Journal of Terramechanics, Vol. 22, No. 3, pp. 147–156, 1985.
Printed in Great Britain.

0022–4898/86$3.00+0.00
Pergamon Press Ltd.


# BASIC STUDY OF PROBABILISTIC APPROACH TO PREDICTION OF SOIL TRAFFICABILITY — STATISTICAL CHARACTERISTICS OF CONE INDEX

K. KOGURE*, Y. OHIRA* and H. YAMAGUCHI*

**Summary**—The principal purpose of this study is to present some of the fundamental data which might be used in the prediction of soil trafficability. These data are based on some sets of random cone index measurements which were taken from the areas of interest. In all cases, it was assumed that the individual measurements of the cone index being investigated would be independent. It is shown that the cone indices at critical layer depth can be regarded as normal random variables. The discussion in this study may be valid for the probabilistic approach of soil trafficability

## NOMENCLATURE

$f$ — Degrees of freedom
$f_i$ — Observed frequencies
$F_i$ — Theoretical frequencies
$K$ — Sample size
$M$ — Sample size
$n$ — Sample size
$N$ — Population size
$P$ — Population proportion having a certain characteristic (Probability of encountering an element with the characteristic on a single trial)
$P_r(a<\chi<b)$ — Probability that the random variable $x$ lies in the interval $a$ and $b$
$S^2$ — Sample variance
$t$ — Random variable of the $t$-distribution
$V$ — Coefficient of variation
$x$ — Random variable
$x_i$ — Sample elements
$x_o$ — Some specified value
$x$ — Sample mean
$\alpha$ — Level of significance
$\mu$ — Population mean
$\sigma^2$ — Population variance
$\Phi()$ — Denotes the normal distribution function
$\chi$ — Random variable of the Chi-square distribution

## INTRODUCTION

WE MAY divide the statistical methodology into two broad classes [1]. In one, we can place the routine collection of large amounts of data and description of these data using certain mathematical methods such as percentage, averaging, graphing and others. In the second class, we place a methodology which has been developed for making predictions, or drawing

*Department of Civil Engineering, The National Defense Academy, Yokosuka, Japan.

inferences, from a given sample of observations about a large population. In this study, we shall primarily deal with the first class of problems.

The Waterways Experiment Station has conducted a number of studies on the relationship between the average soil strength measurement and the size of the area measured, the accuracy of soil moisture prediction methods, and various other terrain evaluation considerations [2] [3].

It is readily apparent that a great deal of personal judgement is involved in determining soil trafficability. For example, the current procedure does not explicitly state what cone index measurement (mean, median, smallest, largest, or some other mathematically determined number) should be used for comparison to the vehicle cone index [1]. It is known from sampling theory, however, that if a sample has been chosen from a population in such a way as not to be a random sample, then the validity of an estimate of a population parameter made from it is questionable. Hence as a first step the random variation of the cone index is considered in this study.

## FACTORS INFLUENCING VARIATIONS

It is very important to establish a method to treat systems of soil trafficability prediction. Uncertainties are the main obstacles in treating systems, and therefore dealing with uncertainties is desired in off-road mobility.

Analyses of the random variation of the strength properties of material such as concrete and steel show that a number of probability functions may be used to describe the variation [4]. Although the various distribution functions can be derived theoretically, the controlling factors that are responsible for the variation in soil properties are not well understood at present. Hence, it is necessary to determine the appropriate probability function empirically by fitting it to the experimental data [5]. The necessary data should consist of a large number of test results on a given site. In this study the test considered is the cone penetration test.

The reliability of the prediction method of soil trafficability is sharply influenced by the accuracy of the cone penetration test. In other words, this accuracy is concerned with the extent and number of cone index measurements. It is well known that the data of the results of cone penetration test show large variations. These variations depend on many kinds of factors, which cause unreliability of the prediction of soil trafficability. Generally, two kinds of variations are contained in those data; one is due to the physical nature of soil, and the other is attributed to technical problems [6]. That is, all natural soils show variations in properties from point to point. Variations due to many kinds of technical errors also contaminate the data. Variations of the results of cone penetration tests cannot be eliminated completely. This suggests the necessity of probabilistic approaches in the field of off-road mobility. From this point of view, variations in cone indices are investigated.

## SOIL PROPERTIES OF MEASURED AREAS

Cone indices were measured in eight areas. The physical soil properties of the measured areas are shown in Table 1.

These values in Table 1 were measured at the soil layer between the depths of 25 and 35 cm below the ground surface. Figure 1 is the plasticity chart. The point obtained in area B is located above line A and the soil is inorganic clay of high plasticity. The point obtained in area E is located below line A and the soil is inorganic silt of high compressibility. The points

TABLE 1. PHYSICAL SOIL PROPERTIES OF THE MEASURED AREAS

| Area | Water content $w$ (%) | Wet density $\rho$ (g/cm$^3$) | Atterberg limit LL (%) | PL (%) | PI (%) | Mechanical analysis >2.0 (mm) (%) | 2.0–0.074 (mm) (%) | 0.074> (mm) (%) |
|---|---|---|---|---|---|---|---|---|
| A | 25.1- 29.3 | 1.37-1.46 | 34 | 11 | 13 | 4 | 11 | 85 |
| B | 36.9- 40.1 | 1.28-1.38 | 65 | 27 | 38 | 5 | 11 | 84 |
| C | 20.5- 28.8 | 1.37-1.47 | 36 | 19 | 17 | 1 | 10 | 89 |
| D | 26.9- 27.3 | 1.34-1.39 | 43 | 22 | 21 | 0 | 7 | 93 |
| E | 138.0-170.0 | 1.01-1.16 | 120 | 72 | 48 | 3 | 9 | 88 |
| F | 22.4- 25.6 | 1.44-1.61 | 29 | 16 | 13 | 5 | 34 | 61 |
| G | 26.8- 32.1 | 1.32-1.44 | 34 | 20 | 14 | 0 | 4 | 96 |
| H | 22.5- 27.3 | 1.43-1.70 | 32 | 21 | 11 | 7 | 12 | 81 |

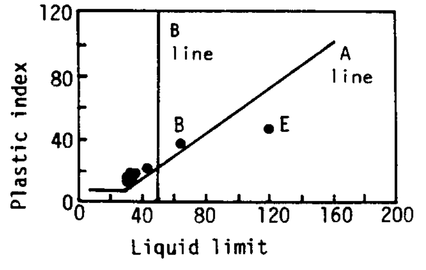


FIG. 1. Expression by plasticity chart.

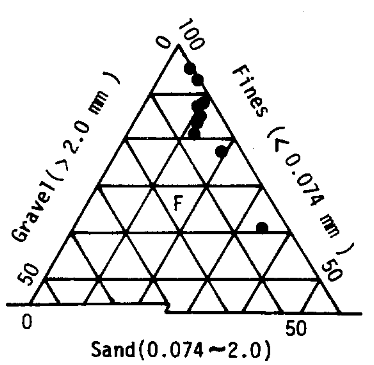


FIG. 2. Expression by classification chart.

representing soils not in areas B or E are located above line A and these soils are inorganic clay of medium plasticity.

Figure 2 is the soil classification chart in which each of the three coordinate axes pertains to one of the grain-size fractions, designated as gravel (>2.0 mm), sand (0.074–2.0 mm) and fines (<0.074 mm). All the points obtained in the soils from the measured areas are located in the region of fine-grained soil which is indicated by the symbol F in the chart.

## VARIATIONS IN CONE INDEX

The outlines of the variations of the cone index in areas A and B are shown in Figs. 3 and 4. It is obvious from these figures that it is difficult to divide the soil layer into more than two

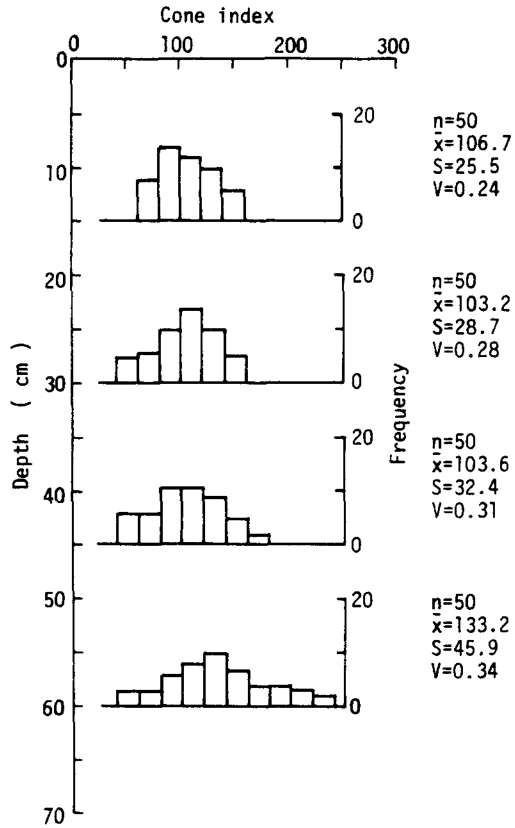


FIG. 3. Outline of variation of cone index (area A).

different parts in the engineering sense. Therefore, we cannot help regarding this kind of soil layer as a homogeneous layer. The criteria for a homogeneous soil layer have been discussed [6].

The soil layer which is considered to be most pertinent in determining soil trafficability is between about 15 and 40 cm depth. Cone indices of the 15–40 cm deep soil layers at the areas A and B were collected, and the variation of the collected cone indices was plotted in the form of the frequency distribution in Figs. 5 and 6, with the class interval chosen at 20. In connection with Figs. 5 and 6 the arithmetic mean $\bar{x}$, together with the S.D. $S$, and coefficient of variation $V$ have been calculated for each sample of size $n$. In addition, the equation of the theoretical normal curve corresponding to the above values has been calculated and a plot superimposed on each histogram. It was assumed in the latter calculations that the frequency distribution of the cone index is non-skewed and that the Gaussian distribution equation may be directly applicable. Although not strictly true, this assumption makes for a greatly simplified analysis and does not involve any significant error in calculating the standard deviation.

Cone index data in the other areas were similarly collected and investigated. The variations of cone indices in areas E and H were similarly indicated in the form of the frequency distribution in Figs. 7 and 8, with the class interval chosen at 10.

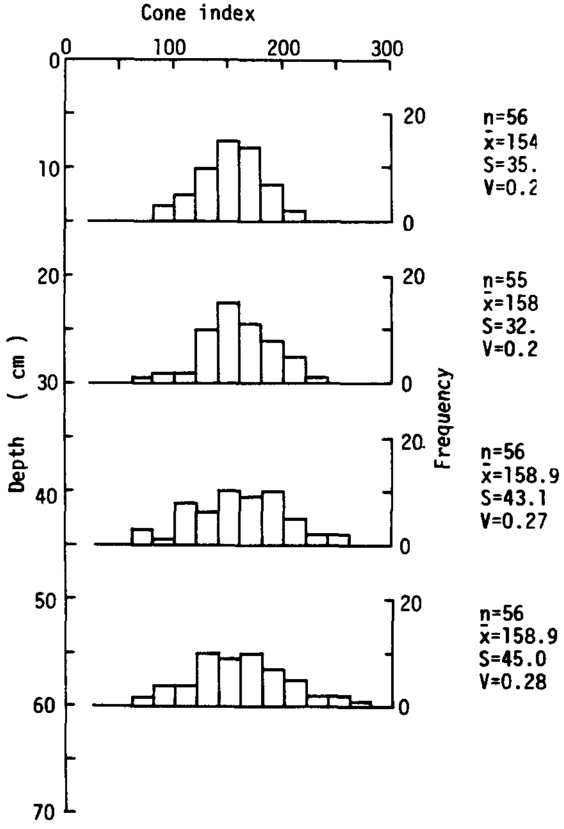


FIG. 4. Outline of variation to cone index (area B).

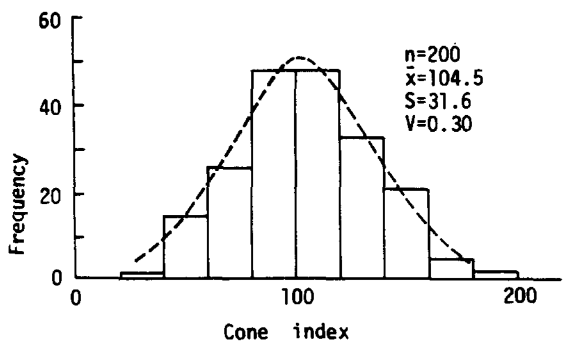


FIG. 5. Frequency distribution of cone index (between the depths of 15 and 40 cm, area A).

## DISCUSSION

The assumption that random variable values such as strength, density, water content, and other properties of soil from the population are described by normal distribution does not seem to be too far from being realistic [7, 8]. The normal density function is completely determined when we know the values of two parameters (population mean and standard

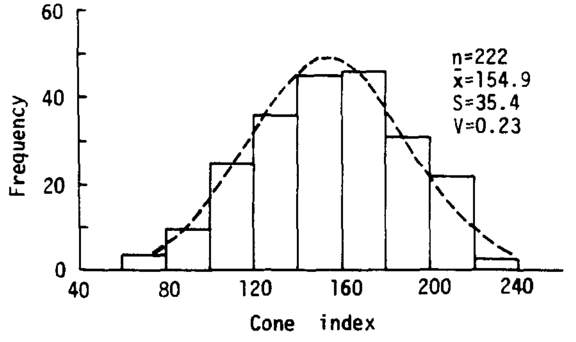


FIG. 6. Frequency distribution of cone index (between the depths of 15 and 40 cm, area B).

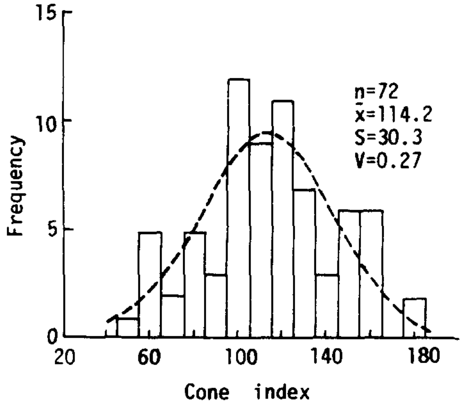


FIG. 7. Frequency distribution of cone index (between the depths of 15 and 40 cm, area E).

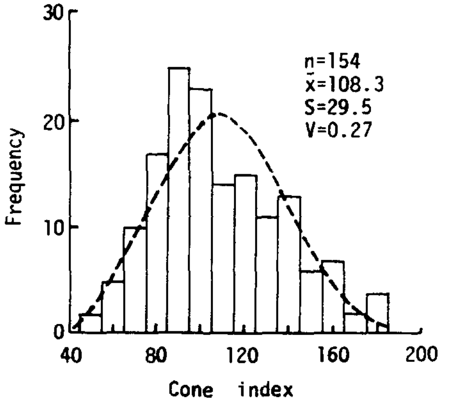


FIG. 8. Frequency distribution of cone index (between the depths of 15 and 40 cm, area H).

deviation) which describe it. Therefore, we must first find some method or procedure which will enable us to estimate these values with some degree of confidence.

The sample mean $\bar{x}$ is

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n} \tag{1}$$

where $x_1, x_2, \ldots . x_n$ is a random sample of size $n$ drawn from the population. $\bar{x}$ is an unbiased estimate of the population mean $\mu$. The variance $S^2$ of the sample is

$$S^2 = \sum_{i=1}^{n} \frac{(x_i - x)^2}{n - 1}. \tag{2}$$

The sample variance $S^2$ is an unbiased estimate of the population variance $\sigma^2$.

We also know from the Central Limit Theorem that the means of sample are also normally distributed.

$$t_{\alpha/2} = \frac{\sqrt{n}\,(x-\mu)^2}{S}. \tag{3}$$

The random variable $t$ is the random variable described in Student's-Distribution with $n$-1 degrees of freedom. Therefore, by rearranging equation (3) we can get

$$\mu = \bar{x} \pm t_{\alpha/2}\sqrt{\frac{S^2}{n}}$$

$$P_r\left\{\bar{x} - t_{\alpha/2}\sqrt{\frac{S^2}{n}} < \mu < \bar{x} + t_{\alpha/2}\sqrt{\frac{S^2}{n}}\right\} = 1 - \alpha. \tag{4}$$

With respect to $\mu$, we may say that we are $100(1\text{-}\alpha)$ % confident that $\mu$ will lie within the limits shown in equation (4).

If $S^2$ is the variance of a random sample of size $n$ from the normal population $N(x; \mu, \sigma^2)$, then $(n\text{-}1)S^2/\sigma^2$ has a Chi-square distribution with $n$-1 degrees of freedom. We can, then, state with a probability of 1-$\alpha$ that this random variable assumes a value between $\chi^2_{1-\alpha/2,\, n-1}$ and $\chi^2_{\alpha/2,\, n-1}$ or with a degree of confidence of 1-$\alpha$ that for a given sample

$$P_r\left\{\frac{(n-1)S^2}{\chi^2_{\alpha/2,\,n-1}} < \sigma^2 < \frac{(n-1)\,S^2}{\chi^2_{1-\alpha/2,\,n-1}}\right\} = 1 - \alpha. \tag{5}$$

This 1-$\alpha$ confidence interval for $\sigma^2$ may be converted into a 1-$\alpha$ confidence interval for $\sigma$ by taking square roots. The above information could be used to determine the reliability of the estimates.

When a theoretical distribution model is assumed from the shape of the frequency diagram or the plot of the data on a probability paper, the goodness of fit of the assumed model must be statistically tested. In this study, it is assumed that the values of random variable (cone index) may be described by the normal distribution, and the goodness of fit of the normal distribution is tested by Chi-square method.

Consider the random sample of the size $n$ drawn from the population in which we are interested. Let $f_1, f_2, \ldots f_k$ be the theoretical frequencies by the assumed normal distribution, and we use the following quantity $\chi^2$.

$$\chi^2 = \sum_{i=1}^{K} \frac{(f_i - F_i)^2}{F_i} . \tag{6}$$

The values of equation (6) have an approximate Chi-square distribution with $f = K - q$ degrees of freedom. Here $q$ is the number of the statistics which is used to determine the theoretical frequencies. The sample size $n$, sample mean $\bar{x}$ and sample standard deviation $S$ are used to determine the theoretical frequencies for the normal distribution and the number of degrees of freedom is $f = K - 3$. For the assumed theoretical distribution, if

$$\chi^2 = \sum_{i=1}^{K} \frac{(f_i - F_i)^2}{F_i} > \chi^2_{1-\alpha, f} \tag{7}$$

where $\alpha$ is the desired significance level of the test, then the hypothesis is rejected. If $\chi^2 < \chi^2_{1-\alpha, f}$, then the hypothesis cannot be rejected, and fairly close agreement between the observed and expected frequencies makes it reasonable to accept the hypothesis that the data come from a normal population. In other words, the normal distribution provides a good fit.

The results of the Chi-square tests for 5% risk (i.e. 95% confidence) are given in Table 2. The mean value, standard deviation and coefficient of variation are also shown in Table 2. The theoretical normal distribution curves were shown in Figs. 5–8. Figures 5–8 and Table 2 show that the cone index at depth of critical layer in homogeneous soil layer can be regarded as normal random variable, and it is important that the values of the coefficient of variation $V$ of these data are restricted in a narrow range from about 0.2 to 0.4 irrespective of the large difference in their mean values $\bar{x}$ and standard deviation $S$. These values of $V$ correspond with the values of soil strength reported in past studies [8, 9].

In the preceding discussions, we have considered the procedure for determining limits on the mean and variance of a normally distributed population, from which we may make a decision. In making a decision based on the preceding procedure, the primary difficulty is the determination of the confidence that we might have in our decision. Therefore, we need to

TABLE 2. $\chi^2$-TESTS OF CONE INDEX AND OTHERS

| Area | Sample size ($n$) | Mean ($\bar{x}$) | Standard deviation ($S$) | Value of $\chi^2$ ($\chi^2$) | Degrees of freedom ($f$) |
|---|---|---|---|---|---|
| A | 200 | 104.5 | 31.8 | 1.233 | 4 |
| B | 222 | 154.9 | 35.4 | 1.113 | 4 |
| C | 237 | 109.5 | 38.8 | 23.659* | 13 |
| D | 124 | 167.7 | 28.4 | 3.733 | 7 |
| E | 72 | 114.2 | 30.3 | 4.073 | 4 |
| F | 134 | 69.1 | 23.9 | 13.165 | 7 |
| G | 198 | 127.8 | 37.9 | 28.276** | 12 |
| H | 150 | 108.3 | 29.5 | 14.889 | 9 |

*was tested for = 1%
**was tested for = 0.5%

find a method which enables us to establish limits on the probability statements and, to which, we can attach a specified degree of confidence;

(1) First, we would like to know what is the probability of encountering an element less than or equal to some specified value $x_o$. For a normally distributed population, this probability can be written in the following manner

$$P_r(x \leqslant x_o) = \Phi\left(\frac{x_o - \mu}{\sigma}\right) = \int_{-\infty}^{x_o} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx. \tag{8}$$

This information could be used to determine, after a known number of measurements, the probability of a vehicle encountering a point which it could not negotiate.

(2) Second, we would like to know what is the probability of encountering at most $M$ elements less than or equal to $x_o$ out of a possible $K$ elements, where $M < K$. If we let $P = \Phi[(x_o-\mu)/\sigma]$ then we find that this probability could be written in the following manner;

$$P_r(u \leq M) = \sum_{u=0}^{M} \frac{K!}{u!\,(K-u)\,!} P^u (1-P)^{K-u}. \tag{9}$$

It is apparent, from the two preceding probability statements, that the first problem is to find bounds on the quantity $P = \Phi[(x_o-\mu)/\sigma]$ for a given confidence.

## SUMMARIZING REMARKS

There are two kinds of variation in the cone index data as well as in other soil properties. The first is due to the physical nature of soil and the second is attributed to technical problems. The variations due to many kinds of technical errors are mixed up in the data. Engineering efforts should be made in order to decrease the variations caused by technical problems, but usually it is very difficult to separate these two kinds of variations. Variations in the results of cone penetration tests cannot be eliminated completely. Therefore, it is important to investigate the total variabilities containing the two kinds of variations and the method to treat them.

The mean, standard deviation, coefficient of variation and others have been calculated, and the equation of the theoretical normal curve corresponding to the above values has been calculated. The cone index can be regarded as normal random variable in a uniform soil layer in an engineering sense, and the values of the coefficient of variation range from about 0.2 to 0.4. The assumption of the normal distribution allows for a greatly simplified analysis.

## REFERENCES


[1] GARY M. GARDNER, Forecasting soil strength measurements from a random sample, Davidson Laboratory, Stevens Institute of Technology, Report SIT-PL-71-1526 (1971).

[2] C. A. CARLSON and A. R. MCDANIEL, Forecasting trafficability of soils; Variability of physical properties of loess soils Warren County, Mississippi, Technical memo. No. 3-331, Waterways Experiment Station (1967).

[3] P. H. T. BECKETT, Terrain evaluation for off-road mobility; Criticism and models of terrain evaluation procedures for large areas. European Research Office, U.S. Army (1969).

[4] A. M. FREUDENTHAL, Safety and probability of structural failure. *Trans. ASCE* **121**, 1337 (1956).

[5] H. W. TIEN and M. K. LELAND, The probability of foundation safety. *Proceedings of ASCE*, SM5, pp. 213–231 (1967).

[6] M. MATSUO and K. KURODA, A stochastic study on some properties and failure probability for unsaturated soils. *Trans. JSCE* **4**, (3), 22–33 (1972).

[7] M. G. BEKKER. *Introduction to Terrain-Vehicle Systems*, The University of Michigan Press, Ann Arbor (1969).
[8] J. A. HOOPER and F. G. BUTLER, Some numerical results concerning the shear strength of London clay. *Geotechnique* **16**, 282–304 (1966).
[9] M. MATSUO and K. KURODA, Probabilistic approach to design of embankments. *Soil and Foundations* **14** (2), 1–17 (1974).