Engineering Geology 195 (2015) 214–224

Contents lists available at ScienceDirect

Engineering Geology

journal homepage: www.elsevier.com/locate/enggeo

Predicting tunnel squeezing with incomplete data using Bayesian networks

Xianda Feng, Rafael Jimenez *

Technical University of Madrid, ETSI Caminos, C. y P, C/Profesor Aranguren s/n, 28040 Madrid, Spain

ARTICLE INFO

Article history:
Received 24 October 2014
Received in revised form 17 June 2015
Accepted 21 June 2015
Available online 27 June 2015

Keywords:
Tunnel squeezing
Bayesian networks
Naïve Bayes classifier
Incomplete data
Cross-validation
Sensitivity analysis

ABSTRACT

Tunnel squeezing or time-dependent large deformations due to creep are common in tunnels constructed in weak rock masses at large depth or subjected to high horizontal in situ stresses in tectonically active regions. Squeezing can produce tunnel collapses, budget overruns and construction delays, and being able to predict squeezing is therefore important. This study presents a novel application of Bayesian networks (BNs) to predict squeezing. In particular, we employ a Naïve Bayes classifier based on five parameters – support stiffness (*K*), Rock Tunneling Quality Index (*Q*), tunnel depth (*H*), tunnel diameter (*D*), and strength–stress ratio (*SSR*) – about which information is commonly available at early design stages. The Naïve Bayes classifier is "learned", using the Expectation Maximization algorithm, with a database of 166 tunneling case histories from 7 countries compiled by the authors which is provided as Supplementary material. Then, the Junction Tree algorithm is employed for "belief updating"; i.e., to predict the probability of tunnel squeezing for a given set of (probably incomplete) evidence. The model is validated using 10-fold cross-validation and also using an additional set of case-histories that had not been originally employed to learn the network. Results show that, when compared with other available criteria, the error rate of our BN is among the lowest, but with the advantage that it is able to provide predictions even with incomplete data. Results of a sensitivity analysis to assess the importance of input parameters on the squeezing outcome are also presented. And, finally, a web-based implementation of the proposed BN is provided to improve the ease-of-use of our approach.

1. Introduction

The ISRM Commission on Squeezing Rocks in Tunnels defined squeezing as "the time-dependent large deformation, which occurs around the tunnel, and is essentially associated with creep caused by exceeding a limiting shear stress" (Barla, 1995). A commonly accepted deformation threshold to define squeezing occurrence is $\varepsilon = 1\%$ (Chern et al., 1998; Jimenez and Recio, 2011; Sakurai, 1983), where ε is the "normalized convergence" defined as the ratio (in %) between tunnel closure and tunnel diameter. Note, however, that such 1% limit is "only an indication of increasing difficulty", so that tunnels with much higher strains have been completed without stability problems (Hoek, 2001).

Squeezing in rock tunnels often occurs in soft (weak) rock masses at great depth; it is also common in tunnels subjected to high horizontal in situ stresses in tectonically active regions like the Himalayas (Jimenez and Recio, 2011; Steiner, 2000; Sunuwar, 2007). Squeezing may cause tunnel collapses, budget overruns and construction delays. Therefore, many correlations – based on case histories, closed-form solutions, or numerical models – have been proposed to predict squeezing; a summary of these studies is presented in Table 1. Analytical or numerical solutions to time-dependent (or creep) deformations (see, e.g. Guan et al., 2008; Phienwej et al., 2007; Sterpi and Gioda, 2009) are outside the scope of our analysis and, therefore, they are not reviewed herein.

This paper uses Bayesian networks (BNs) to predict the occurrence of squeezing. BNs were introduced by Pearl (1986) to more easily deal with conditional dependency relationships between the (observable or unobservable) random variables of a statistical model, and they have been shown to have advantages to deal with inference, classification, and decision making problems (Aguilera et al., 2011). As a result, they are becoming increasingly popular in fields such as environmental science (Aguilera et al., 2011; Uusitalo, 2007), ecology (Landuyt et al., 2013), water resources management (Batchelor and Cain, 1999), and agriculture (Cain et al., 2003); and they have also been employed in geotechnical engineering (Huang et al., 2012; Jimenez-Rodriguez and Sitar, 2006; Medina-Cetina and Nadim, 2008; Peng et al., 2014; Schubert et al., 2012; Song et al., 2012; Sousa and Einstein, 2012; Špačková et al., 2013; Xu et al., 2011; Zazzaro et al., 2012; Zhang et al., 2011).

This study aims to develop a Bayesian network and, specifically, a Naïve Bayes classifier, to probabilistically predict the occurrence of squeezing in rock tunnels for which (sometimes incomplete) information is available. The main advantage of BNs with respect to previous approaches, such as those listed in Table 1, is that BNs are particularly useful for dealing with two situations that are common with tunnel squeezing: (i) they can deal with incomplete input information, i.e., missing data (Uusitalo, 2007); and (ii) they are able to provide

* Corresponding author.
E-mail address: rafael.jimenez@upm.es (R. Jimenez).

http://dx.doi.org/10.1016/j.enggeo.2015.06.017
0013-7952/

Table 1
A summary of previous approaches to predict tunnel squeezing.

Proposed by	Approaches	Required parameters	Sources	Type	Eq. #
Jethwa et al. (1984)	$N_c = \sigma_{cm}/\gamma H \leq 2.0$; $\sigma_{cm} = 2c_p\cos\phi_p/(1 - \sin\phi_p)$	σ_{cm}, γ, and H	–	Semi-empirical	(1)
Singh et al. (1992)	$H \geq 350Q^{1/3}$	H and Q	39 case histories	Empirical	(2)
Goel et al. (1995)	$H \geq 270\ N^{0.33} \cdot B^{-0.1}$ with $N = (Q)_{SRF\,=\,1}$	H, N, and B	72 cases histories	Empirical	(3)
Aydan et al. (1993)	$\alpha = \sigma_c/\gamma H \leq 2.0$	σ_c, γ, and H	Cases from Japan tunnels	Semi-empirical	(4)
Barla (1995)	$\sigma_{cm}/\gamma H \leq 1.0$	σ_{cm}, γ, and H	–	Semi-empirical	(5)
Bhasin and Grimstad (1996)	$\sigma_\theta/\sigma_{cm} \geq 1.0$ with $\sigma_{cm} = 0.7\gamma Q^{1/3}$	σ_{cm} and σ_θ	Tunnel case histories	Semi-empirical	(6)
Hoek and Marinos (2000)	$\varepsilon = (0.002–0.0025p_i/p_0) \cdot (\sigma_{cm}/p_0)^{(2.4p_i/p_0 - 2)} \geq 1\%$ with $\sigma_{cm} = 0.0034m_i^{0.8} \cdot \sigma_{ci} \cdot [1.029 + 0.025\exp(-0.1m_i)]^{GSI}$	p_i, p_0, and σ_{cm}	Monte Carlo simulations	Semi-empirical	(7)
Hoek (2001)	$\sigma_{cm}/p_0 = \sigma_{cm}/\gamma H \leq 0.35$	σ_{cm}, γ, and H	Cases from 16 tunnels	Semi-empirical	(8)
Hoek (2001)	$\varepsilon\% = 0.15(1 - p_i/p_0) \cdot (\sigma_{cm}/p_0)^{-(3p_i/p_0 + 1)/(3.8p_i/p_0 + 0.54)} \geq 1(\%)$	p_i, p_0, and σ_{cm}	Finite-element models	Semi-empirical	(9)
Jimenez and Recio (2011)	$H \geq 424.4Q^{0.32}$	H and Q	62 case histories	Empirical	(10)
Dwivedi et al. (2013)	$\varepsilon = (0.0191\sigma_v Q^{-0.2}) / (K + 1) + 0.0025 \geq 1\%$ with $\sigma_v = 0.027H$	σ_v, Q, and K	63 case histories	Empirical	(11)

Notation: N_c (or α) competency factor (also called "strength stress ratio (SSR)"), σ_{cm} rock mass uniaxial compressive strength (MPa), γ rock mass specific weight (MN/m^3), H overburden or depth of tunnel (m), c_p rock mass peak cohesion (MPa), ϕ_p rock mass peak friction angle (degree), Q Rock Tunneling Quality Index, N rock mass number (or stress-free Q), SRF stress reduction factor, B tunnel span or diameter (m), σ_c uniaxial compressive strength of intact rock (MPa), σ_θ tangential stress (MPa), ε percentage strain (ratio of tunnel closure to tunnel diameter), p_0 in situ vertical stress at tunnel depth (MPa), p_i internal support pressure (MPa), m_i Hoek–Brown constant, GSI Geological Strength Index, σ_v vertical in situ stress (MPa), K support stiffness (MPa).

good estimates based on limited data sets (Kontkanen et al., 1997; Sen et al., 2012). In other words, they are still able to make predictions even in cases in which – as often happens in tunnel engineering practice – information is scarce or incomplete at early design stages. This also has advantages when data are gathered to construct a database of case histories, as case histories from different authors or from different sources tend to emphasize different aspects of the tunnel project. To conduct our analyses, a database of squeezing case histories was compiled and employed to "learn" the Naïve Bayes classifier that predicts, with the aid of the Junction Tree algorithm, the probabilities of squeezing. Finally, the predictions of the proposed Naïve Bayes classifier have been validated using 10-fold cross-validation, as well as an additional set of tunnel case histories that was not included in the original database; and a sensitivity analysis is conducted to assess the importance of the input parameters considered on the squeezing outcome.

2. Parameters employed for analysis and database description

2.1. Parameters related to tunnel squeezing

Based on the literature review conducted (see the list of references), and also on the parameters considered by common methods to predict squeezing (Table 1), five main parameters that might influence squeezing are identified and employed in our analyses: tunnel depth (H), Rock Tunneling Quality Index (Q), tunnel span or diameter (B or D), support stiffness (K), and stress strength ratio (SSR). In addition, there are other parameters – rock mass strength (σ_{cm}), specific weight of rock mass (γ), and support pressure (p_i) – that, when available, are included in the database, although they are not included in the developed BN. The reasons are that (i) some of them (such as γ) are not very variable and their influence within typical ranges of variability is limited; and (ii) some of them (such as σ_{cm} and p_i) are not commonly available at design stage.

2.2. Description of the database

We conducted a literature review of rock tunneling case histories in which the occurrence (or absence) of squeezing has been reported. The database contains 166 cases from 30 projects located in 7 countries, among which 109 are squeezing cases and 57 are non-squeezing cases. The database includes fields for all the eight parameters affecting squeezing that were discussed above, although some of the case histories included in the database are "incomplete", i.e., they do not report information about all fields. However, note that we are still able to "extract" information from them, given the ability of the BN approach to learn from case histories with "incomplete" information.

Fig. 1 shows the histograms, cumulative distribution functions (CDF's), and additional statistics – number of known data and missing data, maximum and minimum values, means and standard deviations – of the five parameters considered to predict squeezing with the BN (i.e., H, Q, D, K, and SSR). It can be seen in Fig. 1 that our database contains data covering a wide range of values for these five parameters, hence having in principle a wide range of applicability that is, of course, limited to the range of available input data. Our database is made available to readers as Supplementary material to this article (Appendix A.)

2.3. Inputs in the BN

2.3.1. Tunnel depth (H)

Almost all methods to predict squeezing (Table 1) consider tunnel depth (H) or in-situ stress (often estimated as $p_0 = \gamma H$). This indicates that tunnel depth is an important parameter to predict tunnel squeezing. Values of tunnel depth are commonly reported in the literature, and all the H values corresponding to case histories in our database are known (none is missing).

2.3.2. Rock Tunneling Quality Index (Q)

The Rock Tunneling Quality Index (or Q-system proposed by Barton et al., 1974) has been often used as input to predict tunnel squeezing (Basnet, 2013; Jimenez and Recio, 2011; Singh et al., 1992). Q values for many histories (136 out of 166) could be collected from the literature but, for 28 cases, RMR values are reported instead of Q values. Then, the empirical correlation proposed by Barton (1995), i.e., $Q = 10^{(RMR\,-\,50)\,/\,15}$, has been used to estimate Q values based on RMR. In addition, there are two tunnels with unknown Q or RMR values.

2.3.3. Tunnel span or diameter (B or D)

The size of the tunnel, as given by its span or diameter (B or D), also influences squeezing (Goel, 1994). Therefore, we consider tunnel diameter – or, when the tunnel is non-circular (Dwivedi et al., 2013), its "equivalent diameter" given by $D = \sqrt{4A/\pi}$, with A being its cross-sectional area – as one of the parameters to predict tunnel squeezing with the BN. Tunnel diameters are commonly reported in the literature, and only one of them is unknown (missing) in all case histories considered in the database.

2.3.4. Support stiffness (K)

Installing an adequate support at an appropriate time may reduce tunnel deformation. In other words, deformation of squeezing tunnels is influenced by the support systems installed within the tunnel, so that, for instance, Dwivedi et al. (2013) has recently employed support stiffness as one of the parameters to estimate the deformation of squeezing tunnels. In this study, we also use support stiffness to predict squeezing with the BN. About a third of the case histories in our