Cost-benefit analysis for selection of pile tests

Roger Skirrow, Rocky Wang
Technical Standards Branch, Alberta Transportation, Edm, Alb, Canada

ABSTRACT
The design of piles using limit states design (LSD) methods incorporates the use of geotechnical reduction factors (GRF). The value of the GRF varies depending on the method used to estimate or measure pile capacity. A low GRF value is assigned to an empirical-based method while a high GRF value is assigned to full-scale field test method. If a designer knows in advance, or specifies, the method of pile capacity verification they can take advantage of the applicable GRF and potentially reduce the number or length of piles if a high GRF value is used, relative to a design based purely on empirical methods. However, since full-scale load tests are costly in relation to pulse-velocity, PDA or empirical methods, there exists an opportunity for designers and owner to optimize the pile design, pile capacity verification and pile quality assurance program. This paper describes a cost-benefit analysis for both end-bearing piles and friction piles. Based on the number of sites and piles in a project, a nomogram may be used to choose the optimal pile capacity quality assurance testing method. A case study demonstrates the methodology.

RÉSUMÉ
La conception des piles employant des méthodes de la conception d'états de limite (LSD) incorpore l'utilisation des facteurs géotechniques de réduction (GRF). La valeur ou le GRF change selon la méthode employée pour estimer ou mesurer la capacité de pile. Des valeurs plus basses de GRF sont assignées aux méthodes empirique-basées et à la valeur plus élevée de GRF assignées aux méthodes complètes d'essai sur le terrain. Si un concepteur sait à l'avance, ou indique, la méthode de vérification de capacité de pile ils peuvent tirer profit du GRF applicable et potentiellement réduire le nombre ou la longueur de piles si une valeur élevée de GRF est employée, relativement à une conception purement empirique-basée. Puisque les essais complets de charge sont coûteux par rapport à l'impulsion-vitesse, au PDA ou aux méthodes empiriques, là existe une occasion pour que les concepteurs et le propriétaire optimise la conception de pile et le programme de garantie de la qualité de pile. Cet article décrit une analyse coûts-avantages pour des piles d'extrémité-roulement et des piles de frottement. Basé sur le nombre d'emplacements et de piles dans un projet, un abaque peut être employé pour choisir la méthode d'essai optimale de garantie de la qualité de capacité de pile. Une étude de cas démontre la méthodologie.

1 INTRODUCTION

Piled foundations support the majority of bridge structures in Alberta. Most bridge pile foundations consist of driven pipe or H-Pile or timber piles. During preliminary design, the ultimate axial load capacity of piles is usually estimated using semi-empirical equations based on static analysis methods and soil mechanics principles (Tschebotarioff, 1973). The allowable or working stress on each pile is then determined by applying a factor of safety to the ultimate pile capacity. Soil parameters used in these calculations are estimated using engineering judgement, laboratory testing, and correlations with in-situ test results (usually SPT or CPT-based). In the case of driven piles the prediction of pile driveability done prior to construction is made using engineering judgement, empirical energy equations, pilot piles and wave equation analysis (WEAP programs).

In the case of drilled piles the as-built axial load capacity is estimated through a review of the soil profile encountered during drilling; in the case of driven piles the load capacity is estimated using driving resistance formulae often supplemented with set and compression values provided from the pile monitoring program. On rare occasion Alberta Transportation has used full-scale load tests to directly measure pile load capacity. High strain dynamic testing, such as the Pile Dynamic Analyzer (PDA) is now used routinely on department projects and has provided a cost-effective means of estimating pile capacity and near real-time monitoring of construction QA/QC. Other pile capacity 'measurement' techniques such as the Osterberg Cell and Statnamic testing are now accepted in Alberta industry, and are being introduced for use on large scale Alberta Transportation projects.

In concurrence with the introduction of the PDA method, major agencies are moving away from allowable stress design (ASD) or working stress design (WSD) methods for substructure design and away from load factor design (LFD) methods for superstructure design. The trend in design methodology is toward the use of Limit States Design (LSD) methods (FHWA, 2001). During this current transition period many agencies continue to use the WSD methodology, and determine allowable loads by application of a Factor of Safety. The LSD method is based on Load and Resistance Factor Design (LRFD) which uses geotechnical resistance factors (GRF). In the absence of what might be termed codified GRFs that are specific to Alberta soils and bedrocks, many practitioners in essence use a GRF that is the inverse of a familiar Factor of Safety. This practice should be discouraged, but this is not the subject of this paper. It is expected that local or regional GRF values will be developed as the LSD method matures in Canada.

In relation to pile designs based on semi-empirical equations there are clear incentives to base designs on more reliable and defensible design methods. The use of higher GRFs associated with PDA tests or higher order pile capacity measuring methods can result in significant project cost savings. They can be the result of a reduction in the number of piles, the length of piles or the size of piles. PDA testing, in particular can offer a cost effective design optimization method, capacity prediction tool, and QA/QC program. However, there are costs associated with inclusion of a PDA test protocol in a given piling project and the overall benefit to the project may not justify PDA, or more elaborate test methods.

2 PILE TESTS AND GEOTECHNICAL RESISTANCE FACTORS

Under the framework of WSD, safety factors are applied to a given pile or pile group according to engineering experience and various established engineering or building codes (Linkins, 2004). However, associating a GRF value with a given pile capacity prediction method requires a well developed database, thorough statistical analysis and ongoing calibration, plus incorporation of site-specific pile testing results. The incorporation of in-situ pile testing results can be achieved by means of a Bayesian approach (Tang, 1997). If collected data is "normal", then a graphical estimate approach can be used (Selvin, 1976).

Table 1 presents a condensed list of GRFs for National Building Code of Canada (NBCC, 2005) and Canadian Highway Bridge Design Code (CHBDC, 2000) GRFs.

For demonstration purposes 5 classes of pile monitoring were used, as follows:

Method 1 - Conventional Monitoring, e.g. Gates, with an assigned resistance factor 0.35;

Method 2 - Conventional Monitoring w/Field Measurements, with an assigned resistance factor 0.40;

Method 3 - Dynamic Test, e.g., PDA, with an assigned resistance factor 0.50;

Method 4 - Dyn/Static Test, e.g., Statnamic, with an assigned resistance factor 0.55;

Method 5 - Static Test, e.g., classic load test or O-Cell, with an assigned resistance factor 0.60.

Further assumptions were incorporated into the analysis, including: for any case, when Method 4 and/or Method 5 are planned for an individual site, then at least one such pile testing should be undertaken, and; no increase in resistance factor was assigned to correspond to increase in testing frequency.

Table 1. Geotechnical resistance factors, Φ, to axial load of deep foundation

Description	NBCC (2005)	CHBDC (2000)
Analysis using dynamic method (no field measurements)	NA	0.4
Semi-empirical analysis using laboratory and in-situ test data	0.4	
Analysis using dynamic monitoring results	0.5	
Analysis using static loading test results	0.6	

3 DETERMINATION OF OPTIMUM NUMBER OF PILE TESTS

An acceptance sampling plan is defined as determination of an optimal combination of *n* and *r* for a given sample size, *N*, where *n* is the pile number submitted for testing, inspection or to be proofed, and *r* the number of defective piles recognized after inspection. The selection of *n* and *r* should be agreed upon by both the contractor and the owner of the project, in order to share risk and manage costs.

If the total pile number projected is *N* for a project or worksite, then the "Percentage of Test (%)" is *n*/*N* for that type of pile test. Note that each of tested piles is classified as "acceptable" or "defective" after the testing, say, following "acceptance sampling by attributes". The attributes or classification criterion should be formulated by the owner or owner's agent (consulting engineer). As a general criterion, if more than *r* defective piles are discovered from the all tested piles of *n*, then the quality of constructed pile engineering project will be rejected by the owner.

Assuming that among all piles of size *N*, the actual fraction of defective is *p*; the total number of defective piles in the sample *n* is described by hyper-geometric distribution. If the lot size *N* is very large or *n* is small as compared to *N*, then the probability that there are at most *r* defective piles can be approximated by a cumulative binomial distribution function P_a, or we can say that the probability of acceptance is the probability that *r*, the number of defectives, is less than or equal to c, the acceptable number.

The cumulative binomial distribution function B(*r*; *n*, *p*) is referred to as the OC curve (Operating Characteristic curve). There are two useful Microsoft EXCEL built-in functions available to assist in this aspect, namely, HYPGEOMDIST and BINOMDIST. The cumulative binomial distribution (accepting probability) is,

$$B(x; n, p) = \sum_{x=0}^{r} \binom{n}{x} p^{x} (1-p)^{n-x} \qquad [1]$$

Generally, there are conflicting interests between the contractor and the owner, which need to be resolved, preferably before the piling contract is released and accepted. The contractor hopes to have a low probability of rejection for a batch of piles, in which the actual fraction of defective piles p is less than p_1,; however the owner does not want to have a high probability of accepting a lot if p exceeds p_2, the minimum fraction of defective piles sufficient to define the lot as poor quality work.

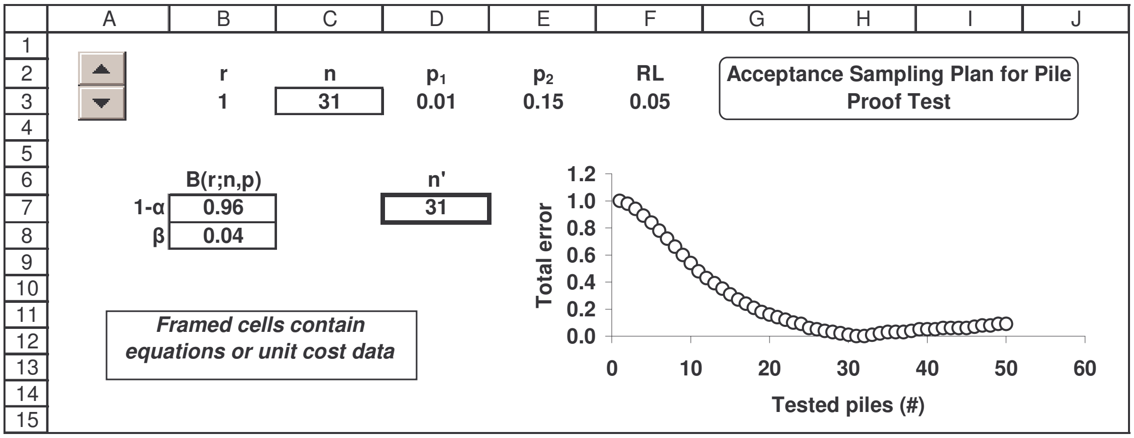
These two simultaneous equations considering both owner and contractor are nonlinear and the two unknowns n and r are whole numbers so there is no simple, direct solution. A spreadsheet as shown in Figure 1 uses Data->Tables tool and iteration feature in Excel, which are part of a suite of commands sometimes called What-If analysis, to solve the equations. Note that symbol "RL" represents a risk level for both contractor and owner, say, $\alpha = \beta =$ RL; Cell C3 equals to Cell D7; Cell B7=ROUND(BINOMDIST(r_,n,p1_,TRUE),2); Cell B8 =ROUND(BINOMDIST(r_,n,p2_,TRUE),2); Cell D7 =INDEX(O9:O58,MATCH(MIN(P9:P58),P9:P58,0)), which refers to a data table. The SpinButton object is linked to Cell B3 in order to adjust r value so that the calculated risk levels are within the RL's limit. This example shows if a 5% risk for both sides, 1% to 15 % of tested piles being the specified limits, are assumed, then the optimal inspection plan should have n=30, r=1, or 30 piles should be tested and will be accepted if no more than one pile is defective.

If setting up Acceptance Sampling Planning by "Variables", then "Since the measured data are more fully utilized here than in sampling by attributes, the sample size required to achieve the same degree of quality control can be significantly smaller than that of sampling by attributes." (Ang et al., 1975). An alternative simplifying approach is to use the following formula,

$$n_{\min} = \left[\frac{\delta_x \cdot t_\alpha\left(n_{\min}-1\right)}{\Delta_x}\right] \qquad [2]$$

where δ_x is the c.o.v. of x; t_α is one-tailed Student's t-distribution with free degree of $n_{\min}$-1; Δ_x is relative error of x. Iteration is needed to solve the equation above. NCHRP (2004) discusses the same issue but shows no details of a solving process.

A common practice is a certain number of production piles are tested to ensure the piles being satisfied as constructed. For example, for Test 1 and Test 2, 100% of production piles are tested; for Test 3, 10%; for Test 4, 2%; for Test 5, 1%. However, for any case and any test method, at least one pile has to be tested. The following cost-benefit analysis complies with this criterion.



	A	B	C	D	E	F	G	H	I	J
1										
2		r	n	p_1	p_2	RL	**Acceptance Sampling Plan for Pile Proof Test**			
3		1	31	0.01	0.15	0.05				
4										
5										
6		B(r;n,p)		n'						
7	1-α	0.96		31						
8	β	0.04								
9										
10										
11										
12										
13										
14										
15										

Framed cells contain equations or unit cost data

Figure 1. Determination of optimal inspection plan by spreadsheet

4 COST-BENEFIT ANALYSIS OF PILE TESTS IN DESIGN PROCESS

A spreadsheet was developed to review potential piling scenarios, with the purpose of providing guidance on the scope of piling project required in order to ascertain which testing method provided an economically optimal QC/QA advantage. For example, potential savings for PDA testing method may occur for several reasons:

(1) PDA testing satisfies the code requirements for dynamic monitoring of field conditions and thus the GRF of 0.5 may be used, an increase from conventional pile design and monitoring practice which provides a GRF of 0.4. This increase in GRF will allow for fewer or shorter piles to be used, which will reduce costs for pile driver set-up, splicing and the like;

(2) PDA testing of a statistically large sample of piles will provide the designers confidence with their estimates of the ultimate pile capacity, and may permit them to increase their estimates, also resulting in the need for fewer or shorter piles;

(3) PDA testing often predicts a pile capacity that is greater than the initial pile capacity prediction based on semi-empirical methods. This will permit the designer to use fewer or short piles;

(4) PDA testing is less costly, less disruptive and less intrusive than full-scale pile load tests. Contractor down time is reduced and equipment that is already on site can be used to perform the test.

The spreadsheet assumes, for demonstration and simplicity sake, that the ultimate pile capacity as determined by all design and field proofing methods is identical, in which case the savings are a result of the use of a higher GRF only. The spreadsheet uses reasonable estimates of costs for the various test methods and 2007 prices for piles in the department. Graphs produced by the spreadsheet are shown in Figures 2 through 5, differentiating between piles that act primarily in end bearing, and shaft friction, respectively, and reduction. On the figures 'sites' refers to portions of the project that are treated as independent sites, for example a three span bridge project will have 4 sites.

Our analysis has shown that in general if there are less than 10 piles involved for a single site, there is likely no economic or practical benefit to be gained from PDA testing, and conventional pile driving to refusal conditions are satisfactory. Essentially the cost to do the testing, potential to delay the project, added complexity on small projects and similar issues outweighs the potential for pile optimization.

For both the end-bearing and shaft-friction analysis there is a limit of tested pile number, beyond which the more expensive static and Statnamic testing methods can conversely save cost. The greater the number of piles on a site the more saving can get by PDA. However, cost-benefit analysis should be carried out on the basis of individual project in order to obtain an accurate cost-benefit answer.

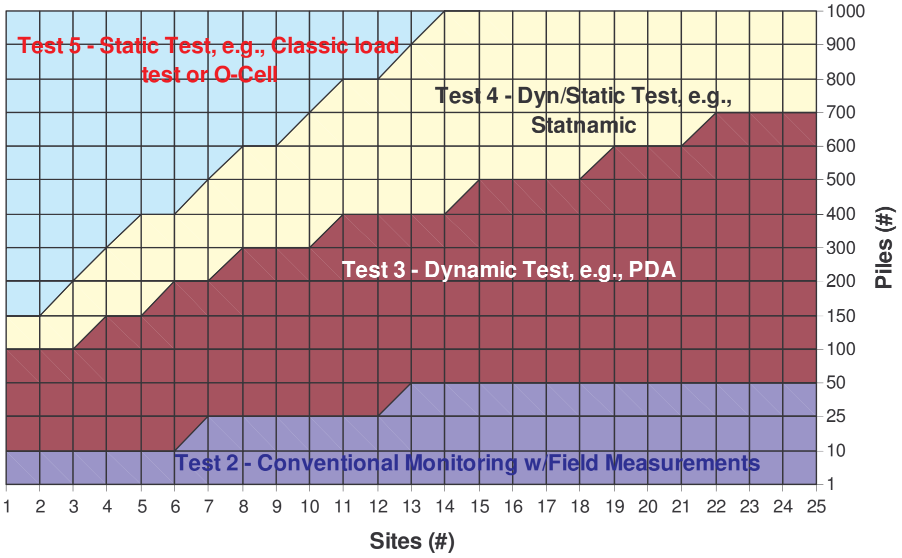


Figure 2. Nomogram for choosing the optimal H-pile capacity quality assurance testing method - End bearing case

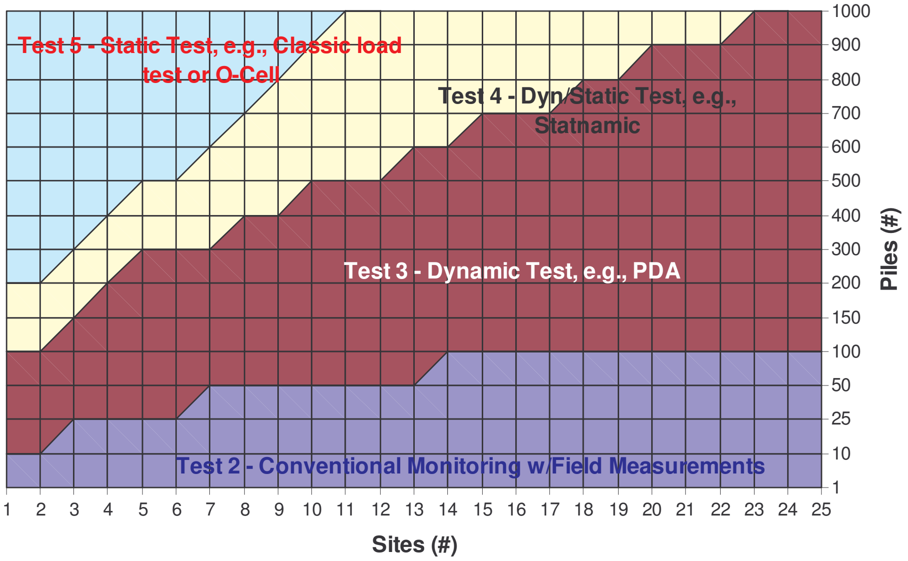


Figure 3. Nomogram for choosing the optimal pipe-pile capacity quality assurance testing method – Friction Case

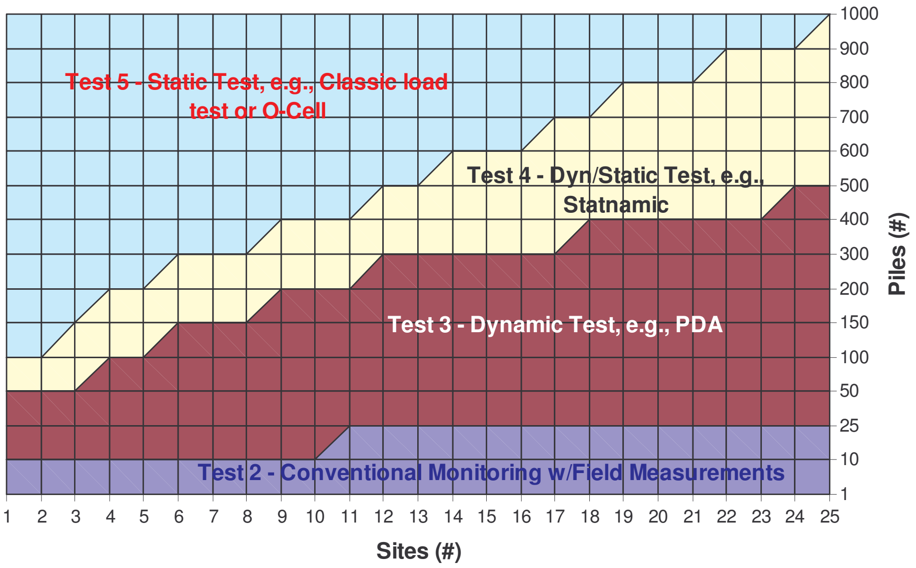


Figure 4. Nomogram for choosing the optimal pipe-pile capacity quality assurance testing method – End bearing case

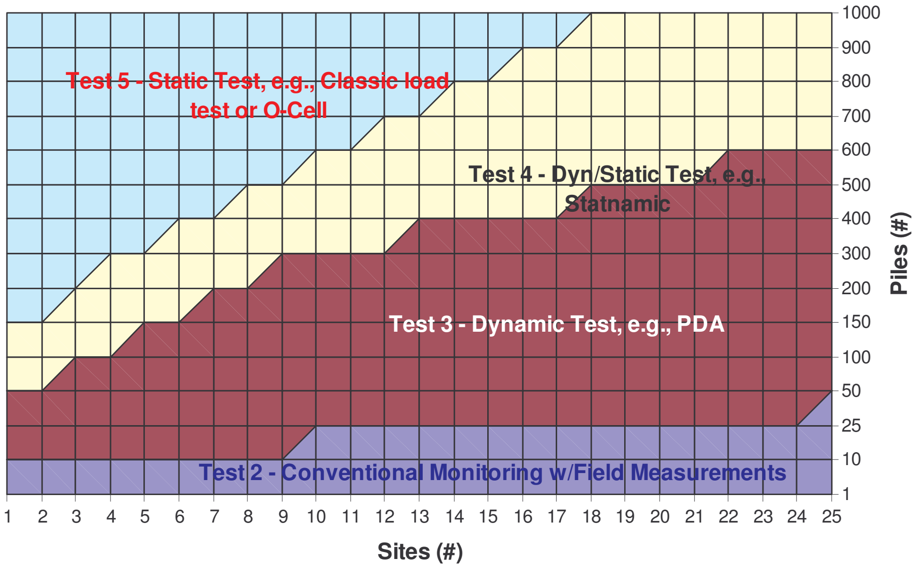
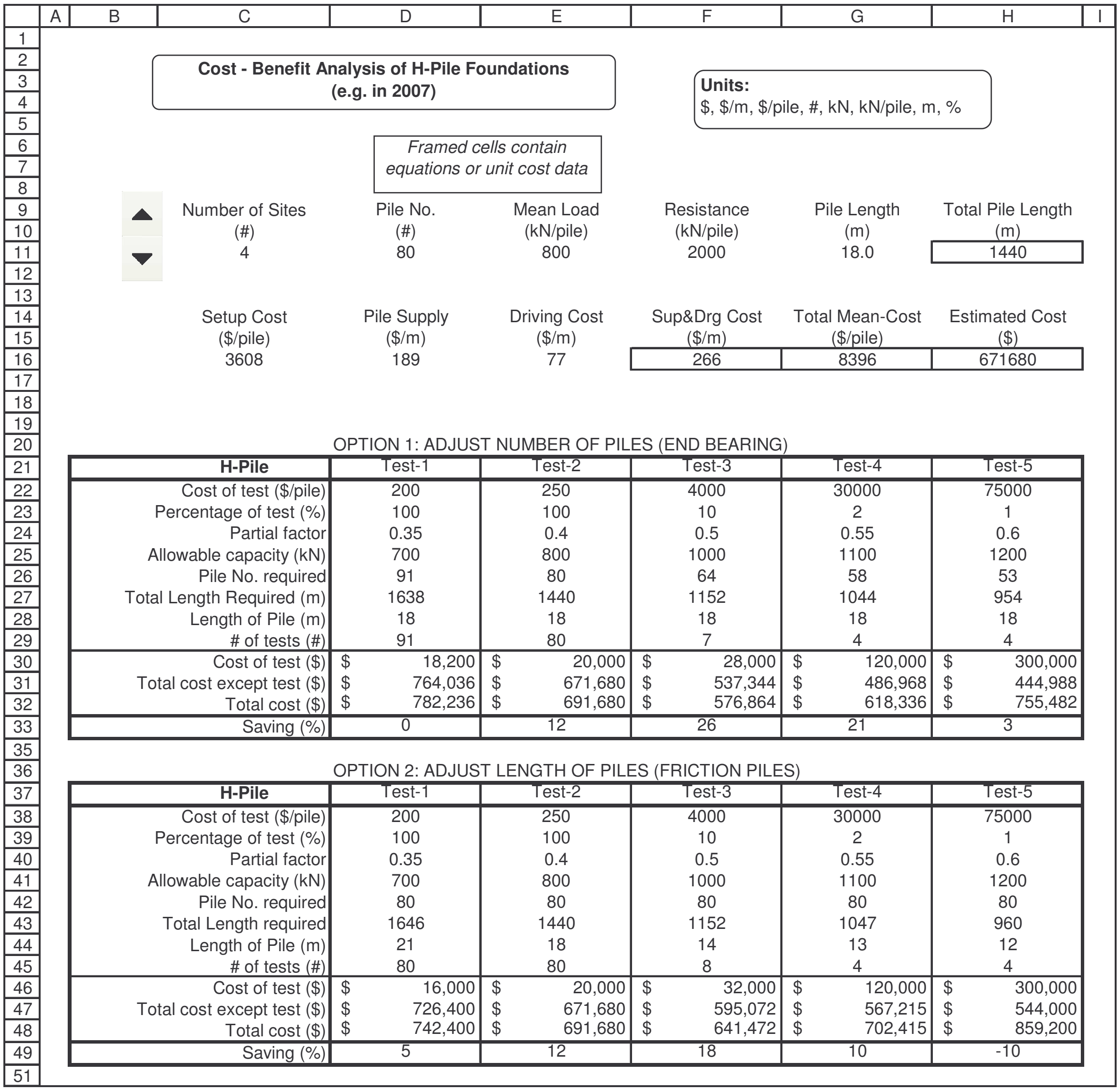


Figure 5. Nomogram for choosing the optimal pipe-pile capacity quality assurance testing method - Friction case

5 EXAMPLE

Figure 6 shows a discussion example. Based on a bridge project with 4 sites and a total of 80 piles. Using Method 1 and 2, 100% of the production piles are evaluated, using Method 3, 4 and 5 the percent of piles tested drops to 10, 2 and 1% respectively, however in our criteria for Method 4 and 5 at least one test is done per site, therefore 4 tests are included for Method 4 and 5. Typical pile construction costs from Alberta Transportation projects from 2007 were included in the analysis. These values can be adjusted to local conditions. Similarly values were used for costing each QA/QC method, which can also be adjusted to local conditions. For this particular case PDA testing, Method 3, when applied to 10% of the piles will provide the greatest cost saving to the project, determined to be about a 26%, savings for primarily end-bearing piles, and 18% for primarily shaft friction piles, inclusive of testing costs. The difference is explained by the greater savings resulting from fewer pile set-ups versus the savings from reduction in pile length. If this analysis is done during the design stage it would be possible to further optimize the pile length, size group size, etc. Obviously there will be situations where soil conditions do not permit reduction in pile length for shaft friction piles, in which case the pile size could be reduced, or fewer piles used.



Cost - Benefit Analysis of H-Pile Foundations
(e.g. in 2007)

Units:
$, $/m, $/pile, #, kN, kN/pile, m, %

Framed cells contain equations or unit cost data

Number of Sites (#)	Pile No. (#)	Mean Load (kN/pile)	Resistance (kN/pile)	Pile Length (m)	Total Pile Length (m)
4	80	800	2000	18.0	1440

Setup Cost ($/pile)	Pile Supply ($/m)	Driving Cost ($/m)	Sup&Drg Cost ($/m)	Total Mean-Cost ($/pile)	Estimated Cost ($)
3608	189	77	266	8396	671680

OPTION 1: ADJUST NUMBER OF PILES (END BEARING)

H-Pile	Test-1	Test-2	Test-3	Test-4	Test-5
Cost of test ($/pile)	200	250	4000	30000	75000
Percentage of test (%)	100	100	10	2	1
Partial factor	0.35	0.4	0.5	0.55	0.6
Allowable capacity (kN)	700	800	1000	1100	1200
Pile No. required	91	80	64	58	53
Total Length Required (m)	1638	1440	1152	1044	954
Length of Pile (m)	18	18	18	18	18
# of tests (#)	91	80	7	4	4
Cost of test ($)	$ 18,200	$ 20,000	$ 28,000	$ 120,000	$ 300,000
Total cost except test ($)	$ 764,036	$ 671,680	$ 537,344	$ 486,968	$ 444,988
Total cost ($)	$ 782,236	$ 691,680	$ 576,864	$ 618,336	$ 755,482
Saving (%)	0	12	26	21	3

OPTION 2: ADJUST LENGTH OF PILES (FRICTION PILES)

H-Pile	Test-1	Test-2	Test-3	Test-4	Test-5
Cost of test ($/pile)	200	250	4000	30000	75000
Percentage of test (%)	100	100	10	2	1
Partial factor	0.35	0.4	0.5	0.55	0.6
Allowable capacity (kN)	700	800	1000	1100	1200
Pile No. required	80	80	80	80	80
Total Length required	1646	1440	1152	1047	960
Length of Pile (m)	21	18	14	13	12
# of tests (#)	80	80	8	4	4
Cost of test ($)	$ 16,000	$ 20,000	$ 32,000	$ 120,000	$ 300,000
Total cost except test ($)	$ 726,400	$ 671,680	$ 595,072	$ 567,215	$ 544,000
Total cost ($)	$ 742,400	$ 691,680	$ 641,472	$ 702,415	$ 859,200
Saving (%)	5	12	18	10	-10

Figure 6. Project example

6 COMMENTS TO PDA AS QA/QC TOOL

The PDA was developed primarily to assist in evaluating the ability of pile driving equipment to install piles to the desired depth without damage. It measures driving stresses and is therefore useful to prevent pile overstressing. The ability to predict load capacity also can be used to assess variation in pile capacity across a site and hence provide guidance for designers. In some cases the PDA has been used to confirm pile termination depths when deep borehole information was not available.

When used as a QA/QC tool various agencies suggest that a minimum of 5% to 10% of production piles should be monitored dynamically using PDA methods. An issue arises if the PDA determined load capacities are significantly greater than, or less than, the design ultimate load capacities. In this situation it may be necessary to undertake static load tests to calibrate the PDA CAPWAP analysis; however a thorough review of all inputs into the PDA software should precede any such venture and an accounting of pile setup is required. A PDA QA/QC program may specify testing at initial strike for 10% of production piles, and again at restrike on a sample of previously tested piles. This will help determine if pile set-up or relaxation has occurred.

In general the procurement of PDA testing services is the responsibility of the contractor since they control their schedule and equipment availability. Details regarding frequency of testing, documentation of results, and the like are outlined in the contract document. A test pile program that utilizes PDA testing has the greatest opportunity to optimize and rationalize the pile design, and thereby give the greatest benefit to the project design and cost balance.

7 SUMMARY

A spreadsheet methodology has been presented to assist designers to select an appropriate level of sampling and testing for pile acceptance. This information was further developed and incorporated into a second spreadsheet analysis that provides a method for optimizing the selection of pile capacity monitoring methods based on number of piles, variable GRFs and costing data. The results show that for typical small piling projects of 10 or fewer piles there may be no justification for advanced pile capacity monitoring programs, beyond field monitoring and conventional good construction practices. As more piles are involved at a site the optimization process suggests that more elaborate pile capacity methods provide better value.

The decision of which test method to be used during pile monitoring should be considered during the design phase of the project, since it is easier to adjust designs at that stage, then during construction. Consistent with the findings presented by Linkins (2004), the paper shows that significant piling project cost savings can be achieved through the use of a well planned, and statistically defensible testing program. As is the case for many investments it often costs money to make money, or in this case to save money.

REFERENCES

Ang, Alfredo H-S., and Tang, Wilson H. 1975. *Probability Concepts in Engineering Planning and Design*, Vol.1 - Basic Principles, John Wiley & Sons. Inc., pp.360-378.

NCHRP (National Cooperative Highway Research Program, Transportation Research Board of the National Academies, USA) Report 507. 2004. *Load and Resistance Factor Design (LRFD) for Deep Foundations*, (See also: http://www.engmath.dal.ca/risk/pubs/NCHRP507.pdf).

Kenneth, R.B., John, R.D., Jose, L.C., and Garland, L. 2002. Proven success for driven pile foundations, *ASCE 2000 International Deep Foundations Congress*, Orlando, Florida, USA, pages 9 (See also: http://www.bechtel.com/assets/files/PDF/BIP/22249.pdf).

Likins, G. E. 2004. Pile testing – selection and economy of safety factors, *Current Practices and Future Trends in Deep Foundations*, Geotechnical Special Publication No. 125, DiMaggio, J. A., and Hussein, M. H., Eds, American Society of Civil Engineers: Reston, VA, pages 239-252 (See also: http://www.pile.com/reference/ggg2004/14-Likins.pdf).

Federal Highway Administration. 2001. *Load and Resistance Factor Design (LRFD) for Highway Bridge Substructures*, Publication No. FHWA HI-98-032, pages 592.

Tschebotarioff, G.P. 1973. *Foundations, Retaining and Earth Structures*, McGraw-Hill Book Company, pages 642.

Tang, W.H. 1997. Correlation, multiple RV's, and system reliability, in: Fenton, G.A., ed. *Probabilistic Methods in Geotechnical Engineering, Short Course Notes, Uncertainty'96 (Madison, Wisconsin)*, GeoLogan'97 (Logan, Utah), 1997, 95 p. (See also: http://www.engmath.dal.ca/risk/georisk/part3.pdf).

Selvin, S. 1976. A graphical estimate of the population mean from censored normal data, *Appl. Statist.*, 25(1): 8-11(See also: http://www.jstor.org/stable/2346512?seq=1).