Structural Safety 61 (2016) 22–42

Contents lists available at ScienceDirect

Structural Safety

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

A hybrid approach combining uniform design and support vector machine to probabilistic tunnel stability assessment

Xiang Li [a,*], Xibing Li [a], Yonghua Su [b]

[a] School of Resources and Safety Engineering, Central South University, Changsha, Hunan 410083, China
[b] College of Civil Engineering, Hunan University, Changsha, Hunan 410082, China

ARTICLE INFO

Article history:
Received 18 December 2013
Received in revised form 4 March 2016
Accepted 6 March 2016
Available online 7 April 2016

Keywords:
Tunnel stability
Probabilistic assessment
Limit state function (LSF)
Response surface method (RSM)
Support vector machine (SVM)
Uniform design (UD)

ABSTRACT

Applications of the reliability-based method to stability evaluation of tunnel structures have become an ever-increasing concern over recent years. One critical challenge in conducting such a task is the implicit nature of the limit state function (LSF). To address this issue, the focus of this study is on, among others, the use of response surface method (RSM) by considering both the selection of the sampling method and the choice of the response surface form (as two major factors affecting the RSM's performance). In this context, the current paper develops for tunnel-reliability analysis a hybrid approach combing an experimental design called uniform design (UD) and a regression device known as support vector machine (SVM). For the proposed hybrid approach, the UD is used to generate sampling points and then the SVM is employed to construct the response surface approximating the original inexplicit LSF. Such an approach integrates the merits of both UD and SVM used for complex nonlinear modelling. Three carefully selected tunnel examples are illustrated: one for a typical tunnel under relatively simplified tunnelling conditions and the other two for real-life tunnels. Comparisons are made to validate the computational accuracy and efficiency of the present approach. In particular, for the tunnel example where the LSF is known only implicitly through the numerical analyses (which is the scenario of many real-world applications in tunnel community), the obtained results further demonstrate the efficiency of this approach: it can be much more economical to achieve reasonable accuracy than the conventional RSMs when a small number of sampling data is used. Such comparisons made in this work verify the application potential of the developed hybrid approach for probabilistic tunnel stability assessment involving the implicit LSF.

1. Introduction

In geotechnical engineering analysis and design, the existence of inherent uncertainties and their importance for the problems of decision making, risk evaluation and management have widely been recognized since the pioneering work reported in the 1960's [1,2]. The conventional deterministic approach to assessing the performance of geotechnical structures involves calculation of a factor of safety. Since such a method cannot explicitly and sufficiently characterize uncertainties and may sometimes be inclined to yield misleading results, it is desirable to employ a more logical and realistic treatment within a probabilistic framework (also called the reliability-based method) to cope with uncertainties in geotechnical structures [3]. The assessment of tunnel stability (or safety including strength and serviceability) is one important geotechnical subject dominated heavily by uncertainties. One of the earliest suggestions to utilize the probabilistic principles in tunnels was made in 1983 [4]. Since the beginning of the 1990's, a number of probabilistic studies of tunnels and other structures related to underground excavations have been published in the literature and the readers may refer to [5–14], for example. These valuable contributions set the foundation for illustrating the implementation and benefits of reliability-based method in this field.

For the stability evaluation in tunnel community, owing to the innate complexity (e.g., the ground-support interaction) and the multiple effects of diverse kinds of factors (e.g., rock mass conditions, excavation techniques and support structures, see some classical monographs [e.g., 15–18]), the mechanical model (corresponding to deterministic modelling) exhibits to a large extent a highly nonlinear behavior in many situations. As such, it may hence be not possible in the subsequent reliability analysis

* Corresponding author.
E-mail addresses: xli_xiangli@csu.edu.cn, xiangli_hnugeotech@126.com (X. Li), xbli@csu.edu.cn (X. Li), syh5327@hnu.edu.cn, yonghuasu@126.com (Y. Su).

http://dx.doi.org/10.1016/j.strusafe.2016.03.001
0167-4730/

to obtain an explicit, closed-form limit state function (LSF); instead, the resulting LSF can only be expressed implicitly. In this environment, estimating the probability of failure by the direct computation of a multifold integral could often become intractable, since the computational challenge in determining the integral lies in the multiple evaluations of LSF [19,20]. Among other various approximation methods, the so-called "fast probability integration" method [21] (i.e., the first- and second-order reliability methods) may be viewed as one of the most extensively used tools. When employing such a method for tunnel structures, calculation of the LSF's derivatives (which is essential to the reliability scheme) can be readily performed for the simple and explicit LSF [e.g., 22], whereas it could be hindered by the complicated and implicit LSF. To circumvent this problem, one possible way is to resort to the Monte-Carlo technique [e.g., 11], whose results obtained are known to be quite accurate except for the tremendous computational cost. With some improvements (like importance sampling [23], directional simulation [24], antithetic variates [25] and conditional simulation [26]) and further rampant growth in technology and availability of computational resources in the near future, wide application of this technique may be no longer infeasible for many practical problems. Yet the calculation at present is still quite time-consuming [20,27]. Several other methods could also be pursued for the purpose. They can mainly be divided into three categories: (a) method where the computation of the derivatives is avoided, including the point element method [28,29] and the spreadsheet algorithm [13,14]; (b) method where estimation of the derivatives is conducted via simple approximations, involving the rational polynomial approximation [30] and the difference approximation (proposed by the author and co-workers [31]); and (c) method where the original complex and implicit LSF is replaced by the simple and explicit function (called the response surface method (RSM)), typically the polynomial-based RSM [e.g., 7,32–34], the artificial neural network (ANN)-based RSM [e.g., 9,35,36] and more recently, the Hermite polynomial-based RSM [e.g., 37].

For the aforecited RSM, it has become one well-established class of methods to solve probabilistic tunnel stability problems with the implicit LSF. Generally, the RSM's performance is largely dependent on two factors: one is relevant to the sampling method for the location and number of data points selected to identify the response function and the other is pertinent to the response function shape adopted for fitting. Considering in tunnel-reliability analysis with complicated and inexplicit LSF, the polynomial- and ANN-based RSMs are currently the most two representatives, we take such two types of RSM as an example to elucidate, respectively, the two factors as follows:

First, for the former factor (i.e., the sampling method): in the context of polynomial-based RSM, the conventional factorial designs (e.g., the central composite design) may lead to the unacceptably high computational efforts with increasing the number of random variables for complex systems and even become more time consuming than direct Monte-Carlo method [38,39]. On the other hand the widely used interpolation scheme in sampling for the complicated LSF is hindered by the choice of an arbitrary parameter, whose extremely low variations may, however, trigger wild and unexpected fluctuations of the calculation results [40]. It is worth noting that the analyses of tunnel-reliability in [32] indicated another difficulty when using this interpolation scheme with the symmetrical pattern, and thus developed a modified sampling strategy in a nonsymmetrical manner. In the context of ANN-based RSM, one disadvantage of the frequently used random method in sampling ([9,41,42]) is that the randomly selected sampling points without uniformly covering the design space may lead to erroneous results, particularly when the number of variables is large and the number of sampling points is relatively small [27,43]. Additionally, since there is no-well defined criterion to determine the optimal network architecture, it is relatively difficult to identify the number of sampling points for building a good ANN model [44]. At this point, a case study for tunnel excavations in [45] proved the major effect of the number of sampling points used for the ANN model on how to reliably and completely predict the unknown relationship of weak geological zones.

Secondly, for the latter factor (i.e., the response surface): in the polynomial-based RSM, the main limitations reside in like the existence of false design points [46], the difficulty in handling a large number of random variables, especially for mixed or statistically dependent ones [9,41], the biased approximation of results for cases without conforming to the true LSF's nonlinearities [47] and the severe oscillations with increasing the polynomial order [20]. Particularly, the probabilistic stability analysis for tunnels in [32] showed that using the polynomial model fails to converge to the design point, and thus had to suggest a sophisticated function by modifying the polynomial model. In the ANN-based RSM, some issues are needed to be further examined like the difficulty in designing the network architecture, under- or over-fitting, local minimum and less generalization ability [48]. Several of them were documented in [49] when the ANN model had application in tunnelling for the ground movement prediction.

Considering the potential inadequacies of the current RSMs in some applications and meanwhile the intrinsic complexity of tunnel stability issue, examination of a suitable approach that provides results with reasonable accuracy and also alleviates the computational cost as effectively as possible could always be desirable for the RSM's application in tunnel structures. Motivated by this, we propose in the RSM framework a hybrid approach combining uniform design (UD; an experimental design corresponding to the sampling method) and support vector machine (SVM; a regression device corresponding to the response surface). The reason for developing such a hybrid approach is attributed to the appealing properties of UD and SVM both in modelling the complex nonlinear relationship. By integrating the merits of both UD and SVM, we here make an attempt to suggest an RSM to handle the implicit LSF in tunnel-reliability analysis. Although for either UD or SVM, so far each has been found for a range of engineering applications, the hybridization of UD and SVM in the RSM context for probabilistic tunnel stability assessment has yet to be investigated [50].

The remainder of this paper is framed as follows. In Section 2, some basic concepts of UD and SVM are, respectively, summarized, and this summary is followed by the concise presentation of the analysis procedure for the proposed hybrid approach. In the sequel, our emphasis is placed on its detailed applications in the context of tunnel structures in Section 3, where three selected tunnel examples are illustrated to demonstrate the accuracy and efficiency of our approach. Then Section 4 reviews the results obtained from the three examples, respectively. The conclusions are finally given in Section 5.

2. Proposed hybrid approach combining UD and SVM

2.1. Allocation of sampling points by UD

To find an approximate model for the implicit LSF in tunnel-reliability analysis, some merits of the UD provide us the most important motivation to use such a sampling strategy. They are concisely listed as follows (see [51,52] for detailed information): (a) help users in modelling with a small number of experiments; (b) accommodate the largest possible number of levels (i.e., representatives values taken) for each factor among many experiment designs; and (c) impose no strong assumption on the underlying model (i.e., the UD's performance is robustness against changes of the underlying model).