Bayesian Network Characterization of Shear Wave Velocity for Sands

Man Kong Lo[1], Xiao Wei[2], Siau Chen Chian[3] and Taeseo Ku[4]

[1]*Post-doctoral Research Fellow, Dept. of Civil and Environmental Engineering, National Univ. of Singapore. Email: ceelmk@nus.edu.sg*
[2]*Post-doctoral Research Fellow, Dept. of Civil and Environmental Engineering, National Univ. of Singapore. Email: ceeweix@nus.edu.sg*
[3]*Associate Professor, Dept. of Civil and Environmental Engineering, National Univ. of Singapore. Email: sc.chian@nus.edu.sg*
[4]*Assistant Professor, Dept. of Civil and Environmental Engineering, National Univ. of Singapore. Email: ceekt@nus.edu.sg*

Abstract: Shear wave velocity is an important parameter in geotechnical engineering involving dynamic analysis and liquefaction evaluation. However, proper characterization of the shear wave velocity remains an area of difficulty, because shear wave velocity is a complicated function of a variety of parameters. In this regard, this paper attempts to model the association of shear wave velocity of sand with various sand parameters using Bayesian network. Bayesian network is a flexible method to construct a joint probability distribution of sand parameters, however training and prediction of Bayesian network is relatively complicated due to its non-linear nature. To train the Bayesian network, a database of shear wave velocity is compiled. The effects of particle shape (roundness and sphericity) and gradation are included. The prediction of Bayesian network is performed by the Just Another Gibbs Sampler (JAGS) package, with an application to the Natori river sand site, Japan. The case study shows that index properties and site-specific data can be integrated in a coherent manner to reduce the prediction uncertainty.

Keywords: bayesian network, shear wave velocity, sand, gibbs sampler, site investigation.

1. Introduction

Shear wave velocity (V_s) of sand is an important parameter in geotechnical engineering, which is required for dynamic analysis and liquefaction assessment. It is also directly related to small-strain shear modulus (G_{max}), which is an essential parameter in ground deformation analysis. Shear wave velocity is a complicated function of many factors such as confining stress, void ratio, stress history, fines content, mineralogy, particle shape, etc. Among them, confining stress is identified as a significant factor (Hardin and Richart 1963; Wichtmann and Triantafyllidis 2009). In particular, V_s depends on the effective stresses in the direction of polarization and propagation of the shear waves (Roesler 1979). The following average stress model can be adopted to model the stress dependency (Yu and Richart Jr 1984; Yan and Byrne 1990; Ku et al. 2016):

$$V_s = \alpha \left(\frac{\sigma_\perp' + \sigma_\parallel'}{2} \right)^\beta \qquad (1)$$

where α (m·s^{-1}) is the shear wave velocity at 1 kPa, and β reflects the sensitivity of V_s to the average of the effective stress in the direction of polarization ($\sigma_\perp'$) and propagation ($\sigma_\parallel'$) of the shear waves. The unit of stress is kPa.

Previous studies (Cho et al. 2006; Payan et al. 2016) suggest that α and β may depend on the shape of sand particle. Two important parameters that reflect particle shape are sphericity and roundness (Krumbein and Sloss 1963) (Fig. 1). Sphericity (S) indicates the similarity between the particle's length, height and width, which is a ratio between the radius of the largest inscribed sphere in the particle to the radius of the smallest circumscribed sphere to the particle. Roundness (R) indicates the smoothness of the particle surface, which is the average radius of curvature of surface features to the radius of the largest inscribed sphere.
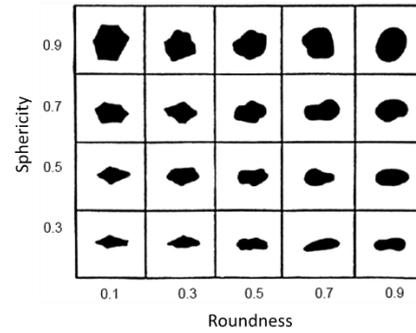

Figure 1. Particle shape identification chart (adopted from Krumbein and Sloss (1963)).

A systematic approach to incorporate the associations between the sand parameters (i.e., V_s, α, β, R, S) is to construct a joint probability distribution. A flexible approach to construct a joint probability distribution is by Bayesian network. Bayesian network can represent a probability distribution through an acyclic graph, where each node is a conditional probability distribution of a sand parameter. The links that converge to a node represent the parameters on which the distribution is conditioned. If the Bayesian network has N nodes, i.e., $\boldsymbol{X} = \{x_1, \dots x_N\}$, then the joint distribution is given by (Bishop 2006):

$$P(\boldsymbol{X}) = \prod_{n=1}^{N} P(x_n | \text{parent}[x_n]) \qquad (2)$$

where $\text{parent}[x_n]$ denotes the parent nodes of x_n. Implementing a Bayesian network involves two important stages: training and prediction. Training refers to learning the structure (i.e., links) of the network using a database. If all the parameters in the network are discrete, then the structure can be learnt directly from the data without any prior assumptions (Scutari and Denis 2014). However, the sand parameters have continuous values, and the training

process is non-trivial in terms of algorithm. This study will attempt to establish the network structure through investigating the empirical relationships proposed in previous literatures. This ensures the network structure matches with the geotechnical understanding.

Prediction of the Bayesian network involves evaluating the probability distribution of the unobserved nodes given the observed nodes, i.e., P (unobserved nodes| observed nodes). In this study, the unobserved nodes correspond to shear wave velocity at untested locations. Prediction can be performed by Markov Chain Monte Carlo (MCMC) algorithm, which can draw samples from an arbitrary probability distribution. This study will apply a MCMC algorithm called Gibbs sampler (Geman and Geman 1984). The prediction process will be illustrated through a case study in Japan.

2. Database for shear wave velocity

A database for shear wave velocity is compiled for establishing the Bayesian network, which is summarized in Table 1. The data are extracted from Cho et al. (2006) and Payan et al. (2016). In Cho et al. (2006), V_s of the sand is measured in an oedometer cell fitted with bender elements. In Payan et al. (2016), a modulus Stoke-type resonant apparatus was used to measure G_{max}. In this study, the reported G_{max} values are converted to V_s through $G_{max} = \rho V_s^2$, while assuming the density of sand (ρ) is $1800 \, \text{kg} \cdot \text{m}^{-3}$. In both studies, sphericity and roundness are determined by observing individual grains under an optical microscope, and then the observed geometry is compared to the identification chart (Fig. 1). The values reported in Table 1 are the average value based on 30 sand grains. For some cases, maximum and minimum void ratios (e_{max} and e_{min}) are measured together with sphericity and roundness. Also in this database, only the records with coefficient of uniformity (C_u) smaller than 3.5 are included.

3. Establishing the Bayesian network

Based on empirical relationships reported in previous literature, the following Bayesian network (Fig. 2) is proposed to model the association between the various sand parameters. $V_{s,1},\ldots,V_{s,L}$ represent the shear wave velocities at different locations, under a sand type with a given α and β value. The links of the network will be examined in detail in the rest of this section.

Inspection of α and β against R and S (Fig. 3) indicates that as the roundness and sphericity of sand particle decreases, α decreases and β increases. Cho et al. (2006) suggested that the packing is looser for irregular sand particles, while the contacts between particles become more deformable. This leads to lower shear wave velocity (α) and higher sensitivity to confining stress (β). The following multiplicative model is fitted through the least square error criteria:

$$\alpha = 92.3 R^{0.392} S^{0.179} \tag{3}$$

$$\beta = 0.205 R^{-0.149} S^{-0.207} \tag{4}$$

Table 1. V_s database (Cho et al. 2006; Payan et al. 2016).

| Sand type	D_{50}	C_u	R	S	e_{max}	e_{min}	α	β
1K9 crushed	0.3	3.4	0.2	0.4	1.16	--	35	0.35
1O2 crushed	0.25	2.9	0.25	0.8	0.83	--	--	--
1O6 crushed	0.21	2.8	0.3	0.7	0.77	--	--	--
2L6 crushed	0.28	3.5	0.25	0.8	0.84	--	--	--
30UB-70UBL	0.59	1.99	0.31	0.57	--	--	50.5	0.275
3C7 crushed	0.26	3.2	0.25	0.8	0.85	--	--	--
3P3 crushed	0.27	2.2	0.2	0.7	0.95	--	41	0.28
50UB-50UBL	0.54	1.96	0.36	0.61	--	--	55.1	0.26
5U1 crushed	0.32	3.5	0.15	0.7	0.84	--	42.6	0.266
6F5 crushed	0.25	3.3	0.25	0.8	0.91	--	--	--
70UB-30UBL	0.49	2.01	0.41	0.65	--	--	59.8	0.25
7U7 crushed	0.3	3.2	0.2	0.8	0.79	--	--	--
8M8 crushed	0.38	3.3	0.2	0.7	0.97	--	55.7	0.262
9C1 crushed	0.52	2.3	0.25	0.7	0.91	--	54	0.297
9F1 crushed	0.33	3.5	0.2	0.8	0.9	--	41.8	0.31
ASTM 20/30	0.6	1.4	0.8	0.9	0.69	--	72.7	0.223
ASTM graded	0.35	1.7	0.8	0.9	0.82	0.5	--	--
Blasting	0.71	1.9	0.3	0.55	1.025	0.698	--	--
Blue sand 1	1.66	1.41	0.24	0.51	--	--	54.5	0.265
Blue sand 2	1.94	2.8	0.24	0.51	--	--	45.4	0.29
Daytona Beach	0.23	1.4	0.62	0.7	1	0.64	--	--
Fraser River	0.3	1.9	0.25	0.5	1.13	0.78	--	--
Glass beads	0.32	1.4	1	1	0.72	0.542	--	--
Margaret river	0.49	1.9	0.7	0.7	0.87	--	93.2	0.219
Michigan dune	0.33	1.5	0.77	0.87	0.8	0.56	--	--
Nevada	0.15	1.8	0.6	0.85	0.85	0.57	56.3	0.242
Newcastle	0.33	1.94	0.64	0.73	--	--	72.9	0.24
Ottawa # 90	0.27	2.2	0.4	0.6	1.1	0.73	--	--
Ottawa #20/30	0.72	1.2	0.9	0.9	0.742	0.502	--	--
Ottawa #20/70	0.53	2.4	0.76	0.81	0.78	0.47	--	--
Ottawa #45	0.57	2.1	0.45	0.68	1.11	0.75	--	--
Ottawa #60/80	0.21	2.4	0.65	0.78	0.85	0.55	--	--
Ottawa F-110	0.12	1.7	0.7	0.7	0.848	0.535	--	--
Sandboil	0.36	2.4	0.55	0.7	0.79	0.51	--	--
Syncrude Tailings	0.18	2.5	0.47	0.62	1.14	0.59	--	--
Ticino	0.58	1.5	0.4	0.8	0.99	0.574	70.7	0.231
Uniform Sydney	0.36	1.18	0.61	0.76	--	--	90.9	0.215
White	0.24	1.69	0.71	0.76	--	--	69.9	0.24
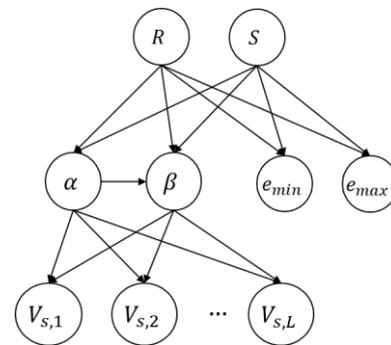
Figure 2. Proposed Bayesian network for sand parameters.
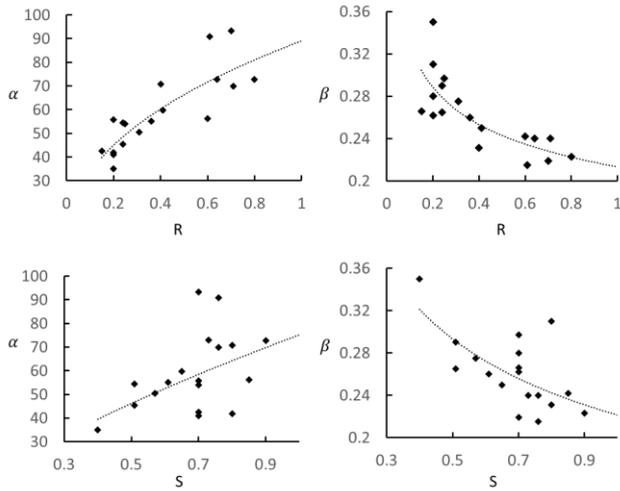
Figure 3. α and β against roundness and sphericity.

18 data points are used, and the R square values for α and β are 0.68 and 0.72 respectively. In Cho et al. (2006), a regularity parameter (i.e., $(R+S)/2$) is adopted to fit α and β, which assumes α and β are equally sensitive to R and S. However, the multiplicative model (Eq. 3 and Eq. 4) suggests that α is more sensitive to roundness, while β has similar sensitivity to both roundness and sphericity. The residuals of α and β (i.e., original value minus predicted value) are shown in Fig. 4. The residuals are assumed to be bivariate normal distributed, with standard deviation $\sigma_\alpha = 9.25 \mathrm{m \cdot s^{-1}}$ and $\sigma_\beta = 0.0191$. The residuals are correlated with correlation coefficient $\rho_{\alpha\beta} = -0.613$. This residual correlation is reflected in the Bayesian network (Fig. 2) through the link joining node α and node β.
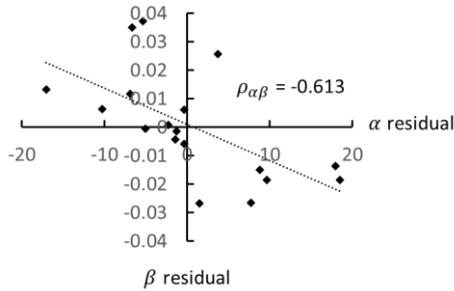

Figure 4. α and β residuals against roundness and sphericity.

Next, the relationship between e_{max}, e_{min} with R and S is investigated (Fig. 5). Such relationship is useful because normally e_{max} and e_{min} are measured, rather than roundness and sphericity. Fig. 5 shows that the smaller R and S will increase e_{max} and e_{min}, indicating that irregular particle shape leads to looser packing. The following linear model is fitted:

$$e_{max} = 1.515 - 0.012R - 0.822S \qquad (5)$$
$$e_{min} = 0.885 - 0.224R - 0.209S \qquad (6)$$

with the standard deviation of the residuals are $\sigma_{emax} = 0.078$ and $\sigma_{emin} = 0.061$. Again, the residuals are assumed

normally distributed. Eq. 5 indicates that e_{max} depends differently on R and S, suggesting that e_{max} is primarily controlled by sphericity rather than roundness. Statistically, R is not a significant parameter to e_{max} based on its p-value. Therefore, it suggests that regularity adopted by Cho et al. (2006) is not optimal to characterize the relationship between particle shape and e_{max}. But, e_{min} depends nearly the same on R and S, according to Eq. 6.

Finally, the standard deviation of V_s residual is assigned to be $5 \mathrm{m \cdot s^{-1}}$, and the distributions of R and S nodes are assumed uniform with ranges (0.1, 1) and (0.3, 1) respectively.
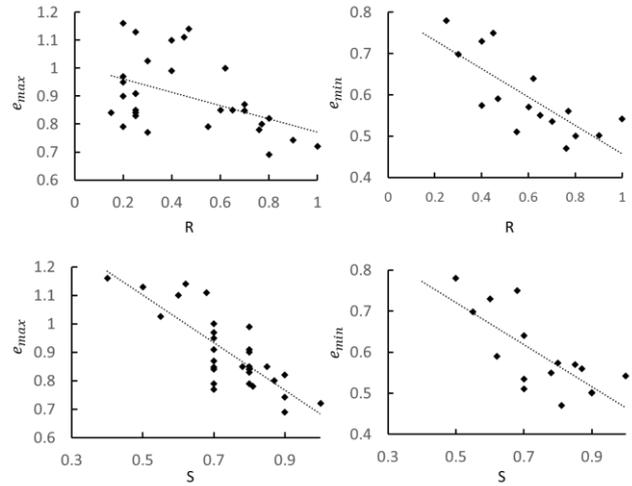

Figure 5. e_{max} and e_{min} against roundness and sphericity.

4. Prediction using the Bayesian network

The trained Bayesian network can be used to predict the shear wave velocity at unobserved locations. If some nodes are observed (such as index properties, in-situ testing), they can be fixed. Through the association between the nodes (i.e., links), the nodes that correspond to shear wave velocity at unobserved locations can be updated to their posterior distribution. The updating is implemented by Gibbs sampler (Geman and Geman 1984). Gibbs sampler aims to generate samples from the posterior distribution, based on the conditional probability distribution of each node. Bayesian network utilizes the Markov blanket property to obtain the conditional probability distribution. Markov blanket of a node consists of its parents, its children, and all other nodes sharing its children:

$$P(x_n | X \text{ except } x_n) \propto P(x_n | \mathrm{parent}[x_n]) \prod_{Y \in \mathrm{children}[x_n]} P(Y | \mathrm{parent}[Y]) \qquad (7)$$

This study uses an open source package called Just Another Gibbs Sampler (JAGS), which is developed by Martyn Plummer (Plummer 2015). Through the JAGS interface, Bayesian network can be defined conveniently, and then Gibbs sampling can be executed. During Gibbs sampling, only one out of k Gibbs samples is retained,

where k is the thinning interval. This is to eliminate the autocorrelation in the Gibbs samples.

5. Application on Natori river sand site

The prediction capability of the Bayesian network is illustrated via a case study in Japan, the Natori river sand site (Suzuki et al. 2003; Mimura 2003; Mayne 2006). The data of Natori river sand is not included when training the Bayesian network. Fig. 6 shows a boring log of the site. The fine and coarse sand layers are overlain by silt and clay layers. The sands are of Holocene origin, and mostly compose of quartz, feldspar and mica. A soil sample was extracted using ground freezing method at a depth of 8.4m (Mimura 2003), and Table 3 shows its testing data. G_{max} of the sample is measured by a resonant column test, while the peak friction angle (ϕ') is measured by a drained triaxial compression test. The sand practically does not contain fines, as the fines content (FC) is very low. The G_{max} profile is also obtained using the PS logging technique (Suzuki et al. 2003). Assuming the soil is fully saturated under the groundwater table, G_{max} data is converted to V_s using a density of 1888 kg·m^{-3}, obtained based on the void ratio (e_0) and specific gravity (G_s) of the sample. The V_s profile is used to validate the Bayesian network prediction.
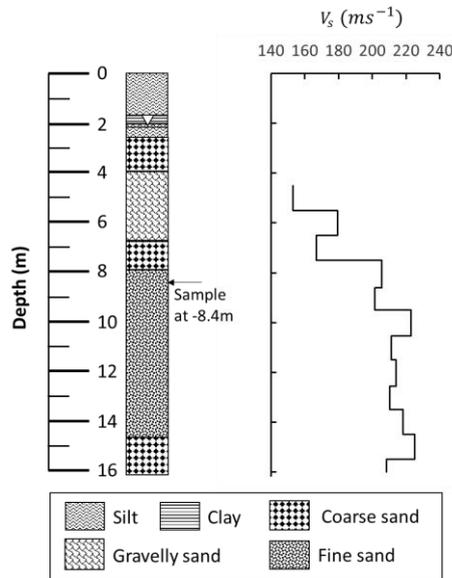

Figure 6. Boring log and shear wave velocity profile of Natori river sand site.

Table 3. Information of the soil sample.

Depth	D_{50}	C_u	e_0	e_{max}
8.4m	0.22	2	0.857	1.167
e_{min}	G_{max}	ϕ'	G_s	FC
0.765	78MPa	40.9º	2.65	0.23%

The prediction of V_s will be from the depth of 5 m to 16 m, with 0.5 m interval. Therefore the total number of V_s nodes is 23. Two prediction cases will be considered. For prediction case 1, only e_{max} and e_{min} are known and their nodes are fixed. For case 2, in addition to e_{max}

and e_{min}, the V_s node that corresponds to depth 8.5 m is fixed to 203.2 m·s^{-1}, which is converted from the G_{max} value of the sample. To assign the lateral effective stress (i.e. $\sigma'_\perp$ in Eq. 1 assuming shear wave propagates downward), the lateral stress coefficient at rest (K_0) is assigned as 0.345, based on Jaky's formula: $K_0 = 1 - \sin\phi'$, and assume the soil is normally consolidated, as the stress history of the site is not clear. Uncertainty of K_0 poses a limitation to this analysis, and direct measurement of K_0 from self-boring pressuremeter, total stress cell or hydraulic fracture should be used if it is available. For Gibb's sampling in both prediction cases, the Markov chain length is 100000, with a thinning interval of 10.

Based on the samples generated by Gibb's sampler, the mean and standard deviation (SD) of the sand parameters are evaluated, which are summarized in Table 4. The SD of V_s is the average SD of all the V_s nodes. Meanwhile, the prediction interval of V_s is displayed in Fig 7, which consists of mean $\pm$ one standard deviation. The actual V_s profile is also plotted for validation. For the depths where the V_s value jumps (such as depth 5.5m in Fig. 6), the average of two V_s values are taken.

For prediction case 1, Bayesian network indicates that the sand grain is quite irregular with small roundness and sphericity, since both e_{max} and e_{min} are large. The prediction interval envelopes the actual V_s profile, but the mean prediction has a systematic bias, and underestimate the true value. Also the width of the prediction interval is relatively large (34.2 m·s^{-1}). This is due to the large uncertainty of α, which is solely derived based on e_{max} and e_{min}.

For prediction case 2, when a piece of site-specific data is included, the uncertainty of α decreases by half, and the uncertainty of V_s decreases significantly from 34.2 m·s^{-1} to 7.36 m·s^{-1}. The prediction bias is also corrected. The prediction interval broadly envelopes the true profile, except at the shallow region (depth 5-7 m) that corresponds to gravelly sand layer. Gravelly sand may exhibit different V_s behavior than fine sand (Chen et al. 2019), and the database in this study does not contain gravelly sand. The V_s prediction at that region is expected to improve if the Bayesian network further incorporates the behavior of gravelly sand.

Table 4. Statistics of the Gibbs samples.

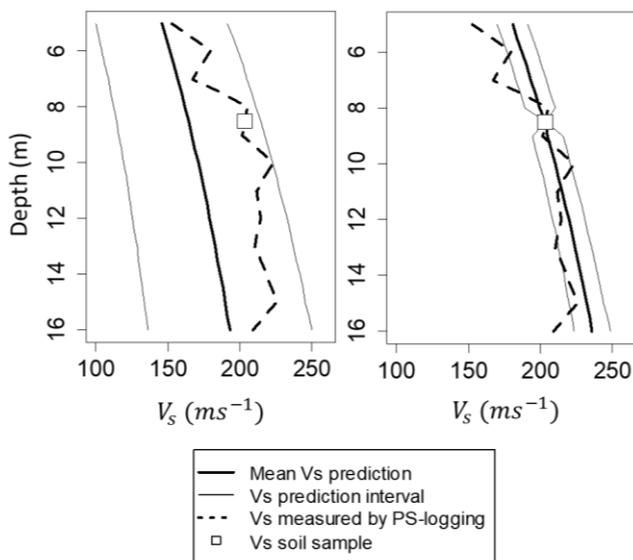|	Case 1	Case 2
Mean of α	49.8	64.7
SD of α	13.5	7.18
Mean of β	0.301	0.284
SD of β	0.031	0.027
Mean of R	0.33	0.46
SD of R	0.17	0.17
Mean of S	0.42	0.40
SD of S	0.07	0.07
SD of V_s (m·s^{-1})	34.2	7.36

Figure 7. Shear wave velocity prediction for case 1 (left) and case 2 (right).

6. Conclusion

This study proposes a Bayesian network approach to model the shear wave velocity of sand in a probabilistic manner. To establish the Bayesian network, a database of shear wave velocity is compiled. In the Bayesian network, the V_s parameters (i.e., α, β) are associated with roundness and sphericity of soil particles, which in turn are associated with maximum and minimum void ratio. Prediction of the Bayesian network is performed by Gibbs sampler, which is implemented in the Just Another Gibbs Sampler package. In the Natori river sand case study, the index parameters (e_{max} and e_{min}) and site-specific data (e.g., soil sample) can be integrated coherently in the Bayesian network. The case study also shows that including a single site-specific data can correct the prediction bias, and significantly reduces the prediction uncertainty of shear wave velocity.

References

Bishop, C.M. 2006. *Pattern Recognition and Machine Learning*. Springer.

Chen, G., Zhou, Z, Sun, T., Wu, Q., Xu, L., Khoshnevisan, S. and Ling, D. 2019. Shear Modulus and Damping Ratio of Sand–Gravel Mixtures Over a Wide Strain Range. *Journal of Earthquake Engineering*, 23(8): 1407–40.

Cho, G.C., Dodds, J. and Santamarina, J.C. 2006. Particle Shape Effects on Packing Density, Stiffness, and Strength: Natural and Crushed Sands. *Journal of Geotechnical and Geoenvironmental Engineering*, 132(5): 591–602.

Geman, S. and Geman, D. 1984. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6: 721–41.

Hardin, B.O. and Richart Jr, F.E. 1963. Elastic Wave Velocities in Granular Soils. *Journal of Soil Mechanics and Foundations Division*, 89(1): 33–65.

Ku, T., Subramanian, S., Moon, S.-W. and Jung, J. 2016. Stress Dependency of Shear-Wave Velocity Measurements in Soils. *Journal of Geotechnical and Geoenvironmental Engineering*, 143(2): 4016092.

Krumbein, W.C. and Sloss, L.L. 1963. *Stratigraphy and Sedimentation*. Freeman and Company, San Francisco.

Mayne, P W. 2006. The Second James K. Mitchell Lecture Undisturbed Sand Strength from Seismic Cone Tests. *Geomechanics and Geoengineering: An International Journal*, 1(4): 239–57.

Mimura, M. 2003. Characteristics of Some Japanese Natural Sands – Data from Undisturbed Frozen Samples. *Characterization and Engineering Properties of Natural Soils*, 2: 1149–68.

Payan, M., Khoshghalb, A., Senetakis, K. and Khalili, N. 2016. Effect of Particle Shape and Validity of Gmax Models for Sand: A Critical Review and a New Expression. *Computers and Geotechnics*, 72: 28–41.

Plummer, M. 2015. *JAGS Version 4.0.0 User Manual*. http://ftp.tw.freebsd.org/distfiles/mcmc-jags/jags_user_manual.pdf.

Roesler, S.K. 1979. Anisotropic Shear Modulus Due to Stress Anisotropy. *Journal of the Geotechnical Engineering Division*, 105(7): 871–80.

Scutari, M. and Denis, J.B. 2014. *Bayesian Networks: With Examples in R*. CRC press.

Wichtmann, T. and Triantafyllidis, T. 2009. Influence of the Grain-Size Distribution Curve of Quartz Sand on the Small Strain Shear Modulus Gmax. *Journal of Geotechnical and Geoenvironmental Engineering*, 135(10): 1404–18.

Yan, L. and Byrne, P. M. 1990. Simulation of downhole and crosshole seismic tests on sand using the hydraulic gradient similitude method. *Canadian Geotechnical Journal*, 27(4): 441–460.

Yu, P. and Richart Jr, F. E. 1984. Stress ratio effects on shear modulus of dry sands. *Journal of Geotechnical Engineering, American Society of Civil Engineers*, 110(3): 331–345.