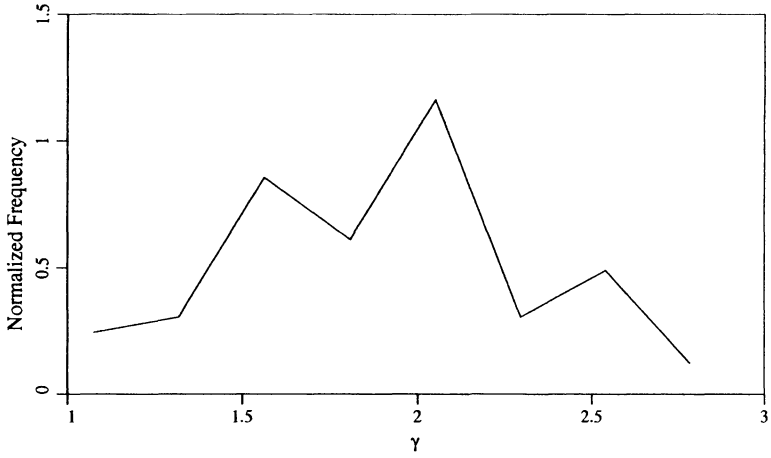

Figure 10. Distribution of ML estimates of γ via the wavelet basis over all 67 CPT sites.

Figure 11 shows a histogram of γ estimates over the 67 sites using the periodogram maximum likelihood estimator discussed in Fenton (1999a). The average estimated γ value is now 1.9 ± 0.38 with an overall range from 1.04 to 2.70. The periodogram estimates agree very well with those obtained using the wavelet basis, and have somewhat superior precision.

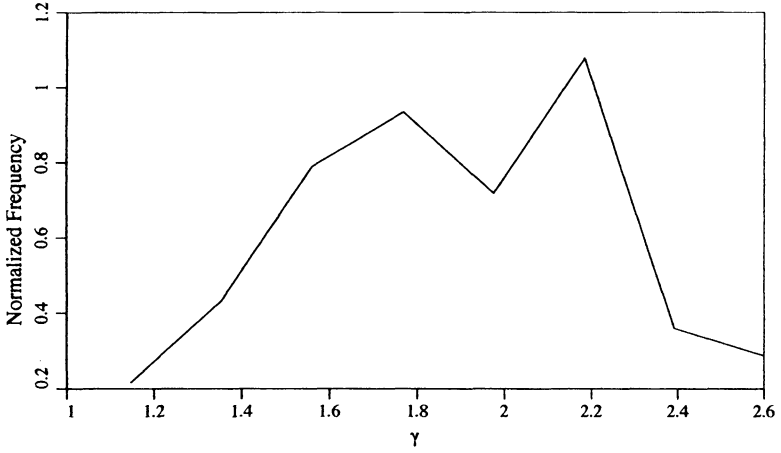

Figure 11. Distribution of ML estimates of γ obtained using the periodogram over all 67 sites.

In the following, the periodogram estimator will be used alone, the wavelet estimator being viewed primarily as corroborative evidence on the estimate of γ.

The spectral exponent γ may be equivalently expressed by the fractal dimension $D_f = (5 - \gamma)/2$ or the Hurst coefficient of fractional Brownian motion, $H = (\gamma - 1)/2$. The latter coefficient can lead to some confusion, however, since the Hurst coefficient for fractional *Gaussian* motion, the derivative of the Brownian motion, which applies when $0 \leq \gamma < 1$, is $H = (\gamma + 1)/2$.

The spectral intensity parameter G_o was found to have an average value of 0.45 and a standard deviation of 0.42. This parameter is related to the process variance and both are considered to be site specific and also *very* affected by the type of data transform used. The result quoted here is for the $\ln q_c$ transformation. Over the 67 CPT data sites, the average process variance was estimated (by method of moments) to be 0.84 and a standard deviation of 0.93 and the mean of the $\ln q_c$ data was estimated to be 1.02 ± 0.66.

Armed with these parameters, which can be summarized as $\hat{\mu}_X = 1.02$, $\hat{G}_o = 0.45$, and $\hat{\gamma} = 1.9$, one can presumably create a stochastic model for the $X = \ln q_c$ process. Then letting $q_c = \exp\{X\}$ recovers the desired process. There are, however, a few more details to work out. Recall that the fractal process has infinite variance, and for $\gamma \geq 1$ will behave as if non-stationary. A process with infinite variance is physically unrealizable, and the apparent non-stationarity is inconvenient since it introduces an origin issue. The stochastic model must be modified in order to become physically realizable and useful. For this purpose a lower frequency cutoff, ω_o, is introduced. An appropriate spectral density function is of the form illustrated in Figure 12. The lower frequency cutoff can logically be taken to be $2\pi/D$, where D is the soil depth (Fenton, 1999b). This suggestion is perhaps particularly appropriate if the spectral intensity at the target site is unknown. However, if one or more CPT soundings are made at the target site and the spectral intensity estimated, then the cutoff frequency can be estimated by matching the area under the spectral density function shown in Figure 12 to the estimated process variance.

According to theory, the variance can be obtained from the spectral density function according to

$$\sigma_x^2 = \int_0^\infty G(\omega)\, d\omega = \int_0^{\omega_o} \frac{G_o}{\omega_o^\gamma}\, d\omega + \int_{\omega_o}^\infty \frac{G_o}{\omega^\gamma}\, d\omega = G_o \left(\frac{\gamma}{1-\gamma} \right) \omega_o^{1-\gamma} \tag{7}$$

which can be inverted to solve for the required lower cutoff frequency in terms of the estimated variance and spectral parameters,

$$\omega_o = \left(\frac{\hat{G}_o \hat{\gamma}}{\hat{\sigma}_x^2 (\hat{\gamma} - 1)} \right)^{1/(\hat{\gamma}-1)} \tag{8}$$

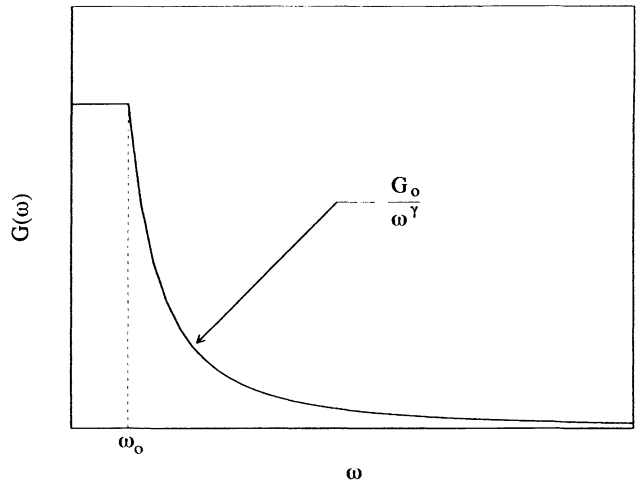

Figure 12. Truncated spectral density function for fractal processes with $\gamma > 1$.

Because of the fractal nature of the data, the cutoff frequency should be selected for a particular sample length D. For example, the CPT sounding shown in Figure 1 has length $D = 30.7$ m and parameter estimates $\hat{\gamma} = 2.23$, $\hat{G}_o = 0.203$, and $\hat{\sigma}_x^2 = 1.335$. For these values, the cutoff frequency, is given by Eq. (8) to be $\omega_o = 0.351$ rad/m. This corresponds to a cutoff frequency almost twice that suggested in the absence of knowledge of the spectral intensity. The frequency cutoff really controls how stationary realizations of the model appear. A smaller cutoff frequency leads to realizations with more pronounced apparent 'trends' over the soil depth, although the 'trends' in this case are actually random. Note that 'trends' of this sort were actually observed in the NGES data, as discussed in Section 2.

A realization of the resulting random process, with truncated spectral density function

$$G(\omega) = \begin{cases} \hat{G}_o/\omega_o^{\hat{\gamma}} & \text{if } 0 \le \omega < \omega_o \\ \hat{G}_o/\omega^{\hat{\gamma}} & \text{if } \omega > \omega_o \end{cases} \tag{9}$$

using the parameters found above for the CPT data of Figure 1, is shown in Figure 13. This realization was produced using the Fast Fourier Transform method (Fenton, 1994). Aside from its increased high frequency content, it has very much the same statistical nature as seen in Figure 1. Note that it is not expected to be identical since it is merely one possible realization. It does, however, include the high-frequency content essential to the fine-scale self-similarity of a true fractal process. If such detail is not desired, it can be eliminated by placing an upper bound on the spectral density function of Eq. (9), logically related to the sampling interval, or by passing each realization through a low-pass filter.

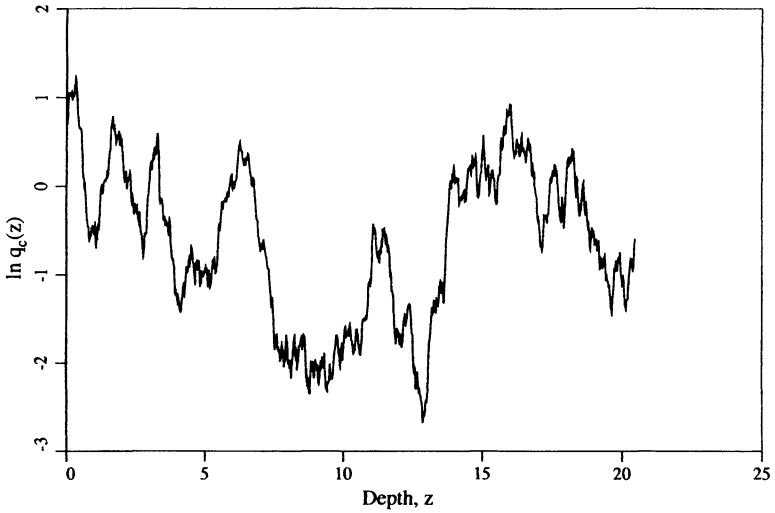

Figure 13. Simulated ln q_c record at the same site as in Figure 1 using $\gamma = 1.9$.

Using the result that the scale of fluctuation is proportional to $G(0)$ (Vanmarcke, 1983), an equivalent scale of fluctuation can be associated with the lower cutoff frequency applied to the fractal model, as follows

$$\theta^* = \frac{\pi \hat{G}_o}{\hat{\sigma}_x^2 \omega_o^{\hat{\gamma}}}. \tag{10}$$

Interestingly, the equivalent scale of fluctuation for the sounding shown in Figure 1 is computed to be $\theta^* = 4.9$ m, while its maximum likelihood estimate under the Gauss-Markov model is $\hat{\theta} = 94.5$ m, representing a theoretical long-range component implicit in this model. Over all 67 sites, the average scale estimated under the Gauss-Markov model was 51 m, with a very large standard deviation of 178 m. Note in the expression for θ^* the inverse proportionality to ω_o raised to the power γ, indicating that modest changes in the estimated values for ω_o or γ can yield dramatically different estimates for the scale of fluctuation.

Once the cutoff frequency, G_o and γ are known, the model is fully specified, at least as far as the second moment is concerned. Unfortunately, no simple closed form expression for the covariance function corresponds to the truncated fractal spectral density function of Eq. (9), and so this quantity is most easily obtained by numerical integration using the Wiener-Khinchine relationship

$$C(\tau) = \int_0^\infty G(\omega) \cos(\omega\tau)\, d\omega \tag{11}$$

if it is desired. Because the integrand above contains the cosine function, it alternates in sign. Numerically, Eq. (11) is therefore prone to errors due to so-called catastrophic

cancellation (loss of accuracy in the difference between two large numbers). The same integral can be written

$$C(\tau) = \frac{1}{\tau} \sum_{k=0}^{\infty} \int_0^{\pi/2} \Delta_G(k, u) \cos u \, du \tag{12}$$

where, for $\alpha_{ku} = 2\pi k + u$,

$$\Delta_G(k, u) = G\left(\frac{\alpha_{ku}}{\tau}\right) + G\left(\frac{\alpha_{ku} + 1.5\pi}{\tau}\right) - G\left(\frac{\alpha_{ku} + \pi}{\tau}\right) - G\left(\frac{\alpha_{ku} + 0.5\pi}{\tau}\right) \tag{13}$$

which, still involving differences, can at least be analytically approximated when the sum of the first two terms is very similar to the sum of the second two terms. The remaining integration in Eq. (12) can be performed, in the usual way, for example by using Gaussian quadrature.

Summary

The primary result of this paper lies in the inference, based on various ways to characterize second-order properties of random functions, that the vertical variation of CPT q_c data at the NGES sites appears to be fractal in nature. If CPT soundings are available at the target site for which the stochastic soil model is to be applied then all of the fractal parameters, including the spectral exponent, can be estimated at the site. In that case, the major use of this study is in establishing basis for the use of the fractal model and an indication of the variability of the estimators.

Alternatively, if data taken at a similar target site are sufficient only to establish the mean and variance of the desired soil property, then an *a-priori* second moment model for the vertical soil variability would have $\gamma = 1.9$, a lower frequency cutoff equal to about $2\pi/D$, where D is the soil depth, and a spectral intensity G_o computed so that the area under the spectral density function is equal to the estimated variance (see Eq. 7).

Clearly there is much work yet to be done on the inferential characterization of soils. For one, this study considers a hypothetical single site, a composite of the 5 sites in the NGES program and ignores information about site-specific geology, layering, and soil type (clay v. sand). The issues related to the choice of a lower cutoff frequency need additional clarification as this parameter is still somewhat arbitrary, related to the depth of the soundings. Also needed is a good error model to distinguish between the real soil behaviour and measurement error. Because such a model involves the estimation of additional parameters, an even more extensive database may be required in order to make confident inferences. It is felt, however, that the methodologies and reasoning set out here lay the groundwork for additional inferential studies on one-dimensional (vertical) variation in soil properties. In three dimensions, fractal models will presumably still apply, so that researchers can concentrate on estimation issues.

References

Benoit, J. and Lutenegger, A.J. eds. (Apr 2000). *Proceedings of the Geo-Institute Specialty Conference on Performance Confirmation of Constructed Geotechnical Facilities*, University of Massachusetts–Amherst, Geotechnical Special Publication No. 93.

Beran, J. (1994). *Statistics for Long-Memory Processes*, in *Monographs on Statistics and Applied Probability*, Chapman & Hall, New York, NY.

Fenton, G.A. (1999a). "Estimation for stochastic soil models," *ASCE J. Geotech. Geoenv. Eng.*, **125**(6), 470–485.

Fenton, G.A. (1999b). "Random field modeling of CPT data," *ASCE J. Geotech. Geoenv. Eng.*, **125**(6), 486–498.

Fenton, G.A. (1994). "Error evaluation of three random field generators," *ASCE J. Engrg. Mech.*, **120**(12), 2478–2497.

O' Neill, M. (2002), "Spatial Variability at University of Houston NGES," in *Probabilistic Site Characterization at the National Geotechnical Experimental Sites*, ASCE Geotechnical Special Publication.

Priestley, M.B. (1981). *Spectral Analysis and Time Series*, Vol. 1: Univariate Series, Academic Press, New York, NY.

Tumay, M.T. (Feb. 1998). "In Situ Testing at the National Geotechnical Experimentation Sites (Phase 2)", Federal Highway Administration Contract No. DTFH61-97-P-00161, Louisiana Transportation Research Center, Louisiana State University, Baton Rouge, LA.

Vanmarcke, E.H. (1983). *Random Fields: Analysis and Synthesis*, The MIT Press, Cambridge, Massachusetts.

Wornell, G.W. (1996). *Signal Processing with Fractals: A Wavelet Based Approach*, in *Prentice Hall Signal Processing Series*, Prentice Hall, Upper Saddle River, NJ.

Notation

The following symbols are used in this paper:

a, b, r = constants

$\hat{C}(\tau_j)$ = estimated covariance function at discrete lag τ_j

D = CPT sample length

D_f = fractal dimension

f_s = side friction from CPT data, kPa

$G(\omega)$ = one-sided spectral density function

$\hat{G}(\omega)$ = sample one-sided spectral density function

G_o = spectral intensity parameter

$\hat{G}_o$ = estimated spectral intensity parameter

H = Hurst or self-similarity coefficient

m = scale (dilation) index for wavelet basis

n = number of observations in a sample of $X(z)$

q_c = cone resistance from CPT data, kPa

$\hat{V}(\tau_j)$ = estimated variogram

X_i = random value of transformed CPT data at z_i

$X_{i,j}$ = average of $X_j, X_{j+1}, \ldots, X_{j+i}$

X_j^m = random wavelet coefficient

x_i = observed value of transformed CPT data at z_i

z = depth coordinate
z_i = discrete points along z
γ = spectral exponent in fractal model
$\hat{\gamma}$ = estimated spectral exponent
$\hat{\gamma}(i)$ = sample discrete variance function
Δz = incremental distance between observations
$\Delta_G(k, u)$ = differencing operation on $G(\omega)$
ϵ = residual random process
θ = scale of fluctuation
$\hat{\theta}$ = estimated scale of fluctuation
θ^* = equivalent scale of fluctuation
μ_x = mean of $X(z)$
$\hat{\mu}_x$ = estimated mean of $X(z)$
μ = pore pressure from CPT data, kPa
$\hat{\rho}(\tau_j)$ = estimated correlation function at discrete lag τ_j
σ_x^2 = variance of $X(z)$
$\hat{\sigma}_x^2$ = estimated variance of $X(z)$
τ = separation distance
τ_j = discrete separation distance, = $j\Delta z$
d = representative length
ω = frequency or wavenumber
ω_o = lower frequency cutoff

A New Approach to Site Characterization Using Generalized Regression Neural Networks

C. H. Juang[1] and T. Jiang[2]

Abstract

Site characterization is an important task in geotechnical engineering practice. The ultimate goal in site characterization is to be able to estimate in situ soil properties at any half-space point at a site based on limited tests. This estimate may be a point estimate or in terms of some statistical parameters. Geostatistical and random field methods have been applied with various degrees of success. This paper presents a new approach, based on artificial neural network, for site characterization. Emphasis is placed on application of generalized regression neural networks for site characterization. The results show that neural network approach has a potential to be a practical tool for site characterization.

Introduction

The ultimate goal in site characterization is to be able to estimate the *in situ* soil properties at any half-space (subsurface) point at a site based on limited number of tests. In the conventional approach, the engineer characterizes a site based on limited test results and interprets them in terms of working soil profiles. These working profiles represent a simplified model of the in situ soil properties. The task of establishing a working profile requires generalization of soil properties based on limited test data. In this paper, a new approach using neural network technology is developed for the very task of soil property generalization.

---

[1]Professor, Department of Civil Engineering, Clemson University, Clemson, SC 29634-0911. E-mail: hsein@clemson.edu.

[2]Research Assistant, Department of Civil Engineering, Clemson University, Clemson, SC 29634-0911. E-mail: taoj@clemson.edu.

Mapping and predicting in situ soil properties is a very challenging task (Tabba and Yong, 1981a, b). Uncertainty about the properties of in situ soils may come from various sources including spatial variability, measurement noise, measurement and model bias, and statistical error due to limited measurements (Baecher, 1986). To better characterize in situ soil properties and to provide a basis for subsequent probabilistic analysis, probability theory and geostatistics have been attempted. The classical work on random field for probabilistic site characterization has been documented by VanMarcke (1983). Characterization of spatial variability of soils using geostatistics has been reported (Soulie et al. 1990; Chiasson et al. 1995; DeGroot, 1996). Random field theory and geostatistics in site characterization using the National Geotechnical Experimentation Sites (NGES) data have been reported (Fenton, 1998; Kulatilake, 1998; Wu, 1998).

The approaches taken by the exiting probabilistic and geostatistics methods for site characterization may be viewed as a calibration process. Limited soil test data are used to calibrate a theory or model, i.e., to estimate some *undetermined* coefficients in the model or to derive some statistical parameters that characterize the site. In the present study, a different approach to site characterization based on artificial neural network (ANN) is taken. The neural network approach, which does not require a specific physical model to begin with, focuses on learning of patterns in the data.

Generalized Regression Neural Networks

ANN technology has been applied to address many civil engineering problems (Adeli, 2001). This technology has been shown to be effective in two general types of data analysis: classification and function approximation. The problem of site characterization is treated as a task of function approximation in this paper. In essence, an unknown function, $u = f(x,y,z)$ where u is a soil parameter, x, y, and z are the coordinates of a half-space (subsurface) point, and f is the unknown function, is to be approximately characterized based on limited test data of u. More specifically, the problem of site characterization is seen as a task to obtain an approximation of this unknown function f.

An artificial neural network *learns* from examples (i.e., input and output pairs). The goal of network learning is to *generalize* the relationship between the input and the output. To achieve the goal of generalization, the network needs to be trained and tested. Thus, *training* and *testing* are two major tasks in the development of a neural network. Generally, all available data for the development of a neural network is separated randomly into two subsets, a training data subset and a testing data subset. A neural network with the desired topology and training criteria is then preceded with the task of training, followed by the task of testing. In many cases, another data are obtained to further validate the trained and tested network.

Two types of neural networks, *backpropagation* (BP) network and *radial basis* (RB) networks, are commonly used for function approximation. BP networks use backpropagation algorithm (Rumelhart et al., 1986) to gradually reduce the prediction error of the trained network to meet a training goal of error. BP networks with an input layer, a sigmoid hidden layer, and a linear output layer are capable of approximating any function with a finite number of discontinuities (Demuth and Beale, 1999). Note that BP networks perform *global mapping*, meaning each and every input activates all neurons connected to it and causes an output. In some problems, such as site characterization, soil property at a half-space point could be more accurately estimated based on the soil property at its neighboring half-space points than that at far away locations. In such cases, it may be desirable to perform *local mapping*, meaning only the input near a *receptive field* (i.e., a range centered at that input) produces an activation of neurons. The concept of local mapping is analogous to kriging operation in geostatistics. In the present study, generalized regression neural network, which is a variant of RB network (Specht, 1991), is utilized to approximate the unknown function $u = f(x,y,z)$.

Generalized regression neural network (GRNN) consists of three layers, the input layer, the hidden layer and the output layer. In a GRNN, the hidden layer is a nonlinear *local* mapping. This layer contains *radial-basis* neurons that use the Gaussian transfer function (Figure 1). The Gaussian function is centered over each receptive field in the input space. If an input vector lies in a receptive field, the corresponding local neurons in the hidden layer will be activated. On the other hand, if an input vector lies outside a receptive field, the corresponding hidden neurons will not be activated.

In a GRNN design (Figure 2), each training vector has a corresponding neuron in the hidden layer. The weight matrix W_R in the hidden layer, a radial basis layer, is simply a collection of all training vectors presented to the network. When a new input vector is presented to the network, the similarity in terms of distance (*dist*) between this new input vector with each of those stored in W_R is calculated:

$$dist = |X - W_R^j|, \quad \text{for } j = 1, Q \tag{1}$$

where Q = number of neurons in the hidden layer; X = input vector, *dist* = distance between X and W_R^j. The calculated distance is then adjusted by the bias, b:

$$b = 0.8326/s \tag{2}$$

$$n_1 = dist \times b \tag{3}$$