Characterizations of Indiana Soil Properties

by

Y-K T. Lo[1] and G. P. McCabe[2]

Technical Report #84-5

Department of Statistics
Purdue University

February, 1984

[1]Y-K T. Lo, Geotechnical Engineer, Woodward-Clyde Consultants, 202 North Golden Circle Drive, Santa Ana, CA 92705

[2]G. P. McCabe, Professor of Statistics, Department of Statistics, Purdue University, West Lafayette, Ind. 47907

Characterizations of Indiana Soil Properties

Y-K T. Lo[1] and G. P. McCabe[2]

Abstract

The Indiana Geotechnical Data Bank was established to collect engineering soil data in the state and to provide information for engineering planning and design works. It contains nearly 10,000 sets of soil data.

The soil data are grouped by using physiographical regions, engineering classifications, pedological soil classifications, or a combination of these. Robust statistical methods are employed to examine the sample distributions of the data and to determine distribution parameters. Various sample comparison methods are applied to refine the grouping units. Regression analysis is used to predict design parameters from index properties, locations and parent materials.

For the engineering soil analysis in this study, robust statistical methods are preferred to conventional parametric methods. Grouping units of soil data and their distribution parameters are established. A good homogeneity of soil characteristics is obtained with these groupings. The equations developed to predict soil design parameters from index properties are reasonably accurate. Finally, practical uses of this data bank for typical highway projects are suggested.

Introduction

The computerized Indiana Geotechnical Data Bank was established to collect pedologic and engineering soil information from private consulting firms, soil testing laboratories and from tests conducted by Indiana Department of Transportation. A data storage and retrieval system is used to generate information for building highways, bridges and other facilities in the state.

The methodology used in forming this bank and its development are discussed in the reports by Goldberg (1) and Lo (4). In addition, the nature and some uses of the Bank are described in papers by Goldberg, et al (2) and Lo and Lovell (5). The Bank is user oriented and flexible in the sense that both data in the same categories and data from other categories can be added as needed.

For a typical highway subsurface investigation, data generally is taken at limited depths and only soil classification tests are involved. Bridge locations receive more detailed attention. In general the geotechnical data are widely scattered in area and limited to simple index properties.

---

[1] Y-K T. Lo, Geotechnical Engineer, Woodward-Clyde Consultants, 202 North Golden Circle Drive, Santa Ana, CA 92705
[2] G. P. McCabe, Professor of Statistics, Department of Statistics, Purdue University, West Lafayette, Ind. 47907

Visual textural classification and color in a moist condition are available for all samples. If the sample is to be classified according to the Unified System or the AASHO System, grain size distribution and or Atterberg limits are required as well.

Other variables in the data bank include
organic content,
natural water content,
in-place densities (wet or dry),
specific gravity,
compaction curves,
California bearing ratio (CBR) for compacted and soaked soils,
unconfined undrained strength and other strength parameters,
compressibility parameters,
standard penetration test (SPT) values particularly for sand,
depth to the water table and depth to bedrock, if encountered,
boring location - township, range and section.

Pedologic maps are used to identify the soil association and the series. The usual generalized ratings (available for pedologically classified soils) are assigned to samples as appropriate. These ratings include: erosion class, natural drainage class, permeability, flooding potential, frost heave susceptibility, shrink-swell potential, and pH. It is also possible to access the regional geomorphic (physiographic) location of boring, and to group data according to these units.

Data can be retrieved for analysis by any description of location, depth or type. They may be simply printed out and examined qualitatively, or subjected to statistical analysis.

Descriptive Statistical Analysis

To describe the variability of selected soil characteristics, the frequency distributions of these characteristics are examined. Robust estimates, such as median, confidence interval of median, percentiles, and interquartile range (7) are used to describe the sample distributions in this study. The sample distributions found were frequently non-normal. Medians were used to characterize center of the sample distribution. If the distribution is symmetric, the mean and the median coincide. With highly skewed distributions the median is preferred, since it seems to represent the concept of a center better than the mean. The interquartile range (IR), defined as the difference between the 75th and the 25th percentiles, is used as a measure of population variability rather than standard deviation.

The following characteristics were described using these measures:

(1) Topographic characteristics:

The topographic characteristics are examined within physiographic regions. Table 1 shows the distribution of ground water elevation for the total state and for some of the 13 physiographic regions within the state.

(2) Relationships between the remolded soil characteristics and AASHTO classification within a physiographic region:

To develop these kinds of relationships on a regional basis, the following procedures are used.

(a) Examine the distribution of AASHTO classification units within a given physiographic region and select three or four most probable AASHTO classification units as the representative soil groups in the region.
(b) Examine the distributions of both visual texture and Unified classification units for each of the selected AASHTO classification units within the specific region. Select the most probable texture and Unified classification units as the representatives for soil identification and correlation.
(c) Apply appropriate statistical methods to obtain the estimates for sample distributions of the following remolded soil characteristics: natural dry density, (NATDD), specific gravity (SPECGR), shrinkage limit (SL), maximum dry density (MAXDD), optimum moisture content (OMC), CBR value at 100% maximum dry density (CBR SO1), and CBR value at 95% maximum dry density (CBR SO2) for each classification unit selected above. A typical result is illustrated in Table 2.

(3) Statistical soil profiles

The statistical soil profiles were generated according to the pedological soil associations, because these are the only soil grouping units which are reasonably large and grossly homogeneous. The following is an outline of the procedures used to generate a statistical soil profile for a given soil association. More detail is given in Lo (4).

(a) Examine the counties and physiographic regions with the given soil association and select representative locations.

(b) Apply appropriate statistical methods to assess the distributional parameters of topographic characteristics at the given locations.

(c) Establish the soil profile by examining the distributions of the AASHTO classification units within the given soil association and select the ones with highest frequencies as representatives.

(d) Use 90th percentiles of top and bottom depths from which the sample has been removed as upper and lower depth boundaries of the soil profile for each selected AASHTO classification unit within the given soil association. This was done to eliminate the extremes of depths from which the sample has been removed and, hopefully, to minimize the discontinuity of soil sample versus depth distributions.

(e) Divide the soil profile into sublayers based on the further examinations of the distributrions of selected AASHTO classification units versus depths.

(f) Apply appropriate statistical methods to obtain the estimates of characteristics of engineering soils at each sublayer.

Table 3 shows the statistical soil profile for the soil association Fincastle - Ragsdale - Brookston at the physiographic region of Tipton Till Plain.

Regression Analysis

Regression analysis provide a convenient method for investigating relationships among variables. In this study design parameters, such as compaction and compressibility characteristics, are correlated with soil index properties. The effects of soil location and genesis - i.e. physiographic region and parent material are incorporated into the regression equations using dummy variables (6).

The variables are 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 are well fit, a careful residual analysis is also made. The procedure used to reduce the number of independent variables is to compare the full model and the reduced model with an F-statistic (6).

The regression analysis is discussed in detail and reported by Lo (4,5) and is not further considered here.

Discussion of Results

The descriptive statistics and regression analyses are developed to provide information for future planning. Such information is useful if it provides an accurate picture of what can be expected. Small confidence intervals for the median, small interquartile ranges and large values of the squared multiple correlation coefficient (R^2) are associated with good descriptions and regression analyses. In cases where such results are not obtained, either more data or a different grouping procedure is needed.

The Medians and Soil Variability

In this study soils were grouped by physiographic regions, AASHTO classifications, soil associations, or a combination of these. The sample distributions of soil characteristics were studied according to these groups. The median of the sample distribution of a soil characteristic was used as the central measure.

Since medians rather than means are used, the usual analysis of variance procedures for testing equality of groups are inapplicable. However, nonparametric versions of these methods are available (3) and are used extensively in this study.

Topographic Characteristics Versus Physiographic Regions

Topographic characteristics are grouped according to physiographic regions. The test results verify that the topographic features, such as the ground elevations and ground water elevations vary with physiographic regions (4). In addition, the elevation and ruggedness varies with physiographic region.

Remolded Soil Characteristics Versus Physiographic Regions and AASHTO Classifications

The sample distributions of remolded soil characteristics were studied according to AASHTO classifications for each physiographic region. This is a two-way classification, i.e., the grouping effects are due to a combination of AASHTO classifications and physiographic regions. Results indicate that the remolded soil characteristics vary with both grouping units (4). These analyses also verify the facts that the more plastic soils have lower maximum dry density and higher optimum moisture content values (4).

Soil Characteristics Versus Soil Associations

The layout for the generation of a statistical soil profile involves a breakdown of the data for each soil characteristic by physiographic regions, and then by soil associations. Finally they are grouped into proper layer systems. Results indicate that only some of the soil characteristics vary with soil associations (4). Large variability among soil association units is found in this study (4). This may mean either that: (1) the soil samples are not sufficient for verification, or (2) the grouping unit is not refined sufficiently. In the latter case a subgrouping unit such as soil series may be useful. This subject needs further research.

Soil Series as a Grouping Unit

The results of the multiple comparison tests of soil characteristics versus soil series indicate that the soil characteristics do not vary significantly with soil series (4). Therefore, the soil series as a grouping unit is indicated to be inferior to the soil association. However, wide scatter of the data and relatively small sample sizes of soil characteristics are present. Further study is required to investigate this point in detail.

Applications

The uses of Indiana Geotechnical Data Bank for typical highway projects are described in the report by Lo (4). These examples show the uses of this data base (1) to determine the soil types and topography within a specific area and the engineering characteristics of these soils, (2) to develop a more efficient and economical subsurface exploration program, (3) to test the consistency and reliability of laboratory testing results, and (4) to make recommendations for design and construction of proposed highways and their relevant facilities.

Summary

A computerized data storage and retrieval system has been developed for the State of Indiana. Both conventional and nonparametric statistical methods are employed in the analysis of these data. The studies on the topographic characteristics versus physiographic region were based on a one-way classification. The results give a general impression of the topographic features of a physiographic region and can be used to make comparisons of overall topographic features among regions. The methods can also be applied to smaller areas if topographic data are adequate. The studies on the remolded soil characteristics versus physiographic regions and AASHTO classifications are based on a two-way classification, and show the relationships on a regional basis. The studies on statistical soil profiles show the general subsoil conditions qualitatively and provide estimates of soil characteristics quantitatively by depth for soil associations on a regional basis. Finally, the regression analysis gives good to adequate relations between design parameters and index properties.

Conclusions

1. Topographic features vary with physiographic regions.

2. The remolded soil characteristics can be evaluated and contrasted among physiographic regions and also among AASHTO classifications.

3. The data confirm that the more plastic soils have lower maximum dry density and higher optimum moisture content values.

4. The soil association is the most refined unit presently available for grouping soils to generate soil profiles.

5. Nonparametric robust statistical methods are preferred to conventional statistical methods for soil data analysis.

6. The physiographic region, engineering soil classification, soil association, and a combination of these are used as grouping units. The interquartile ranges (IR's) for most soil characteristics are small and tolerable. In other words, a good homogeneity of soil characteristics is evidenced with these groupings.

7. The data bank is valuable for making recommendations for preliminary design of geotechnical works. It can also be used as a framework against which various testing results can be judged for their consistency and reliability.

ACKNOWLEDGEMENT

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 Administration. Their support is gratefully acknowledged. We thank C. W. Lovell for his guidance throughout the project; R. D. Miles for his assistance, including valuable editorial aid; and T. R. West for valuable geological advice.

References

1. Goldberg, G. D. (1978) Development of the Computerized Geotechnical Data Bank for the State of Indiana. Purdue University, West Lafayette, IN, M.S.C.E. thesis and Joint Highway Research Project Report No. 76-6.

2. Goldberg, G. D., Lovell, C. W. and Miles, R. D. (1978) Use the Geotechnical Data Bank. TRB, Transportation Research Record 702, 140-146.

3. Hollander, M. and Wolfe, D. A. (1973) Nonparametric Statistical Methods. Wiley, New York.

4. Lo, Y-K. T. (1980) Geotechnical Data Bank for Indiana. Purdue University, West Lafayette, IN, Ph.D. thesis and Joint Highway Research Project Report No. 80-7.

5. Lo, Y-K. T. and Lovell, C. W. (1982) Prediction of Soil Properties from Simple Indices. TRB, Transportation Research Record 873, 43-49.

6. Neter, J. and Wasserman, W. (1974) Applied Linear Statistical Models, Irwin, Homewood, Illinois.

7. Tukey, J. W. (1977) Exploring Data Analysis, Addison-Wesley, Reading, Massachusetts.

Table 1. Ground water level with respect to ground level (ft)

Physiographic region	Mean	s.d.	Median	95% C.I. of median	$t_{.25}$*	$t_{.75}$**	I.R.***	Minimum	Maximum
Whole state (2395 cases)	5.40	4.80	4.33	4.13--4.53	2.11	6.73	4.62	0.10	60.00
Tipton Till Plain (834 cases)	5.16	3.41	4.10	3.86--4.37	2.50	6.12	3.62	0.20	20.00
Dearborn Upland (89 cases)	4.93	2.96	4.21	3.54--4.86	2.54	5.86	3.32	0.50	15.00
Norman Upland (83 cases)	7.67	5.52	6.54	5.78--7.23	4.44	8.50	4.06	0.50	17.00
Mitchell Plain (140 cases)	8.54	7.22	6.29	5.61--7.03	4.33	8.84	4.51	1.00	38.00
Crawford Upland (144 cases)	8.77	8.01	6.45	5.70--7.18	4.07	10.00	5.93	0.50	60.00
Wabash Lowland (348 cases)	5.88	5.38	4.12	3.75--4.61	2.54	7.06	4.52	0.10	37.00
Calumet Lacustrine Section (204 cases)	5.84	5.63	4.08	3.45--4.82	1.89	8.07	6.18	0.20	35.00
Valparaiso Moraine (150 cases)	2.26	1.72	1.86	1.55--2.28	0.94	3.31	2.37	0.10	10.00
Steuben Morainal Section (335 cases)	3.98	3.81	2.96	2.59--3.33	1.39	4.92	3.53	0.10	35.00

* 25% quartile ** 75% quartile *** Interquartile Range

Table 2. Characteristics of soil properties obtained from disturbed soil testing within Tipton Till Plain

Physiographic region: Tipton Till Plain
AASHTO classification: A-4
Texture: sandy loam, clay loam
Unified classification: CL, CL-ML

Variable	Mean	s.d.	Median	$t_{.25}$	$t_{.75}$	I.R.	95% C.I. of median	Minimum	Maximum	No. of Cases
NATDD	104.33	21.51	106.17	94.14	120.75	26.61	100.57--111.17	12.40	132.00	41
SPECGR	2.70	0.046	2.69	2.66	2.73	0.07	2.66--2.73	2.64	2.77	13
SL	14.05	3.28	13.23	11.21	15.65	4.44	12.89--13.57	0.10	26.00	552
MAXDD	120.53	7.12	122.26	115.76	126.29	10.53	120.00--124.52	98.20	130.70	49
OMC	12.27	3.04	11.25	10.00	13.43	3.43	10.55--15.95	8.30	24.70	48
CBRSO1	10.38	6.61	8.20	6.30	12.50	6.20	7.00--10.00	3.00	42.10	37
CBRSO2	6.07	2.70	5.00	3.33	7.67	4.34	3.87--6.33	2.00	14.00	39

Table 3. Typical soil profile for soil association.

Fincastle-Ragsdale-Brookston

General description: Nearly level, somewhat poorly drained, silty Fincastle in windblown silt and glacial till, very poorly drained, silty Ragsdale in windblown silts and loamy Brookston in glacial till.

Parent material: Soils formed in moderately thick loess deposits over loamy Wisconsin age glacial till.

Distributions: Physiographic region: Tipton Till Plain
Counties: Tippecanoe, little in Clinton

Ground elevation:

	Mean	s.d.	Median	$t_{.25}$	$t_{.75}$	I.R.	95% C.I. of Median	Minimum	Maxi-mum	N. of Cases
	728.12	70.26	708.33	673.81	788.88	115.07	685.71--761.11	645.90	866.80	47

Ground water level:

Season	Mean	s.d.	Median	$t_{.25}$	$t_{.75}$	I.R.	95% C.I. of Median	Minimum	Maxi-mum	N. of Cases
Total	723.17	69.76	708.33	672.50	785.00	112.50	685.00--700.00	641.70	854.90	47
Winter	846.68	13.85	875.00	850.00	883.33	33.33		826.10	854.90	4
Spring	759.11	62.15	725.00	695.00	806.25	111.25	700.00--800.00	662.60	808.90	17
Summer	653.26	7.16	671.43	653.57	689.29	35.72	650.00--663.00	641.70	663.50	9
Fall	695.18	36.03	700.00	671.87	737.50	65.63	675.00--733.33	642.50	767.30	17

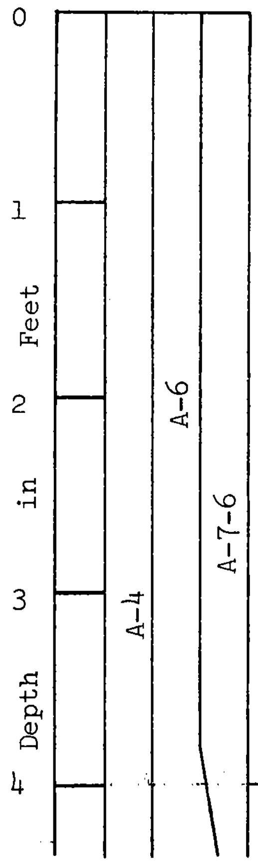
Ground water level with respect to ground level:

Season	Mean	s.d.	Median	$t_{.25}$	$t_{.75}$	I.R.	95% C.I. of Median	Minimum	Maxi-mum	N. of Cases
Total	4.95	2.67	4.00	2.91	6.00	3.09	3.36--5.17	1.60	11.90	47
Winter	9.53	4.36	11.00	4.00	11.33	7.33	10.00--11.90	3.00	11.91	4
Spring	3.33	1.38	3.17	2.42	3.92	1.50	2.50--3.83	1.60	6.00	17
Summer	4.42	0.87	4.86	4.14	5.57	1.43	4.00--5.40	3.00	5.40	9
Fall	5.78	2.40	6.00	3.28	7.28	4.00	3.43--7.14	2.50	10.00	17

Table 3 (continued). Typical soil profile.

Unified classification: CL   Texture: clay and silty loam

PH. 5.10 - 6.00   Organic material: not traceable

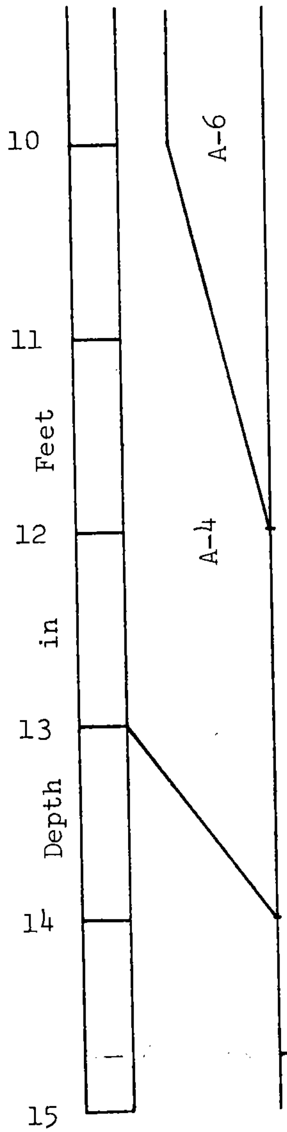


Unified classification: CL, CH   Texture: clay, silty clay and silty clay loam

PH: 5.60 - 6.00   Organic material: not traceable

Variable	Mean	s.d.	Median	$t_{.25}$	$t_{.75}$	I.R.	Min.	Max.	No. of Cases
SPT	5		5						1
SL	15.82	4.18	15.16	12.88	17.89	5.01	2.00	29.00	102
NATMC	23.00	8.11	22.50	17.78	27.67	9.89	8.00	65.00	67
NATDD	93.37	6.50	92.40	85.00	97.50	12.50	84.60	98.40	4
SPECGR	2.72	0.03	2.72	2.70	2.75	0.05	2.67	2.77	6
MAXDD	105.18	13.23	104.00	100.50	117.50	17.00	80.00	125.00	13
OMC	21.65	13.42	20.00	15.50	24.50	9.00	10.00	63.00	13
CBRSO1	7.86	2.14	7.67	6.50	10.33	3.83	4.50	11.30	13
CBRSO2	4.71	1.68	4.67	3.25	5.83	2.58	2.00	9.00	13
Qu	1.43	0.49	1.20				1.10	2.00	3

Table 3.(continued).



Unified classification: CL   Texture: clay, silty clay loam, and sandy loam

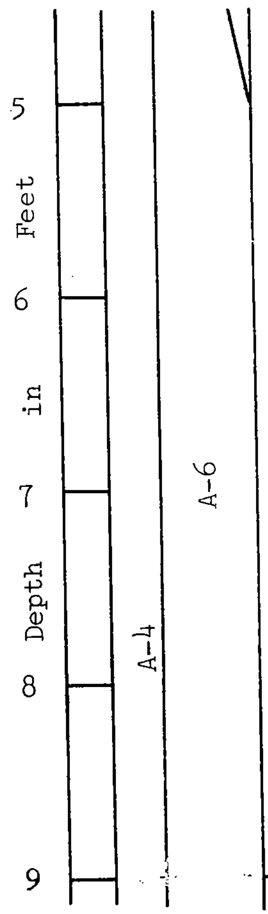
PH: 7.40 - 7.80   Organic material: not traceable

Variable	Mean	s.d.	Median	$t_{.25}$	$t_{.75}$	I.R.	Min.	Max.	No. of Cases
SPT	63	23	63				47	80	2
SL	13.69	3.35	13.50	10.13	16.50	6.37	9.00	18.00	13
NATMC	17.82	6.49	18.12	12.50	25.00	12.50	10.00	27.00	12
NATDD	103.00	2.48	103.33	100.83	105.00	4.17	99.80	106.20	5
SPECGR	2.76	.021	2.76				2.74	2.77	2
MAXDD	128.45	0.50	128.45				128.10	128.80	2
OMC	10.20	0.14	10.20				10.10	10.30	2
CBRS01	9.50	2.12	9.50				8.00	11.00	2
CBRS02	5.50	0.71	5.50				5.00	6.00	2
Qu	0.70	0.28	0.70				0.50	0.90	2

Note:

SPT-Standard penetration resistance
NATMC-Natural moisture content (%)
SPECGR-Specific gravity
OMC-Optimum moisture content (%)
Qu-Unconfined compressive strength (TSF)
SL-Shrinkage Limit (%)
NATDD-Natural dry density (PCF)
MAXDD-Maximum dry density (PCF)
CBRS01-CBR soaked value at 100% maximum dry density
CBRS02-CBR soaked value at 95% maximum dry density

Table 3. (continued).



Unified classification: CL, CL-ML

PH: 6.10 - 6.50

Texture: clay loam and sandy loam

Organic material: not traceable

Variable	Mean	s.d.	Median	$t_{.25}$	$t_{.75}$	I.R.	Min.	Max.	No. of Cases
SPT	7	4	7				5	10	2
SL	13.38	3.32	13	13.00	10.90	4.02	1.90	24.00	45
NATMC	21.50	11.12	18.54	15.00	23.57	8.57	11.00	65.00	34
NATDD	86.13	7.03	85.30	80.00	90.00	10.00	77.60	93.50	4
SPECGR	2.72	0.007	2.72				2.71	2.72	2
MAXDD	125.65	0.50	125.65				125.30	126.00	2
OMC	10.25	0.35	10.25				10.00	10.50	2
CBRSO1	6.60		6.60						1
CBRSO2	3.90		3.90						1
Qu	0.67	0.115	0.65				0.60	0.80	3