed process simulation cannot ensure, that the final time/cost estimates do not include uncertainties. On the contrary, the detailed process simulation typically tend to sim-
ulate an ideal progress of the construction process and they just can assess the time/costs to optimistically (close to the mode as shown in Figure 1).

2.2 Interval and percentile estimates

Many authorities have realized in the last years that neglecting the uncertainty of the time and cost estimates is misleading. For example in the Netherlands, the estimator is asked to give an interval estimate of the construction time and cost. The accuracy of this estimate (the width of the interval) depends on the phase of the project: the interval can be wider in early phases of the project, later the accuracy of the estimate must increase. The estimates are based primarily on expert judgement.

The British authorities recommend adjusting the cost estimates for bias and risk (HM Treasury, 2003). This adjustment should be based on analysis of projects constructed in the past. Extensive statistical analyses of the construction cost escalations were made (Flyvbjerg and COWI, 2004). These allowed determining the uplifts of the deterministic estimate for different reference project classes. For example, to obtain a cost estimate of tunnel or bridge construction projects with 80% probability of not being exceeded (i.e., the 80th percentile), the deterministic expert estimate should be increased by 55%.

This study of Flyvbjerg and COWI (2004) is the first attempt known to the authors, which quantifies the uncertainty in cost estimates based on analysis of data from previous projects. The utilized approach differs from the one presented later in Section 3 in two main points: First, Flyvbjerg and COWI (2004) do not study the actual cost related to a unit length (or unit volume) of the infrastructure but they examine the probability distribution of cost overrun. In our opinion, this approach is useful for descriptive analysis of the common practice, but a direct projection of the former cost overruns onto future projects should be carefully considered. Assuming that the method will be systematically used in the practice and the practitioners will become more aware of the uncertainties in the estimates, they might start estimating even the deterministic costs more realistically. In such a case, the present study of the cost overruns will not be valid anymore. Additionally, the application of a prescribed proportional adjustment of the cost estimate disadvantages the more conservative estimators. At the end, this approach might rather motivate the estimators to even larger underestimation of costs, as they will know that their estimates will anyway be increased by several percent. Second, the specifics of individual projects (e.g., geology, geometry and layout of the infrastructure, geographical location) are not considered in the analysis of Flyvbjerg and COWI (2004). The guidance leaves considerable space for expert judgement on how to include these project specifics and the benefits of the statistical approach might thus be reduced.

2.3 Probabilistic models

A well-known model for probabilistic quantification of variability of the tunnel construction processes is the Decision Aids for Tunnelling (DAT), developed in the group of Prof. Einstein at MIT. It has been applied to several projects, an overview of them is given for example in Min (2008). DAT uses Monte Carlo (MC) simulation for probabilistic prediction of construction time, costs and consumption of resources. It takes into account the geotechnical uncertainties as well as the uncertainties in the construction process. The updating of the model predictions with observations from the construction was carried out by Haas and Einstein (2002). In the published applications of the DAT model, the coefficients of variation of the total construction time and cost are typically less than 5%. This computed uncertainty is lower than the one observed in practice, e.g., in Flyvbjerg et al. (2002).

More recent models allow one to include both the common variability of the construction process and extraordinary events: Isaksson (2002) and Isaksson and Stille (2005) suggest an analytical solution for a probabilistic estimation of tunnel construction time and cost and apply the model to a case study of the Grauhölz tunnel in Switzerland. The model considers the correlations in the construction performance and costs as well as the uncertain geotechnical conditions. Grasso et al. (2006) and Moret (2011) present an estimate of the construction time and costs using an extended version of the DAT model. The latter applies the model to a case study of a new rail line in Portugal. The model includes the correlations in the construction performance and costs. Steiger (2009) suggests a model combining Bayesian network for representation of geotechnical uncertainties and Monte Carlo simulation for modelling of the construction process. However, the model of geology is not connected with the model of construction process and a significant part of the uncertainty is thus not captured. Špačková and Straub (2013) suggest a model using Dynamic Bayesian Networks (DBNs). It models both the common variability and the extraordinary events; it includes the uncertainty in the geotechnical conditions and the dependencies in the construction performance, which are due to the influence of common factors (human, organizational and other external factors). The model also allows more efficient updating of the predictions with observations made during the construction than the MC based models. Last but not least, the DBN is a graphical tool, which allows an understandable communication of the model assumption with other experts, who are not necessarily specialists in probabilistic modelling.

Applications of the above mentioned models commonly rely on expert estimates of the input
3 STATISTICAL ANALYSIS OF DATA

Methods for statistical analysis of data from tunnels constructed in the past were introduced in (Špačková et al., 2013) and (Špačková, 2012). Here we shortly present the main findings and selected results. Only data on construction time are analysed, information on construction costs were not available.

It is proposed to categorize the performance of the excavation process in three classes: (1) Normal performance, where the excavation round is commonly finished within one day. (2) Small disturbances of the process associated with delays in the order of a few days. (3) Extraordinary events, corresponding to cases when the excavation stopped for longer than 15 days.

The unit cost and activity durations or advance rates are commonly described using uniform, triangular or beta distribution (Min, 2003; van Dorp, 2005; Yang, 2007; Said et al., 2009) — see Figure 2. Triangular and uniform distributions are especially popular, because the experts feel generally comfortable in assessing the boundary values resp. mean/mode of the variables. Studies analysing the data from construction projects, however, show that other probabilistic models, such as lognormal or Weibull distribution, are more suitable (Wall, 1997; Chou, 2011).

![Figure 2. Illustration of selected probabilistic models for representation of unit time (costs).](image)

In our analysis, none of these models represent the data accurately. In Špačková et al. (2013) we therefore proposed a combined probability distribution (see Figure 2), which better captures the different phenomena (1 and 2) influencing the construction process and we analyse the extraordinary events (3) separately. The statistics of normal performance (1) and small disturbances (2) can be assessed from the excavation performance observed in three Czech tunnels (see Section 3.1). For extraordinary events (3), statistical analysis is only meaningful if it is based on a larger dataset including a large number of tunnel projects, as presented in Section 3.2.

3.1 Unit time

The unit time is here defined as the time spent on excavating a segment with length of 5 m under the normal performance and small disturbances. Data on the excavation progress from three tunnels built in the Czech Republic, denoted as TUN1 – TUN3, were used for the analysis of unit time – for more details see Špačková et al. (2013). In all three tunnels, inspections of geotechnical conditions at the tunnel heading and controls of construction performance were made regularly. From these records we obtained the following data:
- Date of the inspection
- Position of the main tunnel heading at this time
- Classification of the geotechnical conditions in the vicinity of the tunnel heading to ground classes.

Only the progress of the main tunnel heading is studied because it has decisive influence on the overall excavation performance. As an example, the excavation progress in tube 2 of TUN2 is depicted in Figure 3. The variability of the observed unit time per 5 m in different locations along the tunnel is depicted in Figure 4.

![Figure 3. Observed construction progress tube 2 of the tunnel TUN2 (all ground classes).](image)

![Figure 4. Observed unit time per 5 m in different positions of the tube 2 of the tunnel TUN2 (all ground classes).](image)
The probabilistic model of the unit time must include both the normal performance (1) and small disturbances (2). The PDF of the unit time including both (1) and (2) is thus defined as:

\[ f_T(t) = \Pr(\text{normal p.}) \cdot f(t|\text{normal p.}) + \Pr(\text{small d.}) \cdot f(t|\text{small d.}) \]  

(1)

where \( f(t|\text{normal p.}) \) is PDF of unit time under normal performance, \( f(t|\text{small d.}) \) is PDF of unit time under small disturbances and \( \Pr(\text{normal p.}) \) and \( \Pr(\text{small d.}) \) are probabilities of normal performance and small disturbance, respectively. The left bounded lognormal distribution describes well the normal performance, which has mean close to zero, relatively small variance and is slightly skewed. The beta distribution is suitable for the small disturbances, which have much higher variance and following the definition in our model framework are bounded between 0 and 15 days. The PDF for tunnel TUN2 and ground class 5 is shown in Figure 5.

![Figure 5. PDF of unit time per 5 m fitted to data from both tubes of tunnel TUN2, ground class 5.](image)

The analysis shows that the small disturbances significantly influence the mean and variance of the observed unit time. According to our experience, most experts would only estimate the normal performance and would thus underestimate the unit time.

3.2 Failure rate and delay due to construction failures

The failure rate is defined as the number of failures (extraordinary events) per unit length of the tunnel tube. Expert estimations of probabilities of rare events are commonly not reliable. They can be strongly biased by recent experiences (either positive or negative) of the expert and by many other factors (Lin and Bier, 2008; Goodwin and Wright, 2010).

To determine failure rate based on data, one must know the total number of failure events and the total length of excavated tunnels, ideally separately for individual ground classes, tunnel types etc. Because failures are rare events, data collected from a large number of tunnels would be needed. A rough estimate of failure rate based on global data from years 1999-2004 gathered from Sousa (2010), HSE (2006), Seidenfuss (2006), and Stallmann (2005) is made and compared with estimates using data from the Czech Republic from past c. 20 years - e.g. Barták (2007), Srb (2011) and Špačková (2012). The comparison is shown in Table 1. Because no information is available on the geological conditions and other features of the included tunnels, only the unconditional failure rate can be assessed.

<table>
<thead>
<tr>
<th>Total length of mined tunnels</th>
<th>Number of failures</th>
<th>Failure rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global data</td>
<td>5750 km</td>
<td>63</td>
</tr>
<tr>
<td>Czech data</td>
<td>60 km</td>
<td>14</td>
</tr>
</tbody>
</table>

The presented estimates show a huge spread and can only serve as a basis orientation for critical expert estimation. It must be further noted, that the global databases of failures are likely to be incomplete and that the failure rate estimated from global data is thus possibly too low.

Project delays resulting from failures of the tunnel construction process are analysed by Sousa (2010), using data from sixty-four failures. The data and the fitted shifted exponential PDF are shown in Figure 6.

![Figure 6. Distribution of delay caused by one failure.](image)

4 APPLICATION EXAMPLE

Using the DBN model introduced in Špačková and Straub (2013), the construction time for a Korean road tunnel (denoted as Dolsan A) with a length of 610 m is predicted probabilistically. The case study was originally published in Min (2003), where the DAT model (see Section 2.3) was used
for estimate the construction time and costs. The
DBN model inputs were based on the analysis of
data from the Czech tunnel TUN1 (reference tun-
nel). For more details on the case study and on
the DBN model we refer to Špačková and Straub
(2013).
The resulting PDF of construction time without
consideration of extraordinary events is depicted
in Figure 7. The result can be compared with the
estimate of Min (2003), which assessed the mean
value equal to 195 days and standard deviation of
3.39 days. The surprisingly good fit of the mean
value can be considered a coincidence, given the
significant differences between the reference
tunnel (TUN1) and the Dolsan A tunnel. However,
a comparison of the standard deviations shows
that utilization of real data and consideration of
uncertainties, which were not considered in the
original DAT model, leads to much higher uncer-
tainty in the predicted excavation time as com-
pared to Min (2003). A resulting standard devia-
tion of 23 days seems to more realistically reflect
the uncertainty of construction time estimates in
the design phase.
In Figure 8, the PDF of the predicted total excava-
tion time for the whole tunnel, including the ex-
traordinary events, is depicted. It shows the huge
influence of extraordinary events on the predic-
tion. The mean value increases by almost 40 %
and the increase in standard deviation is even
more significant. The reason is that the failure
rates considered in the DBN model are possibly
overly conservative and should be revised. How-
ever, the example does show the significant influ-
ence of the extraordinary events on the construc-
tion time estimates and it demonstrates the
importance of an accurate estimate of the failure
rate.

5 DISCUSSION AND CONCLUSIONS
There are three approaches to represent and
communicate construction time and cost esti-
mates: (a) deterministic estimates, which do not
include the uncertainty, (b) interval/percentile
estimates which take into account the uncertainty
but do not provide a full probability distribution
and (c) fully probabilistic estimates. As was
shown in the literature and practice, the first ap-
proach is likely to lead to wrong expectations and
decisions. However, even if the probabilistic
approach is used, it does not ensure a correct result,
if the inputs of the probabilistic model such as
unit time (advance rates) and unit costs are not
assessed realistically. To obtain such realistic es-
timates, it is desirable to assess the inputs based
on analysis of the available data. While it is quite
common to analyse the data to provide the aver-
age advance rates and unit costs (Section 2.1), the
probabilistic estimates typically rely on expert
assessments (Section 2.3). The only analysis of
uncertainty in the costs estimates known to the
authors is provided by Flyvbjerg and COWI
(2004) who examine the costs overruns. We
argue that this approach is suitable for describing
the present state but it might be problematic for
predictive purposes (see Section 2.2 for discus-
sion).
We therefore pursue an alternative approach to
statistical analysis of data from past projects
(Section 3), which examines the uncertainty in
the construction time per unit length of the tun-
nel. The method takes into account the depen-
dence of the unit time on the geotechnical con-
tions, tunnel geometry and other factors
influencing the construction performance. It uses
essentially the same approach as the determinis-
tic analyses of data commonly applied in present
practices, but it quantifies the variability of the
unit time, not only its mean. An analogous
approach can also be used for analysis of unit costs.
Probabilistic models have not been widely ac-
cepted in the practice so far. A first reason is that
there was no real demand for quantitative model-
ling of uncertainties and risk, because decision makers were not used to work with such information. A second reason is that the existing models often did not provide a realistic estimate of the uncertainties and they therefore did not gain acceptance among practitioners. However, this situation seems to have been changing in recent years and both the demand and the reliability of the model results have increased. The systematic statistical analysis of construction data presented in this paper is a step toward a realistic prediction of construction time and costs uncertainties.

REFERENCES


[14] HSE (Health & Safety Executive), 2006. The risk to third parties from bored tunnelling in soft ground – research report 453. Health & Safety Executive, HSE Books, Sudbury, Suffolk, GB.


