Managing risk in geotechnical engineering – from data to digitalization

Kok-Kwang Phoon[1], Jianye Ching[2], and Yu Wang[3]

[1]Department of Civil and Environmental Engineering, National University of Singapore, Singapore.
E-mail: kkphoon@nus.edu.sg
[2]Department of Civil Engineering, National Taiwan University, Taipei, Taiwan.
E-mail: jyching@gmail.com
[3]Department of Architecture and Civil Engineering, City University of Hong Kong, Hong Kong.
yuwang@cityu.edu.hk

Abstract: If you scan a page from a soil report, this is called digitization. If you deploy digital technologies, both software such as building information modeling and machine learning and hardware such as autonomous drones and additive manufacturing, to support new and more collaborative forms of project delivery, this is called digitalization. Data lies at the heart of this transformation that is targeted at re-valuing infrastructure from a "brick and mortar" asset to a service for the interests of the end-users. There is a need to view the value of data completely differently from how they are routinely used in current practice. In particular, there is a need to treat data as assets in themselves, over and above their conventional roles as inputs to a physical model or as monitoring data to trigger interventions. This paper explores the availability and nature of geotechnical data and presents two recent advances made in this direction for a specific but important task of estimating soil/rock properties (compressive sampling and Bayesian machine learning). Data-driven decision making does not imply taking the engineer out of the entire life cycle management chain. It is intended to support rather than to replace human judgment.

Keywords: Risk; geotechnical data; digitalization; compressive sampling; Bayesian machine learning.

1 Introduction

The principal finding in the report of the National Research Council (1995) concerning the role of probability in geotechnical engineering is that "probabilistic methods, while not a substitute for traditional deterministic design methods, do offer a systematic and quantitative way of accounting for uncertainties encountered by geotechnical engineers, and they are most effective when used to organize and quantify these uncertainties for engineering designs and decisions". There is no debate that the geotechnical engineer has to grapple with many sources of uncertainties, including natural geologic variabilities. There is however a long drawn discussion on whether these uncertainties and the associated risks arising from consequences of a decision made in the face of uncertainties can be treated more formally. For example, the partial factors of safety in Eurocode 7 (EN 1997−1:2004) are not explicitly calibrated according to reliability principles described in Annex C of the head Eurocode (EN 1990:2002) or Annex D of ISO2394: 2015 at this point in time. Notwithstanding the unique features of geotechnical practice, Phoon (2017) opined that reliability methods can handle complex real world information (cross and/or spatially correlated multivariate data) and information imperfections (sparse, uncertain and/or incomplete information) more effectively than relying on empiricism and judgment alone. In particular, it is sensitive to data while the conventional factor of safety or partial factors of safety are not. He further clarified that "reliability analysis is not a panacea for all uncertainties affecting design calculations based on the factor of safety or geotechnical practice in general. Reliability analysis is merely one of the many mathematical methods routinely applied to model the complex real-world for engineering applications. It is susceptible to abuse in the absence of sound judgment in the same manner as a finite element analysis. The importance of engineering judgment clearly has not diminished with the growth of theory and computational tools. However, its role has become more focused on those design aspects that remain outside the scope of theoretical analyses." In short, data-driven decision making supports rather than replaces human judgment.

In the meantime, the Institution of Civil Engineers (ICE) strongly urged the civil engineering industry to engage in digital transformation with greater urgency. The ICE State of the Nation Report in 2017 looks at how advances in digital technologies and data are transforming how we design, deliver and operate infrastructure and recommends the following:

1. Need to view the value of data differently - "bodies of data on built assets are becoming increasingly important, and need to be managed as significant assets in themselves …",
2. Need to consider infrastructure as a service - "putting the end-user first should prompt us to embrace the full value of new technologies and data estates ...",
3. Need to keep pace with rapid advances. The report cited 64% of firms operating in Europe & the Middle East are rated as either 'industry following' or 'behind the curve' in terms of technology adoption, and

Proceedings of the 7th International Symposium on Geotechnical Safety and Risk (ISGSR)
Editors: Jianye Ching, Dian-Qing Li and Jie Zhang

Published by Research Publishing, Singapore.
ISBN: 978-981-11-2725-0; doi:10.3850/978-981-11-2725-0_SL-cd

4. Need for infrastructure and construction industries to work with other industries – "Need to collaborate and coordinate … with the technology and manufacturing industries if we are to keep pace with these advances, and seize the moment."

Gerbert et al. (2016) pointed out that the construction sector in general is "ripe for change: labor productivity in construction has been stagnating for decades, and companies have been slow to adapt and innovate". At present, it is accurate to say geotechnical risks are largely managed by the factor of safety at the design stage and the observational approach (Peck 1969) at the construction stage. In fact, the design phase and the construction phase in geotechnical engineering may not be as distinct as those in structural engineering. For example, it is not uncommon to adjust rock bolt spacing as tunneling progresses, but it is unheard of to adjust column spacing as each story is erected in a building. This is not a difference in tradition, but a fundamental difference in risk management to address qualitatively different design conditions. Although the geotechnical engineering profession has been very successful in making safe decisions based on a hybrid strategy combining site and observational data, modeling, testing, precedents, experience, and judgment, this strategy is now fundamentally out of alignment with broad sweeping trends disrupting all industries due to the advent of digital technologies. For example, the factor of safety remains effectively the same since it was presented as early as 1948 in the classic text "Soil Mechanics in Engineering Practice" (Terzaghi and Peck 1948). There is no rational mechanism to adjust the factor of safety (or partial factors of safety) based on the amount of data collected at a given site. The Eurocode 7 (EN 1997−1:2004) adopts a notion of a characteristic value that can be adjusted (mostly empirically) based on site data. The design value is a function of the characteristic value and a fixed partial factor. In this sense, the design value depends on site data. Nonetheless, the design value is an input and its effect on performance (an output) cannot be assessed by judgment alone. One key advantage of the reliability index is that it is sensitive to data at the performance level. A design parameter that is estimated with more precision would result in a higher reliability index. Alternately, one can say that a more economical design can meet the desired target reliability index when more precise information is available or vice-versa (Ching et al. 2014a).

Risk-informed decision making needs data (Gransberg et al. 2018). It may be possible to do this informally using extremely limited data complemented by experience and judgment. Geotechnical practice is one such example, although we are none the wiser if our decisions are extremely safe or optimal for a particular site. Leaving aside the debate between deterministic and reliability approaches (Phoon 2017), the emerging limitation for the former in the face of digital transformation is that it does not quite know what to do with more data, beyond reducing it to a single number (average, cautious estimate, worst credible estimate, etc.), transforming it as an input to a physical model, or treating it as a simple trigger to activate interventions. There may be other applications, but arguably none is truly transformative when compared to developments in other industries and none is founded on capturing the best value from data as a core objective. Alternate approaches will need more data and will need clever and efficient algorithms to extract the most value out of data for decision making. The authors view the reliability approach as a good start, but it is unlikely to be the best when data scarcity is no longer a feature of geotechnical engineering.

The objectives of this paper are to: (1) clarify if geotechnical engineering is data rich or data poor, (2) examine the general characteristics of its data, and (3) present possibilities on how data can support decision making in its own right. The observations made in this paper are entirely preliminary and restricted to one design decision pertaining to the estimation of soil/rock properties. They are intended to stimulate discussions so that we can keep pace with advances elsewhere.

2 Data Rich or Data Poor?

One common criticism of the reliability approach is that geotechnical information is too scarce for the approach to be meaningfully deployed in practice. For example, Schuppener and Heibaum (2011) remarked that "soil excavations and tests of the mechanical properties of soil never provide enough data to enable a probability calculation to be performed". Macciotta et al. (2019) argued that there was not enough information for quantitative risk assessment to guide decision-making for adoption of rock fall protection strategies. Chilès and Delfiner (1999) noted that volume fractions for cores, cuttings, and logging at a Brent Field site in the North Sea are 1×10^{-9}, 7×10^{-9}, and 1×10^{-6}, respectively. This "curse of small sample size", a phrase coined by Phoon (2017), is certainly more conspicuous in geotechnical engineering. Nonetheless, there are two aspects that are generally not highlighted in this debate. First, the effect of sample size can be formally modeled as a statistical uncertainty. The National Research Council (1995) clarified this common misconception: "the lack of a large data set does not preclude the use of probability theory. Probability theory can be used to evaluate the uncertainties involved in working with meager information". Prästings et al. (2018) also emphasized this advantage: "From a Bayesian point of view, one would rather have highly uncertain – but probabilistic – estimates of the geotechnical properties than no estimates at all". Second, generic databases can be large, even when the constituent site-specific databases are small.

Phoon et al. (2016) and Ching et al. (2016a) provided useful overviews of generic univariate and multivariate databases on soil/rock properties, respectively. Table 1 shows a summary of these databases, labeled as (geo-material type)/(number of parameters of interest)/(number of data points). For example, the CLAY/10/7490 database consists of 7490 data points for ten clay parameters from 251 studies carried out in 30 countries. The clay parameters cover a wide range of overconsolidation ratio (OCR) (but mostly 1~10), a wide range of sensitivity (S_t) (sites with S_t = 1~ tens or hundreds are fairly typical), and a wide range of plasticity index (PI) (but mostly 8 ~ 100). Most data points are classified as clays (some are sensitive or organic clays) on the Robertson's soil classification chart. Some data points are classified as clayey silts or silt mixtures, and few are classified as sand mixtures or sands. This line of research has inspired comparable databases to be assembled in the literature recently (Müller et al. 2014; Liu et al. 2016). The availability of SPM2 (Soil Properties Manual version 2) as a freeware will hopefully encourage more data sharing and further enrichment of these databases to cover more parameters and/or more site conditions - http://140.112.10.150/fmanalysis.html?view=spm2 (Phoon and Ching 2017). The ISSMGE TC304 launched a database sharing initiative (304dB) recently to hasten the pace of machine learning research (http://140.112.12.21/issmge/tc304.htm?=6).

Another source of information frequently collected comes from pile load tests. The performance databases for other geotechnical structures (in addition to piles) are available, but less commonly reported in the literature. A comprehensive survey of these databases was carried out by Phoon & Tang (2019). Table 2 includes further updates. The following geotechnical structures are covered: (1) shallow and deep foundations, (2) offshore spudcans, (3) mechanically stabilized earth and soil nail walls, (4) pipes and anchors (plate, helical, and shoring), (5) slopes and base heave, (6) cantilever walls, and (7) braced excavations. Details are given elsewhere (Phoon and Tang 2019). Another ongoing database sharing project called DINGO (Databases to Interrogate Geotechnical Observations) was reported by Hancock (2018).

For soil/rock properties, the most basic design decision in geotechnical practice is to estimate their values from other test results, typically field test results. Empirical transformation (or regression) models such as those shown in Figure 1 are widely used for this purpose. They are based on generic databases covering multiple sites such as those presented in Table 1, because there are insufficient data in one site to establish a purely site-specific or local model. Transformation uncertainty (scatter about the regression line) is an intrinsic characteristic of these empirical models. A second characteristic that is well understood but does not feature in the actual estimation of soil/rock properties is site effect as shown in Figure 2. If one were to accept these observations, namely: (1) a generic database is large, (2) a site-specific database is small, and (3) there are site differences, one could readily imagine research questions where data-driven algorithms can add value to routine decision making:

1. How to characterize "site differences" based on sparse data from a routine project?
2. How to adapt a generic database so that it is more relevant to a specific site?

The above questions would apply to other design decisions. For example, Section 7.4.1 "Design methods" in Eurocode 7 (EN 1997−1:2004) recommends different design approaches for pile foundations:

1. The results of static load tests, which have been demonstrated, by means of calculations or otherwise, to be consistent with other relevant experience;
2. Empirical or analytical calculation methods whose validity has been demonstrated by static load tests in comparable situations;
3. The results of dynamic load tests whose validity has been demonstrated by static load tests in comparable situations;
4. The observed performance of a comparable pile foundation, provided that this approach is supported by the results of site investigation and ground testing.

More site-specific estimates of soil/rock properties and the associated uncertainties in these estimates would clearly contribute to the second approach. For other design approaches, relevant databases could be compiled and exploited in the same way. In fact, the distinction between different approaches diminishes when different databases could be combined to support decision making even more holistically. Phoon and Tang (2019) opined that there is "potential to apply new deep learning methods to identify 'similar' load test data from a generic database to supplement limited site-specific load test data. By doing so, 'site-specific' model factors can be derived. This effort will contribute to a broader agenda to digitalize foundation design for 'precision construction', where 'site-specific' model factors and soil parameters can possibly customize design to a particular site and even a particular location in a site". This direction of inquiry is closer in spirit to digitalization and clearly transforms existing practice more fundamentally than reliability-based design. It is evident that an outcome such as more site-specific property estimates will be useful for any design approach, deterministic or otherwise and will impact a design more directly than probability distributions.

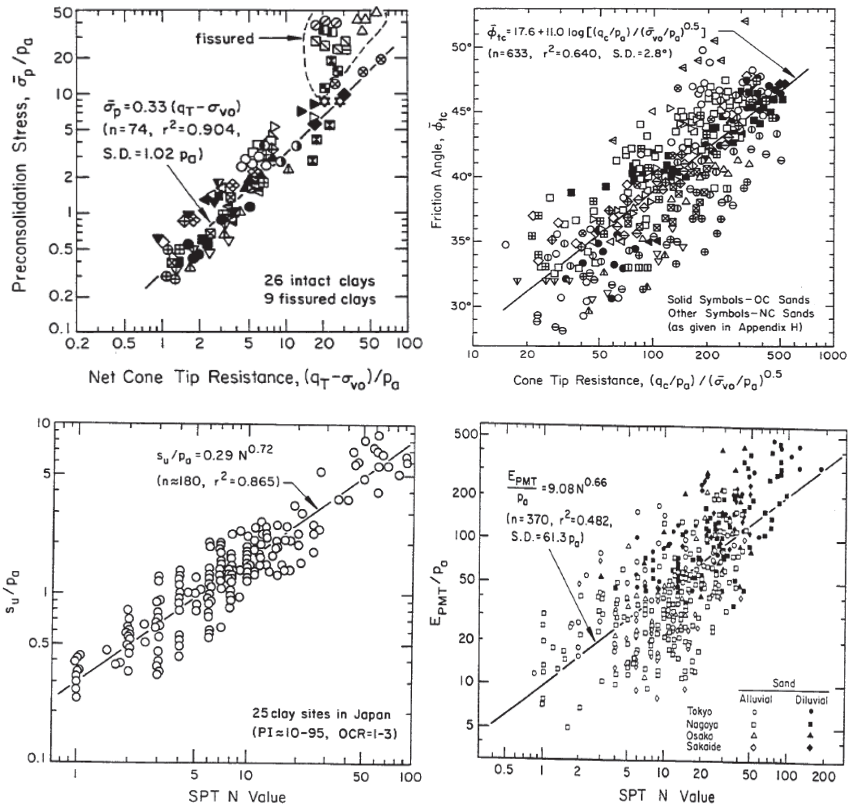

Figure 1. Examples of transformation models in EPRI EL-6800 (Kulhawy and Mayne 1990).

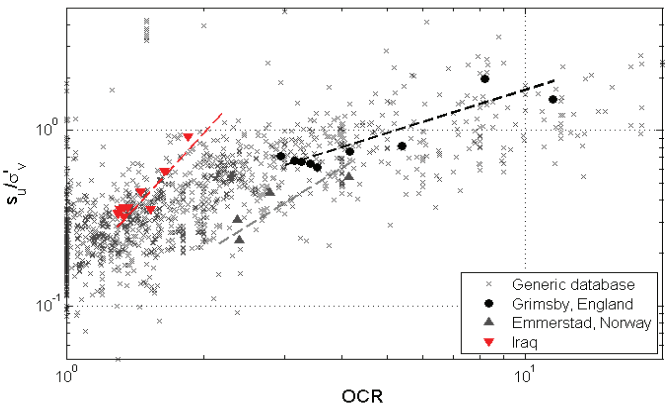

Figure 2. Example of site-specific effects in the correlation between normalized undrained shear strength (s_u/σ'_v) and overconsolidation ratio (OCR) (Ching and Phoon 2019a).

Table 1. Summary of some soil/rock databases (Phoon and Ching 2017).

Database	Reference	Parameters of interest	# Data points	# Sites/studies	Range of parameters		
					OCR	PI	S_t
CLAY/5/345	Ching and Phoon (2012)	LI, s_u, s_u^{re}, σ'_p, σ'_v	345	37 sites	1~4	—	Sensitive to quick clays
CLAY/6/535	Ching et al. (2014b)	s_u/σ'_v, OCR, $(q_t–\sigma_v)/\sigma'_v$, $(q_t–u_2)/\sigma'_v$, $(u_2–u_0)/\sigma'_v$, B_q	535	40 sites	1~6	Low to very high plasticity	Insensitive to quick clays
CLAY/7/6310	Ching and Phoon (2013, 2015a)	s_u from 7 different test procedures	6310	164 studies	1~10	Low to very high plasticity	Insensitive to quick clays
CLAY/10/7490	Ching and Phoon (2014)	LL, PI, LI, σ'_v/P_a, S_t, B_q, σ'_p/P_a, s_u/σ'_v, $(q_t–\sigma_v)/\sigma'_v$, $(q_t–u_2)/\sigma'_v$	7490	251 studies	1~10	Low to very high plasticity	Insensitive to quick clays
F-CLAY/7/216	D'Ignazio et al. (2016)	s_u^{FV}, σ'_v, σ'_p, w_n, LL, PL, S_t	216	24 sites	1~7.5	Low to very high plasticity	Insensitive to quick clays
FG/KSAT-1358	Feng and Vardanega (2019a, b)	e, k_{sat}, LL, PI	1358	33 studies	Fat clay, lean clay, elastic silts, and silts. e = 0.19 – 8.57; k_{sat} = 1.44×10-13 – 7.5×10-6; LL = 22 – 675; PI = 5 – 625.9		
J-Clay/5/124	Liu et al. (2016)	M_r, q_c, f_s, w_n, γ_d	124	16	Soft to stiff clayey soils and silty clay soils with high variability of the strength and stiffness characteristics M_r = 12.54~95.82 MPa, q_c = 0.22~3.93 MPa, f_s = 0.03~0.14 MPa, w_n (%) = 6.91~78.11, γ_d=10.47~19.92 kN/m^3		
SAND/7/2794	Ching et al. (2017a)	D_{50}, C_u, D_r, σ'_v/P_a, ϕ', q_{t1}, $(N_1)_{60}$	2794	176 studies	1~15	D_{50} = 0.1~40 mm, C_u = 1~1000+ D_r = -0.1~117%	
ROCK/9/4069	Ching et al. (2018)	n, γ, R_L, S_h, σ_{bt}, I_{s50}, V_p, σ_c, E	4069	184 studies	γ = 15~35 kN/m^3, n = 0.01~55% σ_c = 0.7~380 MPa, E = 0.03~120 GPa		

Note: LL = liquid limit; PL = plastic limit; PI = plasticity index; LI = liquidity index; w_n = natural water content; M_r = resilient modulus; q_c = cone tip resistance; f_s = sleeve friction; γ_d = dry density; e = void ratio; k_{sat} = saturated hydraulic conductivity; D_{50} = median grain size; C_u = coefficient of uniformity; D_r = relative density; σ'_v = vertical effective stress; σ'_p = preconsolidation stress; s_u = undrained shear strength; s_u^{FV} = undrained shear strength from field vane; s_u^{re} = remoulded s_u; ϕ' = effective friction angle; S_t = sensitivity; OCR = overconsolidation ratio, $(q_t-\sigma_v)/\sigma'_v$ = normalized cone tip resistance; $(q_t-u_2)/\sigma'_v$ = effective cone tip resistance; u_0 = hydrostatic pore pressure; $(u_2-u_0)/\sigma'_v$ = normalized excess pore pressure; B_q = pore pressure ratio = $(u_2-u_0)/(q_t-\sigma_v)$; P_a = atmospheric pressure = 101.3 kPa; q_{t1} = $(q_t/P_a)\times C_N$ (C_N is the correction factor for overburden stress); $(N_1)_{60}$ = $N_{60}\times C_N$ (N_{60} is the N value corrected for the energy ratio); n = porosity; γ = unit weight; R = Schmidt hammer hardness (R_L = L-type Schmidt hammer hardness); S_h = Shore scleroscope hardness; σ_{bt} = Brazilian tensile strength; I_s = point load strength index (I_{s50} = I_s for diameter 50 mm); V_p = P-wave velocity; σ_c = uniaxial compressive strength; E = Young's modulus.

Table 2. Summary of performance databases for some geotechnical structures (updated from Table 1, Phoon and Tang 2019).

Geotechnical structure	Database/reference	Data source	Test type	Geomaterial	N
Shallow foundations	UML-GTR ShalFound07 (Paikowsky et al. 2010)	Global	Laboratory/field	Cohesionless	549
	UML-GTR RockFound07 (Paikowsky et al. 2010)	Global	Field	Rock	122
	Akbas (2007)	Global	Field	Cohesionless	400
	Mayne and Dasenbrock (2018)	Global	Field	Sand	130
	Patra et al. (2012a, b)	—	Laboratory	Sand	192
	Okamura et al. (1997)	Japan	Centrifuge	Sand over clay	31
	Tang and Phoon (2017)	—	Centrifuge	Dense sand	53
	Samtani and Allen (2018)	USA/Europe	Field	Cohesionless	71
Offshore spudcans	Teh (2007)	NUS	Centrifuge	Sand over clay	14
	Hossain (2014)	UWA	Centrifuge	Clay with sand	14
	Hossain and Randolph (2010)	UWA	Centrifuge	Layered clay	42
	Lee (2009)	UWA	Centrifuge	Sand over clay	35
	Hu (2015)	UWA	Centrifuge	Sand over clay	32
	Ullah (2016)	UWA	Centrifuge	Clay-sand-clay	27
	Tang and Phoon (2019a)	NUS and UWA	Centrifuge	Clay with sand	128
Drilled shafts (vertical load)	Ng et al. (2001)	Hong Kong	Field	Rock/saprolite	38
	AbdelSalam et al. (2015)	Egypt	Field	Various	318
	Asem et al. (2018)	Global	Field	Soft rock	190
	DSHAFT (Garder et al. 2012)	Iowa, USA	Field	Various	38
	Motamed et al. (2016)	Las Vegas Valley	Field	Caliche	41
	Stark et al. (2017)	Illinois, USA	Field	Weak rock	155
	TxDOT (Moghaddam et al. 2018)	Texas	Field	Various	27
	Tang et al. (2019)	Global	Field	Various	320
Drilled shafts (lateral load)	EPRI (Chen and Kulhawy 1994)	Global	Field	Clay/sand	88
	Chen and Lee (2010)	Global	Field	Clay/sand	99
	Chen et al. (2011)	Global	Field	Clay/sand	40
	Marcos and Chen (2013)	Global	Field	Gravel	24
Augered cast-in-place piles	Reddy and Stuedlein (2017)	USA	Field	Cohesionless	112
	McVay et al. (2016)	Florida, USA	Field	Various	78
Driven piles	AAU-NGI (Augustesen 2006)	Global	Field	Various	420
	Zhang et al. (2006)	Hong Kong	Field (static/dynamic)	Weathered granite	1514
	Long et al. (2009)	Wisconsin, USA	Field (dynamic)	Various	316
	PILOT (Roling et al. 2011)	Iowa, USA	Field	Various	275
	PSU (Smith et al. 2011)	Global	Field	Various	322
	Long and Anderson (2014)	Illinois, USA	Field (dynamic)	Various	111
	ZJU-ICL (Yang et al. 2016)	Global	Field	Sand	117
	Long (2016)	Wisconsin, USA	Field (static/dynamic)	IGM	215
	Lehane et al. (2017)	Global	Field	Various	120
	Adhikari et al. (2018)	Wyoming, USA	Field	Soft rock	25

Table 2 (continued).

Geotechnical structure	Database/reference	Data source	Test type	Geomaterial	N
Driven piles	TxDOT (Moghaddam et al. 2018)	Texas	Field	Various	33
	Tang and Phoon (2018b, 2018c, 2018d)	Global	Field	Various	783
Helical piles	Tang and Phoon (2018a, 2019b)	Canada/USA	Field	Various	1010
Driven cast-in-situ piles	Long (2013)	Wisconsin, USA	Field	Various	182
	Flynn (2014)	United Kingdom	Field	Sand	116
Pile foundations	FHWA DFTLD (Abu-Hejleh et al. 2015)	Mainly in USA	Field	Various	1567
	Dithinde et al. (2011)	South Africa	Field	Various	174
	IFSTTAR (Burlon et al. 2014)	France	Field	Various	174
	Niazi (2014)	Global	Field	Various	330
	Galbraith et al. (2014)	Ireland	Field	Various	175
	AUT-CPT (Moshfeghi and Eslami 2018)	Global	Field	Various	466
	WBPLT (Chen et al. 2014)	Global	Field	Various	613
	LADOTD (Rauser and Tsai 2016)	Louisiana, USA	Field (static/dynamic)	Various	1465
	Nanazawa et al. (2019)	Japan	Field	Various	441
Micropiles	Almeida and Liu (2019)	Canada	Field	Ontario soils	47
Foundations	EPRI (Kulhawy et al. 1983)	USA	Field	Various	804
Mechanically stabilized earth walls	Huang and Bathurst (2009)	—	Laboratory	Cohesionless	318
	Miyata and Bathurst (2012a)	Japan	Laboratory/in situ	Cohesionless	652
	Miyata and Bathurst (2012b)	Japan	Laboratory	Various	503
	Miyata et al. (2014)	Japan	Laboratory	N/A	362
	Miyata and Bathurst (2015)	Japan	Field	Various	520
	Miyata and Bathurst (2019)	Global	In situ	Cohesionless	113
	Allen and Bathurst (2018)	—	Field	Various	378
	Miyata et al. (2018)	—	In situ/laboratory	Various	202
	Wood et al. (2012a, b)	Texas, USA	Laboratory	Cohesionless	650
Soil nail walls	Lazarte (2011)	—	Field	—	166
	Cheung and Shum (2012)	Hong Kong	Field	CDG/CDV	913
	Lin et al. (2017)	Global	In situ	—	123
	Liu et al. (2018)	—	In situ	—	95
	Yuan et al. (2019)	China	In situ	Various	144
Multi-anchor walls	Miyata et al. (2011)	Japan	In situ	Various	28
Slopes	Travis et al. (2011)	Global	Field	Various	157
	Bahsan et al. (2014)	—	Field	Clay	43
Excavations (base heave)	Wu et al. (2014)	Global	In situ	Cohesive	24
Pipes	White et al. (2008)	—	Small/full-scale	Sand	61
	Stuyts et al. (2016)	—	Small/full-scale	Sand	108
	Ismail et al. (2018)	—	Small scale/centrifuge	Sand	143
Plate anchors	White et al. (2008)	—	Small/full-scale	Sand	54

Table 2 (continued).

Geotechnical structure	Database/reference	Data source	Test type	Geomaterial	N
Plate anchors	Stuyts et al. (2016)	—	Small/full-scale	Sand	192
Helical anchors	Tang and Phoon (2016)	—	Laboratory	Cohesive	78
			Field	Cohesive	25
Shoring anchors	Chahbaz et al. (2019)	Beirut	Field	Clay/marl/limestone	70
Cantilever wall	Phoon et al. (2009)	—	Centrifuge	Sand	20
Excavation (stability)	Marsland (1953)	—	Small-scale	Loose/dense sand	23
			Large-scale		10
Excavation (wall displacement)	Long (2001)	Global	Field	Various	296
	Moormann (2004)	Global	Field	Soft soil	530
	Wang J. et al. (2010)	Shanghai	Field	Soft soil	300
	Wu et al. (2013)	Taipei	Field	Soft clay	22

Note: CDG = completely decomposed granite; CDV = completely decomposed volcanic; IGM = intermediate geomaterial; N = number of load tests; NUS = National University of Singapore; UWA = University of Western Australia; ZJU = Zhejiang University; ICL = Imperial College London.

The key theoretical difficulty here is that the characteristics of geotechnical data are more challenging than scarcity. The entire literature on reliability and risk management exist, because geotechnical data are uncertain and this uncertainty is magnified and its formal treatment possibly restricted by scarcity. The notion that decision making in geotechnical engineering is a matter of "calculated" risk is well appreciated for many years, although the actual "calculation" remains steeped in empiricism. Casagrande (1965)'s concept of "calculated risk" embodies the following two elements:

1. The use of imperfect knowledge, guided by judgment and experience, to estimate the probable ranges for all pertinent quantities that enter into the solution of a problem;
2. The decision on an appropriate margin of safety, or degree of risk, taking into consideration economic factors and the magnitude of losses that would result from failure.

"Imperfect knowledge" has been interpreted as uncertain knowledge, but there are other characteristics that are arguably of comparable importance as explained in the following section. The topic of geotechnical risk has since been covered by at least three Terzaghi Lectures [Robert Whitman (1981). Evaluating calculated risk in geotechnical engineering; Suzanne Lacasse (2001). Protecting society from landslides - the role of the geotechnical engineer; and John Christian (2003). Geotechnical engineering reliability: How well do we know what we are doing?] and one Rankine Lecture [Suzanne Lacasse (2015). Hazard, Risk and Reliability in Geotechnical Practice]. It is safe to say that management of "uncertain geotechnical truth" (Baker 2010; Spross et al. 2018) is more of an art than science in practice. It is timely to examine the role of data in geotechnical risk management with these methodological advancements in mind.

3 Characteristics of Geotechnical Data

It is overly simplistic to say that geotechnical data are always scarce. The previous section clearly points out that this is true only for site-specific data. One can ponder if this will remain true even at the site level in the face of fast developing digital technologies. It is safe to say that the volume, variety, and velocity of data will continue to increase and the demand to manage data as assets in themselves will increase. Even at this point in time, the amount of generic data from multiple sites is certainly much larger than what is shown in Tables 1 and 2. Data from past projects are frequently left unattended, because engineers do not know what to do with them! The authors venture to suggest that ideal data (site-specific data directly suitable for design) may be scarce, but less ideal data from other sites are voluminous. One may argue against the presence of big data in geotechnical engineering by appealing to site-specificity, but we are undoubtedly in possession of big indirect data (BID). BID will encompass any data that are potentially useful but not directly applicable to the decision at hand. A generic database will be one type of BID.

Besides possible scarcity, geotechnical data are generally multivariate as shown in Table 3. It is uneconomical to mobilize equipment just to conduct a single test. In addition, genuine multivariate data are rarely collected in a site investigation program, because it is not cost effective to conduct multiple tests in close proximity. There is an obvious tradeoff between conducting different tests in different locations and conducting different tests in the same location. The former strategy collects more information on the spatial variability of the site. The latter strategy collects information on the cross-correlations among all tests. In practice, it is common to adopt an intermediate strategy involving conducting different test combinations at different depths and locations. The grayed out cells in Table 3 denote absent measurements. Hence, geotechnical data are typically "incomplete".

Table 3. Site investigation results for a silty clay layer at a Taipei site (Ou and Liao 1987).

				Test results						
Depth (m)	s_u (kN/m^2)		s_u(mob) (kN/m^2)	LL (Y_1)	PI (Y_2)	LI (Y_3)	s'_v/P_a (Y_4)	s'_p/P_a (Y_5)	s_u(mob)/s'_v (Y_6)	q_{t1} (Y_9)
12.8	UU	55.2	46.9	30.1	9.1	1.20	1.26	1.71	0.37	3.35
14.8	VST	50.7	52.9	32.8	12.8	1.43	1.43		0.36	3.34
16.1	UU	61.9	51.7	36.4	14.5	1.24	1.54		0.33	3.15
17.8	UU	54.2	42.8	41.9	18.9	0.90	1.68	1.79	0.25	2.74
18.3	VST	59.5	59.3				1.72		0.34	2.76
20.2	UU	73.1	60.5	38.1	17.3	0.70	1.88		0.32	2.73
22.7	VST	63.3	64.4	37.0	16.0	0.58	2.08		0.31	2.97
24.0	UU	82.2	67.5	38.0	16.2	0.75	2.19	2.19	0.30	2.80
26.6	UU	98.1	82.1	34.8	13.8	0.80	2.41		0.34	3.92

Figure 2 illustrates that site effects do exist in an important transformation model that relates the normalized undrained shear strength to the overconsolidation ratio. Although site effects are well known, they are mainly characterized in research studies through a testing programme that is more detailed that what is routinely carried out in practice and for rather distinctive geo-materials. Kulhawy and Mayne (1990) pointed out that "comprehensive characterization of the soil at a particular site would require an elaborate and costly testing programme, well beyond the scope of most project budgets". To the knowledge of the authors, no one has characterized site effects based on more routine data such as those shown in Table 3 commonly collected at a project level. In practice, site effects are broadly appreciated based on geology, soil mechanics, and experiences at comparable sites, rather than characterized quantitatively through a detailed multivariate analysis of the site data. The typical caveat included in design guides would include a general statement such as "caution must always be exercised when using broad, generalized correlations of index parameters or in-situ test results with soil properties. The source, extent, limitations of each correlation should be examined carefully before use to ensure that extrapolation is not being done beyond the original boundary conditions. 'Local' calibrations, where available, are to be preferred over the broad, generalized correlations" (Kulhawy and Mayne 1990). Notwithstanding this sensible caveat, the engineer is typically left with no recourse but to use these generalized correlations in the absence of "local" versions. Hence, BID is already routinely used in practice in the form of Figure 1. One could surmise that it has some real value.

Phoon (2018) suggested that the characteristics of geotechnical data can be succinctly described as *MUSIC*: Multivariate, Uncertain and Unique, Sparse, and InComplete. The "unique" and "incomplete" characteristics have not received the attention they deserve in the literature, although they are surely present to different degrees in geotechnical databases. Table 3 is a site-specific example of a *MUSIC* database. Each row (record) in a *MUSIC* database is treated as independent. This assumption is reasonable if the depth interval between each record is larger than the spatial correlation length. Ching and Phoon (2019b, 2019d) extended *MUSIC* to *MUSIC-X* to account for spatial correlation between two records measured in close proximity. The symbol "*X*" is adopted to foreground the spatial/temporal dimension in *MUSIC* data. Spatial variability is a well-recognized characteristic in many geo-disciplines such as geostatistics. Spatial variation is used in the broad sense where stratigraphic changes and other variable geologic features are included. Other characteristics may emerge as property databases grow to incorporate other sources of data. It will be illustrated in the next section that an in-depth understanding of these data characteristics is needed to develop data-driven algorithms that will bring more value to practice.

4 Data-Driven Algorithms

4.1 Compressive sampling

Compressive sampling (or sensing, CS) is a novel sampling paradigm in digital signal processing to reconstruct a signal (e.g., an image with 1000 × 1000 = 1 million pixels) from a small number of measurements on that signal (Candès et al. 2006; Donoho 2006; Candès and Wakin 2008; Wang and Zhao 2016; Comerford et al. 2016, 2017). In the context of signal processing, Table 3 is a 9 × 9 matrix and can be considered as an image with 9 × 9 = 81 pixels and missing values at 9 pixels. Then, the *MUSIC-X* problem associated with Table 3 becomes a problem of how to estimate or recover the 9 missing values or how to add a new row to Table 3 at a new given depth. Indeed, many geotechnical data are images, such as geology maps and subsurface geological cross-section, and direct measurements on the image are often sparse and only taken at a limited number of locations. In linear algebra, a 2D image with $N_{x1} \times N_{x2}$ pixels, such as the color map shown in Figure 3, can be represented by a matrix **F** with a dimension of $N_{x1} \times N_{x2}$ and expressed as a weighted summation of $N_{x1} \times N_{x2}$ number of 2D basis functions, such as cosine or wavelet functions (Zhao et al. 2018):

$$\mathbf{F} = \sum_{t=1}^{N_{x1} \times N_{x2}} \mathbf{B}_t^{\mathrm{2D}} \omega_t^{\mathrm{2D}} \tag{1}$$

in which $\mathbf{B}_t^{\mathrm{2D}}$ is the t-th 2D basis function that is independent of **F**, while ω_t^{2D} is the weight corresponding to $\mathbf{B}_t^{\mathrm{2D}}$. In the context of CS, most images are compressible, suggesting that only a small number of basis functions is necessary to properly represent the image and that the magnitudes of most ω_t^{2D} are almost zero or trivial except several non-trivial ones (i.e., coefficients with significantly large magnitudes). Therefore, once the non-trivial coefficients ω_t^{2D} can be identified and estimated using sparse measurements **Y**, signal **F** can be approximately reconstructed. The relation between **Y** and ω_t^{2D} is expressed as (Zhao et al. 2018):

$$\mathbf{Y} = \boldsymbol{\Psi}_{x_1} \mathbf{F} \boldsymbol{\Psi}_{x_2} = \sum_{t=1}^{N_{x1} \times N_{x2}} \mathbf{A}_t^{\mathrm{2D}} \omega_t^{\mathrm{2D}} \tag{2}$$

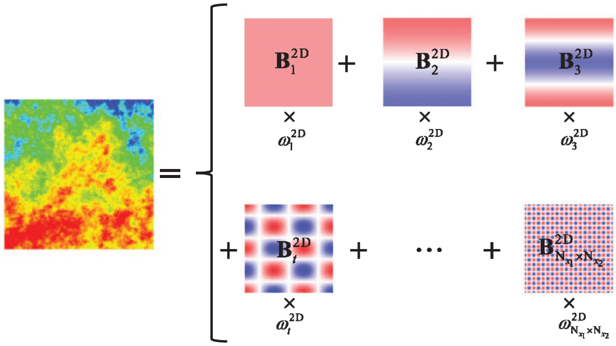

Figure 3. Representation of a 2D image in compressive sampling (Zhao et al. 2018).

where $\mathbf{\Psi}_{x1}$ and $\mathbf{\Psi}_{x2}$ are problem-specific measurement matrices, reflecting the locations of elements of **Y** in **F** along row and column directions, respectively. Although Eq. (2) is underdetermined, the non-trivial coefficients in ω_t^{2D} can be obtained using several existing methods, including non-probabilistic methods, such as orthogonal matching pursuit (Pati et al. 1993; Wang and Zhao 2016), and Bayesian methods (Ji et al. 2008, 2009; Wang and Zhao 2017). Once the non-trivial coefficients in ω_t^{2D} are properly estimated, the ω_t^{2D} can be approximated as $\hat{\omega}_t^{2D}$ by setting those trivial elements of ω_t^{2D} as zero. Then, the 2D image of interest **F** can be approximated as:

$$\hat{\mathbf{F}} = \sum_{t=1}^{N_{x1}\times N_{x2}} \mathbf{B}_t^{2D}\hat{\omega}_t^{2D} \tag{3}$$

When Bayesian methods is used to estimate $\hat{\omega}_t^{2D}$ (Ji et al. 2008, 2009; Wang and Zhao 2017; Huang et al. 2016; Zhao et al. 2018), both the best estimate and covariance of $\hat{\mathbf{F}}$ are obtained. As an illustration, Figure 4a shows a 2D vertical cross-section with a thickness of 10.20m (in depth direction x_1) and a length of 20.44m (in horizontal direction x_2). A resolution of 0.04m is adopted in this example for both x_1 and x_2 directions, leading to 256×512 = 131, 072 data points in total. Suppose that 10×4 data points, as shown by open circles in Figure 4a, are taken as the measured data **Y** and used together with their corresponding locations to recover the complete 2D cross-section with 131, 072 data points. Figure 4b shows the best estimate of the 2D cross-section obtained from Bayesian CS or BCS (Zhao et al. 2018). A similar spatially varying trend can be observed in Figures 4a and 4b, even when only 40/131,072 = 0.03% of the original data are used as input to BCS. To examine the uncertainty associated with the BCS results, Figure 4c shows standard deviation, SD, obtained from BCS multiplying a factor of 1.96, and Figure 4d shows the absolute residuals between the original 2D data (i.e., Figure 4a) and the BCS best estimates from **Y** (i.e., Figure 4b). The 1.96 SD surface shown in Figure 4c are generally larger than most residuals shown in Figure 4d, while some residuals at locations far away from the measurement data **Y** are larger than the 1.96 SD. These observations imply that many residuals fall within the region defined by the mean±1.96 SD (i.e., approximate 95% confidence interval or "credible interval" in Bayesian statistics parlance).

CS is data-driven and the CS results improve as the data quantity increases. Figures 5b, 5c, and 5d show the BCS best estimate for three different measurement number scenarios of 30×15, 50×25 and 100×50, respectively, together with the 10×4 scenario in Figure 5a. When compared with the 10×4 scenario in Figure 5a, the best estimate for the 30×15 scenario (see Figure 5b) is more similar to the original 2D data as shown in Figure 4a. As the number of measurement data further increases, the best estimate of 2D data (see Figures 5c and 5d) gradually approaches to the original complete 2D data (see Figure 4a), and the BCS SD is reduced to almost zero, as shown by Zhao et al. (2018).

Although Bayesian method can be used to provide both best estimate and uncertainty of ω_t^{2D} in Eq. (1) (Wang and Zhao 2017; Zhao et al. 2018), the fundamental principle of CS [see Eq. (1) and Eq. (2)] is non-probabilistic and philosophically different from the commonly used geostatistical methods, such as kriging. In kriging, function types for both trend function and auto-covariance function are generally pre-assumed, and extensive measurement data, which are often not available in geotechnical practice, are needed for validating the assumptions and stationarity and for estimating the parameters required in the trend function and auto-covariance function. When the measurements are sparse, it is extremely challenging to examine the stationarity assumption, properly select the suitable function types for trend function and auto-covariance function, or accurately estimate the required parameters, such as correlation length. In contrast, CS and BCS are non-parametric, and they do not need pre-assumed trend function or auto-covariance function [see Eq. (1) and Eq. (2)], therefore bypassing all the difficulties mentioned above for kriging. Wang Y. et al. (2017) and Zhao et al. (2018) performed comparative studies between kriging and BCS for 1D and 2D data, respectively, and showed that BCS performs much better than kriging for sparse measurements and that BCS and kriging have similar performance for extensive measurements.

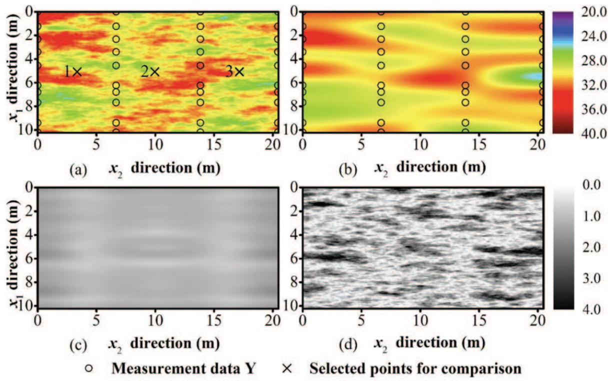

Figure 4. Comparison between the original 2D data and that estimated from 10×4 measurement: (a) Original 2D data; (b) best estimate of spatially varying 2D data; (c) 1.96 standard deviation of estimated results; and (d) absolute residuals between (a) and (b) (Zhao et al. 2018).

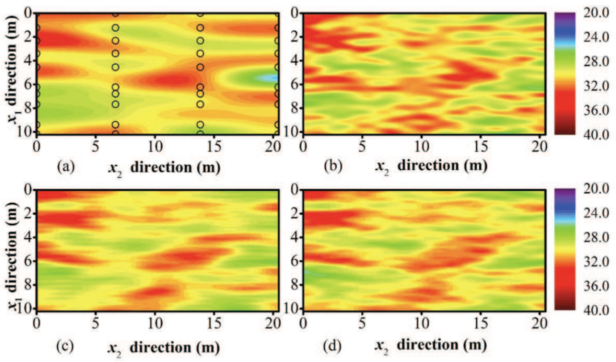

Figure 5. Best estimate of spatially varying 2D data under different number of measurements scenarios: (a) 10×4 (b) 30×15 (c) 50×25 and (d) 100× 50 (Zhao et al. 2018).

The BCS results can be used together with Karhunen–Loève (KL) expansion to generate random field samples (RFSs) directly from sparse measurements (Wang Y. et al. 2018). KL simulation of RFSs generally requires the mean of the random field of interest and deterministic orthogonal eigen-functions and eigenvalues corresponding to the covariance function or covariance matrix (Phoon et al. 2002). On the other hand, BCS provides both the best estimate (i.e., the mean of the random field) and the covariance matrix for the signal of interest directly from sparse measurements. Wang Y. et al. (2018) developed a BCS-KL random field generator to simulate RFSs directly from sparse measurements and offered a Bayesian perspective of random field modeling of site-specific spatial variability (Wang Y. et al. 2019a). The BCS-KL generator is non-parametric and data-driven. No pre-determined function forms are needed for marginal probability density function or covariance function of the random field. Therefore, the BCS-KL generator is readily applicable to non-Gaussian and non-stationary RFSs, including RFSs with non-stationary auto-covariance structure (Montoya-Noguera et al. 2019) and RFSs with unknown trend function without de-trending (Wang Y. et al. 2019b). In addition, the BCS-KL generator may be readily extended to simulate cross-correlated bivariate RFSs (Zhao and Wang 2018).

4.2 Bayesian machine learning

Ching and Phoon (2019c) proposed a Bayesian machine learning method to construct a site-specific distribution function for a *MUSIC* database such as that shown in Table 3. Each database consists of m soil parameters (Y_1, Y_2, … , Y_m) (columns of Table 3) at n different depths (z_1, z_2, …, z_n) (rows of Table 3). Note that site data are typically multivariate (m > 1) and incomplete (grayed out cells in Table 3). The observed data are denoted by $\mathbf{Y}^o$ and unobserved data denoted by $\mathbf{Y}^u$. Because soil parameters can be highly non-normal, Ching and Phoon (2015b) adopted an analytical transformation based on the Johnson distribution to convert (Y_1, …, Y_m) to approximately normal data. The approximately normal data are denoted by $\mathbf{x} = (X_1, \ldots, X_m)^T$, where "T" refers to vector/matrix transpose. A key assumption made in Ching and Phoon (2019c) is that $\mathbf{x}$ at a certain depth follows the multivariate normal PDF:

$$f\left(\mathbf{x} \mid \boldsymbol{\mu}_s, \mathbf{C}_s\right) = \left|\mathbf{C}_s\right|^{-\frac{1}{2}} \left(2\pi\right)^{-\frac{m}{2}} \exp\left[-\frac{1}{2}\left(\mathbf{x}-\boldsymbol{\mu}_s\right)^T \mathbf{C}_s^{-1} \left(\mathbf{x}-\boldsymbol{\mu}_s\right)\right] \tag{4}$$

The multivariate normal PDF has mean vector = $\boldsymbol{\mu}_s$ and covariance matrix = $\mathbf{C}_s$; the subscript "s" is to highlight that $\boldsymbol{\mu}_s$ and $\mathbf{C}_s$ are "site-specific". Because site-specific data are sparse (small n), it is technically challenging to estimate $\boldsymbol{\mu}_s$ and $\mathbf{C}_s$ using conventional methods such as matching moments or maximizing likelihood. It is also very challenging to estimate the statistical uncertainties associated with $\boldsymbol{\mu}_s$ and $\mathbf{C}_s$, which are significant for a typical set of site-specific data and will dominate other uncertainties when n is sufficiently small. Ching and Phoon (2019c) developed a novel Gibbs sampler to overcome this long standing challenge. The key idea is to treat $\boldsymbol{\mu}_s$, $\mathbf{C}_s$, and $\mathbf{x}^u$ (transformed from $\mathbf{Y}^u$) as unknown random quantities and to sequentially sample one random quantity at a time from distributions conditioned on the rest of the quantities and the observed data $\mathbf{x}^o$ (transformed from $\mathbf{Y}^o$). Simulation is practical because these conditioned distributions are available in closed-form for suitably chosen conjugate priors. There is room to further improve efficiency, but this is possibly the first practical proposal to tackle all aspects of *MUSIC*, particularly incompleteness in the presence of sparsity and high random dimensions.

Consider properties at a new depth (x_{new}) that does not appear in the training data previously used in the Gibbs sampler. Based on the total probability theorem, the conditional multivariate PDF $f(\underline{x}_{new}|\mathbf{X}^o)$ is a mixture of multivariate normal PDFs:

$$f\left(\underline{x}_{new} \mid \mathbf{X}^o\right) = \int f\left(\underline{x}_{new} \mid \underline{\mu}_s, \mathbf{C}_s\right) \cdot f\left(\underline{\mu}_s, \mathbf{C}_s \mid \mathbf{X}^o\right) \cdot d\underline{\mu}_s d\mathbf{C}_s \approx \frac{1}{T-t_b}\left[\sum_{t=t_b+1}^{T} N\left(\underline{x}_{new} \mid \underline{\mu}_{s,t}, \mathbf{C}_{s,t}\right)\right] \tag{5}$$

where ($\underline{\mu}_{s,t}$, $\mathbf{C}_{s,t}$) are the GS samples at time step = t; t_b is the end of the burning-period; and T is the total number of GS time steps or samples. Figure 6 illustrates the shape of $f(\underline{x}_{new}|\mathbf{X}^o)$, the histogram of the mean of X_1, and the histogram of the correlation coefficient for two, ten, and one hundred data points simulated from a bivariate normal distribution (X_1, X_2) with mean = 0 and covariance matrix = identity. In general, $f(\underline{x}_{new}|\mathbf{X}^o)$ is not a multivariate normal distribution. It is flat or uninformative when n = 2, because there is almost no site data to "learnt" from. The histogram of the mean covers a wide range and the histogram of the correlation coefficient is not too far from a uniform distribution as to be expected.

The simulation of a site-specific probability distribution appears very complicated to the average engineer, but it can support a critical design decision on how to choose soil/rock properties at a particular site by "learning" from site-specific data alone. An appreciation of geology tempered by experience and judgment remain important as a reality check, but such a machine learning method is clearly of immediate value to routine practice when applied judiciously to complement the expertise of the engineer. It is not appropriate to ask an engineer to process *MUSIC* data by judgment alone. For example, Ching and Phoon (2019a) developed a similarity index (S) based on $f(\underline{x}_{new}|\mathbf{X}^o)$ to identify records from a generic database that are "similar" to those from a specific site. Figure 7 illustrates a target site in Onsøy, Norway (red solid squares), and how records from another site in Norway (Drammen) are identified as "similar" (S > 1) (black solid circles) or "dissimilar" (S < 1) (black open circles) based on this concept. The Drammen and Onsøy sites are roughly 50 km apart with comparable geologic origins (Lacasse et al. 1981; Lacasse and Lunne 1982). Ching and Phoon (2019b, 2019d) generalized Eq. (5) to predict properties in a new location based on all available information, by conditioning on the other test results at the same depth using both parameter cross-correlation and conditioning on the data measured at nearby depths through spatial correlation. The insights provided by these complex algorithms are surely beyond the reach of judgment.

It is also noteworthy that the applicability of the proposed GS method is independent of the nature of the data. It can be used to construct the site-specific PDF model for clays, sands, or rocks. Namely, it is a machine learning framework that is purely driven by data. Bayesian machine learning methods such as Bayesian network (Heckerman et al. 1995), Bayesian neural network (MacKay 1995), Gaussian processes (Rasmussen and

Williams 2006), relevance vector machine (Tipping 2001), Bayesian deep learning (Wang and Yeung 2016), Bayesian model class selection (Beck and Yuen 2004; Yuen 2010), and Bayesian simulation (MacKay 1998; Gilks et al. 1996; Doucet et al. 2001) have made significant advancement in recent years. The GS method proposed in the current study belongs to Bayesian simulation methods.

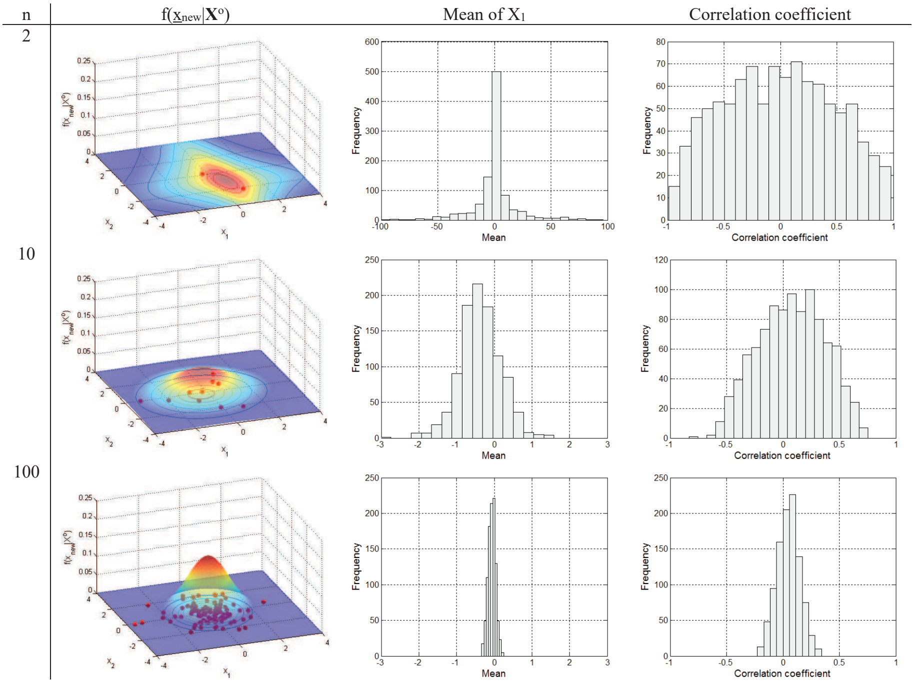

Figure 6. Site-specific probability distribution $f(\underline{x}_{new}|\mathbf{X}^o)$ and the histogram of the correlation coefficient "learnt" from two, ten, and one hundred measured data points simulated from a bivariate normal distribution (X_1, X_2) with mean = 0 and covariance matrix = identity.

5 What Next?

A taxonomy of methods based on the type/amount of data available could help guide future development in data-driven algorithms and strengthen a virtuous cycle of data collection hardware developing hand in hand with algorithms. Hand (2014) said: "In general, when building statistical models, we must not forget that the aim is to understand something about the real world. Or predict, choose an action, make a decision, summarize evidence, and so on, but always about the real world, not an abstract mathematical world: our models are not the reality - a point well made by George Box in his oft-cited remark that 'all models are wrong, but some are useful'". Hence, it is not fruitful to ask whether a probability model is right or wrong (our community has been embroiled in this question for many years), but to judge a model by its ability to help us make economic decisions in the real world.

In fact, why do we need a model at all? One answer is that we do not have sufficient data to make a decision without mediation by a model. The simplest probability model is to assume data are independent and identically distributed (i.i.d.). Limited data are needed to characterize this model, but it clearly deviates from a reality that exhibits spatial variability. The random field model is a closer match to this reality, but it cannot be applied in its most general non-stationary form because we do not have sufficient site investigation data for statistical characterization. The current practice is to assume a trend function can be removed from the data and the residuals are second-order stationary within a typical site. The reason for this assumption is that pairs of measurements regardless of where they are measured can be used to estimate the autocorrelation function. Needless to say, there is no trend, no stationary residuals, and no autocorrelation function in reality. These concepts exist purely within the stationary random field model. However, it can produce useful outcomes, such as estimating the values at unmeasured locations using kriging or general regression (Yuen and Ortiz 2016, 2018; Yuen et al. 2016). These predictions produced by the stationary random field model are closer to reality than those produced by the i.i.d. model (which are simply equal to the mean).

However, trend removal can be difficult (Ching et al. 2016b, 2017b; Ching and Phoon 2017). Estimation of random field parameters is also computationally challenging (Tian et al. 2016; Xiao et al. 2018; Wang H. et al. 2018). Fine details of the autocorrelation function such as sample path "smoothness" are important (Ching and Phoon 2019e). Characterization of site stratigraphy is a major missing feature of past random field studies until quite recently (Wang Y. et al. 2013; Ching et al. 2015; Li et al. 2016; Qi et al. 2016; Wang X. et al. 2016; Wang H. et al. 2017; Wang X. et al. 2018; Cao et al. 2019; Wang H. et al. 2019; Wang X. et al. 2019). More discussions are found elsewhere (Juang et al. 2018).

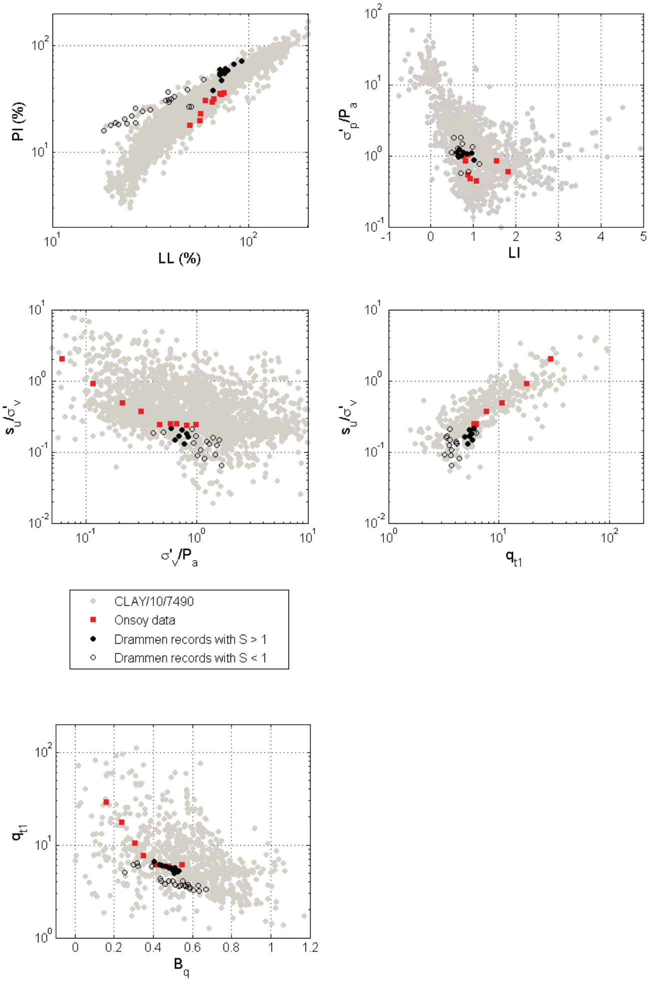

Figure 7. Automatic detection of records from a generic database CLAY/10/7490 that are "similar" to those from a specific site in Onsøy, Norway (Ching and Phoon 2019a).

Compressive sampling is not derived from the random field model, but originates from signal processing. Some attempts have been made to apply compressive sampling without detrending (Wang Y. et al. 2019a) and without assuming stationarity (Wang Y. et al. 2019b) as discussed in the preceding section. In fact, when sufficient data are available say in the form of training images, multiple point methods that consider more than two-point autocorrelation information are being explored in geostatistics (Mariethoz and Caers 2015). These methods are regarded as closer to "model free" in the sense that they are not founded on probability theory. The level of abstraction is certainly higher than going from parametric to non-parametric statistics, but how does one quantify veracity of the outcome in the absence of a probabilistic basis? Bayesian methods have been adopted to carry out compressive sampling for this reason (Ji et al. 2008, 2009; Wang and Zhao 2017; Huang et al. 2016; Zhao et al. 2018). The conventional wisdom is that big data can be characterized by 4Vs: volume, velocity, variety and veracity.

It suffices to say that no data-driven algorithm exist that can deal with the complex subsurface reality in its 3D entirety and for the full range of *MUSIC-X* characteristics. This line of inquiry is likely to be very active in the near future with the strong interest in machine learning. The authors venture to suggest "**Seven Es**" to guide the development of such algorithms that will be of value to practice, promotes data exchange, robust, maintains alignment with current knowledge and experience, and engages engineering judgment in a meaningful way:

1. **Essence**: Data is the essence and therefore, algorithms must be data-centric besides value-centric. More precise understanding of the data characteristics in the geotechnical environment is needed. An algorithm-centric strategy requires data to fit its assumptions. This is only possible if new data acquisition hardware is developed alongside.
2. **Economic value**: Focus on monetizing data. Remember the adage: "all models are wrong, but some are useful".
3. **Exchange**: The industry is more likely to share and exchange data if client confidentiality can be respected. This requires development of suitable data anonymization methods.
4. **Extremes:** Identification of outliers and/or robustness of algorithms against outliers are fundamental issues that one should be mindful of given their potential impact on the outcomes (Yuen and Mu 2012; Mu and Yuen 2019). The authors suggest that *MUSIC* can be re-interpreted to cover extremes: Multivariate, Uncertain and Unique, Sparse, Incomplete, and potentially *Corrupted*.
5. **Errors:** An engineer can make a more informed decision if both bias and precision of the outcomes can be provided. Biased and imprecise data will produce biased and imprecise outcomes. It is not sufficient to provide the most likely outcomes, because an engineer needs to manage risks. Responsible risk management is a core element of our professional ethics.
6. **Extrapolation:** Need to watch out for over-fitting and to caution users when extrapolation occurs.
7. **Explanation:** It is judicious to establish a degree of connection with the existing body of knowledge and experience. Correlation is not the same causality. Engineers cannot "understand" outcomes delivered purely by a black-box algorithm and cannot meaningfully "agree" or "disagree" with such outcomes.

6 Concluding Remarks

Digitization is the process of converting information to a digital format. This is more or less taking place in tandem with the growth of computing. Digitalization is the deployment of digital technologies to transform an entire industry. Data lies at the core. But deeper insights must be gleaned, beyond applying data as inputs in a physical model to predict responses or as direct measurements of responses to support the observational approach, to produce sufficient value for decision making so that data can be viewed as assets in themselves. This paper explores the availability and nature of geotechnical data and presents two recent advances made in this direction for a specific but important task of estimating soil/rock properties (compressive sampling and Bayesian machine learning). Data-driven decision making does not imply taking the engineer out of the entire life cycle management chain. It is intended to support rather than to replace human judgment.

Gerbert et al. (2016) concluded that the construction sector "has finally set out on the digital pathway, and a profound transformation — long overdue — now seems inevitable. The sector as a whole is bound to benefit; so, too, is society at large as well as the international economy". It cautioned against staying still: "Individual companies that continue to ignore the digital wave will struggle to survive. For adopters, speed matters: there is only a narrow window of time during which digital savvy provides a significant competitive advantage over the average industry participant. If companies want to contribute to redefining the competitive landscape, they need to seize the opportunity soon".

Acknowledgments

The first author is grateful to ISSMGE TC304 for the invitation to deliver the 4th Suzanne Lacasse Lecture in the 7th International Symposium on Geotechnical Safety and Risk (ISGSR2019), 11-13 Dec 2019, Taipei, Taiwan. The first author was privileged to spend 5 months in the Norwegian Geotechnical Institute between May and September 2004, from which he developed a richer understanding of geotechnical risk management in practice from many conversations with Dr Suzanne Lacasse and his host, Dr Farrokh Nadim. The authors would like to thank Dr Chong Tang for his valuable comments including updating Tables 1 and 2 and editorial assistance. The following experts generously contributed their time to improve this paper: Zijun Cao, Stefan Larsson, Dian-qing Li, He-Qing (Max) Mu, Anders Prästings, Hui Wang, Ka-Veng Yuen, Kelvin, Limin Zhang. Lastly, the authors acknowledge the use of soil data from the ISSMGE TC304 database open sharing initiative (304dB) (http://140.112.12.21/issmge/Database_2010.htm).

References

AbdelSalam, S., Baligh, F., and El-Naggar, H.M. (2015). A database to ensure reliability of bored pile design in Egypt. *Proceedings of the Institution of Civil Engineers – Geotechnical Engineering*, 168(2), 131-143.

Abu-Hejleh, N., Abu-Farsakh, M., Suleiman, M., and Tsai, C. (2015). Development and use of high-quality databases of deep foundation load tests. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2511, 27-36.

Adhikari, P., Gebreslasie, Y., Ng, K., Sullivan, T., and Wulff, S. (2018). Static and dynamic analysis of driven piles in soft rocks considering LRFD using a recently developed electronic database. *Installation, Testing, and Analysis of Deep Foundations (GSP 294)*, 83-92. Reston, VA: ASCE.

Akbas, S. (2007). *Deterministic and probabilistic assessment of settlements of shallow foundations in cohesionless soils*. Ph.D. thesis, Cornell University.

Allen, T. and Bathurst, R.J. (2018). Application of the simplified stiffness method to design of reinforced soil walls. *ASCE Journal of Geotechnical and Geoenvironmental Engineering*, 144(5), 04018024.

Almeida, A. and Liu, J.Y. (2018). Statistical evaluation of design methods for micropiles in Ontario soils. *DFI Journal – The Journal of the Deep Foundations Institute*, 12(3), 133-146.

Asem, P., Long, J., and Gardoni, P. (2018). Probabilistic model and LRFD resistance factors for the tip resistance of drilled shafts in soft sedimentary rock based on axial load tests. *Innovations in Geotechnical Engineering: Honoring Jean-Louis Briaud (GSP 299)*, 1-46. Reston, VA: ASCE.

Augustesen, A. (2006). *The effects of time on soil behaviour and pile capacity*. Ph.D. thesis, Aalborg University.

Bahsan, E., Liao, H.J., Ching, J., and Lee, S.W. (2014). Statistics for the calculated safety factors of undrained failure slopes. *Engineering Geology*, 172, 85-94.

Baker, C.N. (2010). Uncertain geotechnical truth and cost effective high-rise foundation design. *Art of Foundation Engineering Practice (GSP 198)*, 1-43. Reston, VA: ASCE.

Beck, J.L. and Yuen, K.V. (2004). Model selection using response measurements: Bayesian probabilistic approach. *Journal of Engineering Mechanics*, 130(2), 192-203.

Burlon, S., Frank, R., Baguelin, F., Habert, J., and Legrand, S. (2014). Model factor for the bearing capacity of piles from pressuremeter test results: Eurocode 7 approach. *Géotechnique*, 64(7), 513-525.

Candès, E.J., Romberg, J.K., and Tao, T. (2006). Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59(8), 1207-1223.

Candès, E.J. and Wakin, M.B. (2008). An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 25(2), 21-30.

Cao, Z.J., Zheng, S., Li., D.Q., and Phoon, K.K. (2019). Bayesian identification of soil stratigraphy based on soil behavior type index. *Canadian Geotechnical Journal*, 56(4), 570-586.

Casagrande, A. (1965). Role of the "calculated risk" in earthwork and foundation engineering. *ASCE Journal of Soil Mechanics and Foundations Division*, 91(SM4), 1–40.

Chahbaz, R., Sadek, S., and Najjar, S. (2019). Uncertainty quantification of the bond stress – displacement relationship of shoring anchors in different geologic units. *Georisk*, under review.

Chen, Y.J. and Kulhawy F.H. (1994). *Case history evaluation of behavior of drilled shafts under axial and lateral loading*. Report EPRI TR-104601. Palo Alto, CA: Electric Power Research Institute (EPRI).

Chen, Y.J. and Lee, Y.H. (2010). Evaluation of lateral interpretation criteria for drilled shaft capacity. *ASCE Journal of Geotechnical and Geoenvironmental Engineering*, 136(8), 1124-1136.

Chen, Y.J., Lin, S.W., and Kulhawy, F.H. (2011). Evaluation of lateral interpretation criteria for rigid drilled shafts. *Canadian Geotechnical Journal*, 48(4), 634-643.

Chen, Y.J., Liao, M.R., Lin, S.S., Huang, J.K., and Marcos, M.C.M. (2014). Development of an integrated Web-based system with a pile load test database and pre-analyzed data. *Geomechanics and Engineering*, 7(1), 37-53.

Cheung, R.W.M. and Shum, K.W. (2012). *Review of the approach for estimation of pullout resistance of soil nails*. GEO Report No. 264. Geotechnical Engineering Office, Civil Engineering and Development Department, Hong Kong.

Chilès, J-P. and Delfiner, P. (1999). *Geostatistics: Modeling Spatial Uncertainty*. John Wiley & Sons, New York.

Ching, J. and Phoon, K.K. (2012). Modeling parameters of structured clays as a multivariate normal distribution. *Canadian Geotechnical Journal*, 49(5): 522-545.

Ching, J. and Phoon, K.K. (2013). Multivariate distribution for undrained shear strengths under various test procedures. *Canadian Geotechnical Journal*, 50(9), 907-923.

Ching, J. and Phoon, K.K. (2014). Transformations and correlations among some clay parameters – the global database. *Canadian Geotechnical Journal*, 51(6), 663-685.

Ching, J. and Phoon, K.K. (2015a). Reducing the transformation uncertainty for the mobilized undrained shear strength of clays. *ASCE Journal of Geotechnical and Geoenvironmental Engineering*, 141(2), 04014103.

Ching, J. and Phoon, K.K. (2015b). Constructing multivariate distribution for soil parameters. *Chapter 1, Risk and Reliability in Geotechnical Engineering*, CRC Press, 2015, 3-76.

Ching, J. and Phoon, K.K. (2017). Characterizing uncertain site-specific trend function by sparse Bayesian learning. *ASCE Journal of Engineering Mechanics*, 143(7), 04017028.

Ching, J. and Phoon, K.K. (2019a). Measuring similarity between site-specific data and records from other sites. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, under review.

Ching, J. and Phoon, K.K. (2019b). Constructing a site-specific multivariate probability distribution using sparse, incomplete, and spatially variable data. *ASCE Journal of Engineering Mechanics*, under review.

Ching, J. and Phoon, K.K. (2019c). Constructing site-specific multivariate probability distribution model by Bayesian machine learning, *ASCE Journal of Engineering Mechanics*, 145(1), 04018126.

Ching, J. and Phoon, K.K. (2019d). Modeling multivariate, uncertain, sparse, and incomplete site investigation data with spatial variation (MUSIC-X). *Proceedings, 7th International Symposium on Geotechnical Safety and Risk (ISGSR 2019)*, in press.

Ching, J. and Phoon, K.K. (2019e). Impact of auto-correlation function model on the probability of failure. *ASCE Journal of Engineering Mechanics*, 145(1), 04018123.

Ching, J., Phoon, K.K., and Yu, J.W. (2014a). Linking site investigation efforts to final design savings with simplified reliability-based design methods. *ASCE Journal of Geotechnical and Geoenvironmental Engineering*, 140(3), 04013032.

Ching, J., Phoon, K.K., and Chen, C.H. (2014b). Modeling piezocone cone penetration (CPTU) parameters of clays as a multivariate normal distribution. *Canadian Geotechnical Journal*, 51(1), 77-91.

Ching, J., Wang, J.S., Juang, C.H., and Ku, C.S. (2015). Cone penetration test (CPT)-based stratigraphic profiling using the wavelet transform modulus maxima method. *Canadian Geotechnical Journal*, 52(12), 1993-2007.

Ching, J., Li, D.Q., and Phoon, K.K. (2016a). Statistical characterization of multivariate geotechnical data. *Chapter 4, Reliability of Geotechnical Structures in ISO2394*, CRC Press/Balkema, 89-126.

Ching J., Wu, S.S., and Phoon, K.K. (2016b). Statistical characterization of random field parameters using frequentist and Bayesian approaches. *Canadian Geotechnical Journal*, 53(2), 285-298.

Ching, J., Li, K.H., Weng, M.C., and Phoon, K.K. (2018). Generic transformation models for some intact rock properties. *Canadian Geotechnical Journal*, 55(12), 1702-1741.

Ching, J., Lin, G.H., Chen, J.R., and Phoon, K.K. (2017a). Transformation models for effective friction angle and relative density calibrated based on generic database of coarse-grained soils. *Canadian Geotechnical Journal*, 54(4), 481-501.

Ching, J., Phoon K.K., Beck, J.L., and Huang, Y. (2017b). Identifiability of geotechnical site-specific trend functions. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 3(4), 04017021.

Comerford, L., Jensen, H.A., Mayorga, F., Beer, M., and Kougioumtzoglou, I.A. (2017). Compressive sensing with an adaptive wavelet basis for structural system response and reliability analysis under missing data. *Computers and Structures*, 182, 26–40.

Comerford, L., Kougioumtzoglou, I.A., and Beer, M. (2016). Compressive sensing based stochastic process power spectrum estimation subject to missing data. *Probabilistic Engineering Mechanics*, 44, 66–76.

Dithinde, M., Phoon, K.K., Wet, M., and Retief, J. (2011). Characterization of model uncertainty in the static pile design formula. *ASCE Journal of Geotechnical and Geoenvironmental Engineering*, 137(1), 70-85.

D'Ignazio, M., Phoon, K.K., Tan, S.A., and Lansivaara, T. (2016). Correlations for undrained shear strength of Finnish soft clays. *Canadian Geotechnical Journal*, 53(10), 1628-1645.

Donoho, D.L. (2006). Compressed sensing. *IEEE Transactions on Information Theory*, 52(4), 1289-1306.

Doucet, A., De Freitas, N., and Gordon, N. (2001). *Sequential Monte Carlo Methods in Practice*. New York, Springer.

EN 1990:2002. *Eurocode – Basis of structural design*. European Committee for Standardization (CEN), Brussels, Belgium.

EN 1997−1:2004. *Eurocode 7: Geotechnical design — Part 1: General rules*. European Committee for Standardization (CEN), Brussels, Belgium.

Feng, S., and Vardanega, P. J. (2019a). Correlation of the hydraulic conductivity of fine-grained soils with water content ratio using a database. *Environmental Geotechnics*, in press.

Feng, S., and Vardanega, P. J. (2019b). A database of saturated hydraulic conductivity of fine-grained soils: probability density functions. *Georisk*, in press.

Flynn, K. (2014). *Experimental investigations of driven cast-in-situ piles*. Ph.D. thesis, National University of Ireland, Galway.

Galbraith, A., Farrell, E., and Byrne, J. (2014). Uncertainty in pile resistance from static load test database. *Proceedings of the Institute of Civil Engineers – Geotechnical Engineering*, 167(5), 431-446.

Garder, J., Ng, K., Sritharan, S., and Roling, M. (2012). *An Electronic Database for Drilled SHAft Foundation Testing (DSHAFT)*. Report No. InTrans Project 10-366, Iowa Department of Transportation.

Gerbert, P, Castagnino, S., Rothballer, C., Renz, A., and Filitz, R. (2016). *Digital in Engineering and Construction*. Boston Consulting Group.

Gilks, W.R., Richardson, S., and Spiegelhalter, D.J. (1996). *Markov Chain Monte Carlo in Practice*. Chapman and Hill, London.

Gransberg, D.D., Loulakis, M., Touran, A., Gad, G., McLain, K., Sweitzer, S., Pittenger, D., Nova, I.C., Pereira, R.T., and Pinto-Nunez, M. (2018). *Guidelines for managing geotechnical risks in design-build projects*. NCHRP Research Report 884. Washington, DC: The National Academies Press.

Hancock, M. (2018). Data sharing: please feed the Dingo. https://www.geplus.co.uk/features/data-sharing-please-feed-the-dingo/10030557.article.

Hand, D.J. (2014). Wonderful examples, but let's not close our eyes. *Statistical Science*, 29, 98-100.

Heckerman, D., Geiger, D., and Chickering, D. M. (1995). Learning Bayesian networks: The combination of knowledge and statistical data. *Machine Learning*, 20(3), 197-243.

Hossain, M.S. (2014). Experimental investigation of spudcan penetration in multi-layer clays with interbedded sand layers. *Géotechnique*, 64(4), 258-277.

Hossain, M.S. and Randolph, M.F. (2010). Deep-penetration spudcan foundations on layered clays: centrifuge tests. *Géotechnique*, 60(3), 157-170.
Huang, B.Q. and Bathurst, R.J. (2009). Evaluation of soil-geogrid pullout models using a statistical approach. *Geotechnical Testing Journal*, 32(6), 489-504.
Hu, P. (2015). *Predicting punch-through failure of a spudcan on sand overlying clay*. Ph.D. thesis, The University of Western Australia, Perth, Australia.
Huang, Y., Beck, J.L., Wu, S., and Li, H. (2016). Bayesian compressive sensing for approximately sparse signals and application to structural health monitoring signals for data loss recovery. *Probabilistic Engineering Mechanics*, 31 (46), 62–79.
Ismail, S., Najjar, S.S., and Sadek, S. (2018). Reliability analysis of buried offshore pipelines in sand subjected to upheaval buckling. *Proceedings, Offshore Technology Conference (OTC)*, Houston, Texas. OTC-28882-MS.
ISO2394: 2015. *General principles on reliability for structures*. International Organization for Standardization, Geneva, Switzerland.
Juang, C. H., Zhang, J., Shen, M., and Hu, J. (2018). Probabilistic methods for unified treatment of geotechnical and geological uncertainties in a geotechnical analysis. *Engineering geology*, 249, 148-161.
Ji, S., Dunson, D., and Carin, L. (2009). Multitask compressive sensing. *IEEE Transactions on Signal Processing*, 57(1), 92–106.
Ji, S., Xue, Y., and Carin, L. (2008). Bayesian compressive sensing. *IEEE Transactions on Signal Processing*, 56(6), 2346–2356.
Kulhawy, F.H. and Mayne, P.W. (1990). *Manual on estimating soil properties for foundation design*. Report EL-6800, Electric Power Research Institute, Palo Alto, California.
Kulhawy, F.H., O'Rourke, T.D., Stewart, J.P., and Beech, J.F. (1983). *Transmission line structure foundations for uplift-compression loading: load test summaries. Appendix to EPRI final report EL-2870*. Report No. EL-3160-LD. Palo Alto, CA: Electric Power Research Institute (EPRI).
Lacasse, S., Jamiolkowski, M., Lancellotta, R., and Lunne, T. (1981). In situ characteristics of two Norwegian clays. *Proceedings, 10th International Conference on Soil Mechanics and Foundation Engineering*, Stockholm, Vol. 2, 507 - 511.
Lacasse, S. and Lunne, T. (1982). Penetration tests in two Norwegian clays. *Proceedings, 2nd European Symposium on Penetration Testing*, Amsterdam, 661-670.
Lazarte, C.A. (2011). *Proposed specifications for LRFD soil-nailing design and construction*. NCHRP Report 701. Washington, DC: Transportation Research Board.
Lee, K.K. (2009). *Investigation of potential spudcan punch-through failure on sand overlying clays*. Ph.D. thesis, The University of Western Australia, Perth, Australia.
Lehane, B.M., Kim, J.K., Carotenuto, P., Nadim, F., Lacasse, S., Jardine, R.J., and Van Dijk B.F.J. (2017). Characteristics of unified databases for driven piles. *Proceedings, 8th International Conference of Offshore Site Investigation and Geomechanics*, vol. 1, 162-191. London, UK: Society for Underwater Technology.
Li, Z., Wang, X., Wang, H., and Liang, R. Y. (2016). Quantifying stratigraphic uncertainties by stochastic simulation techniques based on Markov random field. *Engineering geology*, 201, 106-122.
Lin, P.Y., Bathurst, R., and Liu. J.Y. (2017). Statistical evaluation of the FHWA simplified method and modifications for predicting soil nail loads. *ASCE Journal of Geotechnical and Geoenvironmental Engineering*, 143(3), 04016107.
Liu, H.F., Tang, L.S., Lin, P.Y., and Mei. G.X. (2018). Accuracy assessment of default and modified Federal Highway Administration (FHWA) simplified models for estimation of facing tensile forces of soil nail walls. *Canadian Geotechnical Journal*, 55(8), 1104-1115.
Liu, S., Zou, H., Cai, G., Bheemasetti, B.V., Puppala, A.J., and Lin, J. (2016). Multivariate correlation among resilient modulus and cone penetration test parameters of cohesive subgrade soils. *Engineering Geology*, 209, 128–142.
Long, J. (2013). *Improving agreement between static method and dynamic formula for driven cast-in-place piles in Wisconsin*. Report No. 0092-10-09, Wisconsin Department of Transportation.
Long, J. (2016). *Static pile load tests on driven piles into intermediate-geo materials*. Report No. WHRP 0092-12-08, Wisconsin Department of Transportation.
Long, J. and Anderson, A. (2014). *Improved design for driven piles based on a pile load test program in Illinois: phase 2*. Report No. FHWA-ICT-14-019, Illinois Department of Transportation.
Long, J., Hendrix, J., and Jaromin, D. (2009). *Comparison of five different methods for determining pile bearing capacities*. Report No. WisDOT 0092-07-04, Wisconsin Department of Transportation.
Long, M. (2001). Database for retaining wall and ground movements due to deep excavations. *ASCE Journal of Geotechnical and Geoenvironmental Engineering*, 127(3), 203-224.
Macciotta, R., Gräpel, C., Keegan, T., Duxbury, J., and Skirrow, R. (2019). Quantitative risk assessment of rock slope instabilities that threaten a highway near Canmore, AB, Canada: managing risk calculation uncertainty in practice. *Canadian Geotechnical Journal*, in press.
MacKay, D.J.C. (1995). Probable networks and plausible predictions - a review of practical Bayesian methods for supervised neural networks. *Network: Computation in Neural Systems*, 6(3), 469-505.
MacKay, D.J.C. (1998). Introduction to Monte Carlo methods. *Learning in Graphical Models*. MIT Press.
Marcos, M.C.M. and Chen, Y.J. (2013). Evaluation of lateral interpretation criteria for drilled shaft capacity in gravels. *Geotechnical and Geological Engineering*, 31(5), 1411-1420.
Mariethoz, G. and Caers, J. (2015). *Multiple-point geostatistics - stochastic modeling with training images*. John Wiley & Sons, West Sussex, UK.

Marsland, A. (1953). Model experiments to study the influence of seepage on the stability of a sheeted excavation in sand. *Géotechnique*, 3(6), 223-241.

Mayne, P. and Dasenbrock, D. 2018. Direct CPT method for 130 footings on sands. *Innovations in Geotechnical Engineering: Honoring Jean-Louis Briaud (GSP 299)*, 135-146. Reston, VA: ASCE.

McVay, M., Wasman, S., Huang, L., and Crawford, S. (2016). *Load and resistance factor design (LRFD) resistance factors for auger cast in place piles*. Florida Department of Transportation.

Mesri, G. and Huvaj, N. (2007). Shear strength mobilized in undrained failure of soft clay and silt deposits. *Advances in Measurement and Modeling of Soil Behavior (GSP 173)*, 1-22. Reston, VA: ASCE.

Miyata, Y. and Bathurst, R.J. (2012a). Analysis and calibration of default steel strip pullout models used in Japan. *Soils and Foundations*, 52(3), 481-497.

Miyata, Y. and Bathurst, R.J. (2012b). Reliability analysis of soil-geogrid pullout models in Japan. *Soils and Foundations*, 52(4), 620-633.

Miyata, Y. and Bathurst, R.J. (2015). Reliability analysis of geogrid installation damage test data in Japan. *Soils and Foundations*, 55(2), 393-403.

Miyata, Y. and Bathurst, R.J. (2019). Statistical assessment of load model accuracy for steel grid-reinforced soil walls. *Acta Geotechnica*, 14(1), 57-70.

Miyata, Y., Bathurst, R.J., and Konami, T. (2011). Evaluation of two anchor plate capacity models for MAW systems. *Soils and Foundations*, 51(5), 885-896.

Miyata, Y., Bathurst, R.J., and Allen, T.M. (2014). Reliability analysis of geogrid creep data in Japan. *Soils and Foundations*, 54(4), 608-620.

Miyata, Y., Yu, Y., and Bathurst, R.J. (2018). Calibration of soil-steel grid pullout models using a statistical approach. *ASCE Journal of Geotechnical and Geoenvironmental Engineering*, 144(2), 04017106.

Moghaddam, R.B., Jayawickrama, P.W. Lawson, W.D., Surles, J.G., and Seo, H. (2018). Texas cone penetrometer foundation design method: qualitative and quantitative assessment. *DFI Journal – The Journal of the Deep Foundations Institute*, 12(2), 69-80.

Moormann, C. (2004). Analysis of wall and ground movements due to deep excavations in soft soil based on a new worldwide database. *Soils and Foundations*, 44(1), 87-98.

Montoya-Noguera, S., Zhao, T., Hu, Y., Wang, Y. and Phoon, K.K. (2019). Simulation of non-stationary non-Gaussian random fields from sparse measurements using Bayesian compressive sampling and Karhunen-Loève expansion. *Structural Safety*, 79, 66-79.

Moshfeghi, S. and Eslami, A. (2018). Study on pile ultimate capacity criteria and CPT-based direct methods. *International Journal of Geotechnical Engineering*, 12(1), 28-39.

Motamed, R., Elfass, S., and Stanton, K. (2016). *LRFD resistance factor calibration for axially loaded drilled shafts in the Las Vegas Valley*. Report No. 515-13-803, Nevada Department of Transportation.

Mu, H.Q. and Yuen, K.V. (2019). Bayesian learning-based data analysis of uniaxial compressive strength of rock: relevance feature selection and prediction reliability assessment. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, in press.

Müller, R., Larsson, S., and Spross, J. (2014). Extended multivariate approach for uncertainty reduction in the assessment of undrained shear strength in clays. *Canadian Geotechnical Journal*, 51(3), 231-245.

Nanazawa, T., Kouno, T., Sakashita, G., and Oshiro, K. (2019). Development of partial factor design method on bearing capacity of pile foundations for Japanese specifications for highway bridges. *Georisk*, in press.

National Research Council (1995). *Probabilistic Methods in Geotechnical Engineering*. National Academies Press, Washington, DC.

Ng, C.W.W., Yau, T.L.Y., Li, J.H.M., and Tang, W.H. (2001). New failure load criterion for large diameter bored piles in weathered geomaterials. *ASCE Journal of Geotechnical and Geoenvironmental Engineering*, 127(6), 488-498.

Niazi, F.S. (2014). *Static axial pile foundation response using seismic piezocone data*. Ph.D. thesis, Georgia Institute of Technology, USA.

Okamura, M., Takemura, J., and Kimura, T. (1997). Centrifuge model test on bearing capacity and deformation of sand layer overlying clay. *Soils and Foundations*, 37(1), 73-88.

Ou, C.Y. and Liao, J.T. (1987). *Geotechnical Engineering Research Report*. GT96008, National Taiwan University of Science and Technology, Taipei.

Paikowsky, S., Canniff, M., Lesny, K., Kisse, A., Amatya, S., and Muganga, R. (2010). *LRFD design and construction of shallow foundations for highway bridge structures*. NCHRP Report 651. Washington, DC: Transportation Research Board.

Pati, Y.C., Rezaiifar, R. and Krishnaprasad, P.S. (1993). Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. *Proceedings, 27th Asilomar Conference on Signals, Systems and Computers*. Pacific Grove, California, USA.

Patra, C., Behara, R., Sivakugan, N., and Das, B. (2012a). Ultimate bearing capacity of shallow strip foundation under eccentrically inclined load, part I. *International Journal of Geotechnical Engineering*, 6(3), 343-352.

Patra, C., Behara, R., Sivakugan, N., and Das, B. (2012b). Ultimate bearing capacity of shallow strip foundation under eccentrically inclined load, part II. *International Journal of Geotechnical Engineering*, 6(4), 507-514.

Peck, R.B. (1969). Advantages and limitations of the observational method in applied soil mechanics. *Géotechnique*, 19(2), 171–187.

Phoon, K.K. (2017). Role of reliability calculations in geotechnical design. *Georisk*, 11(1), 4-21.

Phoon, K.K. (2018). Editorial for Special Collection on Probabilistic Site Characterization. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 4(4), 02018002.

Phoon, K.K. and Ching, J. (2017). Better correlations for geotechnical engineering. *A Decade of Geotechnical Advances, Geotechnical Society of Singapore (GeoSS)*, 73-102.

Phoon, K.K. and Tang, C. (2019). Characterization of geotechnical model uncertainty. *Georisk*, 13(2), 101-130.

Phoon, K.K., Huang, S.P., and Quek, S.T. (2002). Simulation of second-order processes using Karhunen–Loève expansion. *Computers and Structures*, 80(12), 1049–1060.

Phoon, K.K., Liu, S.L., and Chow, Y.K. (2009). Characterization of model uncertainties for cantilever walls in sand. *Journal of GeoEngineering*, 4(3), 75-85.

Phoon, K.K., Prakoso, W.A., Wang, Y. and Ching, J. (2016). Uncertainty representation of geotechnical design parameters. *Chapter 3, Reliability of Geotechnical Structures in ISO2394*, CRC Press/Balkema, 49-87.

Prästings, A., Spross, J., and Larsson, S. (2018). Characteristic values of geotechnical parameters in Eurocode 7. *Proceedings of the Institution of Civil Engineers: Geotechnical Engineering*, in press.

Qi, X.H., Li, D.Q., Phoon, K.K., Cao, Z.J., and Tang X.S. (2016). Simulation of geologic uncertainty using coupled Markov chain. *Engineering Geology*, 207, 129-140.

Rasmussen, C.E. and Williams, C.K. (2006). *Gaussian Processes for Machine Learning* (Vol. 1). Cambridge: MIT Press.

Rauser, J. and Tsai, C. (2016). Beneficial use of the Louisiana foundation load test database. *Transportation Research Board 95th Annual Meeting*, 1-17. Washington, DC: Transportation Research Board.

Reddy, S. and Stuedlein, A. (2017). Ultimate limit state reliability-based design of augered cast-in-place piles considering lower-bound capacities. *Canadian Geotechnical Journal*. 54(12), 1693-1703.

Roling, M., Sritharan, S., and Suleiman, M. (2011). *Development of LRFD procedures for bridge pile foundations in Iowa. Vol. 1: An electronic database for pile load tests (PILOT)*. Report No. IHRB Project TR-573, Iowa Department of Transportation.

Samtani, N.C. and Allen, T.M. (2018). *Implementation report – expanded database for service limit state calibration of immediate settlement of bridge foundation on soil*. Report No. FHWA-HIF-18-008. Washington, DC: Federal Highway Administration (FHWA).

Schuppener, B. and Heibaum, M. (2011). Reliability theory and safety in German geotechnical design. *Proceedings, 3rd International Symposium on Geotechnical Safety & Risk, Federal Waterways Engineering and Research Institute*, Germany, 527-536.

Smith, T., Banas, A., Gummer, M., and Jin, J. (2011). *Recalibration of the GRLWEAP LRFD resistance factor for Oregon DOT*. Report No. FHWA-OR-RD-11-08, Oregon Department of Transportation.

Spross, J., Olsson, L., and Stille, H. (2018). The Swedish Geotechnical Society's methodology for risk management: a tool for engineers in their everyday work. *Georisk*, 12(3), 183-189.

Stark, T., Long, J., Baghdady, A., and Osouli, A. (2017). *Modified standard penetration test-based drilled shaft design method for weak rocks (phase 2 study)*. Report No. FHWA-ICT-17-018, Illinois Department of Transportation.

Stuyts, B., Cathie, D., and Powell, T. (2016). Model uncertainty in uplift resistance calculations for sandy backfills. *Canadian Geotechnical Journal*, 53(11), 1831-1840.

Tang, C. and Phoon, K.K. (2016). Model uncertainty of cylindrical shear method for calculating the uplift capacity of helical anchors in clay. *Engineering Geology*, 207, 14-23.

Tang, C. and Phoon, K.K. (2017). Model uncertainty of Eurocode 7 approach for bearing capacity of circular footings on dense sand. *ASCE International Journal of Geomechanics*, 17(3), 04016069.

Tang, C. and Phoon, K.K. (2018a). Statistics of model factors and consideration in reliability-based design of axially loaded helical piles. *ASCE Journal of Geotechnical and Geoenvironmental Engineering*, 144(8), 04018050.

Tang, C. and Phoon, K.K. (2018b). Evaluation of model uncertainties in reliability-based design of steel H-piles in axial compression. *Canadian Geotechnical Journal*, 55(11), 1513-1532.

Tang, C. and Phoon, K.K. (2018c). Statistics of model factors in reliability-based design of axially loaded driven piles in sand. *Canadian Geotechnical Journal*, 55(11), 1592-1610.

Tang, C. and Phoon, K.K. (2018d). Characterization of model uncertainty in predicting axial resistance of piles driven into clay. *Canadian Geotechnical Journal*, in press.

Tang, C. and Phoon, K.K. (2019a). Evaluation of stress dependent methods for the punch-through capacity of foundations in clay with sand. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, in press.

Tang, C. and Phoon, K.K. (2019b). Statistical evaluation of model factors in reliability calibration of high displacement helical piles under axial loading. *Canadian Geotechnical Journal*, in press.

Tang, C., Phoon, K.K., and Chen, Y.-J. (2019). Statistical analyses of model factors in reliability-based limit state design of drilled shafts under axial loading. *ASCE Journal of Geotechnical and Geoenvironmental Engineering*, in press.

Teh, K.L. (2007). *Punch-through of spudcan foundation in sand overlying clay*. Ph.D. thesis, National University of Singapore, Singapore.

Terzaghi, K. and Peck, R.B. (1948). *Soil Mechanics in Engineering Practice*. John Wiley.

Tian, M., Li, D.Q., Cao, Z.J., Phoon, K.K., and Wang, Y. (2016). Bayesian identification of random field model using indirect test data. *Engineering Geology*, 210, 197-211.

Tipping, M.E. (2001). Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, 1, 211-244.

Travis, Q., Schmeeckle, M., and Sebert, D. (2011). Meta-analysis of 301 slope failure calculations. I: Database description. *ASCE Journal of Geotechnical and Geoenvironmental Engineering*, 137(5), 453-470.

Ullah, S.N. (2016). *Jackup Foundation Punch-Through in Clay with Interbedded Sand*. Ph.D. thesis, The University of Western Australia, Perth, Australia.

Wang, H. and Yeung, D.Y. (2016). Towards Bayesian deep learning: A framework and some existing methods. *IEEE Transactions on Knowledge and Data Engineering*, 28(12), 3395-3408.
Wang, H., Wellmann, J. F., Li, Z., Wang, X., and Liang, R. Y. (2017). A segmentation approach for stochastic geological modeling using hidden Markov random fields. *Mathematical Geosciences*, 49(2), 145-177.
Wang, H., Wang, X., Wellmann, J. F., and Liang, R. Y. (2018). Bayesian Stochastic Soil Modeling Framework Using Gaussian Markov Random Fields. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 4(2), 04018014.
Wang, H., Wang, X., Wellmann, F., and Liang, R.Y. (2019). A Bayesian unsupervised learning approach for identifying soil stratification using cone penetration data. *Canadian Geotechnical Journal*, in press.
Wang, J., Xu, Z., and Wang, W. (2010). Wall and ground movements due to deep excavations in Shanghai soft soils. *ASCE Journal of Geotechnical and Geoenvironmental Engineering*, 136(7), 985-994.
Wang, X., Li, Z., Wang, H., Rong, Q., and Liang, R. Y. (2016). Probabilistic analysis of shield-driven tunnel in multiple strata considering stratigraphic uncertainty. *Structural safety*, 62, 88-100.
Wang, X., Wang, H., Liang, R.Y., Zhu, H., and Di, H. (2018). A hidden Markov random field model based approach for probabilistic site characterization using multiple cone penetration test data. *Structural Safety*, 70, 128-138.
Wang, X., Wang, H., Liang, R. Y., and Liu, Y. (2019). A semi-supervised clustering-based approach for stratification identification using borehole and cone penetration test data. *Engineering Geology*, 248, 102-116.
Wang, Y. and Zhao, T. (2016). Interpretation of soil property profile from limited measurement data: a compressive sampling perspective. *Canadian Geotechnical Journal*, 53(9), 1547-1559.
Wang, Y. and Zhao, T. (2017). Statistical interpretation of soil property profiles from sparse data using Bayesian Compressive Sampling. *Géotechnique*, 67(6), 523-536.
Wang, Y., Akeju, O.V., and Zhao, T. (2017). Interpolation of spatially varying but sparsely measured geo-data: A comparative study. *Engineering Geology*, 231, 200-217.
Wang, Y., Huang, K., and Cao, Z. (2013). Probabilistic identification of underground soil stratification using cone penetration tests. *Canadian Geotechnical Journal*, 50(7), 766-776.
Wang, Y., Zhao, T., and Phoon, K.K. (2018). Direct simulation of random field samples from sparsely measured geotechnical data with consideration of uncertainty in interpretation. *Canadian Geotechnical Journal*, 55(6), 862-880.
Wang, Y., Zhao, T., and Cao, Z. (2019a). Bayesian perspective on ground property variability for geotechnical practice. *Proceedings, 7th International Symposium on Geotechnical Safety and Risk (ISGSR 2019)*, Taipei, Taiwan, 11–13 December 2019.
Wang, Y., Zhao, T., Hu, Y. and Phoon, K.K. (2019b). Simulation of random fields with trend from sparse measurements without de-trending. *ASCE Journal of Engineering Mechanics*, 145(2), 04018130.
White, D., Cheuk, C., and Bolton, M. (2008). The uplift resistance of pipes and plate anchors buried in sand. *Géotechnique*, 58(10), 771-779.
Wood, T., Jayawickrama, P., Surles, J., and Lawson, W. (2012a). *Pullout resistance of MSE reinforcements in backfills typically used in Texas: Vol. 2, test reports for MSE reinforcements in Type B (sandy) backfill*. Research Report: FHWA/TX-13/0-6493-3, Vol. 2, Texas Department of Transportation.
Wood, T., Jayawickrama, P., Surles, J., and Lawson, W. (2012b). *Pullout resistance of MSE reinforcements in backfills typically used in Texas: Vol. 3, test reports for MSE reinforcements in Type A (gravelly) backfill*. Research Report: FHWA/TX-13/0-6493-3, Vol. 3, Texas Department of Transportation.
Wu, S.H., Ching, J., and Ou, C.Y. (2013). Predicting wall displacements for excavations with cross walls in soft clay. *ASCE Journal of Geotechnical and Geoenvironmental Engineering*, 139(6), 914-927.
Wu, S.H., Ou, C.Y., and Ching, J. (2014). Calibration of model uncertainties in base heave stability for wide excavations in clay. *Soils and Foundations*, 54(6), 1159-1174.
Xiao, T., Li, D.Q., Cao, Z.J., and Zhang, L.M. (2018). CPT-based probabilistic characterization of three-dimensional spatial variability using MLE. *ASCE Journal of Geotechnical and Geoenvironmental Engineering*, 144(5), 04018023.
Yang, Z., Jardine, R., Guo, W., and Chow, F. (2016). *A comprehensive database of tests on axially loaded piles driven in sand*. Academic Press.
Yuan, J., Lin, P.Y., Huang, R., and Que, Y. (2019). Statistical evaluation and calibration of two methods for predicting nail loads of soil nail walls in China. *Computers and Geotechnics*, 108, 269-279.
Yuen, K.V. (2010). Recent developments of Bayesian model class selection and applications in civil engineering. *Structural Safety*, 32(5), 338–346
Yuen, K.V. and Mu, H.Q. (2012). A novel probabilistic method for robust parametric identification and outlier detection. *Probabilistic Engineering Mechanics*, 30, 48-59.
Yuen, K.V. and Ortiz, G.A. (2016). Bayesian nonparametric general regression. *International Journal for Uncertainty Quantification*, 6(3), 195-213.
Yuen, K.V. Ortiz, G.A., and Huang, K. (2016). Novel nonparametric modeling of seismic attenuation and directivity relationship. *Computer Methods in Applied Mechanics and Engineering*, 311, 537-555.
Yuen, K.V. and Ortiz, G.A. (2018). Multi-resolution Bayesian nonparametric general regression for structural model updating. *Structural Control and Health Monitoring*, 25(2), e2077.
Zhao, T. and Wang, Y. (2018). Simulation of cross-correlated random field samples from sparse measurements using Bayesian compressive sensing. *Mechanical Systems and Signal Processing*, 112, 384-400.
Zhao, T., Hu, Y. and Wang, Y. (2018). Statistical interpretation of spatially varying 2D geo-data from sparse measurements using Bayesian compressive sampling. *Engineering Geology*, 246, 162-175.
Zhang, L.M., Shek, L.M.P., Pang, H.W., and Pang, C.F. (2006). Knowledge-based design and construction of driven piles. *Proceedings of the Institution of Civil Engineers: Geotechnical Engineering*, 159 (3), 177-185.