Statistical Considerations in Pile Testing

Considérations Statistiques à propos des Essais sur les Pieux Forés

K. PREISS Department of Mechanical Engineering, Ben-Gurion University of the Negev, Beer Sheva, Israel
J. SHAPIRO Department of Mathematics, Ben-Gurion University of the Negev, Beer Sheva, Israel

SYNOPSIS When using integrity tests for quality control of piles or diaphragm wall sections, the responsible engineer must decide how many piles or elements to test. This paper provides a statistical guide for that decision.

INTRODUCTION

Non-destructive integrity testing of cast in place piles or diaphragm walls is gradually becoming more widely used. The most common methods are the sonic method (Levy, 1970), the vibration method (Davis and Dunn, 1974), both developed in France, and the gamma ray method (Preiss and Caiserman, 1975), developed in Israel. Other methods have been tried, but have not yet achieved widespread use; the number of piles tested by the sonic method is by now probably over 10,000, and by the gamma ray method over 3,000.

When using an integrity test to monitor the quality of the substructural work, the responsible engineer is often faced with the problem of deciding how many piles to test. The number decided upon is a matter of judgement. This paper provides a statistical aid to that judgement.

THE PHILOSOPHICAL BASIS OF THE STATISTICAL METHOD

When manufacturing a product, such as piles, transistors, or whatever, one has to assure the quality of the production process. If one is quite sure that the process always produces a good product, then the product need not be tested. This paper deals with the testing of the product, not the process.

Lot sizes in statistical sampling vary; for instance for transistors the lot size may be many thousands, and the sample size many hundreds. For pile testing the lot size is usually in the range 20 to 200; these are small lots, and so as a result the approach behind the method is as follows:

1. If one is sure that the foundations are sound, without using integrity testing, then the test is not needed at all. For instance, the responsible engineer may be willing to rely only on the supervision during construction. It should be noted that on some projects where integrity testing was not used, faults were found after the superstructure had been partially or completely constructed, leading to substantial delay and expense (Baker and Khan, 1971; Feld, 1963; Szechy, 1961; Lambert, 1973).

   How common are unsound piles? For bored piles, there are statistics available. Preiss and Caiserman (1978) report for 2069 piles and diaphragm wall elements on 79 projects, that 6% of the piles were unsound. Davis and Dunn (1974) report on 717 piles on 5 projects, and found 9.7% to be unsound. The few groups engaged in integrity testing find approximately 5% to 10% of piles to be unsound, on the average.

2. If one cannot rely altogether on the production process, and if one needs to be <u>certain</u> that <u>all</u> piles are sound, then <u>all the</u> piles <u>must</u> be tested. If even one <u>pile</u> is not tested, then there is some non-zero probability that that pile happens to be defective. If the amount of resource which may be devoted to testing is sufficient to test all the piles, then all the piles can indeed be tested. Usually, however, it is either too expensive or not feasible to test all the piles. If all the piles are tested, the method described here is irrelevant.

3. If not all the piles are tested, then the responsible engineer has to decide how many to test. There are many factors involved in this decision, some of which are difficult to express explicitly. For instance, he may decide to test all the piles under a part of the project especially sensitive to differential settlement, such as a high chimney stack, but may require less testing under a less sensitive part of the project, such as a flexible framed structure. In addition, he may require more tests in parts of the site with suspected poor soil conditions, or for piles with a suspicion of poor materials or workmanship. This paper presents statistical parameters which can aid the decision making process.

EFFECT OF THE LOT SIZE ON THE STATISTICAL METHOD

Let N be the total number of piles in a group or lot

N_u be the number of unsound piles in the lot (this quantity is of course unknown)

n be the number of piles in the sample tested

n_u be the number of unsound piles found in the sample n.

If the N_u unsound piles are randomly distributed among the N piles, then the probability p of any one pile being unsound is given by

$$p = N_u/N \qquad (1)$$

N_u is unknown and p is therefore unknown.

The best estimate of p available is

$$\hat{p} = n_u/n \qquad (2)$$

The estimate of the standard derivation of this estimate of $\hat{p}$ is given by (Beyer, 1966)

$$\sigma_{\hat{p}} = \sqrt{\frac{\frac{n_u}{n}\left(1 - \frac{n_u}{n}\right)}{n} \frac{(N - n)}{(N - 1)}} \qquad (3)$$

It can be seen from equation (3) that the standard deviation of the estimate of p decreases rapidly as the sample size n increases (regardless of lot size N), and will be zero when the sample size n is equal to the lot size N. A simple example will show that in the context of piles rather large samples in relation to lot size are needed.

Suppose that N = 100 and N_u = 6, n = 25 piles are sampled and of these n_u = 1 is found defective. The best estimate $\hat{p}$ of p is 1/25 = 0.04 and the standard error of this estimate from equation (3) is

$$\sigma_{\hat{p}} = \sqrt{\frac{\frac{1}{25}\left(1 - \frac{1}{25}\right)}{25} \frac{75}{99}} = 0.034$$

We note that $\sigma_{\hat{p}}$ is almost equal to $\hat{p}$ itself giving a large margin of error in estimating p. The best estimate is that there are 4 faulty piles, but that is a very inaccurate estimate. One often encounters such an approach, in which a statement is made that since x% of faulty piles were found in the first, say 25 piles, it is likely that there be x% unsound piles in the whole lot. Such thinking is applicable when the sample size is large, but one must be very careful when using such an approach on sample sizes less than several hundred, as pointed out by Weltman (1980). This point is further analysed later.

Statistical aids to decision making in pile testing must assume small sample sizes. The relevant approach is given in the next section.

THE STATISTICAL METHOD APPLICABLE

Small samples can yield significant information if one is testing for a continuously variable parameter with a known form of distribution, for example, concrete strength. In the case of piles, one is measuring a property which can have only two values, a pile is either defective or sound. For such a situation, called attribute testing, lot sizes of less than one hundred or so, which are usual, are small lot sizes from the statistical point of view. In order to test a lot one must select a number of unsound piles one is willing to tolerate in the sample. This is called the rejection number r. If more than r unsound piles are found then the lot is defined to be below standard. If a lot is found to be below standard, then from the statistical point of view each pile should be tested and the faulty piles repaired or replaced because it is usually impractical to abandon the site and make a new lot of piles (such a course of action is practical in testing high-quality electronics or machine parts).

Having selected the values of lot size N, sample size n, and rejection number r, then the question to ask is: "if in fact the number of unsound piles in the lot is N_u, then what will be the probability that the sampling strategy will lead to the lot being accepted (or rejected)?" This is calculated from the hypergeometric distribution (Beyer, 1966) as follows:

$$P_{accept} = \sum_{k=o}^{k=r} \frac{\binom{N-N_u}{n-k}\binom{N_u}{k}}{\binom{N}{n}} \qquad (4)$$

where the expression $\binom{i}{j}$, which gives the number of ways j items can be chosen from among i items, is defined by

$$\binom{i}{j} = \frac{i!}{j!(i-j)!}, \text{ for all } j \leq i \qquad (5)$$

For j > i, $\binom{i}{j}$ = 0. The probability of rejection P_{reject} is given by

$$P_{reject} = 1 - P_{accept} \qquad (6)$$

THE OPERATING CHARACTERISTIC CURVES

Figures 1 to 3 are operating characteristic curves for various values of N, n and r. These plots give values of P_{accept} (or P_{reject}) against N_u/N. Note that the values of N_u and N can only be integers, and as a result the results are shown as points, not as continuous curves.

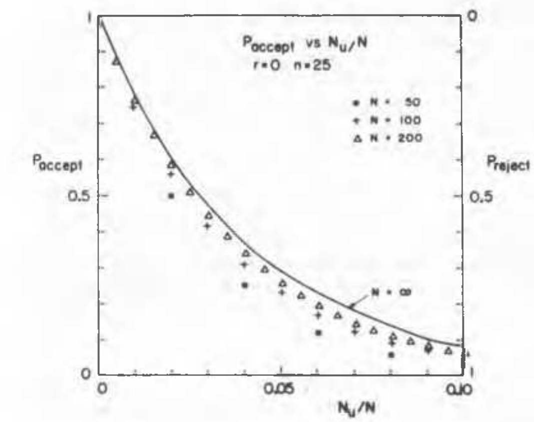

Fig. 1 Illustrates that probability of acceptance depends only weakly on lot size N, for fixed sampling parameters.

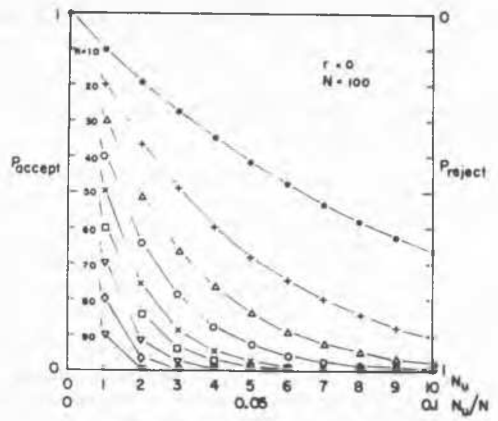

Fig. 2 Probability of acceptance of lots with N_u unsound piles, for various sample sizes n, for lot size N = 100 and rejection number r = 0.

TABLE 1 N = 100, n = 25, r = 0

N_u	0	1	2	3	4	5	6	7	8	9	10
p_{accept}	1	0.75	0.56	0.41	0.31	0.23	0.16	0.12	0.08	0.06	0.05

Figure 1 shows the operating characteristic curve for n = 25, r = 0, for various values of N. It can be seen that the probability of accepting a lot hardly depends on the values of N. Some of the values plotted on Figure 1 are shown in Table 1.

In other words, for a lot size of 100, if the testing strategy states that 25 piles be tested and the lot accepted if no piles are found in the sample, then there is a 75% chance of accepting the lot when there is 1 unsound pile, a 23% chance of accepting the lot when there are 10 unsound piles, and so on. That is, if this is the sampling strategy used, then in 75 out of 100 identical sites the work will be accepted with 1 unsound (and untested, hence unfound) unsound pile, and so on.

Figure 2 shows the operating characteristic curve for N = 100 and r = 0, for various values of n. This figure shows that the larger the sample size, for a given value of N_u, the less the probability of accepting the lot. For instance, let us assume that the lot can be accepted if $N_u \leq 2$ but should be rejected if $N_u > 2$. Relevant observations from Figure 2 are shown in Table 2.

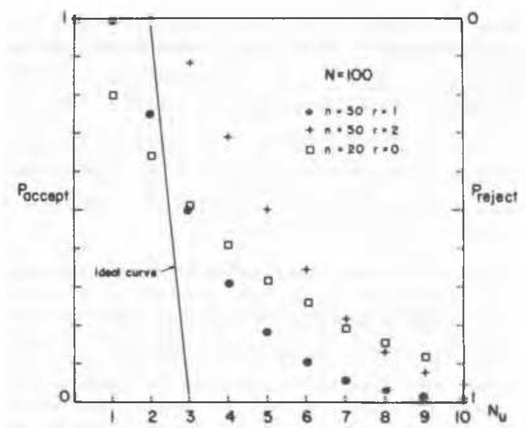

Fig. 3 Given a requirement that the lot of N = 100 piles be accepted if there are 2 or less unsound piles, else accepted, the figure shows the ideal operating characteristic curve, and some approximations to it.

TABLE 2 N = 100, r = 0, N_u = 3

n	10	20	30	40	50	60	70	80	90	100
p_{accept}	0.73	0.51	0.34	0.21	0.12	0.06	0.025	0.007	0.0007	0

These results show that if 10 piles are tested, there will be a 73% chance of accepting a lot with 3 unsound piles (a lot which should be rejected according to the assumptions of the example). This probability drops as n increases, but is still no less than 12% even if half the lot is tested. This point can be further elaborated using Figure 3.

Figure 3 shows, for N = 100, the operating characteristic curve ideally wanted, for this example. This curve is attainable if n = N and r = 2, a situation in which all items are tested, but which is often not possible. Figure 3 gives some operating characteristic curves for various values of the parameters. Of course, the curve found is not the ideal curve, and there is some probability of rejecting when in fact 2 or less unsound piles occur (a situation which should be accepted) or of accepting when there are more than 2 unsound piles. By choosing various values of parameters a curve which is not too far from the ideal curve is sought; as n increases the curve found gets closer to the ideal curve.

CONCLUSIONS AND COMMENTS

A number of conclusions can be drawn.

a) The sample size required depends only weakly on the lot size. The sample size should be chosen from assumptions about what the probability of acceptance should be, for a given value of N_u/N.

b) If n piles are tested and n_u found unsound, then the estimate that for the whole job the number of unsound piles is given by $N_u = N\, n_u/n$, is very unreliable. It implies large samples, an assumption not usually reasonable for piles.

c) If the probability of acceptance is to be zero for all values of N_u greater than zero, then all the piles must be tested (see the results for n = 100 in Figure 2).

d) One may wish to decide that the lot can be accepted if there is a low probability that there be some limiting value $N_{u\ell}$ or more unsound piles in the lot. Reasonable values of $N_{u\ell}/N$ for piles are likely to be in the region of up to 5%, and not in the region more than 10%. The probability of accepting a lot with more than $N_{u\ell}$ piles should be in the region of 5% rather than 20% or more. In order to attain such values, the sample size must be of the order of 50. That means that in small lots all the piles should be tested; in larger lots a large portion should be tested.

e) Weltman (1980) concludes that "the use of suitable integrity tests to improve selection for load tests is clearly a rational approach". One should add to Weltman's conclusion, that to be reasonably reliable from the statistical point of view, the sample size should be sufficiently large, as discussed above.

f) It may be that sample sizes are of the order of 20 or 30, possibly with no unsound piles found in the sample. The statistical curves show that such results of testing on a small sample cannot be used to infer that all the lot is sound. If for some reason the sample cannot be larger, then quality assurance must keep a check on the production process, because checking the product, from such a small sample, is by itself insufficient.

g) The results in this paper assume that the probability that a given pile is unsound, is randomly distributed throughout the lot. In fact, that may not be the case and the testing should be concentrated in the more difficult piles or those in more critical parts of the structure.

REFERENCES

Baker, C.N., Jr. and Khan, F. (1971). Caisson construction problems and correction in Chicago. ASCE J. Soil Mech.Found.Div., Vol. 97 (SM2), Feb., 417-440.

Beyer, W.H. Editor (1966). Handbook of Tables for Probability and Statistics, Chemical Rubber Company Publishers.

Davis, A.G. and Dunn, C.S. (1974). From theory to field experience with the non-destructive vibration testing of piles. Proc.Instn.Civ. Engrs. London, Part 2, Dec., 571-593.

Feld, J. (1963). Diagnosis of distress and failure in structures, The Engineering Journal EIC, Vol. 46, No. 5.

Lambert, D. (1973). Testing and reliability. The Consulting Engineer, Dec., 43-44.

Levy, J.F. (1970). Sonic pulse method of testing cast in situ concrete piles. Ground Engineering, Vol. 3, No. 3, May.

Preiss, K. and Caiserman, A. (1975). Non-destructive testing of bored piles by gamma ray scattering. Ground Engineering, Vol. 8, No. 3, May.

Preiss, K. and Caiserman, A. (1978). Integrity testing of bored piles and diaphragm walls. Civ.Engr.S.Afr., Vol. 20, No. 8, August, 191-196.

Szechy, C. (1961). Foundation failures, Concrete Publications.

Weltman, A.J. (1980). Pile loading test procecures, CIRIA, London.