13th Australia New Zealand Conference on Geomechanics – Acosta-Martínez & Lehane (Eds.)

Influence of distance-weighted averaging of site investigation samples on foundation performance

M. P. Crisp, M. B. Jaksa & Y.L Kuo
School of Civil, Environmental and Mining Engineering, University of Adelaide, Australia

ABSTRACT: Site investigations, for the purpose of determining the material properties of a variable subsurface soil, are an essential part of civil engineering projects. However, there is little research on how best to interpret site investigation data for a soil model. This paper investigates the potential benefit of weighting soil samples by their distance from pile foundations, using a variety of weighting and averaging techniques. The performance metric is total project cost, where construction, site investigation and failure costs are explicitly quantified through a virtual framework, facilitated through the generation of variable, single-layer virtual soils in a Monte Carlo analysis. It has been found that a saving of up to $1.8 million can be achieved by drilling in 9 locations as compared to one, despite the increased initial investment in soil testing.

1 INTRODUCTION

Site investigations are an essential part of civil engineering works, as they provide insight into the otherwise unknown material properties of subsurface soil profiles. Soils exhibit spatial variability, and so multiple locations must be investigated in order to accurately determine subsurface conditions. Despite this, it is not known how best to convert the aggregation of soil property data, from a site investigation, into the idealised soil models used in practice in the most optimal manner (Crisp et al. 2018a).

Unfortunately, insufficient site investigations, or inappropriate soil property idealisations, henceforth termed reduction methods, can lead to a variety of negative outcomes in engineering projects. Such outcomes include cost overruns and change orders (Boeckmann & Loehr 2016), construction delays (Jaksa 2000), foundation failure (Moh 2004) and foundation overdesign (Clayton 2001). In contrast, studies have shown that there can be considerable financial benefits by conducting investigations beyond the minimal scope (Crisp et al. 2018b; Goldsworthy 2006). Clearly, there is a need to formulate a site investigation optimisation guideline.

This study aims to determine the influence of the number of boreholes and the selection of reduction method on site investigation performance, with a focus on weighting the importance of samples by the distance from a foundation. These inverse distance methods (IDMs) reflect the tendency that soil properties, which are in close spatial proximity to each other, tend to be similar in value, and vice versa. This soil self-similarity with distance is a result of the processes that formed and continually modify the ground over time, and is referred to as autocorrelation. Using IDMs allows each individual foundation to be designed independently, in a separate soil model that more accurately reflects local soil conditions, as opposed to all foundations for a structure being designed from the same model. In theory, this should increase the probability of all foundations having the same settlement, thus decreasing the risk of structural damage through differential settlement, which is a key design constraint in the design of foundations for buildings.

The method used to determine site investigation performance is based on a framework described by Crisp et al. (2018a) and originally proposed by Jaksa et al. (2003). The framework utilises Monte Carlo simulation where, for any given realisation, a variable, 3D, single-layer virtual soil is generated. Complete knowledge of this soil allows for virtual site investigations to be undertaken, along with the corresponding foundation designs. The foundation can then be assessed for differential settlement, using linear-elastic finite element analysis (FEA) in the full virtual soil, which may result in structural damage. By assigning costs to the investigation, construction, and repair due to failure throughout the project, averaged across thousands of Monte Carlo realisations, it is possible to represent the quality of a site investigation by total project cost. The optimal investigation is therefore identified by minimising the total cost objective function. As this metric incorporates both economic and risk-based factors, it is considered an ideal objective function for practicing engineers.

The randomly-generated virtual soil profiles, or random fields, are volumes of soil properties represented by a 3D grid of discrete elements. As linear-elastic FEA is used, the required properties are Young's modulus (E) and Poisson's ratio (v). The fields are generated by local average subdivision (LAS) (Fenton & Vanmarcke 1990), which is a commonly-used algorithm in probabilistic research in geotechnical engineering (Fenton & Griffiths 2008). The soil properties within these random fields can be statistically described by three parameters; the mean, standard deviation, and the scale of fluctuation (SOF) (Vanmarcke 1983). The SOF is analogous to the autocorrelation distance mentioned above, and is defined as the distance over which soil properties exhibit strong similarity. In other words, high SOF values correspond to large pockets of similar material. Within this study, the standard deviation is normalised by the mean to produce the coefficient of variation parameter (COV), which is more useful as the results can be applied to any mean parameter value. The soil properties themselves are generated according to the lognormal distribution, which has been found to be a reasonable representation (Fenton & Griffiths 2008), and ensures that stiffness values are strictly non-negative.

Existing literature on the performance of IDMs is fairly limited, as site investigation performance has traditionally been difficult to quantify. Goldsworthy (2006); Goldsworthy et al. (2007) investigated the influence of various reduction techniques on the performance of site investigations for the design of pad foundations. These included weighted arithmetic averages of soil properties, where the weights are based on the inverse of the distance (ID) between the borehole and foundation, and the square of the inverse distance (I2). It was concluded that IDMs had relatively erratic performance with respect to the number of boreholes. This result was due to cases where borehole locations coincided with footing locations, resulting in the coincident boreholes having infinite weight, with the majority of boreholes being ignored, leading to a loss of information. Furthermore, the ID method resulted in the highest total project cost. It was suggested that IDM performance could be improved in cases where information from all boreholes is considered.

On the other hand, Goldsworthy et al. (2005) found that IDMs had the highest reliability of a range of reduction methods regarding the average design error of pad foundations. However, it should be noted that, contrary to the studies mentioned above, there were no sampling errors included in this analysis, which decreased the degree of realism. Alternative reduction methods examined include the standard arithmetic average (SA), geometric average (GA) and harmonic average (HA) in increasing order of conservatism. These are defined mathematically later in the paper. Use of the more conservative GA and HA techniques, which are low-value dominated, may be more accurate, when compared to the SA, for a variety of reasons. Firstly, it has been shown that soil settlement itself is low-value dominated, with less-stiff elements having a greater influence than the stiffer ones on overall response (Griffiths & Fenton 2009). Secondly, geometric averaging preserves the median of the lognormal distribution; the distribution used in the present study and several others (Fenton & Griffiths 2008). Thirdly, the soil below an infinitely-wide shallow foundation is represented perfectly by the harmonic average, assuming that soil is constant in the horizontal direction, varying only with depth (Fenton & Griffiths 2005). If the soil only varied horizontally, then the arithmetic average would be a perfect representation for the same foundation. As the geometric average lies between both cases, and that soil varies in both the vertical and horizontal directions, then this average could be considered the ideal reduction method. However, the infinitely-wide shallow foundation assumption is not necessarily applicable to a deep foundation of finite size. In terms of non-averaging reduction methods, Crisp et al. (2018b) investigated the technique of taking the 1st quartile of all sample values. It was found that a cost saving of up to $350,000 could be achieved by drilling 4 boreholes instead of one for a 4-columned building.

Currently, no study has examined potential benefits of using IDMs with site investigations with regards to pile design. Furthermore, there is discrepancy in the literature as to the benefits of IDMs, and if they are worth the additional effort over a simple average of soil properties. Finally, while there appears to be some benefit in using the more conservative GA and HA average techniques, no study has explored using weighted versions of these averages, as has been done with the SA in the form of the ID and I2 methods. The aims of this study are therefore:

1. To examine the potential benefit of distance-weighted reduction methods, both with and without inherent sampling errors.
2. To compare the SA, GA and HA reduction methods, both in terms of a global average, and by taking the minimum of the averages within each borehole.
3. To recommend an optimal reduction method and number of boreholes, in the context of pile design.

2 METHODOLOGY

2.1 *Overview*

The author refers readers to Crisp et al. (2018a) for details on the general methodology adopted in the present study, beyond that given in the introduction, as well as the verification of the methodology. Space restrictions limit the ability to provide such information in greater detail than that given. In summary, the results presented here were generated using 8,000 realisations of random soils, along with 3D linear elastic FEA to determine pile settlement. The resulting database was generated using the Phoenix supercomputer (University of Adelaide 2018), which reduced 30 years' worth of simulation to a matter of weeks. Fortunately, the database is generalised in terms of possible soil and structural configurations, allowing for site investigations to be optimised for a wide variety of situations through dynamic post-processing, facilitating the present results with minimal additional computational time.

The costs used in this study include those associated with the structure ($6,157,750), the site investigation ($77/m), and foundation ($200/m). Furthermore, failure costs have been derived from associations between various degrees of structural damage and differential settlement (Day 1999), where damage has associated repair costs (Rawlinsons 2016). The resulting failure cost (C) in terms of differential settlement (δ) is given by the linear equation $C = 1.024 \times 10^9\ \delta - 3.056 \times 10^6$. This function is bounded by a minimum of $0 at 0.0030 m/m, and a maximum of $6,534,400 at 0.0094 m/m; corresponding to negligible damage and demolishing/rebuilding respectively. As mentioned above, details of the above are given by Crisp et al. (2018a).

2.2 *Details of site and structure*

The virtual sites analysed in this study are 60 × 60 × 40 m in the x, y, z (depth) dimensions respectively, and are comprised of cubic elements of dimension 0.25 m. Further information on the process of generating the soils, along with alternative techniques of generating random fields, are given by Fenton & Griffiths (2008). Of the two settlement parameters, v has been set constant to 0.3, as this value is widely-found in nature. On the other hand, E is spatially variable with a mean of 40 MPa, with a COV of 80%, which is considered high. Contrary to most studies, COV is not a variable in this analysis, as the relative performance of inverse-distance weighting would be consistent. On the other hand, the degree of soil self-similarity with distance may have a significant impact on the results. As such, three values of SOF are considered: 1 m (low), 8 m (medium) and 24 m (high). The soils are isotropic, meaning the SOF is constant in all directions.

The adopted structure is 20 × 20 m in plan and consists of 6 storeys, which are supported by 9 piles, arranged in a grid pattern, one beneath each column. As each floor is subject to 8 kPa of combined dead and live load, the total structure weighs 19,200 kN, with no load factoring applied. The piles are spaced at 10 m intervals and are 0.5 m in diameter. The corner, edge and internal piles have average lengths of approximately 1.5 m, 4 m and 8 m respectively, and have been designed to an allowable settlement of 25 mm.

2.3 *Site investigation and reduction methods*

The site investigation consists of 5 sets of boreholes, 1, 4, 9, 16 and 25, arranged in a regular grid pattern over the building footprint. Testing has been undertaken using the standard penetration test (SPT) at 1.5 m intervals with depth, both with and without random errors applied. The errors are unit-mean, lognormally-distributed variables, with variances of 25% bias per borehole, 20% random error per sample, and 40% parameter transformation error, as derived by Goldsworthy (2006). These errors seek to model the uncertainty associated with the SPT. The boreholes were drilled to a depth of 20 m, with a total of 17 samples in each.

In terms of the reduction methods, 3 averages have been used; the SA, GA and HA. In addition, 4 interpretations of each average are considered; the simple global average, an inverse-distance weighting, an inverse-distance-squared weighting, and a minimum method, where the samples within each borehole are averaged with the resulting worst case being adopted. This results in 12 types of reduction methods in combination with two tests. The general equations for the reduction methods for n samples are given in Table 1. Here, x_i refers to an arbitrary sample at some distance from a foundation, and w_i is that sample's weight; either unity for the simple average, the inverse of the distance, or the inverse of the distance squared.

Table 1. Generalised weighted equations for the arithmetic, geometric and harmonic averages.

Average type	Equation
Arithmetic	$\left(\frac{\sum_{i=1}^{n} w_i\, x_i}{\sum_{i=1}^{n} w_i}\right)$
Geometric	$\left(\frac{\sum_{i=1}^{n} w_i\, ln(x_i)}{\sum_{i=1}^{n} w_i}\right)^e$
Harmonic	$\left(\frac{\sum_{i=1}^{n} w_i\, x_i^{-1}}{\sum_{i=1}^{n} w_i}\right)^{-1}$

3 RESULTS

The results of the analysis are shown in Figure 1, with low, medium and high SOF soils shown across rows 1–3, and the arithmetic, geometric and harmonic averages shown across columns 1–3. Each subplot contains costs for the simple average, borehole minimum, inverse-distance (IDM 1) and inverse-distance-squared (IDM 2) interpretations of borehole data. The borehole costs are shown with and without errors, referred to as the SPT (dashed lines) and discrete tests (solid lines), respectively.

3.1 *Influence of testing error*

Upon inspection of Figure 1, it is clear that the IDMs used with the SPT show erratic performance across the various numbers of boreholes, as noted by Goldsworthy et al. (2007). In particular, the set of 9 and 25 boreholes typically have inferior performance compared to adjacent values, most prominently with the GA and HA. This is because the foundation consists of 9 piles. As such, these two borehole sets are the only cases where all piles coincide in location

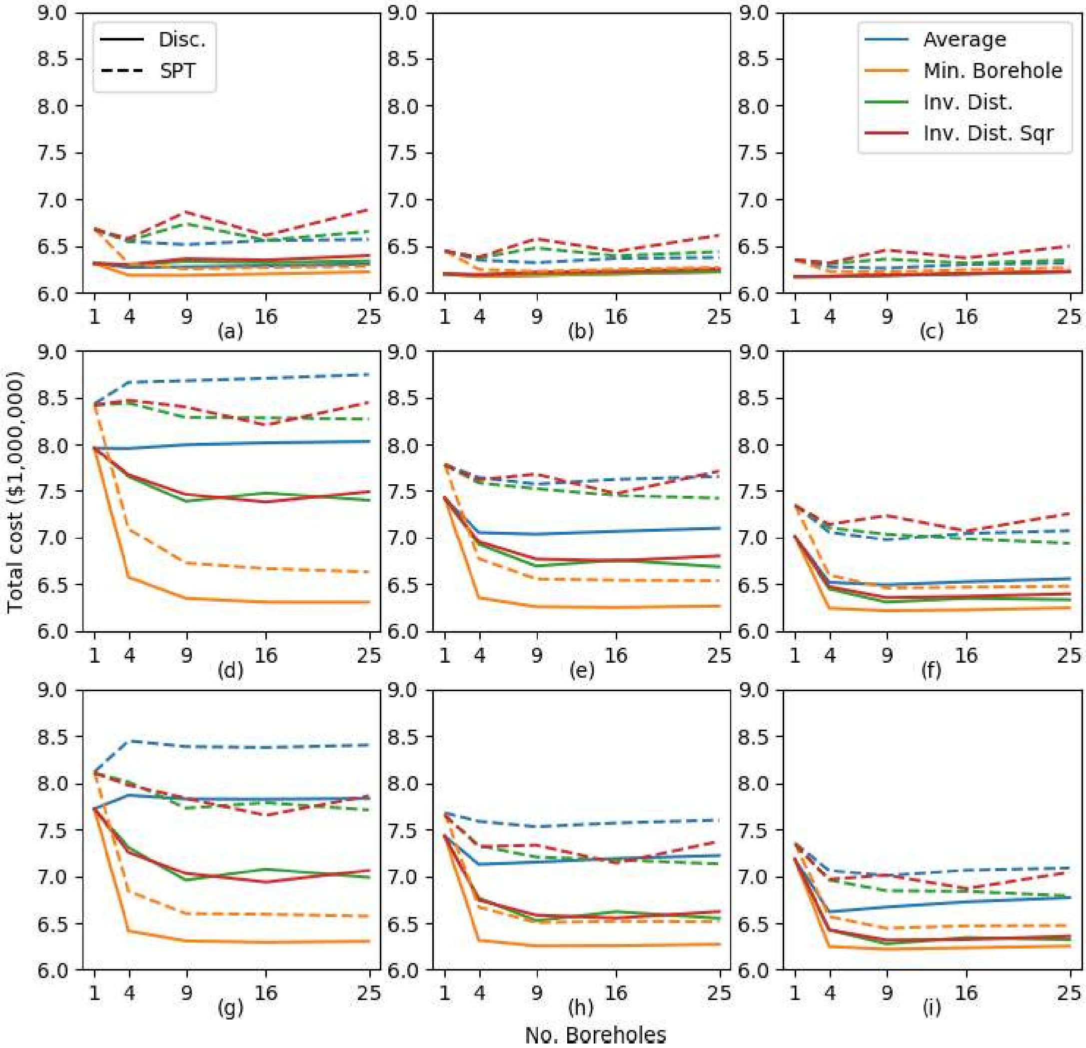

Figure 1. Comparison of different interpretations of sample averages, with and without testing errors, for (a) arithmetic average with low SOF; (b) geometric average with low SOF; (c) harmonic average with low SOF; (d) arithmetic average with medium SOF; (e) geometric average with medium SOF; (f) harmonic average with medium SOF; (g) arithmetic average with high SOF; (h) geometric average with high SOF; and (i) harmonic average with high SOF.

with a borehole. Due to the exponentially-decaying weighting of samples with distance, this effectively causes all boreholes, besides the one coincident with a pile, to be ignored.

It can also be seen that this erratic behaviour is not present, or at least greatly diminished, for the discrete test type. Rather, the site investigation performance generally improves as the number of boreholes increases. This leads to the conclusion that a single borehole located at a foundation is a good representation of the soil around that foundation, in the absence of errors. However, when errors are present, as found in practice, a single borehole is no longer sufficient. In the case of errors, additional boreholes must be conducted so that the increased number of samples can compensate for the overall test inaccuracy.

Comparing the SPT and discrete tests, the former performs consistently poorer across all cases than the latter, due to the presence of errors. The cost increase due to errors can be as high as $700,000 as seen in Figure 1d. However, such errors are unavoidable in practice, and so subsequent discussion will focus primarily on the SPT, unless stated otherwise.

3.2 *Influence of inverse-distance weighting*

Inspecting the different SOF values, it can be seen that the IDMs consistently exhibit the worst performance when used in low SOF soils, across all averages. This is because the fluctuation of soil properties is rapid enough such that the soils appear relatively uniform at a macro scale. In this case, since the soil is generally similar at all locations, there is no advantage to any kind of distance weighting. In contrast, this weighting is detrimental, as it makes sense to weight all samples equally to help overcome testing errors, as discussed in the previous section.

In the case of medium and high SOFs, IDMs exhibit consistently better performance than the global averages for the SA and, to a lesser extent, the more conservative GA. This benefit seems to increase as the SOF increases. The cost savings over the average can be as high as $700,000, as seen in Figure 1g, although they can also be negligible as seen in Figure 1f. Furthermore, the difference between the IDMs and global average decrease as the averaging type becomes more conservative, due to higher design redundancy for individual piles. In other words, inverse-distance weighting is more beneficial for less conservative reduction methods.

The increased benefit in higher SOF soils is logical, as a pile is more likely to be located in a pocket of consistent material, so it makes sense to weight that pocket more heavily. Theoretically, as the SOF increases to infinity, the benefit of distance weighting would begin to diminish due to the soil appearing to be uniform, as described in the case of low SOF. However, as stated previously, the benefit of IDMs does not diminish as SOF increases. Therefore, it is unlikely that soils found in nature would have a high enough SOF for this diminishing effect to occur.

Comparing the IDM 1 and IDM 2 methods across all cases, minor benefit can be obtained by using the former, with a cost benefit of up to $300,000 as seen in Figure 1f. However, they both perform similarly well overall as they give significantly higher weighting to the closest borehole to a foundation.

3.3 *Influence of averaging type*

Regarding the choice of averaging type, site investigation performance generally appears to increase from SA to GA, to HA. In other words, the performance increases as the average becomes more conservative, at least in the case of soils with a medium or high SOF.

The largest discrepancy between the three averages is the SA, where performance actually decreases with additional sampling when using global averaging, which is counter-intuitive. Upon closer inspection, it appears that the average SA investigation results in foundation failure. Therefore, as the number of boreholes increases and the standard deviation of the reduced values decreases, the proportion of safe designs decreases. However, this result may change should a heavier building be used; resulting in a longer mean pile length. There also appears to be minor, if any, improvement with additional sampling when using the GA and HA global averaging. This latter point is due to the presence of testing errors, as there is clear improvement from 1–4 boreholes in the case of the discrete test.

On the other hand, there is no discernible difference between the performance of the three averages when using the borehole minimum method. This consistency suggests that the minimising component of the method is the dominant factor of its strong performance, as opposed to the averaging component within each borehole.

In the case of soils with low SOF, it could be argued that a single borehole is sufficient for all reduction methods except the borehole minimum where four is recommended. Rather than this being a reflection of the investigation quality, it is instead indicative of the foundation performance. As mentioned previously, these soils appear largely uniform at a macro scale. As such, the foundation is unlikely to fail through differential settlement, meaning the failure cost is largely independent of the investigation scale.

3.4 *The minimum borehole method*

An interesting observation is that the minimum borehole method provides the best site investigation performance in terms of lowest total cost, across all cases. This improvement is due to the minimisation of failure costs. The result is surprising, as speculation would have suggested that differential settlement would be minimised by ensuring that each pile settles by the same amount. Theoretically, this consistent settlement would be achieved by an IDM, which considers each pile individually. Instead, each pile set (corner, edge, internal) is given a consistent length according to the worst-case borehole. By increasing the pile lengths to the same amount, there is a risk of increasing differential settlement, should the soil properties at each pile be significantly different. However, it appears that the added conservativism of having longer piles, with reduced total settlement, has overcome this risk quite convincingly.

It is also worth noting that there does not appear to be an optimal number of boreholes with the minimum reduction method, in terms of a clear local minimum cost. In other words, the cost asymptotes as the number of boreholes increases. This suggests that the improved reliability attributed to this reduction method generally compensates for the increased cost of additional sampling. As such, it can be concluded that the borehole minimum method has a good balance of conservatism, which maximises reliability without leading to excessively over-designed foundations. Since this strong performance is seen in soils of all SOFs, this method can be recommended for universal practice over the other methods examined in this study.

Generally speaking, in the case of medium and high SOFs, the total cost has largely plateaued at 9 boreholes. The cost savings by conducting 9 boreholes over one and four boreholes can be as high as $1.8 million and $400,000 respectively, as seen in Figure 1d for the SPT. The exception to this is in soils with a low SOF, as discussed in the previous section.

A potential limitation of using the borehole minimum method is that its performance may depend on the depth of the boreholes, i.e. the number of samples within each borehole. This is because the previously-discussed desirable conservative balance relies on the elimination of excessively-weak samples through the averaging process. If the number of samples in a borehole is small, then this low-value removal may not take place to a sufficient degree, causing the method to be overly-conservative. This is particularly the case with the geometric and harmonic averages due to their low-value weighted nature. As an extreme example, a single zero-valued sample would cause the global average to be zero. Furthermore, the relative improvement of the borehole minimum method over the other methods may depend on the soil COV, which has not been assessed in the present study. Nevertheless, the potential savings are remarkable.

4 CONCLUSION

It has been found that, while there is some apparent benefit to using inverse-distance weighting with site investigations in the majority of soils, the benefit is inconsistent. For example, a single borehole is largely sufficient for soils with a low scale of fluctuation. However, the savings over a global average can be as high as $700,000 when boreholes do not coincide in location with all foundations. When boreholes do coincide, the investigation performance greatly decreases, implying that any given foundation should not be designed from a single borehole in soils with medium and high scales of fluctuation, as additional boreholes are required to overcome errors. Inverse distance methods provide the greater benefits over the global average for less conservative averages.

Similarly, there is no noteworthy improvement with additional sampling when using the global averaging techniques, largely due to inherent errors. In particular, the arithmetic average performs significantly poorer with additional samples. As such, this average should not be used in practice. In general, the more conservative averages performed better.

Of the reduction methods assessed in this study, the borehole minimum technique consistently yielded the best performance, regardless of the choice of averaging method. This led to a saving of up to $1.8 million with the use of 9 boreholes over one, despite the higher initial investment. However, this method has not been assessed for different soil coefficients of variation, or with different borehole depths, where it is possible that performance may vary. Furthermore, it is not currently known whether this method yields the lowest total cost compared to methods not examined here.

It is noted that this analysis has been conducted in a variable, single-layer soil profile. As such, recommended numbers of boreholes given here should be taken as a minimum, as additional boreholes are likely needed to delineate the complex boundaries of multi-layered soils which have not been assessed here. Furthermore, overcoming testing errors to adequately represent soils properties within each layer would also require additional samples. Finally, it is likely that the optimal investigation is related to the structural configuration, a variable which was not considered in the present study due to use of a single building size. As such, these situations should be considered in further work.

REFERENCES

Boeckmann, A. Z., & Loehr, J. E. (2016). *Influence of Geotechnical Investigation and Subsurface Conditions on Claims, Change Orders, and Overruns*.

Clayton, C. (2001). Managing geotechnical risk: time for change? *Proceedings of the Institution of Civil Engineers-Geotechnical Engineering, 149*(1), 3-11.

Crisp, M. P., Jaksa, M. B., & Kuo, Y. L. (2018a). *Framework for the Optimisation of Site Investigations for Pile Designs in Complex Multi-Layered Soil.* Adelaide, Australia. DOI: 10.13140/RG.2.2.23536.71685

Crisp, M. P., Jaksa, M. B., & Kuo, Y. L. (2018b). *Influence of Site Investigation Borehole Pattern and Area on Pile Foundation Performance*. Paper presented at the 12th ANZ Young Geotechnical Professionals Conference, Hobart.

Day, R. W. (1999). *Forensic geotechnical and foundation engineering*: McGraw-Hill New York.

Fenton, G. A., & Griffiths, D. V. (2005). Three-dimensional probabilistic foundation settlement. *Journal of Geotechnical and Geoenvironmental Engineering, 131*(2), 232-239.

Fenton, G. A., & Griffiths, D. V. (2008). *Risk assessment in geotechnical engineering*. Hoboken: Wiley.

Fenton, G. A., & Vanmarcke, E. H. (1990). Simulation of random fields via local average subdivision. *Journal of Engineering Mechanics, 116*(8), 1733-1749.

Goldsworthy, J. S. (2006). *Quantifying the risk of geotechnical site investigations.*

Goldsworthy, J. S., Jaksa, M. B., Fenton, G. A., Griffiths, D. V., Kaggwa, W. S., & Poulos, H. G. (2007). *Measuring the risk of geotechnical site investigations.* Paper presented at the Proc., Geo-Denver 2007.

Goldsworthy, J. S., Jaksa, M. B., Kaggwa, G. S., Fenton, G. A., Griffiths, D. V., & Poulos, H. G. (2005). Reliability of site investigations using different reduction techniques for foundation design. *9th International Conference on Structural Safety and Reliability*, 901-908.

Griffiths, D. V., & Fenton, G. A. (2009). Probabilistic settlement analysis by stochastic and random finite-element methods. *Journal of Geotechnical and Geoenvironmental Engineering, 135*(11), 1629-1637.

Jaksa, M. B. (2000). Geotechnical risk and inadequate site investigations: a case study. *Australian Geomechanics, 35*(2), 39-46.

Jaksa, M. B., Kaggwa, W. S., Fenton, G. A., & Poulos, H. G. (2003). *A framework for quantifying the reliability of geotechnical investigations.* Paper presented at the 9th International Conference on the Application of Statistics and Probability in Civil Engineering.

Moh, Z. C. (2004). *Site investigation and geotechnical failures.* Paper presented at the Proceeding of International Conference on Structural and Foundation Failures.

Rawlinsons, A. (2016). Australian Construction Handbook (34 ed., pp. 1005). Perth, Australia: Rawlhouse Publishing Pty. Ltd.

University of Adelaide. (2018). Phoenix High Performance Computing Technical Information. Retrieved from https://www.adelaide.edu.au/phoenix/training/technical/

Vanmarcke, E. H. (1977). Reliability of earth slopes. *Journal of the Soil Mechanics and Foundations Division, 103*(11), 1247-1265.

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