Engineering Geology 207 (2016) 1–13

Contents lists available at ScienceDirect

Engineering Geology

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

Bayesian approach for probabilistic site characterization assimilating borehole experiments and Cone Penetration Tests

C.H. Wang [a,c,*], B. Harken [b], C.A. Osorio-Murillo [b], H.H. Zhu [c], Y. Rubin [b,c]

[a] School of Railway Transportation, Shanghai Institute of Technology, Shanghai 200235, China
[b] Department of Civil and Environmental Engineering, University of California, Berkeley, CA, 94720, USA
[c] Department of Geotechnical Engineering, Tongji University, Shanghai 200092, China

ARTICLE INFO

Article history:
Received 13 July 2015
Received in revised form 15 January 2016
Accepted 3 April 2016
Available online 13 April 2016

Keywords:
Compression modulus
Bayesian approach
Maximum entropy
Site characterization
Data assimilation

ABSTRACT

This study investigates a Bayesian approach for assimilating data from borehole experiments and Cone Penetration Tests (CPTs) for site characterization (e.g., soil compression modulus). The spatial variability of compression modulus is depicted by random field theory. A Bayesian inverse modeling method is established by combining prior information, in the form of empirical knowledge, with observations, in the form of borehole experiments and in-situ CPT profiles to calculate posterior estimates of compression modulus at unsampled locations. The approach classifies all relevant data as either direct or indirect, and uses geostatistical tools in a Bayesian algorithm to estimate the posteriors. The local uncertainty of indirect data is analyzed using maximum entropy theory and the Bayesian algorithm simulates the posterior prediction of unsampled locations given direct data, plus integrating out the local uncertainties on joint Probability Density Function (PDF) of indirect data. To validate its effectiveness, Bayesian inverse modeling method is applied to a shallow load bearing soil layer located at the National Convention and Exhibition Center in Shanghai. Spatial estimates are refined in comparison to those from Kriging method. It is concluded that the above approach appears promising as a framework for site characterization given multiple types of observation.

1. Introduction

Development and implementation of reliability-based geotechnical design methodologies at present requires probabilistic site characterization. Spatial variability of soil properties has been the subject of much research (e.g., Matheron and Armstrong, 1963; Vanmarke, 1977; Rubin, 2003), and multiple technologies are widely used in site characterization (e.g., Emeriault et al., 2004; Zhao et al., 2006; Rogers, 2006). Challenges to the existing approaches include: (1) the lack of identification of spatial variability of soil properties; (2) the complexity in identifying the uncertainties of different types of observation; and (3) the difficulty in assimilating multiple types of observation in site characterization.

The first problem arises due to the intrinsic spatial variability of soil properties and the fact that only a small proportion of them are observed by geotechnical investigation. Therefore, research interest has been concentrated in interpreting the spatial variability using geostatistics. For example, Vanmarke (1977) utilized the concept of spatial correlated distance in probabilistic modeling of soil profiles, which analyzed variance with spatial scaling for a statistically isotropic soil layer. Following that, Tang (1979) used a variance deduction factor in probability evaluation of CPT vertical penetration resistance. Juang et al. (1999) assessed liquefaction potential of soils employing a reliability-based method. Phoon and Kulhawy (1999) characterized soil properties as random fields which can be described concisely by a Coefficient of Variation (COV) and a scale of fluctuation. Fenton (1999) and Fenton and Griffiths (2007) presented the study into the effect of soil spatial variability on the settlement and ultimate load capacity of piles. Bake and Faber (2008) proposed a method to quantify potential liquefaction by accounting for spatial variability of soil properties and possible future earthquakes in Turkey. Wang et al. (2013) assumed CPT data as random variable in probabilistic identification of underground soil stratification. Zhang et al. (2014) discussed the choice of statistical distributions with limited prior information in slope reliability calculation. And Gong et al. (2014) optimized the site exploration program to improve prediction of tunneling-induced ground settlement in clays.

On the other hand, multiple geophysical technologies (Hubbard and Rubin, 2000; Hou and Rubin, 2006; Porsani et al., 2012) are widely used in geotechnical engineering, which provide more data for site characterization, but the data should be classified in a proper way. For example, lab tests of soil properties from boreholes and in-situ CPTs are obviously

* Corresponding author at: School of Railway Transportation, Shanghai Institute of Technology, Shanghai 200235, China.
E-mail addresses: changhong.wangsir@gmail.com (C.H. Wang), harken@berkeley.edu (B. Harken), carosocali@gmail.com (C.A. Osorio-Murillo), zhuhehua@tongji.edu.cn (H.H. Zhu), rubin@ce.berkeley.edu (Y. Rubin).

http://dx.doi.org/10.1016/j.enggeo.2016.04.002
0013-7952/

different types of observation. Another problem is how to assimilate these multiple types of observation to obtain more accurate estimates of soil properties in site characterization.

Bayesian theory has a well-established role in the field of geotechnical engineering. For example, Lee and Kim (1999) imported an extended Bayesian method for site characterization in a shallow tunneling project. Soil properties were treated as random variables, and a Bayesian framework provided their mode values by combining prior knowledge and field measurements, but the posterior distributions of the target variables were not discussed. Ching and Chen (2007) characterized geotechnical model uncertainty in a Bayesian framework with a transitional Markov chain Monte Carlo sampling method, and Wang et al. (2012) used centrifuge data in Bayesian updating of a semi-empirical mechanics model. However, the spatial variability was not taken into account. Wang et al. (2010), Cao and Wang (2013) and Cao et al. (2014) discussed a Bayesian approach for probabilistic site characterization that depicts the spatial variability of soil properties explicitly by combining prior information and geotechnical observations, which provided the most probable number of layers, thicknesses, and soil properties, but concentrated on a single type of observation.

The present study establishes a Bayesian inverse modeling method for assimilating multiple types of information including prior knowledge, borehole experiments and in-situ CPTs, then estimates the compression modulus at unsampled locations in a statistically isotropic soil layer. Key elements in the method include: review of data availability and data classification, depiction of the spatial variability of compression modulus using random field theory, and assimilating the two types of observation for point estimation within a Bayesian framework. The paper is structured as follows: Section 2 introduces the background and the observed data of the project. Section 3 presents the Bayesian algorithm along with the depiction of local uncertainty of indirect data based on maximum entropy theory. In Section 4, the algorithm is applied to the background project to explore the effect of data assimilation of a statistically isotropic soil layer and mapping site characterization of compression modulus. The paper ends by concluding that the Bayesian method is more flexible in the multiple types of data assimilation and more accurate in the target estimation.

2. Site description and observations

The National Convention and Exhibition Center in Shanghai is the largest single-block building in the world. It is 1.5 km from the Shanghai Hongqiao Transportation Hub. The primary facilities include exhibition areas *A*, *B*, *C*, and *D* and a transportation center, as shown in Fig. 1. In this paper we focus on the site characterization of exhibition region *A* which is leaf-shaped, with the largest dimension equal to about 350 m.

Information available from the engineering geological report includes the interfaces separating the various soil layers, ground water level and soil properties such as coefficient of cohesion, internal friction angle, compression modulus, and void ratio (shown in Fig. 2). The silty clay layer (i.e., geological symbol ②$_1$, the third soil layer from the ground surface, average thickness is about 2.0 m) is selected as the objective of study, because it is the load bearing soil layer of the shallow foundation of the exhibition structure. Compression modulus is identified as the target variable, since its spatial variability in the horizontal plane is a leading cause of differential settlement in large continuous foundation projects (Baecher and Christian, 2006). Geotechnical observations consisted of $n_a = 21$ borehole logs and $n_b = 109$ in-situ CPTs, as shown in Fig. 3. Each borehole provides a laboratory value for the compression modulus directly, and penetration resistance of CPTs is sampled each 0.1 m along the drill bar. Table 1 lists several samples for both types of observation. Due to the impacts of soil interfaces, five samples at the top and bottom of each CPT profile are not taken into account.

2.1. Data classification

The intrinsic spatial variability of compression modulus is represented by a random field function: $Z(\boldsymbol{x})$, where the random field is denoted by Z and the bold letter $\boldsymbol{x}$ denotes the Euclidian coordinates vector. A

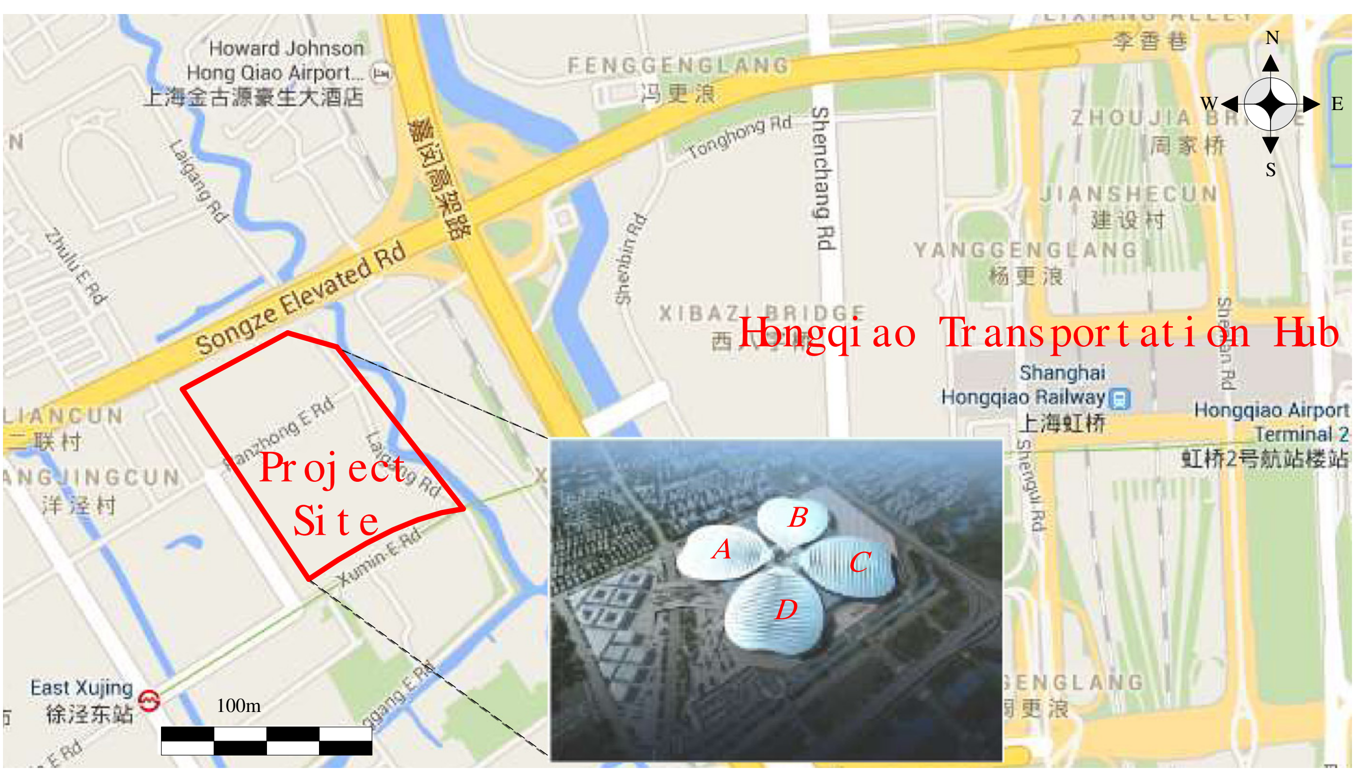

Fig. 1. Location of the background project.