

# Geomechanical parameters updating in an underground work

T. Miranda & A. Gomes Correia

Department of Civil Engineering, University of Minho Guimarães, Portugal

L. Ribeiro e Sousa

Department of Civil Engineering, University of Porto, Porto, Portugal

**ABSTRACT:** In geotechnical engineering, and in the particular case of underground works, a great number of uncertainties arise due to the lack of knowledge of the involved formations and their variability. Geomechanical parameters are one of the main issues in the underground works design. In the initial stages, the available information about the rock masses characteristics is scarce. As the project advances to other stages more and more information from different sources becomes available and can be used for updating the geomechanical model. Bayesian methodologies use probability as the main tool to deal with uncertainty and manage to reduce it using new data via the Bayes theorem. In this work, a part of a developed Bayesian framework to the updating of the deformability modulus ( $E$ ) in an underground structure is presented. Assuming  $E$  as a random variable, data from LFJ tests is used to obtain a posterior and less uncertain distribution of  $E$ . This approach led to good results and considerable uncertainty reduction and increased reliability. The developed Bayesian framework constitutes a rational and structured way of dealing with data with different sources and uncertainty levels.

## 1 INTRODUCTION

The geomechanical parameters determination is an exercise of subjective nature. The inherent uncertainty about their real value hinders the establishment of a deterministic set of values for the parameters. In practice, for each geotechnical horizon, a range of values is assigned for the parameters based on the geotechnical survey and, in the case of rock masses, by application of the empirical classification systems.

In the initial stages, the available information about the rock masses is limited. However, the construction of geotechnical models is a dynamic process and, as the project advances, it can be updated as new data is gathered.

Data can have different sources each with its own precision and reliability. This fact transforms the updating process subjective and dependent on the geotechnical engineer experience. Nowadays, it lacks a methodology to consistently treat the problem of the geomechanical model updating in order to reduce the subjectivity of this procedure.

The characteristics of the Bayesian methods of data analysis makes them well suited for geotechnical purposes where uncertainty is present at several levels and data is compiled in different stages and with different properties.

In Figure 1 a general scheme for the deformability modulus ( $E$ ) calculation and update during preliminary and construction stages is presented. It consists in a Bayesian framework where  $E$  is considered a random variable with a given distribution function. Uncertainty about the parameter is translated by its standard deviation which can be reduced as more data is obtained through *in situ* tests, mapping of the tunnel front and backanalysis data.

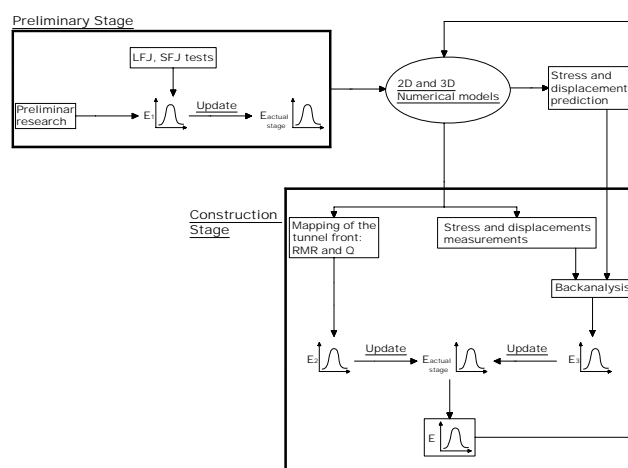


Figure 1 – Scheme of the updating process (adapted from Faber, 2005)

In this work, a part of this general Bayesian framework for the geotechnical model updating is presented. It is shown how data from preliminary geotechnical survey can be updated using reliable *in situ* tests. More specifically information about  $E$  is available by application of the empirical systems and then updated using the results of LFJ tests. Real data from the Venda Nova II powerhouse complex was used for the update process (LNEC, 1983).

This geomechanical parameter was considered a random variable with a normal distribution. Calculations were performed considering mean and variance unknown with prior knowledge based on analytical solutions and empirical systems application.

## 2 BAYESIAN METHODS

### 2.1 Bayesian data analysis and uncertainty

Uncertainties may be represented in terms of mathematical concepts from the probabilistic theory (Ditlevsen and Madsen, 1996; Einstein, 2006). In many cases it is enough to model the uncertain quantities by random variables with given distribution functions and parameters estimated on the basis of statistical and/or subjective information (Faber, 2005). The principles and methodologies for data analysis that derive from the subjective point of view are often referred to as Bayesian statistics. Its central principle is the explicit characterization of all forms of uncertainty in a data analysis problem.

Bayesian techniques allow updating random variables when new data is available using a mathematical process in order to reduce uncertainties. This process can be divided into the three following steps (Ditlevsen and Madsen, 1996):

- 1 – Set up a joint probability distribution for all variables consistent with knowledge about the underlying problem.
- 2 – Calculate the conditional posterior distribution of the variables of interest given new observed data.
- 3 – Evaluate the fit of the model to the data analysing if the conclusions are reasonable and how sensitive are the results to the modelling assumptions on step 1.

The posterior distribution is a compromise with reduced uncertainty between the prior information and the one contained in the new data. This compromise is increasingly controlled by the data as the sample size increases in what is sometimes referred to as asymptotic theory (Gelman et al., 2004).

Bayesian methods provide tools to incorporate data and external information into the data analysis process. In a Bayesian approach, the data analysis process starts already with a given probability distribution. Its parameters may be chosen or estimated based on previous experiment results, experience and professional judgement. This distribution is called prior distribution and translates the uncertainty about the parameter value. When additional data becomes available it is used to update this prior distribution into a posterior distribution using the Bayes theorem. Figure 2 resumes this overall process.

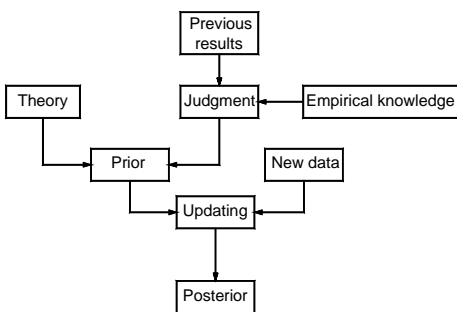


Figure 2 – Scheme of the updating process (adapted from Faber, 2005)

If the prior distribution of a parameter  $\theta$ , with  $n$  possible outcomes  $(\theta_1, \dots, \theta_k)$ , is continuous and the new information  $x$  is available, then the Bayes theorem is translated by:

$$p(\theta | x) = \frac{p(\theta)p(x | \theta)}{\int p(\theta)p(x | \theta)d\theta} \quad (1)$$

where,  $p(\theta)$  is the prior distribution of the possible  $\theta$  values which summarizes the prior beliefs about the possible values of the parameter,  $p(x|\theta)$  is the conditional probability (or likelihood) of the data given  $\theta$  and  $p(\theta|x)$  is the posterior distribution of  $\theta$  given the observed data  $x$ .

The prior and posterior distributions of  $\theta$  are represented by density functions. The joint probability distribution of the data and the parameter is given by  $p(x|\theta)$  which is called the likelihood and is defined by:

$$p(x | \theta) = L(\theta) = \prod_i p(x_i | \theta) \quad (2)$$

Bayes theorem is applied multiplying the prior by the likelihood function and then normalizing, to get the posterior probability distribution, which is the conditional distribution of the uncertain quantity given the data. The posterior density summarizes the total information, after considering the new data, and provides a basis for posterior inference regarding  $\theta$ .

### 2.2 Bayesian inference

The process of Bayesian inference involves passing from a prior distribution  $p(\theta)$  to a posterior distribution  $p(\theta|x)$  using the likelihood function of the data. The consideration of normal likelihood, i.e. that data follows a normal distribution, has the computational advantage of allowing the use of conjugate priors (the posterior distribution follows the same parametric form as the prior). This type of prior distributions has the practical advantage of computational convenience. The obtained results are normally easy to understand and can often be put in analytical form. If information is available that contradicts the conjugate parametric family, it may be necessary to use a more realistic prior distribution (Gelman et al., 2004). The central limit theorem helps to justify the use of the normal likelihood and the results are often acceptable (Ditlevsen & Madsen, 1996).

In the Bayesian approach the parameters of interest are assumed to follow certain probability distributions with one or more unknown distribution parameters. These parameters are also considered to have given distributions with known prior hyperparameters. The hyperparameters are then updated given the data and will be used to infer to the parameter distribution.

In this work a multiparameter model that involves the consideration of both mean and variance as unknowns was used. In the developed Bayesian framework it was considered that both mean and variance of  $E$  were random variables. A normal likelihood was considered together with the conjugate prior.

The natural conjugate prior has the following form:

$$p(\mu | \sigma^2) \propto \left(\frac{n_0}{\sigma^2}\right)^{1/2} \exp\left[-\frac{n_0}{2\sigma_0}(\mu - \mu_0)^2\right] \times \left(\frac{1}{\sigma^2}\right)^{\frac{v_0+1}{2}} \exp\left[-\frac{S_0}{2\sigma^2}\right] \quad (3)$$

This means that the prior is the product of the density of an inverted Gamma distribution with argument  $\sigma^2$  and  $\nu_0$  degrees of freedom and the density of a normal distribution with argument  $\mu$ , where the variance is proportional to  $\sigma^2$ . It is the density of the so-called normal-gamma distribution. Therefore, the prior for  $\mu$  conditional on  $\sigma^2$  is a normal with mean  $\mu_0$  and variance  $\sigma^2/n_0$ :

$$\mu | \sigma^2 \sim N\left(\mu_0, \frac{\sigma^2}{n_0}\right) \quad (4)$$

The prior for  $1/\sigma^2$  is a gamma with hyperparameters  $\nu_0/2$  and  $S_0/2$ :

$$\frac{1}{\sigma^2} \sim \text{gamma}\left(\frac{\nu_0}{2}, \frac{S_0}{2}\right) \quad (5)$$

The appearance of  $\sigma^2$  in the conditional distribution of  $\mu|\sigma^2$  means that  $\mu$  and  $\sigma^2$  are dependent. The conditional posterior density of  $\mu$ , given  $\sigma^2$ , is proportional to  $p(\mu, \sigma)$  with  $\sigma^2$  held constant. It can be shown that:

$$\mu | \sigma^2, x \sim N\left(\mu_1, \frac{\sigma^2}{n_1}\right) \quad (6)$$

where

$$\mu_1 = \frac{n_0}{n_0 + n} \cdot \mu_0 + \frac{n}{n_0 + n} \cdot \bar{x} \quad (7)$$

$$n_1 = n_0 + n \quad (8)$$

The parameters of the posterior distribution combine the prior information and the information contained in the data. For example,  $\mu_1$  is a weighted average of the prior and the sample mean, with weights determined by the relative precision of the two pieces of information. The marginal posterior density of  $1/\sigma^2$  is gamma:

$$\frac{1}{\sigma^2} | x \sim \text{gamma}\left(\frac{\nu_1}{2}, \frac{S_1}{2}\right) \quad (9)$$

where

$$\nu_1 = \nu_0 + n \quad (10)$$

$$S_1 = S_0 + (n-1) \cdot s^2 + \frac{n_0 \cdot n}{n_0 + n} \cdot (x - \mu_0)^2 \quad (11)$$

The posterior sum of squares,  $S_1$ , combines the prior and the sample sum of squares, and the additional uncertainty given by the difference between the sample and the prior mean.

### 2.3 Posterior simulation

The posterior distribution is the fundamental object of Bayesian analysis and contains the relevant information about all values of  $\mu$  and  $\sigma$ . To obtain the complete posterior distributions on the parameters it is normally necessary to use simulation methods.

There are several different algorithms to simulate the posterior distributions. One of the most popular is the Markov Chain Monte Carlo (MCMC). Markov chain simulation is a general method based on a sequential draw of sample values with the distribution of the sampled draws depending only on the last value (Brooks, 1998). In probability theory, a Markov chain is a sequence of random variables  $\theta_1, \theta_2, \dots, \theta_n$  for which, for any time  $t$ , the distribution of  $\theta_t$  depends only on the most recent value,  $\theta_{t-1}$ .

The Metropolis and the Gibbs sampler are particular Markov chain algorithms. The Gibbs sampler is the most popular one and is normally chosen for simulation in conditionally conjugate models, where it is possible to directly sample from each conditional posterior distribution. For parameters whose conditional posterior distribution has standard forms it is better to use the Gibbs sampler otherwise the Metropolis should be used. In this work, the Gibbs sampler was implemented in order to simulate the posterior distributions.

## 3 APPLICATION OF THE BAYESIAN FRAMEWORK TO UPDATE THE DEFORMABILITY MODULUS OF THE ROCK MASS

In this work the general Bayesian framework for the deformability modulus (E) updating was applied in an underground structure. The data was collected in the Venda Nova II project and consisted of 40 LFJ tests performed in a gallery. In each test 4 cycles were conducted resulting in a total of 160 values of E. The mean and standard deviation of the tests were 36.9 GPa and 6.1 GPa, respectively.

The mean and variance of E were considered to be unknown variables. The prior was developed considering the information of 76 values of E calculated using analytical expressions based on the empirical classification systems and takes the following form:

$$\mu | \sigma^2 \sim N\left(38,5; \frac{\sigma^2}{76}\right) \quad (12)$$

$$\frac{1}{\sigma^2} \sim \text{gamma}\left(38,5; \frac{1}{11573,5}\right) \quad (13)$$

Applying Bayes theorem and using the data from the LFJ tests, the conditional posterior distribution for the mean and the marginal posterior for the variance are the following:

$$\mu | \sigma^2 \sim N\left(37,4; \frac{\sigma^2}{236}\right) \quad (14)$$

$$\frac{1}{\sigma^2} \sim \text{gamma}\left(118,5; \frac{1}{14597,6}\right) \quad (15)$$

As the mean is conditional on the variance, prior and posterior estimates for the mean value and standard deviation were obtained by simulation using the Gibbs sampler with 10000 iterations. Using this methodology the parameters for the mean and standard deviation were calculated and are presented in Table 1 as well as the 95% confidence interval for the mean.

Table 1 – Prior and posterior estimates of the mean value of E

Parameter	Prior values (GPa)	Posterior values (GPa)
$\mu$	38.5	37.4
$\sigma(\mu)$	2.02	1.15
$\sigma$	17.5	11.1
$\sigma(\sigma)$	1.45	0.51
95% CI for the mean	35.2-41.8	35.5-39.3

The updated mean value of the mean ( $\mu$ ) is only about 3% lower than the initial guess. This means that the analytical solutions provided a very good estimate of E. The most important aspect is the substantial uncertainty reduction at all levels. The standard deviation of the mean ( $\sigma(\mu)$ ) was reduced from 2.02 GPa to 1.15 GPa, i.e. only 57% of the initial value. The mean of the standard deviation ( $\sigma$ ) suffered a 37% decrease from 17.5 GPa to 11.1 GPa. Finally, the standard deviation of the standard deviation ( $\sigma(\sigma)$ ) was also significantly decreased from 1.45 GPa to 0.51 GPa. To illustrate this fact, Figure 3 shows the prior and posterior probability density functions of the mean value of E considering the mean value of its standard deviation. It can be clearly observed the uncertainty reduction from the prior to the posterior.

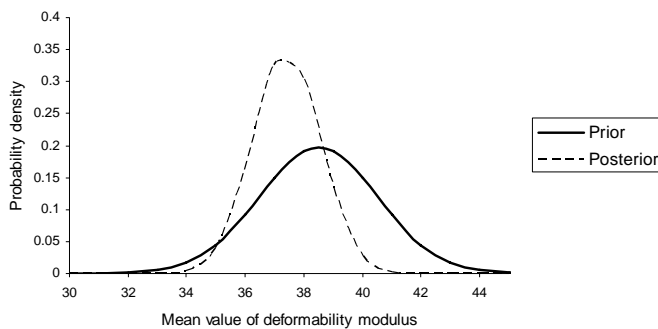


Figure 3 – Prior and posterior probability density functions for the mean value of E

In terms of the inferred population parameters, the mean value and the standard deviation passed from prior values of 38.4 GPa and 19.6 GPa to posterior values of 37.4 GPa and 12.4 GPa, respectively. The 37% decrease on the standard deviation had considerable influence on the 95% confidence intervals. The prior interval ranged from 6.1 GPa to 70.7 GPa and the posterior from 17.1 GPa to 57.7 GPa.

In Figure 4 the prior and posterior probability distributions of E considering the mean values of the mean and standard deviation is presented. Even though the mean value is practically unchanged, the uncertainty about the parameter was clearly reduced using the Bayesian methodology.

From an engineering point of view, the posterior results and inferences can be considered consistent and validate the prior assumptions especially that of normal likelihood.

It is worth underline that, the Bayesian updating procedure did not significantly changed the mean value of E. The preliminary evaluation based on analytical solutions was almost corroborated by the results of the LFJ tests. However, it allowed a very significant decrease in the uncertainty about the parameters.

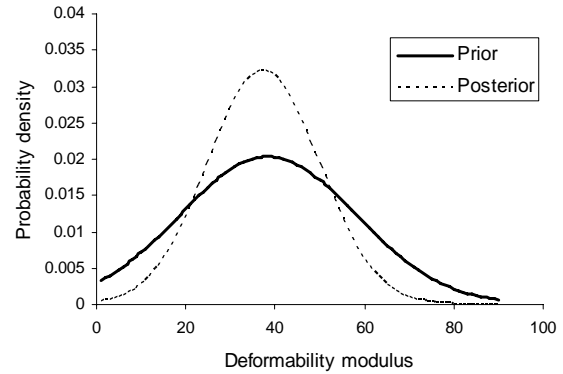


Figure 4 – Prior and posterior probability density functions for E (inferred values for the population)

#### 4 FINAL REMARKS

Bayesian methods have an inherent flexibility introduced by the incorporation of multiples levels of uncertainty and the resultant ability to combine information from different sources. This methodology allows the updating of random variables as new data is collected through Bayes theorem.

This work was an approach on how the Bayesian tools can be used in geotechnics. The developed framework for the E updating showed interesting results especially in the uncertainty reduction. This procedure could be sequentially applied as more information about the rock mass was gathered.

It is believed that the characteristics of the Bayesian data analysis make it well suited to be applied on geotechnical problems where uncertainty is always present at different levels. In geotechnics, the information about the interested formations increases as the project advances for different stages and can be used to update the geotechnical models.

#### ACKNOWLEDGMENTS

The authors wish to acknowledge to Professor Herbert Einstein for reviewing this paper and to the EDP company (Portuguese Electricity Company) for providing the test results. This work was financed by the FCT project POCI/ECM/57495/2004 entitled “Geotechnical Risk in Tunnels for High Speed Trains”.

#### REFERENCES

- Brooks, S. 1998. Markov chain Monte Carlo method and its application. *The Statistician*, 47, Part 1, pp. 69-100.
- Ditlevsen, O. & Madsen, H. 1996. *Structural reliability methods*. John Wiley & Sons. 372p.
- Einstein, H. H. 2006. Use of decision aids for tunnelling. *Geotechnical risk in rock tunnels*. Ed. Campos e Matos, Ribeiro e Sousa, Klerberger and Lopes Pinto. Taylor & Francis group, London, pp. 63-74.
- Faber, M. 2005. *Risk and safety in civil, surveying and environmental engineering*. Lecture notes. 394 p.
- Gelman, A.; Carlin, J.; Stern, H; Rubin, D. 2004. *Bayesian data analysis*. Chapman & Hall/CRC. 668p.
- LNEC. 1983. *LNEC collaboration in the geological/geotechnical studies for the hydraulic circuit of Venda Nova II*. Internal report (in Portuguese), 341p.