Prediction of Soil Properties from Simple Indices

Y-K. T. LO AND C. W. LOVELL

The Indiana geotechnical data bank has been established to collect geotechnical information from the subsurface investigation reports of various geotechnical projects previously conducted in the State. The data bank is user oriented, simple to use, and flexible to accommodate further change as the requirements of users become more clearly defined. Statistical operations on the data, varying from generation of sample distribution parameters of soil characteristics to correlation efforts among variables, are done to aid in evaluating information from new sites or in approximating geotechnical information for preliminary investigations. Correlations between simple soil indices and soil characteristics and properties that are costly to measure directly are discussed. Significant relations were developed for compressibility parameters under sustained loading and for compaction parameters under repeated transient loadings. Little success was achieved in predicting strength parameters from simple indices. The study establishes relations between soil characteristics, which may be valuable in preliminary studies. The information does not replace fuller site investigation, sampling, and testing but does produce a framework against which various test results can be judged for their consistency and reliability.

The need for pedologic and engineering soils information for use in planning, site selection, design, construction, and maintenance of transportation facilities is recognized by civil engineers. Data are necessarily limited in quantity and quality due to economic and time constraints. Although large amounts of detailed soil data are often available from work performed on adjacent or nearby projects, these data are usually not readily accessible for use or their existence is unknown.

Extensive laboratory and field test data have been accumulated by the Indiana State Highway Commission for use in characterizing the engineering properties of Indiana soils. The information is retained in the form of subsurface investigation reports, prepared by or for consulting design firms and government agencies for use in routine soil investigations. In its present voluminous form, most of the information is not very useful. The need existed to make this information more accessible for the engineer interested in detailed information on a site and the engineer interested in soil characteristics over a larger region.

In July 1977, research was initiated to develop and test a computerized information storage and retrieval system for soils in Indiana. The earlier parts of this research effort, which involved developing codes and storage techniques, have been reported elsewhere (1,2).

Phase 2 of the research is the subject of this paper. It involved the storage of 6934 additional data sets, for a total of 9442 data sets as of December 1979. These data sets were from roadway soil boring reports and from those reports on boring for bridge and culvert sites that contained laboratory test data. Additional purposes of phase 2 were to (a) show how to manage the data bank and (b) evaluate the information stored and develop correlations and quantitative values for planning and preliminary design by using statistical methods.

Both conventional and nonparametric statistical methods were used. However, the nonparametric statistical methods appear to fit and explain the varieties of soil characteristics in a superior way. The data were grouped by using physiographic regions, engineering soil classifications, soil associations, or a combination of these. One-way and two-way classification and factorial experiment layouts were used to examine the distributions of the data. Regression analysis was used to investigate the functional relations between design parameters and index properties. The total results are presented in a report by Lo (3). Because the soils data were extremely variable in their characteristics, choosing suitable groupings for study was the most difficult task in the investigation.

PREVIOUS INVESTIGATIONS

Concept of a Geotechnical Data Bank

In 1965, the South Dakota Department of Highways, in cooperation with the U.S. Soil Conservation Service, began a program to collect the accumulated geotechnical data from previous projects (4). The data were then stored in a computerized system and analyzed by using statistical methods for the purposes of planning, location, and preliminary design.

Spradling (5) developed a computerized data storage and retrieval system for the State of Kentucky. The Kentucky Department of Transportation devised an extensive coding system for data that were collected but not suitable for direct computer storage. Due to the completeness of this coding system and its applicability to soil information in general, some of its details were adopted for the Indiana data bank (1).

Recently, the U.S. Department of Transportation (DOT) published a state-of-the-art report that documented basic information on automatic data processing techniques used by eight states (Colorado, Georgia, Illinois, Louisiana, Minnesota, New York, Pennsylvania, and West Virginia) in managing test data for highway materials (6). Three basic data processing techniques--batch information systems, on-line interactive information systems, and on-line interactive laboratory information systems--were in use.

The Prairie Farm Rehabilitation Administration (PFRA) of Canada collected data on the geotechnical properties of glacial till deposits, glacial lake deposits, and alluvial deposits from many of its previous projects in western Canada (7,8). A number of empirical relations between routine classification tests and consolidation and strength characteristics were developed, as discussed below. The correlations were generated primarily to aid in evaluating information from new sites or to approximate geotechnical information for preliminary investigations.

Similar geotechnical data storage and retrieval systems were developed in the following countries: Sweden, Finland, and Denmark (9); France (10); Rhodesia (11); Algeria (12); and South Africa (13).

Empirical Relations in Practice of Geotechnical Engineering

Compressibility Parameters

The laboratory oedometer test is normally used to determine compressibility parameters such as compression index (C_c), recompression index (C_r), and preconsolidation pressure (p_c). The compression index (C_c) has often been correlated with either percentage liquid limit (w_L), percentage natural moisture content (w_n), initial void ratio (e_o), or a combination of these (3). The compression ratio (C_r') is defined as $C_c/(1 + e_o)$. It, too, is correlated with w_L, w_n, e_o, or a combination of these. Rutledge (14) and Fadum (15) showed that, as the natural moisture content increases, the compression ratio increases linearly for normally consolidated clays. Lo (3) gives a summary of available regression equations, together with their geographic region of applicability, for the prediction of C_r'.

By definition, p_c is the greatest effective pressure the soil has carried in the past. Preconsolidation may be caused by a variety of factors (16). The Canadian PFRA (7,8) has described a relation between liquidity index (LI) and log p_c, where as LI increases log p_c decreases linearly for the soils in western Canada. Bjerrum (17) developed a relation between the preconsolidation pressure/overburden pressure ratio and plastic index for late-glacial and postglacial clays.

Compaction Parameters and California Bearing Ratio

Compaction is defined as the densification of a soil by means of mechanical manipulation at constant water content and is measured quantitatively in terms of dry density of the soil. Proctor (18) demonstrated that there is a definite relation between density and moisture content. The characteristic peak in the Proctor curve is known as the maximum dry density (ρ_{dmax}) and its corresponding moisture content as optimum moisture content (OMC).

As the Atterberg liquid limit, plastic limit, or plastic index increases, OMC increases and ρ_{dmax} decreases. Furthermore, as OMC increases, ρ_{dmax} decreases. These relations have been investigated and verified by Woods and Litehiser (19), the U.S. Navy (20), Narayama Murty (21), and PFRA (7,8) for a variety of soils. Lo (3) gives a summary of regression equations, together with their geographic regions of applicability, for the prediction of OMC and ρ_{dmax}.

The California bearing ratio (CBR) test is used to provide a low-deformation measure of strength of compacted subgrade soil and is used with empirical curves to design asphalt pavement structures. In the literature, CBR values have been predicted by means of index properties, strength characteristics, and soil classification units. A relation between the CBR and the group index (GI) was suggested by the Asphalt Institute (22) and later by Gawith and Perrin (23). Both CBR values were measured at 90 percent modified American Association of State Highway and Transportation Officials (AASHTO) maximum dry density. As the GI increases, the CBR decreases. Kassiff, Livneh, and Wiseman (24) found that, for Israel soils, the CBR values increase with decrease in the difference between plastic limit (w_p) and shrinkage limit (SL) and with increase in surcharge.

Robinson and Lewis (25, p. 72) showed that a relation existed between the CBR value and failure load P (in pounds) of a 3-in-square plate pushed into the ground. A rational approach known as the suction method was proposed by Black (26) to estimate CBR values from plasticity and consistency indices. Recently, Black and Lister (27) found that CBR = $c_u/23$, where c_u is the undrained shear strength in kilopascals.

The U.S. Navy (20) correlated the values of typical characteristics such as ρ_{dmax}, OMC, and CBR of compacted materials against the Unified Soil Classification system. The American Hoist and Derrick Company (28) used the AASHTO classification system, along with qualitative descriptions of soil characteristics, to approximate CBR values. Having made a comparison of groups in the AASHTO and Unified Soil Classification systems, Liu (29) proposed approximate relative relations of various groups of both systems to CBR values.

Strength Parameters

The British Road Research Laboratory (30) described a curvilinear relation between unconfined compressive strength (q_u) and w_n for a heavy clay in situ. Peters and Lamb (7) presented an empirical relation between LI and q_u for the soils in western Canada. They showed that, as LI decreases, the unconfined compressive strength increases. Peck, Hanson, and Thornburn (31) suggested a relation between the qualitative terms describing consistency, along with field identification of clays, and the quantitative values of q_u. The National Research Council of Canada (32) also suggested a similar relation for the rough estimate of the undrained shear strength (half of the unconfined compressive strength) for clay soils.

Unconfined compressive strength (q_u) has been correlated with the standard penetration test (SPT)--i.e., the number of blows (N-values) for 1-ft penetration. The U.S. Navy (20) recommended simple relations for clayey silts, CL clays, or varved clays and silts. Terzaghi and Peck (33) also suggested relations among the consistency of clay described in qualitative terms, N-values, and q_u. Sanglerat (34) recommended additional relations.

Summary

A review of the large number of empirical relations documented in the geotechnical engineering literature leads to two principal conclusions:

1. The values of soil parameters are expressed as means.
2. The functional relations among soil characteristics are established with data from a certain region or pooled data from several regions. The regional effects on relations are not commonly investigated or compared.

The random process of soil formation explains the great variation often encountered for a given soil parameter. Therefore, it seems more reasonable to define and use the median, rather than mean, for soil parameter values. The regional effects on functional relation among soil characteristics can be investigated by using the techniques of qualitative variables as regressors associated with analysis of variance. The details of these procedures are discussed and illustrated in the following section.

DEVELOPMENT OF GEOTECHNICAL DATA BANK FOR INDIANA

Goldberg (1,2) initiated the development of the Indiana geotechnical data bank. A total of 2508 data sets were collected and subjected to statistical analyses in an initial research phase. An additional 6934 data sets were added to the bank in the subsequent year and subjected to more detailed statistical analyses. As of January 1980, the Indiana geotechnical data bank contained 9442 data sets. The distribution of data sets throughout the state is presented by Lo (3).

Source and Structure of Data

Both geotechnical and pedological soils information was collected. The geotechnical information was taken from the subsurface investigation reports of various geotechnical projects previously conducted in Indiana, and the pedological soils information was from recent U.S. Department of Agriculture soil survey manuals (35) and general soils maps (36).

A data input form was developed to record the information. Details of the listing and their corresponding coding systems are described by Goldberg (1). Computer programs were written in the Statistical Package for the Social Sciences (SPSS) language available in the Purdue University Computing Center for CDC 6500 and 6600 systems. The SPSS language, like any other recent statistical language, is a "conversational" statistical analysis software. It also has features of data manipulation, data transformation, file definition, and file creation (37). This study has relied heavily on the SPSS language. However, any computer language is merely a means of access to an optimum use of the computer hardware. Therefore, the logical sequences of a program are more important than the program itself.

Methods of Analysis

Two types of prediction equations were used in this work: (a) median models and (b) regression models. To define numerically the variability of selected soil characteristics, the frequency distributions of these characteristics are examined and described. One way to describe the sample distribution is to use the conventional constant mean model (38), which is based on the assumption of normality of the population distribution. This model is characterized by the mean and the standard deviation of the sample distribution. It is shown (39) that in most cases this model is also effective for moderately abnormal distributions of a large sample. In the case of small samples, however, this model may not give accurate approximations. The nonparametric or distribution-free methods are the preferred techniques of inference for nonnormal population (40). These methods make a minimum of assumptions regarding the sample distribution and are generally appropriate for any form of the distribution. They are of high efficiency in comparison with classical techniques, under the assumption of normality, and are often of higher efficiency in other situations (39-41). In this study, the sample distributions were sometimes normal but frequently skewed or bimodal. The median model was used to characterize the sample distribution.

Distributional data are covered in detail by Lo (3) and will not be further considered here.

Regression Models

Regression analysis provides a conceptually simple method for investigating functional relations among variables. In general, the first stage of the analysis is to select the variables to be included in the regression model. This is done based on theory or former examples or by other procedures.

The most thorough approach, known as the all-possible-regression method, is to develop the regression of y (dependent variable) on every subset of the kx variables (independent variables). The major drawback of this method is the amount of computation. Another approach for selecting variables, and the one used in this study, is the stepwise regression method (42,43). It is recommended that the stepwise procedures be applied only to noncollinear data and that the order in which the variables enter or leave the equation not be interpreted as reflecting the relative importance of the variables (42).

In entering the variables to formulate a regression model, a question arises concerning the form of each variable--i.e., whether the variable should enter the model as an original variable x or as some transformed variable such as x^2, log x, or a combination or both. If, from an examination of scatter plots of y versus x, the relation between y and x appears to be nonlinear, appropriate transformations of the data are introduced to produce linearity (42). In this study, all variables, their possible transformations, and their combinations were included in the stepwise procedure for selecting variables as long as they were not collinear.

The variables were selected to minimize the mean square due to error of the prediction. Because a large value of R^2 [square of (multiple) correlation coefficient] or a significant t-statistic does not ensure that the data were well fitted (44), a careful residual analysis was also made. The procedure to reduce the number of independent variables was to compare the full model and the reduced model by using the F-statistic (42).

It was believed that the soil was more homogeneous in a small geologic or pedologic unit. Therefore, the regression models were established on these units. It was often found that for a given dependent variable the regression models generated in this way used different sets of independent variables for various locations. It seems unwise to conclude that these differences are caused by soil differences alone.

The effects of soil location and genesis--namely, physiographic region and parent material--were investigated by means of the statistical technique of using qualitative variables as regressors (42). In order to do so, the qualitative variables were represented by dummy variables that take on only two values, usually zero and one. These two values designated whether the observation belonged in one of two possible categories. Accordingly, the number of these variables required was one less than the number of categories in a grouping unit. Goldberg (1) shows that for Indiana the physiographic regions are coded from 1 to 12 and the parent materials from 1 to 13. The dummy-variable indicators were set up as follows:

$$x_i \begin{cases} = 1 \text{ if soil sample is taken from physiographic region coded i} \\ = 0 \text{ otherwise} \end{cases} \qquad (1)$$

where i = 1,2,3, ..., 11; and

$$z_j \begin{cases} = 1 \text{ if soil sample is derived from parent material coded j} \\ = 0 \text{ otherwise} \end{cases} \qquad (2)$$

where j = 1,2,3, ..., 12.

For the soil sample taken from the physiographic region coded 13, let $x_1 = x_2 = x_3 = \ldots = x_{12} = 0$. For the soil sample derived from the parent material coded 12, let $z_1 = z_2 = z_3 = \ldots = z_{11} = 0$.

Assume that the following relation exists:

$$C_c = a_0 + a_1 w_n + a_2 e_o + a_3 w_L \qquad (3)$$

To investigate how these two soil grouping units affect this relation singly or in combination, the F-statistics were used in making comparisons of the following models:

Model	Equation
1	$C_c = a_0 + a_1 w_n + a_2 e_o + a_3 w_L$
2	$C_c = a_0' + a_1' w_n + a_2' e_o + a_3' w_L + a_4' x_1 + a_5' x_2 + \ldots + a_{15}' x_{12}$
3	$C_c = a_0'' + a_1'' w_n + a_2'' e_o + a_3'' w_L + a_4'' z_1 + a_5'' z_2 + \ldots + a_{14}'' z_{11}$
4	$C_c = a_0''' + a_1''' w_n + a_2''' e_o + a_3''' w_L + a_4''' x_1 + a_5''' x_2 + \ldots + a_{15}''' x_{12} + a_{16}''' z_1 + a_{17}''' z_2 + \ldots + a_{26}''' z_{11}$

Application of Regression Models

The regression models were used to correlate soil design parameters, such as those of compaction, consolidation, and strength, with soil index properties. The following examples illustrate how to apply the procedures described in the previous sections to arrive at the regression model for ρ_{dmax} and for C_c by using index properties.

It was found that ρ_{dmax} correlated well with w_L and w_p. There should also be an examination of whether the interaction between w_L and w_p--i.e., $(w_L \cdot w_p)$--has any significant effect on ρ_{dmax}. Therefore, the following independent variables enter the stepwise regression program for consideration: w_L, w_L^2, w_p, w_p^2, $(w_L \cdot w_p)$, percentage of sand, and percentage of silt.

The F-statistic tests were used to reduce the number of independent variables and produce the final model in the following form:

$$\hat{\rho}_{d\,max} = -0.554\,w_L - 0.0900\,\text{silt} - 0.727\,w_p + 0.008\,49\,w_L w_p + 142.888 \tag{4}$$

where

$|R| = 0.808$,
SD = 4.994, and
n = 601.

For a given predicted ρ_{dmax} $(\hat{\rho}_{dmax})$, about 68 percent of sample observations (measured ρ_{dmax}) fall in the range of $\hat{\rho}_{dmax}$ + 4.994 (pcf) and $\hat{\rho}_{dmax}$ - 4.994 (pcf).

To investigate the effects of physiographic regions and parent materials on Equation 4, the dummy-variable indicators were set up as described previously. The F-statistic tests were then used to examine the value of physiographic regions and parent materials in the model. Neither added significant information (3). Therefore, Equation 4 was the final model for ρ_{dmax}.

For the soil variable of C_c, it was found that effects of both physiographic regions and parent materials did add significant information statistically to the regression model of C_c on index properties. The regression model was found to be

$$\begin{aligned}\hat{C}_c = & -0.151 + 0.003\,26\,w_n + 0.191\,e_o + 0.003\,25\,w_L \\ & + 0.0162\,x_1 - 0.0110\,x_2 + 0.0208\,x_3 + 0.0296\,x_4 \\ & + 0.0120\,x_5 - 0.0110\,x_6 + 0.0365\,x_7 + 0.0351\,x_8 \\ & + 0.0646\,x_9 + 0.0649\,x_{10} - 0.0594\,x_{11} - 0.0245\,z_1 \\ & - 0.0313\,z_2 - 0.009\,87\,z_3 - 0.0917\,z_4 - 0.121\,z_6 \\ & - 0.0292\,z_7 - 0.0667\,z_8 + 0.008\,41\,z_9 - 0.0418\,z_{10} \\ & - 0.008\,84\,z_{11}\end{aligned} \tag{5}$$

where

$|R| = 0.952$,
SD = 0.0670, and
n = 302.

Figures 1 and 2 show the scatter plots with regression lines and 95 percent population confidence intervals for the measured ρ_{dmax} (ρ_{dmax}) and predicted ρ_{dmax} $(\hat{\rho}_{dmax})$ determined by using Equation 4 and the measured C_c (C_c) and predicted C_c $(\hat{C}_c)$ determined by using Equation 5. In the presentation of the data, the solid line represents the best fit line whereas the dashed lines define the boundaries of 95 percent population confidence intervals.

DISCUSSION OF RESULTS

In this study, both median and regression models were developed for statistical forecasting. Predictions are extrapolations into the future of features shown by relevant data in the past. Therefore, a considerable population of values for the dependent variable is required. Another basic requirement for prediction is the existence of a stable data structure. The trend of the data and the statistical variation about the trend must be stable. This can be detected by the confidence intervals and interquartile range for a median model or the standard deviation of estimate and multiple correlation coefficient (|R|) for a regression model. A large difference between the confidence intervals, a large value of interquartile range for a median model, or a large standard deviation of estimate for a regression model indicates that the data structure is not stable. Either more data or a change in grouping unit is needed for better prediction.

Regression Models and Correlations

Regression models were used to correlate soil design parameters such as compaction, consolidation, and strength with index properties. In Table 1, a number of successful regression equations are represented together with their correlation coefficients (R) or multiple correlation coefficients (|R|), standard deviations (or errors) of estimate (SD), and the number of cases (n). The standard deviation of estimate is important, since it represents the variation of estimate $(\hat{y})$; i.e., 68 percent of sample observations (y) fall in the range of $\hat{y}$ - SD and $\hat{y}$ + SD. Specific results are presented below.

Compaction Parameters and CBR

The correlations of ρ_{dmac} and OMC versus plasticity characteristics, as given in Table 1, indicate that, as w_L or w_p increases, OMC increases but ρ_{dmax} decreases. In addition, as OMC increases, ρ_{dmax} decreases. A relation between CBR values at 100 and 95 percent maximum dry densities is also developed and presented in Table 1.

Consolidation Parameters

With either w_L, w_n, or e_o, C_c increases. The quantity of C_r' is a linear function of e_o, as given in Table 1. C_r is usually taken as a fraction of C_c. In this study, it is found that $C_r = A + B\,C_c$, where A is -0.003 27. The standard deviation of A is 0.001 99, B is 0.139, and the standard deviation of B is 0.007 26. Therefore, with 95 percent confidence the C_r will lie in the range of, approximately, $(1/6.5)C_c$ and $(1/8)C_c$ for Indiana soils. p_c was correlated with LI, but this correlation was not a strong one for Indiana soils.

Strength Parameters

The relations of strength and strength parameters with simple index values were examined with little general success (3). However, w_L was found to be a function of w_n, SPT, and location factors--i.e., physiographic regions (x's) and parent material (z's), as given in Table 1.

CONCLUSIONS

A computerized data storage and retrieval system has been developed for the State of Indiana. Both conventional and nonparameteric statistical methods have been used in the analysis of these data. Regression analysis (Table 1) showed certain good to adequate functional relations between design parameters and index properties, as follows:

1. As the liquid limit or plastic limit of a soil increases, the optimum moisture content increases but maximum dry density decreases. This confirms the findings of earlier authors.

2. A unique relation exists between optimum moisture content and maximum dry density, which confirms the findings of earlier authors.

3. The CBR value is a function of plasticity characteristics, and a correlation exists between the CBR value at 100 percent maximum dry density and the CBR value at 95 percent maximum dry density, which confirms the findings of earlier authors.

4. Compression index (C_c) is a function of natural moisture content, initial void ratio, and liquid limit and is significantly influenced by geological factors, which confirms the findings of earlier authors. With 95 percent confidence, the recompression index (C_r) lies in the range of, approximately, $(1/6.5)C_c$ and $(1/8)C_c$ for Indiana soils.

5. The preconsolidation pressure (p_c) is a function of natural water content, initial void ratio, and natural dry density and is significantly influenced by geological factors. But the scatter in the data is large.

6. No definite correlation of strength angle and cohesion intercept versus plasticity characteristics was found for Indiana soils.

7. It is emphasized that the data bank is not proposed as a substitute for fuller site investigation, sampling, and testing but as a framework against which various test results can be judged for their consistency and reliability. It can also be enormously helpful in preliminary assessment of a site or route.

ACKNOWLEDGMENT

The research described in this paper was carried out under the sponsorship of the Joint Highway Research Project of Purdue University, the Indiana State Highway Commission, and the Federal Highway Admin-

Figure 1. Measured maximum dry density versus predicted maximum dry density for Indiana soils.

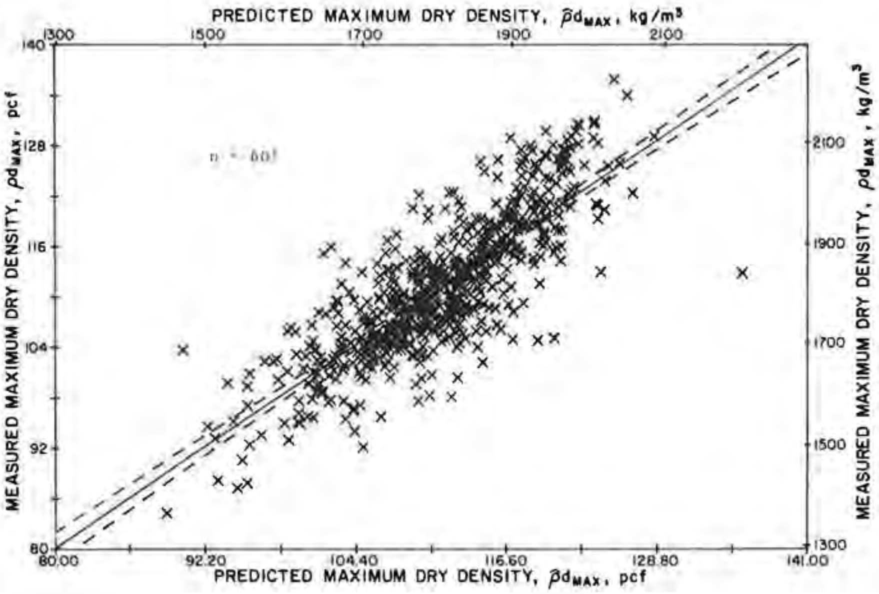

Figure 2. Measured compression index versus predicted compression index for Indiana soils.

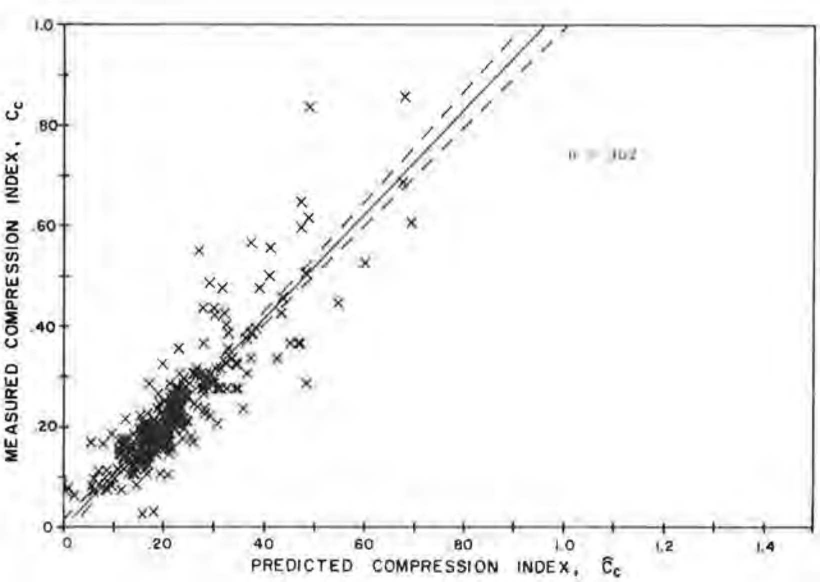

Table 1. Summary of regression equations and correlations.

Parameter	Equation	Statistical Measure
Compaction		
$\rho_{d\,max}$	$\hat{\rho}_{d\,max} = 135.843 - 1.279\,w_p$	R = -0.692 SD = 6.434 n = 601
	$\hat{\rho}_{d\,max} = 128.338 - 0.463\,w_L$	R = 0.744 SD = 5.651 n = 601
	$\hat{\rho}_{d\,max} = 142.888 - 0.554\,w_L - 0.727\,w_p + 0.008\,49\,w_L w_p - 0.0900\,\text{silt}$	|R| = 0.808 SD = 4.994 n = 601
OMC	$\widehat{OMC} = 4.464 + 0.619\,w_p$	R = 0.698 SD = 2.905 n = 596
	$\widehat{OMC} = 7.626 + 0.237\,w_L$	R = 0.794 SD = 2.464 n = 596
	$\widehat{OMC} = 7.457 - 0.0369\,\text{sand} + 0.0174\,\text{silt} + 0.171\,w_L + 0.155\,w_p - 0.413\,x_1 - 0.401\,x_2 - 1.740\,x_3 + 0.968\,x_4 - 0.409\,x_5 + 0.503\,x_6 - 0.186\,x_7 - 0.560\,x_8 + 0.386\,x_9 - 0.463\,x_{10} - 0.752\,x_{11}$	|R| = 0.843 SD = 2.210 n = 596
$\rho_{d\,max}$ versus OMC	$\hat{\rho}_{d\,max} = 150.667 - 3.016\,\text{OMC} + 0.0333\,(\text{OMC})^2$	|R| = 0.906 SD = 3.691 n = 701
CBR at 100 percent $\rho_{d\,max}$ (CBR S01)	$\log \widehat{CBR}\,S01 = 1.204 + 0.145 \log w_L - 0.137 \log PI - 0.149 (\log PI)(\log w_L) - 0.0778\,z_1 - 0.109\,z_2 + 0.112\,z_3 - 0.0505\,z_4 - 0.0858\,z_5 - 0.122\,z_6 - 0.0575\,z_7 - 0.0633\,z_8 - 0.0587\,z_9 - 0.105\,z_{10} - 0.688\,z_{11}$	|R| = 0.620 SD = 0.166 n = 493
CBR S01 versus CBR at 95 percent $\rho_{d\,max}$ (CBR S02)	$\widehat{CBR}\,S02 = 1.339 + 0.433\,\text{CBR S01}$	R = 0.851 SD = 2.010 n = 553
	$\widehat{CBR}\,S02 = 0.051 + 0.667\,\text{CBR S01} - 0.007\,60\,(\text{CBR S01})^2$	|R| = 0.864 SD = 1.515 n = 553
Consolidation		
C_c	$\hat{C}_c = 0.007\,97\,(w_L - 8.16)$	R = 0.829 SD = 0.116 n = 312
	$\hat{C}_c = 0.0126\,w_n - 0.162$	R = 0.925 SD = 0.112 n = 332
	$\hat{C}_c = 0.496\,e_o - 0.195$	R = 0.873 SD = 0.143 n = 335
	$\hat{C}_c = -0.151 + 0.003\,26\,w_n + 0.191\,e_o + 0.003\,25\,w_L + 0.0162\,x_1 - 0.0110\,x_2 + 0.0208\,x_3 + 0.0296\,x_4 + 0.0120\,x_5 - 0.0110\,x_6 + 0.0365\,x_7 + 0.0351\,x_8 + 0.0646\,x_9 + 0.0649\,x_{10} + 0.0594\,x_{11} - 0.0245\,z_1 - 0.0313\,z_2 - 0.009\,87\,z_3 - 0.0917\,z_4 - 0.121\,z_6 - 0.0292\,z_7 - 0.0667\,z_8 + 0.008\,41\,z_9 - 0.0418\,z_{10} - 0.008\,84\,z_{11}$	|R| = 0.952 SD = 0.0670 n = 302
C_r'	$\hat{C}_r' = 0.0125 + 0.152\,e_o$	R = 0.704 SD = 0.0448 n = 333
	$\hat{C}_r' = 0.0249 + 0.003\,w_n$	R = 0.701 SD = 0.0361 n = 325
	$\hat{C}_r' = 0.0294 + 0.002\,38\,w_L$	R = 0.665 SD = 0.0373 n = 309
C_c versus C_r'	$\hat{C}_c = 0.0844 + 9.121\,(C_r')^2$	|R| = 0.948 SD = 0.0928 n = 339
C_r versus C_c	$\hat{C}_r = -0.003\,27 + 0.139\,C_c$	R = 0.743 SD = 0.0173 n = 298
Strength		
LL versus SPT	$w_L = 47.946 + 0.495\,w_n - 9.931 \log SPT + 0.488\,w_n \log SPT - 19.323\,x_1 - 25.121\,x_2 - 30.455\,x_3 - 17.775\,x_4 - 28.875\,x_5 - 24.529\,x_6 - 26.068\,x_7 - 26.272\,x_8 - 24.100\,x_9 - 16.493\,x_{10} - 37.182\,x_{11} + 9.104\,z_1 - 2.930\,z_2 - 3.809\,z_3 - 0.378\,z_4 - 12.178\,z_5 - 5.404\,z_6 - 7.021\,z_7 + 23.069\,z_8 - 0.249\,z_9 + 0.962\,z_{10} + 13.878\,z_{11}$ (in %)	|R| = 0.859 SD = 19.503 n = 533

istration. Their support is gratefully acknowledged. We thank R.D. Miles for his assistance throughout the project, including valuable editorial aid; G.P. McCabe for his statistical advice; and T.R. West for valuable geological advice. W.J. Sisiliano, R.W. Rahn, and C. Andrews of the Indiana State Highway Commission gave valuable guidance throughout this work.

REFERENCES

1. G.D. Goldberg. Development of the Computerized Geotechnical Data Bank for the State of Indiana. Purdue Univ., West Lafayette, IN, M.S.C.E. thesis and Joint Highway Research Project Rept. 76-6, May 1978, 163 pp.
2. G.D. Goldberg, C.W. Lovell, and R.D. Miles. Use

the Geotechnical Data Bank! TRB, Transportation Research Record 702, 1978, pp. 140-146.
3. Y.-K.T. Lo. Geotechnical Data Bank for Indiana. Purdue Univ., West Lafayette, IN, Ph.D. thesis and Joint Highway Research Project Rept. 80-7, Aug. 1980.
4. R.A. Crawford, J.B. Thomas, and M. Stout, Jr. Computerized Soil Test Data for Highway Design. Physical Research Section, South Dakota Department of Highways, Pierre, 1972, 24 pp.
5. D. Spradling. A Soils Data System for Kentucky. Division of Research, Bureau of Highways, Kentucky Department of Transportation, Frankfort, Res. Rept. 441, 1976, 58 pp.
6. Data Transformation Corporation. Materials and Test Data Information Systems in Highway Construction. FHWA, Rept. FHWA-TS-78-221, 1978, 156 pp.
7. N. Peters and K.N. Lamb. Experiences with Alluvial Foundations for Earth Dams in the Prairie Provinces. Canadian Geotechnical Journal, Vol. 16, No. 2, 1979, pp. 255-271.
8. P.J. Rivard and T.E. Goodwin. Geotechnical Characteristics of Compacted Clays for Earth Embankments in the Prairie Provinces. Canadian Geotechnical Journal, Vol. 15, No. 3, 1978, pp. 391-401.
9. S.E. Lundin, O. Stephansson, and P. Zetterlund. Geoteknisk Databank. National Swedish Building Research Institute, Stockholm, Rept. R70, 1973, 158 pp.
10. Lille Regional Laboratory. Establishment of a Card Index for Borings and Tests. Ministry of Public Works, Paris, France, Jan. 1971.
11. A. Holden. Engineering Soil Mapping from Airphotos. Ministry of Roads and Traffic, Salisbury, Rhodesia, July 1969.
12. J.P. Redon. Computerized Inventory of Soil Profiles Throughout Algeria. National Laboratory of Public Works and Buildings, Algiers, Algeria, Sept. 1972.
13. Acquisition and Use of Geotechnical Information. NCHRP, Synthesis of Highway Practice 33, 1976, 40 pp.
14. P.C. Rutledge. Compression Characteristics of Clays and Application to Settlement Analysis. Harvard Univ., Cambridge, MA, thesis, 1939.
15. R.E. Fadum. Observations and Analysis of Building Settlements in Boston. Harvard Univ., Cambridge, MA, thesis, 1941.
16. G.A. Leonards. Engineering Properties of Soils. In Foundation Engineering (G.A. Leonards, ed.), McGraw-Hill, New York, 1962, pp. 145-156.
17. L. Bjerrum. Embankments on Soft Ground. Proc., ASCE Specialty Conference on Performance of Earth and Earth-Supported Structures, Lafayette, IN, Vol. 11, 1972, pp. 1-54.
18. R.R. Proctor. Fundamental Principles of Soil Compaction. Engineering News-Record, Vol. 3, Aug. 31, 1933, pp. 245-248.
19. K.B. Woods and R.R. Litehiser. Soil Mechanics Applied to Highway Engineering in Ohio. Engineering Experiment Station, Ohio State Univ., Columbus, Bull. 99, 1938.
20. Design Manual: Soil Mechanics, Foundations, and Earth Structures. Naval Facilities Engineering Command, U.S. Department of the Navy, Alexandria, VA, NAVFAC DM-7, March 1971, pp. 7-3-8, 7-9-2, and 7-10-1 to 7-10-23.
21. P.L. Narayama Murty. Relationship Between the Plasticity Characteristics and Compaction of Soils. Journal of Indian National Society of Soil Mechanics and Foundation Engineering, Vol. 4, No. 4, 1965, pp. 455-463.
22. Thickness Design: Asphalt Pavement Structures for Highways and Streets. Asphalt Institute, College Park, MD, MS-1, Dec. 1965, 95 pp.
23. A.H. Gawith and C.C. Perrin. Development in the Design and Construction of Bituminous Surfaced Pavements in the State of Victoria, Australia. Presented at International Conference on Structural Design of Asphalt Pavements, Univ. of Michigan, Ann Arbor, 1962, pp. 897-910.
24. G. Kassiff, M. Livneh, and G. Wiseman. Pavements on Expansive Clays. Jerusalem Academic Press, Jerusalem, Israel, 1969, pp. 103-117.
25. P.J.M. Robinson and T. Lewis. A Rapid Method of Determining In Situ CBR Values. Geotechnique, Vol. 8, No. 2, 1958.
26. W.P.M. Black. A Method of Estimating CBR of Cohesive Soils from Plasticity Data. Geotechnique, Vol. 12, No. 4, 1962, pp. 271-282.
27. W.P.M. Black and N.W. Lister. The Strength of Clay Fill Subgrades: Its Predictions in Relation to Road Performance. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1979, 30 pp.
28. Handbook of Soil Compactionology. American Hoist and Derrick Co., St. Paul, MN, 1977, 53 pp.
29. T.K. Liu. A Review of Engineering Soil Classification Systems. HRB, Highway Research Record 156, 1967, pp. 1-22.
30. Soil Mechanics for Road Engineers. British Road Research Laboratory, London, England, 1952.
31. R.B. Peck, W.E. Hanson, and T.H. Thornburn, eds. Foundation Engineering, 2nd ed. Wiley, New York, 1974, pp. 20, 114, 307-330.
32. Canadian Manual On Foundation Engineering. National Research Council of Canada, Ottawa, Ontario, draft, 24 pp.
33. K. Terzaghi and R.B. Peck. Soil Mechanics in Engineering Practice, 2nd ed. Wiley, New York, 1967, pp. 73, 112, 193-216.
34. G. Sanglerat. The Penetrometer and Soil Exploration. Elsevier Publishing Co., Amsterdam, Netherlands, 1972, 464 pp.
35. Soil Survey Manuals. Soil Conservation Service, U.S. Department of Agriculture, 1958 to present.
36. General Soils Maps and Interpretation Tables for the Counties of Indiana. Agriculture Experiment Station/Cooperative Extension Service, Purdue Univ., West Lafayette, IN; Soil Conservation Service, U.S. Department of Agriculture, Nov. 1971.
37. N.H. Nie and others. Statistical Package for the Social Sciences, 2nd ed. McGraw-Hill, New York, 1975, 675 pp.
38. W. Gilchrist. Statistical Forecasting. Wiley, New York, 1976, 300 pp.
39. H.J. Larson. Introduction to the Theory of Statistics. Wiley, New York, 1973, 242 pp.
40. G.W. Snedecor and W.G. Cochran. Statistical Methods, 6th ed. Iowa State Univ. Press, Ames, 1967, 593 pp.
41. M. Hollander and D.A. Wolfe. Nonparametric Statistical Methods. Wiley, New York, 1973, 503 pp.
42. S. Chatterjee and B. Price. Regression Analysis by Example. Wiley, New York, 1977, 228 pp.
43. N.R. Draper and H. Smith. Applied Regression Analysis. Wiley, New York, 1966, 407 pp.
44. F.J. Anscombe. Graphs in Statistical Analysis. American Statistician, Vol. 27, 1973, pp. 17-21.

Publication of this paper sponsored by Committee on Exploration and Classification of Earth Materials.