Tunnelling and Underground Space Technology 69 (2017) 155–161

Contents lists available at ScienceDirect

Tunnelling and Underground Space Technology

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

Gaussian process model of water inflow prediction in tunnel construction and its engineering applications

Shu-cai Li, Peng He*, Li-ping Li, Shao-shuai Shi, Qian-qing Zhang, Jian Zhang, Jie Hu

Geotechnical and Structural Engineering Research Center, Shandong University, Ji'nan 250061, Shandong, China

ARTICLE INFO

Keywords:
Water inrush
Water inflow prediction
Gaussian process regression
Evaluation indices
Engineering applications

ABSTRACT

Due to the extremely complicated hydrogeological environment, significant symptoms of water inrush can not be detected accurately using normal exploratory methods, which produces hundreds of water inrushes occurred during tunnel construction in karst area. This study aims to present a new water inflow prediction technique without considering the relationship between hydrogeological features and water discharge rate. Therefore, the nonlinear regression Gaussian process analysis is applied to develop a model for predicting water inflow into tunnels. In order to meet the requirement of the data format of Gaussian process regression model (GPR), the basic evaluation index system of water inflow into tunnels and corresponding criterion are set up and quantified based on the statistical information of water inrush cases. To verify its feasibility, The GPR model is applied to Zhongjiashan tunnel on Jilian highway in China. The results of the comparisons indicate that the prediction results obtained from the GPR model are generally in a good agreement with the field-observed results. The proposed Gaussian process, on the whole, performs better than the support vector machine (SVM) and artificial neural network (ANN) in predictive analysis of water inflow into tunnels.

1. Introduction

Water inrush is one of the typical geological hazards during tunnel construction, resulting in huge casualties and economic loss. In recent years, a large number of disastrous events associated with water inrush in tunnels have been frequently reported in China and elsewhere (Song et al., 2006; Shahriar et al., 2008; Ge, 2010; Lei, 2011; Zarei et al., 2011; Li et al., 2013). Therefore, it is necessary to accurately predict water inflow into tunnels and take some effective countermeasures to assure the safety of tunnel construction.

Although it is difficult to accurately predict water inrush probability and discharge rate of tunnels, many investigations consisting of analytical, empirical, and numerical approaches have been presented within the last few decades. Three classical analytical methods are often used to estimate water inflow into tunnels (Kong, 2011): (1) Goodman method (Goodman et al., 1965); (2) Heuer and Raymer method (Heuer, 1995; Raymer, 2001), and (3) IMS method (McFeat-Smith et al., 1985). Analytical solutions can be rapid and useful. However, the analytical solutions rely on given hydrogeological assumptions with simple circular or rectangular openings and they are unable to predict water inflow in complex hydrogeological conditions, such as fractured rocks (Li et al., 2009). Besides, the tunnel water inflow problems in various complicated geological conditions could be analyzed using numerical modeling (Hwang and Lu, 2007). Several conceptual models have been proposed so far, including (1) equivalent porous media models, (2) discrete fracture network models, and (3) hybrid models (Berkowitz, 2002; Li et al., 2009). Some other models have also been established to assess the risk of water inrush by using the software RFPA^{2D} and COMSOL (Yao et al., 2012). However, numerical models require a real conceptual model based on the input geological data. Moreover, they either do not account for evaluating the risk of water inrush or fail to profoundly reveal the quantifying relation between water inrush and factors influencing it. Hence, using analytical and numerical tools to predict possible tunnel water inflow often fails because of given hydrogeological assumptions and simplification of these heterogeneous media. In fact, the whole geological condition can be considered as a chaotic system. So, the characteristics of hydrogeology can not be detected accurately using normal exploratory methods in general.

Currently, varied stochastic mathematics methodologies such as attribute mathematical model (Li et al., 2013; Y. Wang et al., 2012), analytic hierarchy process (Xu et al., 2011) and fuzzy extension theory (Li et al., 2015) have been widely used to evaluate the risk of water inrush in tunnels and coal mines. The evaluation results are basically acceptable, though there is subjectivity in weights and criteria of evaluation indices. Additionally, some computational intelligence techniques such as artificial neural network, support vector machines and the

* Corresponding author.
E-mail address: hepenghank@163.com (P. He).

http://dx.doi.org/10.1016/j.tust.2017.06.018
Received 18 February 2016; Received in revised form 14 June 2017; Accepted 19 June 2017

Gaussian process have been applied for landslide displacement (Grelle and Guadagno, 2012; Rohmer and Foerster, 2011; Liu et al., 2014), surrounding rock deformation (Feng et al., 2004; D.D. Wang et al., 2012), surface settlement (Suwansawat and Einstein, 2006; Ovidio et al., 2008; Ibrahim and Sadi, 2013), pipe failure rate (Tabesh et al., 2009; Akbar et al., 2014) and other geotechnical engineering problems (Anthony and Goh, 2007; Ocak and Seker, 2012; Liu et al., 2013). Moreover, the use of artificial neural network and support vector machine for prediction of mine water discharge and water inflow into tunnels were also reported in few related studies (Guo and Ma, 2010; Ren and Xu, 2011; Liu, 2014). These approaches have shown the potential ability of the corresponding methods to analyze the engineering problems. Especially the Gaussian process model, which is suitable for cases characterized by high dimension, small sample population, and nonlinearity, it can apply the mean of the distribution as point predictions to avoid robust point predictions like that in ANN and SVM. However, there is few research on the Gaussian process regression for prediction of water inflow into tunnels.

The objective of this paper is to present the new computational intelligence approach without considering the certain relationship between hydrogeological features and water inflow into tunnels. For this purpose, the nonlinear regression Gaussian process analysis is applied to develop a model for predicting water inflow into tunnels. Meanwhile, the basic evaluation index system of water inflow into tunnels and corresponding criterion are also set up and quantified. To verify its feasibility, The GPR model as well as the other computational intelligence techniques are applied to Zhongjiashan tunnel on Jilian highway in China to predict tunnel discharge rate.

2. Gaussian process model

Consider a data set S of n observations $S = \{(x_i, y_i)|i = 1, \ldots, n\}$,where x_i is a D-dimensional input vector, and y_i is a scalar output or target. This set of input/output pairs will be referred to as sample points or experimental points. For the sake of convenience, the inputs are aggregated into a matrix $X = [x_1, x_2, \ldots, x_n]$. The outputs are likewise aggregated, $y = [y_1, y_2,\ldots,y_n]$. The regression task is, given a new input x_*, to obtain the predictive distribution for the corresponding observation y_* based on S (Kang et al., 2015).

Gaussian process is a stochastic process, which provides a powerful tool for probabilistic inference directly on distributions over functions (e.g. O'Hagan, 1978) and which has gained much attention in recent years (Rasmussen and Williams, 2006). A Gaussian process is the generalization of a Gaussian distribution. While the latter is the distribution of a random variable, the Gaussian process describes a distribution over functions. From function space view, the Gaussian process $f(x)$ can be determined by the corresponding mean $m(x)$ and covariance functions that are defined as follows:

$$m(x) = E(f(x)) \tag{1}$$

$$k(x,x') = E((f(x)-m(x))(f(x')-m(x'))) \tag{2}$$

where k(x, x') is the covariance (or kernel) function evaluated at x and x'.

A Gaussian process $f(x)$ can be represented as

$$f(x) \sim GP(m(x),k(x,x')) \tag{3}$$

Usually, for notational simplicity we will take the mean function to be zero and the offsets and simple trends can be subtracted out before modeling (Snelson, 2007; Kang et al., 2015).

2.1. GPs for regression

In GPs, we assume that the relation between the input vector and the target is given by

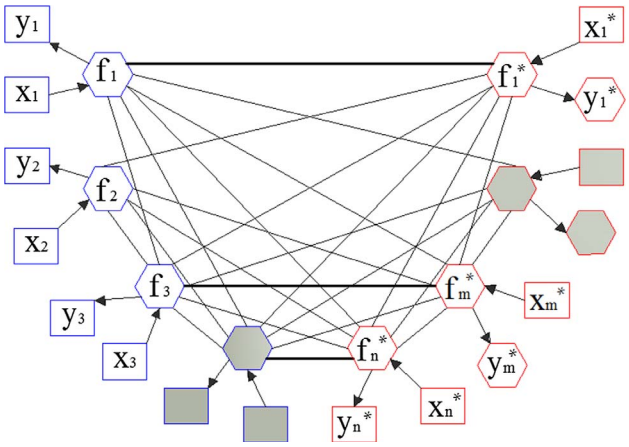

Fig. 1. Graphical model for GPR.

$$y_i = f(x_i) + \varepsilon \tag{4}$$

where $f(x)$ represents an arbitrary regression function while ε is the noise follows an independent, identically distributed Gaussian distribution with zero mean and variance σ^2, that is $\varepsilon \sim N(0, \sigma^2)$.

Furthermore, we assume that $f = [f(x_1), f(x_2), \ldots, f(x_n)]^T$ behaves according to a Gaussian process, that is $P(f|X = N(0,K)$, where K is the covariance matrix with element $K_{ij} = k(x_i, x_j)$.

$$K(X,X) = \begin{pmatrix} k(x_1,x_1) & k(x_1,x_2) & \cdots & k(x_1,x_n) \\ k(x_2,x_1) & k(x_2,x_2) & \cdots & k(x_2,x_n) \\ \vdots & \vdots & \vdots & \vdots \\ k(x_n,x_1) & k(x_n,x_2) & \cdots & k(x_n,x_n) \end{pmatrix} \tag{5}$$

where $K(X,X)$ is the $n \times n$ symmetric and positive definite covariance matrix, any of the matrix measures the correlation between x_i and x_j. The element K_{ij} is the covariance between values of the latent functions $f(x_i)$ and $f(x_j)$, and it encodes about the prior of our knowledge of nonlinear process among latent functions.

Gaussian process regression (GPR) is used to compute the predictive distribution of the function values f^* at test points $X^* = [\ x_1^*; x_2^*; \ldots; x_m^*]$. A graphical model representation of a GP is given in Fig. 1. In the figure, f_i denotes $f(x_i)$. The set of latent function values f_i indexed by the set of indices x_i is fully connected. Each of the connections represents the correlation between two latent variables, which are defined by a covariance function (Yuan et al., 2008; Kang et al., 2015).

The distribution of y conditioned on the values of f is given by an isotropic Gaussian

$$P(y|f,X) = N(f,\sigma_n^2 I) \tag{6}$$

where I is the identity matrix.

From the property of the Gaussian distribution, we can get the marginal distribution of y as

$$P(y|X) = \int p(y|f,X)p(f|X)df = N(0,K + \sigma_n^2 I) \tag{7}$$

The joint distribution of the observed target values and the function values at the test locations under the prior can be written as

$$\begin{bmatrix} y \\ f_* \end{bmatrix} \sim N\left(0,\begin{bmatrix} K(X,X) + \sigma^2 I & K(X,X_*) \\ K(X_*,X) & K(X_*,X_*) \end{bmatrix}\right) \tag{8}$$

where $K(X,X_*)$ is the $n \times 1$ covariance matrix of test point X_* and all the input points X; $K(X_*,X_*)$ is the self-covariance matrix of test point X_*.

After completion of the learning process, the most likely output value corresponding to X_* can be predicted on the basis of the training set according to bayesian theory. The purpose of using bayesian theory is to update the probability distribution by the observed real data, i.e., the most likely predictive distribution f_* can be deduced with the new input X_*, the input values X and the observed target values y of training