# Supporting the Observational Approach in Construction through Bayesian Analysis

C.W. Boon[1] and L.H. Ooi[1]
[1]*MMC-Gamuda KVMRT (T) Sdn Bhd*
*E-mail*: cwboon@kvmrt-ug.com.my

**ABSTRACT:** The validation of design assumptions, as construction works are being carried out, is a vital component in geotechnical engineering. The feedback can be in terms of either the prescribed quality control tests, instrumentation and monitoring, or other site observations. As most of these can be evaluated through systematic logic, the use of data science methods or machine learning procedures is rarely necessary. There are, however, exceptions especially for cases in which: (i) there are several possible causes to the problems which are hard to pinpoint precisely, (ii) the quantum of data is overwhelming, and (iii) there is scatter in the of observed outcomes. Where these features are encountered, it is generally more efficient to process the data using a computer. This paper presents a possible way of interpreting the feedback obtained through observations in construction, via Bayesian programming, which is one of the many methods in machine learning. A case history discussing the performance of ground anchors in a deep excavation project is discussed.

**KEYWORDS:** Observational approach, Bayesian, Probability, Ground anchors, Artificial intelligence.

## 1. INTRODUCTION

The *observational approach* is common vocabulary to geotechnical engineers. Crudely, it suggests that the engineer, through his present observations of the ground response, can make an educated judgment on the next course of action. Conventional tools supporting the *observational approach* are (i) quality control tests and (ii) instrumentation and monitoring. In the specific contexts of *observational approach*, quality control tests saves the engineers time in ruling out the unknowns (to some extent and not completely), and are provisioned in codes of practice and international standards, in which typical recommended testing quantities are normally specified and may vary with the ground variability or the availability of pre-existing information within the project vicinity. An example of a quality control test is the on-site acceptance test for ground anchors. The performance of a single element is examined at the state after installation, to ensure that they are compatible with the engineer's assumptions. Where there are indications where the performance is less than prescribed, corrective measures can be carried out before the structure is loaded fully. On the other hand, instrumentation and monitoring provides an additional layer of validation, to capture deviations from theoretical expectations, during the actual loading. In the discussion here, instrumentation and monitoring includes observations made during construction through visual inspection. Where the performance is better or worse than predicted, the design models may be revised and the required resources for construction may be adjusted in a corresponding manner.

The discussion above bears very close resemblance to the definition of Machine Learning provided by Tom Mitchell (1998), where a computer is said to *learn* from experience *E* with respect to some task *T* and some performance measure *A*, if its performance on *T*, as measured by *A*, improves with experience *E*. It is possible for machine learning to be applied by gathering the experiences from many projects. It is believed, however, that the outcome will fit with typical "rules of thumbs" in engineering practice. As the subject of study in this paper is focussed on the *observational approach*, this definition has to be improvised for the case of a single construction project, so that "experience *E"* is replaced with random variables, i.e. "observations *O*". However, the duration over which the performance is being measured may result in very limited data depending on the constructed element of interest.

Nonetheless, in an environment with large data and a scatter of observations, computers are more capable than humans to process the information objectively.

## 2. BACKGROUND OF BAYESIAN ANALYSIS

In Bayesian statistics, where a prior information on an observation, *O*, is available, the conditional probability of performance *A*, given *O*, is expressed as:

$$P(A|O) = \frac{P(A \cap O)}{P(O)} \tag{1}$$

where *P*( ) denotes probability, and "∩" denotes conjunction. The variables on the right of "|" are variables with known values, and the variables on the left of "|" are being probed given the known variables.

Bayesian statistics have been adopted in various contexts in geotechnical engineering. The main distinction in Bayesian statistics as against the frequentist approach is that Bayesian statistics uses probability theory formally to acknowledge that the probability distribution is conditioned on the observed data, which is finite and may be potentially affected by noise or other data class which may not have been identified for sampling. For example, Houlsby & Houlsby (2009) applied Bayesian statistics to find the best fit of the design strength profiles to measured undrained strength data, given the random variables consisting of the observed data sets and the number of soil layers. In another example, Bayesian updating was used to update the reliability of pile design with load test data (Kay, 1978; Zhang, 2004; Huang, 2016). These two examples attempt to best-fit observations for design purposes, and incorporate observations into design respectively.

In this paper, Bayesian statistics is being used in a different context. Several variables, which had been identified as potentially "important", were studied to establish whether each variable had a strong correlation to the performance outcome. After that, a predictive question is being put forward, by reversing the conditional probabilities. A well-known example used in Bessière et al. (2013) is to predict spam e-mails by gathering *P*(word | spam=true) from compiled e-mail records. The probabilities *P*(spam=true| word) are then established to evaluate whether an incoming e-mail is spam. These probabilities are not the same and can be appreciated more intuitively via the example that the probability of a geotechnical engineer is taller than six feet, i.e. *P*(height>6ft | geotechnical engineer), and the probability that a person, who is six feet tall, is a geotechnical engineer, i.e. *P*(geotechnical engineer | height > 6 ft), are different. The paper shows how Bayesian statistics can have practical application in the context of interpreting wide scatter of information and decision-making.

For cases where there are multiple variables (or observations), i.e. $O_i$ to $O_N$, the computation of past data is time consuming, and methods of simplifications, for instance by assuming that these variables are conditionally independent, have been proposed in the Bayesian programming literature when handling large data. The mathematical details are not included here in this paper, and interested readers are referred to Bessière et al. (2013). Where the variables are not independent, the joint probabilities need to be constructed from the compiled measured data where two variables are true concurrently (Bessière et al., 2013). While the aforementioned is the recommended approach, an example is studied in this paper to provide further insight into the likely error of assuming independence to estimate the performance for non-independent variables.

## 3. APPLICATION OF BAYESIAN ANALYSES

A case history involving the troubleshooting of ground anchor performance is discussed in this paper. It is shown how Bayesian analysis and Bayesian programming could be adopted to support the observational approach.

The case site is a 28 m deep excavation located in the Kuala Lumpur Limestone Formation. The limestone at the site contains cave features, and a high degree of jointing at some areas (see Figure 1). The alluvium soil overburden was retained using secant bored piles (SBPs) with rock socketing of 2.5-5.5m at the secondary reinforced piles (Figure 1), and restrained using ground anchors together with a rock bolt at the SBP toe. The ground anchors are designed with rock socketing, and the strands are greased within a PVC sleeve, looping around the anchor end forming a U-loop (Figure 2). The rockhead varies around the site between approximately 15-28 m below ground level (bgl). A bottom-up construction sequence was adopted. The ground anchors were being installed as the excavation progressed deeper. During excavation and installation of the 3rd row of ground anchors, there were observations that some of the strands in the 1st and 2nd row of ground anchors had loosened. For the loosened anchors which were monitored using load cells, some registered load losses (Figure 3 (a) and (b)), whereas some did not register significant load losses implying that the loads were distributed to the neighbouring strands in the same anchor which had not loosened (Figure 3 (c)).

Following this incident, corrective procedures were implemented:

- To verify if the loosening was due to structural breakage, the loosened strand was restrained at one end, while the other end from the same loop was pulled. It was identified that the strands were structurally sound and there were no breakages;
- The past stressing records were re-visited and reviewed. It was found that the original ground anchor on-site acceptance tests of the loosened ground anchors had met the prescribed requirements (British Standard, 1989). Re-stressing and subsequent lift off tests were carried out on the loosened anchors, and these tests also met the performance requirements;
- More load cells and optical prisms were installed;
- Where load cell readings breached the prescribed working loads, lift-off tests were carried out. Load cells with odd number of strain gauges were changed to even number of strain gauges to avoid eccentric effects which could affect the readings, based on the findings from other project sites (Boon et al., 2015);
- Settlement markers were installed on the secant bored piles to measure if there were settlements on the piles which could potentially lead to relaxation. This hypothesis was ruled out after a few weeks of monitoring;
- Regular visual inspection was carried out to spot for loosening of wedges.

The initial prevailing *hypothesis* was that the root cause of the problem was the lack of fit between the wedges and the holes in the anchor block. The existing wedges were replaced immediately with larger and better fitting wedges, but loosening of some anchors occurred again, while the installation of the deeper 5th row of anchors commenced. It was then *hypothesised* that there was a possibility that the anchor block and wedges were again not fully compatible as there were differences of 1-2° between the tapering of the wedges and the inclinations of the holes in the anchor blocks.

(a)

(b)

Figure 1 Ground anchors in the limestone formation: (a) karst features with caves, (b) localised area which are heavily jointed

(a)

(b)

(c)

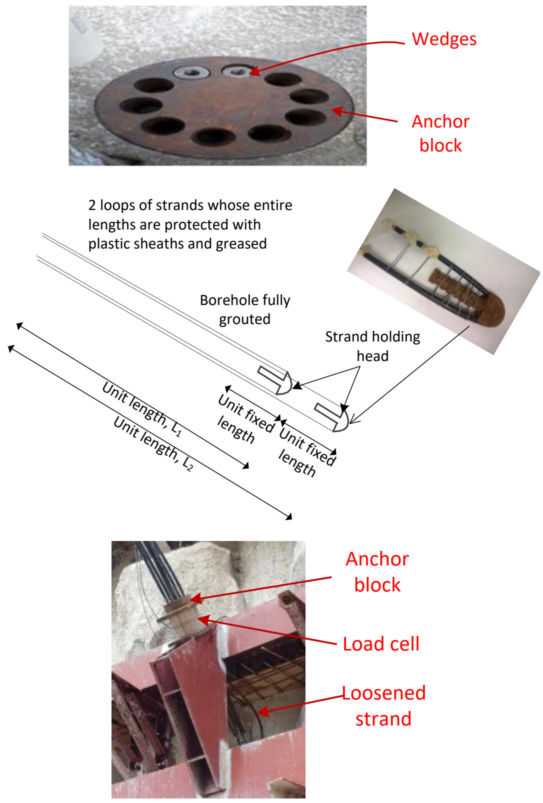


Figure 2 Ground anchor system adopted at the site: (a) image of anchor block and wedges, (b) U-loop system with U-turn end, (c) load cell monitoring where slack could be visually observed in one of the strands due to loosening

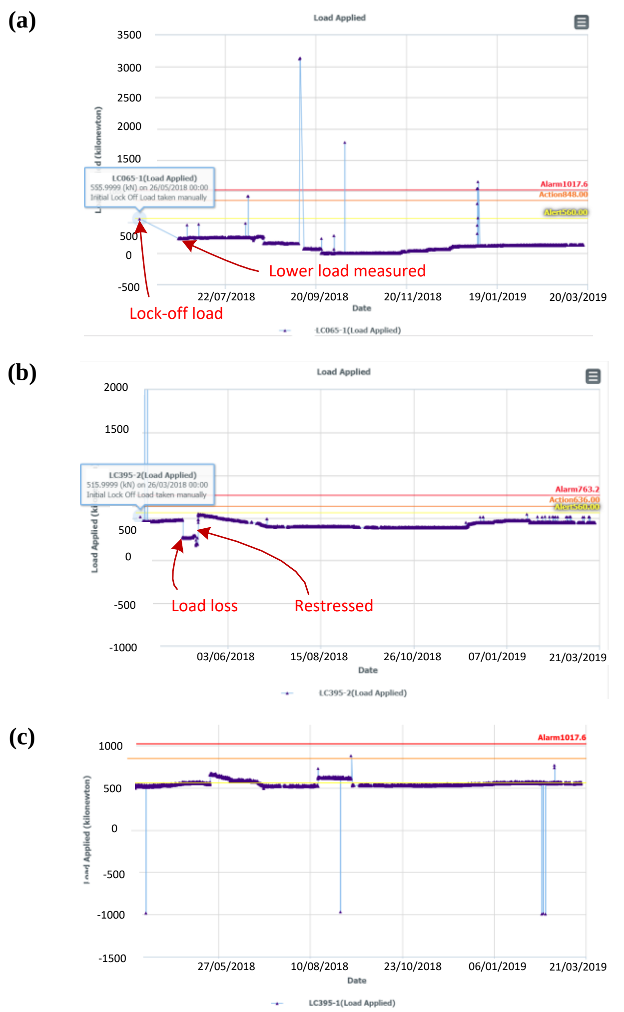


Figure 3 Loosened anchors with load cell measurements: (a) first measurement since lock-off showed loss of loads, (b) abrupt load loss measured, (c) no obvious signs of load loss, even though wedge has loosened. Spikes are noise in the data

At the same time, there were parallel views that the bonding at the rock socket was compromised, and that the slippage of the wedges would unlikely lead to full relaxation of the tensioned strands.

The strategy based on elimination of hypothesis one-at-a-time was difficult to implement, because the ground had to be excavated for the anchors to provide feedback as to whether or not the remedial measure was the correct one. It was obvious that several remedial proposals had to be explored concurrently, as it would be more onerous for the anchors with deeper excavation.

The study discussed in this paper was motivated by the need to identify what was the root cause quickly.

### 3.1. Identification of important parameters through *learning* and checking for *independence*

It was perceived that if the ground anchors had loosened *solely* due to the mechanical compatibility of wedges and anchor blocks, the outcome of loosening would be *independent* of other reasons. This was an important basis for the subsequent interpretation.

In a full machine learning environment, a wealth of variables may be provided to the computer, so that the statistically significant variables which are common among the observations could be identified. For the observation of anchor loosening, the following variables were pre-determined, namely:

- the grout wastage during installation;
- whether or not any large cavities were encountered during drilling which could potentially lead to grout loss;
- length of strands affecting the magnitude of loss in tensioning due to wedge slippage; and
- rockhead depth which may give an indication of past karst action.

From the site records, we compiled the probability of the variables above given that the anchors had loosened, namely:

i. *P*(grout_wastage|loosened)
ii. *P*(cavitiy_size_encountered|loosened)
iii. *P*(length|loosened)
iv. *P*(rockhead in relation to benchmark|loosened)

The results are shown in Figure 4. While there are certain trends in the data, it is important to caution that a high conditional probability does not immediately suggest that there is a high causal link to loosening. Instead, the results have to be studied holistically in terms of the overall probability. The performance *A* is said to be *independent* of the variable (or observation) *O* if:

$$P(A|O) = P(A) \tag{2}$$

The overall probability and conditional probability are overlaid for grout wastage and cavity size in Figure 5. The results show that the relative weight of anchors with higher grout wastage increased when considering only the anchors which had loosened, by comparison to the case in which all the ground anchors were considered (Figure 5 (a)). This is likewise for anchors with large cavities encountered (Figure 5 (b)). The results suggest that the outcome of loosening is statistically not independent of grout wastage and the presence of cavities.

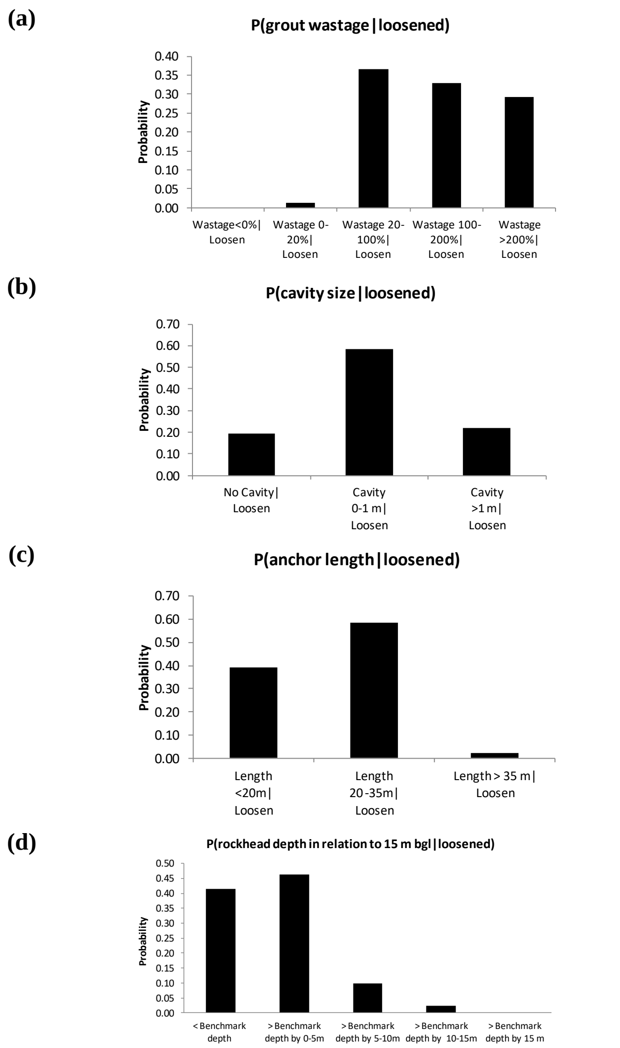


Figure 4 Statistics of (a) grout wastage, (b) cavity size, (c) anchor length, (d) rockhead (using 15 m bgl as the benchmark depth), given that the anchor had loosened


Figure 5 Comparison between overall probability and conditional probability to check for independence

### 3.2. Bayesian analyses of ground anchor loosening

Once the important variables are identified, the conditional probabilities have to be converted for the use of future prediction:

i. *P*(loosen|grout_wastage)
ii. *P*(loosen|cavitiy_size_encountered)
iii. *P*(loosen|length)
iv. *P*(loosen|rockhead)

This step is necessary because the latter indirectly takes into account the statistics contributed by the anchors which did not experience loosening. An example is provided to explain the significance of this step: Anchors which had loosened are all shorter than 50 m. Therefore, *P*(anchor_length<50m|loosened) = 100%. However, *P*(loosen|anchor_length<50m) is not 100%, because there were some anchors shorter than 50m which did not loosen.

The probabilities of *P*(loosen|grout_wastage) and other observations can be calculated directly from the compiled data using rigorous definitions of conventional Bayesian analyses, or through standard procedures set out in Bayesian programming (Bessière et al., 2013). The results are shown in Figure 6. The latter is commonly used in the *artificial intelligence* because answers to complex questions can be computed easily from known probabilities derived from existing observations, e.g. $P(A|O_i \cap O_{ii} \cap O_{iii})$, for example by simplifications which can be exploited where variables are known to be conditionally independent.

Based on the results in Figure 6, anchors encountering a cavity size of greater than 1 m have the highest chance of loosening, i.e. approximately 50%. The other important parameter is grout wastage where there is approximately 40% chance of loosening if the grout wastage is greater than 200%.

The influence of ground anchor length and rockhead depth may appear to be important, but they are believed to be affected by the small sample size for long ground anchors or deep rockhead. Where sample sizes are limited, the predictions are dominated by the limited data, and the problem of *overfitting* may occur (Domingos, 2015; Ng, 2018). The accuracy of the predictions can be examined by running cross-validation tests, i.e. using a sizable fraction of the samples for training, and the remaining samples for testing (Ng, 2018). Another way of overcoming the problem of limited data with Bayesian statistics is by exploiting the use of prior distributions.

The joint probabilities of varying grout wastage and cavity size > 1m on the likelihood of loosening were studied. It was found that for anchors with both cavity size greater than 1m *and* grout wastage greater than 200%, the probability of loosening is 57% (Figure 6 (e)). It is considered here in this study that a probability of greater than 50% is a strong sign that it was likely one of the causes leading to loosening. It is noted that while using Bayesian programming, one needs to be careful of considering whether the observations are independent or non-independent, because it may lead to discrepancies (Figure 6 (e)). Where the variables are non-independent, the variables have to be checked for data instances in which they are fulfilled simultaneously, i.e. $P(A|O_i \cap O_{ii} \cap O_{iii})$, in which case $O_i$, $O_{ii}$ and $O_{iii}$ are true.

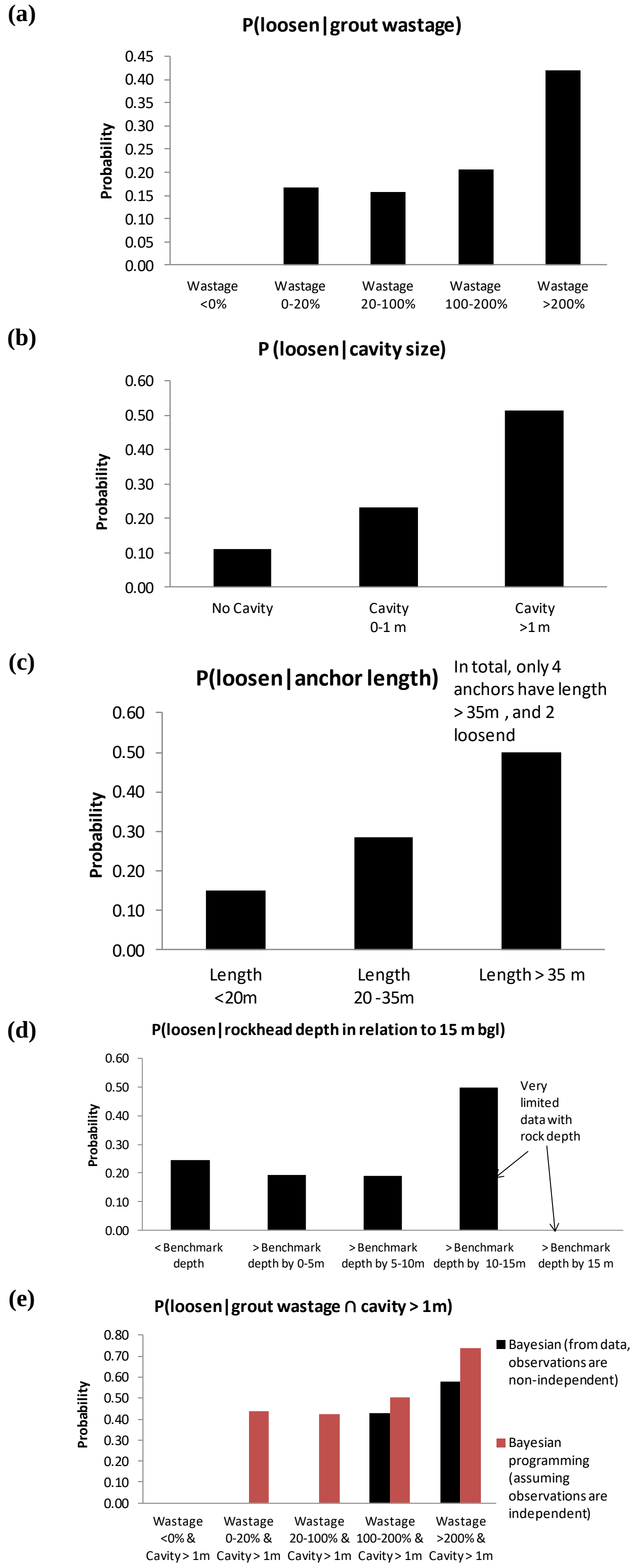


Figure 6 Predictive statistics of ground anchor loosening, given (a) grout wastage, (b) cavity size, (c) anchor length, (d) rock head (using 15 m bgl as th benchmark depth), (e) grout wastage and cavity size > 1 m encountered

### 3.3. Decision-making based on findings from Bayesian analyses

The data suggests that the loosened anchors were correlated to the grout wastage and whether or not large cavities were intersected. It is noted that, for this site, the rock socketing length was re-established from zero once a cavity was intersected. Despite this, the observations suggest that the anchor resistance was compromised at these localised areas.

It was observed that not all of the loops of the problematic anchor had loosened, and that some loops of the same anchor did not lose tensioning. Based on this observation, pull-out tests on the pair of strands originating from the same loosened loop were commissioned in the subsequent incident of loosening.

These tests are referred as *pair tests* subsequently here. It transpired that, of the eight tests which were carried out, none of the loosened loop was able to resist the design load and showed signs of creep at 50%-70% of the loop's working load. Seven out of the eight anchors had the shortest loop loosened. In fact, five out of the eight anchors had grout wastage greater than 200% and cavity size greater than 1 m. During the assessment, the installed ground anchors satisfying these two criteria were found to be around 7% (Figure 7), of which approximately one-fifth of these anchors were downgraded due to pair testing. As the loosening only happened to some of the strands, it was noted that the capacity of the ground anchor was not lost completely. The original factor of safety for the design rock bond strength in relation to the design mobilised resistance was also much larger than required. It was also thought that by pulling the rock bond past its peak strength may have counterproductive effects leading to unrecoverable post peak strengths. Based on these observations and considerations, further pair testing was not pursued further.

The following corrective measures were implemented:

i. Change of anchor block suiting the wedges, and with a higher Rockwell hardness. This is notwithstanding the fact that the existing anchor block and wedges performed satisfactorily when used as a restraint in a laboratory tensile test, as the strand failed at the desired tensile yield strength;
ii. Anchors where the pair tests had failed were downgraded;
iii. The required rock socket was set back by 1.5 m from the rockhead, due to the site specific observations made in the pair tests that the shortest loop was affected. The reason for this was not known, but could be potentially due to (i) the overstressing of the shortest loop as a result of possible initial adoption of incorrect stressing procedures for multi-loop anchors, compounded by unloading and reloading, (ii) grout loss at the limestone interface specific to the site conditions;
iv. More stringent casing of the soil overburden was implemented, ensuring that the casing was socketed into competent rock, to avoid grout loss and collapse of borehole near the interface;
v. Some anchors were lengthened beyond the active wedge, as the rockhead was undulating based on the drilling records for the upper anchors.

After the measures above were implemented, the outcome was that only another seven number anchors had loosened when the site first reached final excavation level (partial extent). Two of these loosened anchors had both grout wastage greater than 200% and cavity size greater than 1 m. These seven anchors were from the initial first and second row anchors, before the risk mitigation measures above were undertaken. One of the anchors had both the anchor block and wedges originating from the same manufacturer, supporting the findings obtained from this Bayesian analysis that the loosening of anchors could also be due to reasons other than the incompatibility between the anchor block and wedges. Finally, when the entire site was excavated fully to the final excavation level, further ten numbers of anchors had loosened, of which six numbers were the anchors which had loosened previously.

It was observed that some of the ground anchors which had loosened were spatially from the same vicinity, i.e. from the topmost row to the lower rows. An exposed rock face, where a few of the ground anchors above it had loosened, is shown in Figure 8. This geological feature could have contributed to the observations of clustered cavities and high grout consumption in the ground anchor installation records, which in turn correlate well with the likelihood of ground anchor loosening. While not all locations had resulted in similar rock exposures, it is possible that there was a reduction in rock uplift resistance around areas with large, persistent and unfilled cavities, especially as more anchors were being installed and stressed.

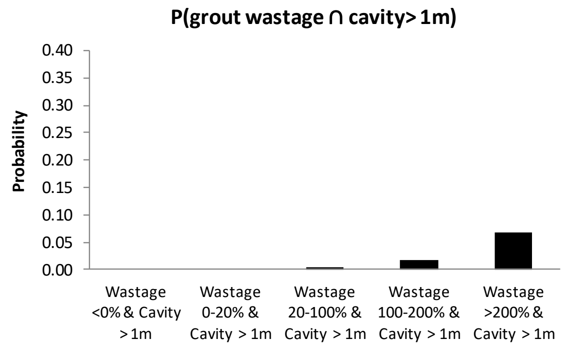


Figure 7 Probability of varying grout wastage and cavity size > 1m encountered in relation to the installed ground anchors

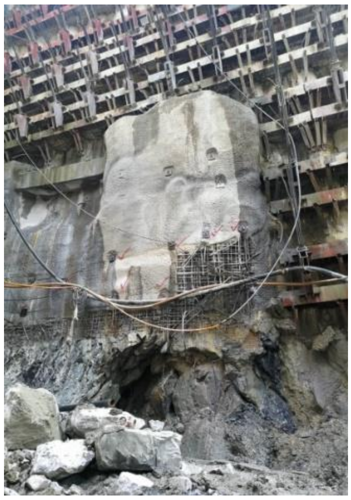

Figure 8 Figure of exposed rock face around which some of the anchors had experienced loosening

## 4. CONCLUSION

In situations where there are several likely causes to the problem, and with overwhelming data, it is formidable to process the information without the aid of a computer. Where computers are used, it is natural and straightforward to include some statistical techniques in the processing of information.

This paper shows how Bayesian statistics or Bayesian programming could be used to the engineer's advantage to interpret the existing observations while troubleshooting ground anchors. The study was not able to quantify the random effects of installation quality or perceived anchor block-wedge compatibility performance, but was able to identify the increased likelihood of anchor loosening for anchors with higher grout wastage and with cavities encountered. At certain locations, it was found that this could be possibly related to geological features based on the rock exposures. The findings in this paper are unique to the geological conditions at this site. Nonetheless, it is believed that a Bayesian analysis approach could be adopted in engineering practice, and extended to other elements of geotechnical works, where required. Similar procedures could be developed further for decision-making problems with conditional dependence using Bayesian networks founded on graphical tree approaches.

While these tools are useful, there is a need to further interpret the results and subsequently strategise on the next course of action using one's experience and judgement. It is qualified that, even with these statistical techniques or machine learning implemented, it may not be possible to obtain a firm predictive answer, as with the case history discussed here. This may be inevitable where there is a scatter of variables associated with the observation of interest. Nonetheless, studying the likelihood of the observation given a few pre-determined variables has found to been a good risk mitigation measure, especially when the method of logical elimination may be limited to one to two iterations in a practical deep excavation timeframe. More often than not, a few variables have to be studied and several mitigation measures targeting at different variables have to be carried out concurrently.

Finally, it is demonstrated here that the main challenge in geotechnical works may not always relate to analyses and calculations, but the evaluation of multiple pieces of information at the same time. There were several prevailing hypotheses which were present in this case history. It was important to be able to examine these hypotheses and question the reliability of the basis or assumptions leading to the hypotheses. The exercise of intellectual caution on the information and assumptions equally lead to the safety of geotechnical works. This has been termed as Type II Factor of Safety in Boon & Ooi (2016) to echo the wisdom from the late Ralph Peck: "until you know what goes on in the field, how people do things and how the movements that occur are related to the loads that you measure, quality of the workmanship and so on, you don't really understand how soil is behaving".

## 5. ACKNOWLEDGEMENT

The discussions undertaken here do not indicate a preference among the different types of analyses by the authors and the institution to which the authors are affiliated.

## 6. REFERENCES

Bessière, P., Mazer, E., Ahuactzin, J.-M., Mekhnacha, K. (2013) "Bayesian Programming", CRC Press, p380.

Boon, C.W., Ooi, L.H., and Low, Y.Y. (2015) "Performance of ground anchors in a Mass Rapid Transit project in Malaysia", 9th International Symposium on Field Measurements in Geomechanics (FMGM), 9-11 September, Dight, P. (Eds.), Sydney, Australia, Australian Centre for Geomechanics, pp621-630.

Boon, C.W., and Ooi, L.H. (2016) "Type II Factor of Safety: Reliability of Information from Instrumentation, Numerical Analysis, Site Investigation and Design", 19th Southeast Asian Geotechnical Conference & 2nd AGSSEA Conference (19SEAGC & 2AGSSEA) Young Geotechnical Engineers Conference, Kuala Lumpur, 30 May 2016, pp40-45.

British Standard. (1989) BS 8081:1989: Code of practice for ground anchorage, BSI.

Domingos, P. (2015) "The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World", Basic Books, p352.

Houlsby, N.M.T., and Houlsby, G.T. (2013) "Statistical fitting of undrained strength data", Géotechnique, 63(14), pp1253-1263.

Huang, J., Kelly, R., Li, D., Zhou, C., Sloan, S. (2016) "Updating reliability of single piles and pile groups by load tests", Computers and Geotechnics, 73, pp221-230.

Kay, J.N. (1978) "Safety factor evaluation for single piles in sand", J. Geotech Eng Div-ASCE, 104(1), pp148-149.

Ng, A. (2018) "Machine Learning Yearning", Deep Learning. AI.

Zhang, L. (2004) "Reliability verification using proof pile load tests", J. Geotech Geoenviron Eng, 130(11), pp1203-1213.

Mitchell, T. M. (1997) "Machine learning", McGraw-Hill, p432.