Automation in Construction 90 (2018) 243–252

Contents lists available at ScienceDirect

Automation in Construction

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

Quantifying cost-effectiveness of subsurface strata exploration in excavation projects through geostatistics and spatial tessellation

Yat Fai Leung*, Wenfei Liu, Yangxue Lei, Shu-Chien Hsu

Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong

ARTICLE INFO

Keywords:
Excavation
Contingency budget
Geotechnical variability
Delaunay triangulation
Voronoi cells

ABSTRACT

A major source of uncertainty in civil engineering projects arises from the geological and geotechnical variability at the project site. This paper presents an approach to quantify such uncertainty, and to rationally incorporate them into estimates of cost contingency. Contrary to the conventional approaches that rely on individual expert's opinion or data from past projects, the proposed approach allows site-specific assessments of the intrinsic geotechnical variability through geostatistical techniques. Such uncertainty can be reduced through geotechnical investigation, and spatial tessellation techniques are proposed to facilitate determination of the optimal locations of new boreholes. Cost-effectiveness of the boreholes can be evaluated based on the corresponding reductions in geotechnical uncertainty and their influence on the budget. The approach is illustrated using a hypothetical excavation scenario, where the project costs are affected by uncertainty in the subsurface strata, particularly the rockhead level across the site. Under the specific site conditions, a Pareto frontier is developed to reveal the relationship between the number of boreholes to be drilled and potential savings in contingency budget. Through this approach, the study promotes better utilization of geotechnical information and rational assessments of project risks associated with their variability, which may lead to improved project planning and resource allocation.

1. Introduction

Uncertainty in geotechnical engineering is a well-known, yet inadequately understood topic in the civil engineering profession. Delay and cost overrun of many large-scale infrastructure projects have been attributed to 'unforeseen' and complex geological and geotechnical conditions. According to a survey of 28 construction projects in the United Kingdom [1], more than 40% of the geotechnical problems encountered during construction arise from uncertainties related to the subsurface strata and the geotechnical properties. To reduce such uncertainties, geotechnical investigation (e.g., rock and soil sampling and testing) can provide additional information about the ground conditions at the site. However, there is currently no quantitative approach to relate this to the level of uncertainty across the site, or to elucidate how the project risks may be reduced through the additional information. Consequently, practitioners often rely on their individual experience or intuition when planning the geotechnical investigation programme. The qualitative nature of this practice makes it difficult to assess the cost-effectiveness of the investigation, or its implications on the overall budget and delivery time of the construction project. The problem can be exacerbated in infrastructure mega-projects, where delay in one part of the project often triggers cascading effects to the entire development plan. From a management standpoint, a cost or time contingency is usually included in the project budget or programme, as a common approach to control the risks of delay and cost overruns due to unforeseen conditions. In fact, the contingency budget or the 'float' of a particular task should be decided according to the level of uncertainty associated with the task. For excavation projects, it is therefore beneficial to quantify the geotechnical uncertainty, which then allows rational planning and apportioning of the risks.

Traditionally, cost contingency is incorporated as a simple percentage addition onto the base (cost) estimate, considering specific project features, past experience and historical data [2]. For large and complex projects, more rational estimates may be obtained either through deterministic or probabilistic approaches [3]. Some of the common deterministic approaches include linear regression models, artificial neural networks (ANN) for more complex problems, and Least Squares Support Vector Machine (LS-SVM) in price variation modelling for construction management. For example, Sonmez et al. [4] proposed a linear regression model to predict the bidding contingency amount for contractors, by focusing on the major influential factors of contingency decisions identified from previous projects, while Thal et al. [5]

* Corresponding author.
E-mail address: andy.yf.leung@polyu.edu.hk (Y.F. Leung).

https://doi.org/10.1016/j.autcon.2018.02.032
Received 22 June 2017; Received in revised form 13 February 2018; Accepted 17 February 2018
0926-5805/

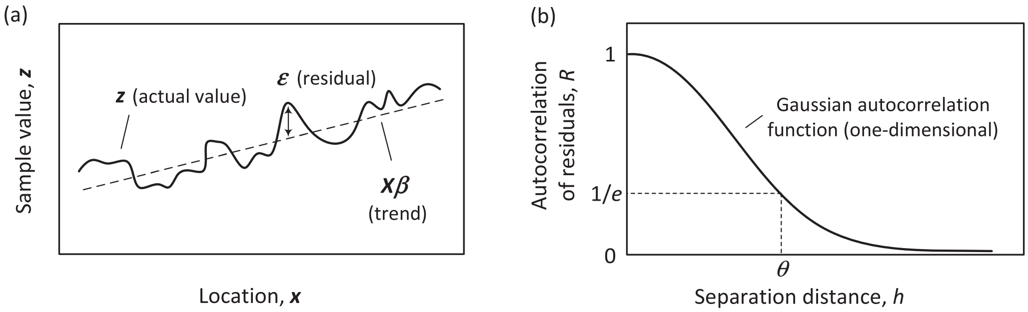

Fig. 1. (a) Trend and residuals of spatial variables (after DeGroot and Baecher 1993); and (b) Autocorrelation of residuals.

developed a multiple linear regression model for similar purposes. Although the regression method outperforms the practice of assigning an arbitrary percentage, a linear relationship may not be able to best fit the available historical data [6]. Therefore, ANN were utilized to perform nonlinear regression for more complex problems. These include the work by [7], who built a back propagation general regression neural networks (GRNN) to determine the cost contingency and allocation strategies at the preliminary stage. Lhee et al. [8] further proposed a two-step ANN-based model adding contingency rate as an intermediate output variable. Meanwhile, Cheng et al. [9] established a hybrid system based on Least Squares Support Vector Machine (LS-SVM) for modelling construction price variations, which can be used for decision making in construction management. Although these previous studies have illustrated the potentials of deterministic approaches, a few major criticisms remain regarding their applications. These include the heavy influence by subjectivity of individual experts, deficiency in accurately quantifying the project risks, and the fact that some of these techniques work like a 'black-box' [10, 11].

Probabilistic approaches were advocated to tackle these deficiencies. For example, Khalafallah et al. [12] proposed the Bayesian Belief Network to quantify project risk and uncertainty level, which allows the determination of the appropriate contingency percentage for construction projects. This approach was further developed by Kim et al. [13] to assess the probability of construction project delays based on case studies in developing countries. Meanwhile, other researchers adopted the Monte Carlo Simulation (MCS) to quantify cost contingency at different risk levels [14]. Since the risk factors in construction projects often contain both 'random' and 'fuzzy' variables [15], a Fuzzy MCS framework was established by Sadeghi et al. [16] to evaluate both components in the estimation of contingency range.

These previous approaches mainly rely on historical data or qualitative experts' opinion and experience [17], with little discussion on the intrinsic source of uncertainty. This paper attempts to quantify a major source of uncertainty in excavation projects, arising from the geotechnical and geological variability at the project site. An automated strategy is proposed to quantify geotechnical uncertainty, and to evaluate its changes with additional boreholes in the project site. It accounts for site-specific geologic features based on the available existing information, which may include irregularly-spaced boreholes revealing variations of subsurface strata in different directions. The quantitative approach enables optimization to be performed to determine the number and locations of sampling points that lead to the most cost-effective investigation programme, with respect to the impacts on time and costs of the tasks. The proposed approach will be demonstrated through the scenario of an excavation project, where the major uncertainty arises from the variations of rockhead level across the site. Such variations heavily influence the quantity of rock materials to be excavated, and hence the planning of project budget and delivery time.

2. Methodology

This study utilizes the geostatistical approach discussed by Liu et al. [18] and Liu and Leung [19] to quantify the geotechnical variability associated with subsurface strata. Meanwhile, spatial tessellation techniques are adopted for the derivation of optimal geotechnical sampling strategies. Their cost-effectiveness can be evaluated through the reductions of uncertainty, and the subsequent implications on the budget and time of the excavation project. The three individual components of the proposed approach are described in the following sections.

2.1. Quantification of geotechnical variability

Liu et al. [18] and Liu and Leung [19] presented the details and verification of an integrated framework established to characterize the spatial variability of geological profiles and geotechnical properties. This will be described briefly herein as it forms the basis of the sampling strategy proposed in this study. In general, the spatial variations of the subsurface strata ($\boldsymbol{z}$) can be represented by a linear mixed model consisting of a large-scale trend ($\mathbf{X}\boldsymbol{\beta}$), and the residual effects ($\boldsymbol{\varepsilon}$) that describe the deviations of the actual values from the trend (Fig. 1):

$$\boldsymbol{z} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{1}$$

where $\mathbf{X}$ is a matrix containing information of the spatial coordinates of sampled points, and $\boldsymbol{\beta}$ represents the trend coefficients. $\boldsymbol{\varepsilon}$ is often observed to be spatially correlated, with greater variations between components ε_i and ε_j associated with larger separation distances between locations i and j. Accordingly, the variance of $\boldsymbol{\varepsilon}$ can be represented by:

$$\mathbf{V} = \sigma_e^2\mathbf{R} + \sigma_n^2\mathbf{I} = (\sigma_e^2 + \sigma_n^2)[s\mathbf{R} + (1-s)\mathbf{I}]$$
$$\text{where } 0 \le s = \frac{\sigma_e^2}{\sigma_e^2 + \sigma_n^2} \le 1 \tag{2}$$

where $\mathbf{I}$ is the identity matrix; σ_n^2 are the random natural effects (white noise effects) which are independent of separation distances; σ_e^2 are the smooth scale variations, or the component of total variance that correlate with separation distance, and such autocorrelation is described by the matrix $\mathbf{R}$. s is referred to as the spatial dependence, and represents the proportion of σ_e^2 within the total variance. Individual components of $\mathbf{R}$ (R_{ij}) describe the correlations between ε_i and ε_j, and the relationship between R and separation distance (h_{ij}) can be modelled by different mathematical functions, such as the exponential, Gaussian (squared exponential), or spherical function, all of which involve a parameter θ that defines the range of correlation. Alternatively, the scale of fluctuation (δ) is another parameter used to define the extent of the correlation [20], and is often taken as the separation distance where the autocorrelation R drops to the value of 0.05. The parameters θ and δ are related to each other according to the adopted correlation function. For example, $\delta \approx \sqrt{\pi}\theta$ for the Gaussian function.

Site-specific characterization of the spatial features mainly involves determination of the trend $\boldsymbol{\beta}$, together with correlation parameters s