Route Geotechnical Characterization and Analysis

By Charles L. Vita,[1] M. ASCE

Abstract: A unified, landform-based probabilistic approach to geotechnical characterization and analysis is outlined for application to ground transportation routes. The approach can explicitly account for uncertainties, geotechnical variability and limitations in available geotechnical data/information. Using Bayesian probability concepts, landform identification and statistical characterization of landform soil property parameters, it can provide site-specific estimates of soil property parameters and geotechnical behavior variables. The methodology is suited to situations where regional or local landform soil property parameters are or can be statistically well known, but site-specific data may generally be either unavailable or sparse. Estimates can be based only on landform information or systematically updated and improved with site-specific data to the extent they are available. Estimates include means, variances, and reliability/probability levels on the estimates, including the probability of any given estimate being above or below a given critical value.

Introduction

Evaluation of potential geotechnical behavior along ground transportation routes traversing geotechnically variable terrain must often address uncertainties associated with sparse site-specific subsurface data, loading conditions, and soil behavior. The purpose of this paper is to outline a systematic approach for: (1) Quantifying these uncertainties; and (2) rationally augmenting sparse site-specific subsurface data with data available from other sites having the same characteristic landforms. These are integrated as illustrated with an example.

The landform-based approach to route characterization and analysis presented here consists of several, major interacting elements: route selection, landform mapping, statistical characterization of landform properties, soil profile development, and analysis of geotechnical behavior. The organization of the paper follows the logical progression of these major elements. After a review of landform basics, concepts and techniques for using landforms to systematically organize and apply geotechnical data needed for analysis and design are presented. The basic theme and premise of the approach is that a practical, efficient way to use geotechnical data needed in analysis and design is organizing and statistically characterizing exploration data for a given route or area on a landform basis such that along a route available, but often sparse, site-specific data can be rationally supplemented using Bayesian probability procedures then used to make site-specific probabilistic predictions of geotechnical behavior.

[1]Vice Pres., R&M Consultants, Inc., 18539 N.E. 184 St., Woodinville, Wash. 98072.

Note.—Discussion open until May 1, 1985. To extend the closing date one month, a written request must be filed with the ASCE Manager of Technical and Professional Publications. The manuscript for this paper was submitted for review and possible publication on July 14, 1983. This paper is part of the *Journal of Geotechnical Engineering*, Vol. 110, No. 12, December, 1984. ©ASCE, ISSN 0733-9410/84/0012-1715/$01.00. Paper No. 19336.

J. Geotech. Engrg. 1984.110:1715-1734.

Geotechnical Aspects of Route Characterization

Landform Basics.—Landforms are elements of the landscape formed by a single geologic process or a combination of associated processes which have: (1) Characteristic surface forms, such as topography, drainage patterns and gully morphology; and (2) typical, recurrent ranges of geotechnical properties. Sand dunes, moraines, floodplains, alluvial fans, and glaciofluvial outwash are all examples of landforms.

From identification of a landform, the characteristic distribution of soil properties as well as some understanding of stratigraphy, soil structure, drainage characteristics, and groundwater conditions can be inferred. Ranges of geotechnical behavior and potential facility response can be deduced from these landform characteristics. The chain of correlation is—characteristic visual forms :: landform :: characteristic distributions of geotechnical properties :: range of geotechnical behavior :: range of potential facility response and behavior. The practical value of this association depends on the dispersion of geotechnical properties in the landform. The sharper the distribution, the more precise geotechnical behavior can be predicted. Sharp distributions are associated with "homogeneous" properties; broad distributions are associated with "heterogeneous" properties. A given landform can have homogeneous or heterogeneous properties or both. For example, sand dunes are relatively homogeneous in all properties, glacial tills relatively heterogeneous in most soil properties, and landforms composed of wind blown silt may have a homogeneous grain size distribution, but heterogeneous moisture content. Distributions of geotechnical properties are variable with landform location because of differences in climate, weathering rates and processes, and predominant bedrock type.

Landform property distributions are also sample volume (site size) dependent. Variability between samples or "sites" tends to decrease with sample volume (site size), depending on the degree of spatial autocorrelation of the property (11).

Landform identification relies heavily on airphoto and other remote sensing techniques (4,5,8,16). Available maps and reports from the literature and unpublished sources are used to assist the remote sensing interpretation. Information from field investigations provides: (1) The data base from which correlations between visual surface features and subsurface conditions can be obtained; and (2) ground truth data for confirmation or modification of the antecedent remote sensing image interpretations; the relationship between remote sensing image interpretation and field investigations is one of successive and mutual refinement.

Route Selection.—The first stage of route characterization is the initial identification of geotechnical conditions along candidate route(s) for the purpose of route selection. This is done using terrain analysis techniques to identify visible features of the ground surface—landforms and associated detail—that can be correlated with geotechnical behavior and subsurface conditions (4,5,8,16). A geotechnical model of proposed or candidate route(s) can be developed by initially mapping observed landforms and associated terrain features in plan to produce a "terrain unit" map and in profile to produce a "landform profile." The route, each point of which will be underlain by one or more landforms, can be represented

J. Geotech. Engrg. 1984.110:1715-1734.

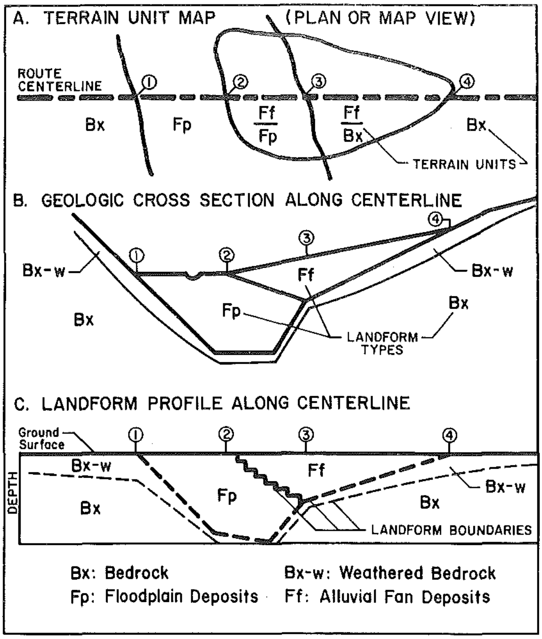


FIG. 1.—Relationships Along a Route Between Terrain Units, A, Landforms, B, and Landform Profiles, C

in geologic profile by a landform profile showing the landforms occurring from the ground surface to a limiting depth (beyond which there is generally little affect of soil or rock conditions on facility performance) and extending continuously along the entire route. Terrain units are constructs used to map landform profiles in plan view. Fig. 1 shows the relationship between terrain unit maps and landforms in geologic cross section and in landform profiles along a route center line.

Given the landforms and terrain features mapped along the proposed route(s), general evaluations and predictions (qualitative and statistical) of engineering behavior can be made and compared against the geotechnical route selection criteria to ascertain the most effective route (after considering all higher level non-geotechnical project/facility constraints).

Mapping for Segment-by-Segment Characterization.—After initial route selection, remote sensing techniques with available information and data from field investigations are used to successively improve and refine terrain unit maps and landform profiles along the route. Kreig and Reger (5) elucidate many of the pertinent considerations in the development of terrain unit maps and landform profiles. Landform profiles are used with field and laboratory data to develop segment-by-segment soil profiles and properties appropriate for engineering analysis and design.

Procedures of Geographic and Informational Updating

Geographic Scales and Bayesian Updating.—Landform data for route

J. Geotech. Engrg. 1984.110:1715-1734.

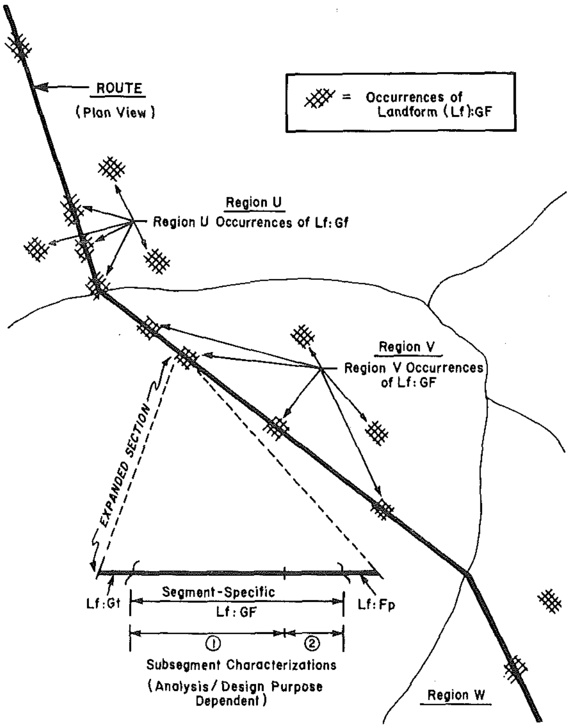


FIG. 2.—Illustration of Geographic Scale of Landform Occurrences and Characterizations for a Route

characterization is available at three (or more) geographic scales, as shown in Fig. 2. In order of decreasing size and data availability but increasing specificity these scales are: (1) For all known occurrences of a landform or landform group (of similar landforms); (2) for all regional occurrences of a landform or landform group within a particular local geologic region, unit, physiographic province or subprovince; and (3) for segment/site-specific occurrences of the landform in a given region. The segment/site-specific scale can be subdivided into subsegment/sites as desired to the meaningful limit supported by available data. Note that the data comprising the property distributions for each geographic scale are made up of the collective data of the geographically smaller scale(s).

In general, landform property distributions are scale and location dependent. The distributions for each geographic scale will differ (although available data will not always be sufficient to show significant differences). It is assumed that the forms (probability density functions: PDF) or basic shapes of the property distributions are not location or scale dependent whereas the statistical parameters (of the PDFs) generally are location and scale dependent (12,14). Landform property parameters that are not strongly location dependent for a given site size are "statistically homogeneous"; parameters that are relatively location dependent are "statistically nonhomogeneous (or heterogeneous)." Statistical homogeneity implies site-to-site uniformity; conversely, statistical heterogeneity implies site-to-site variability. Statistical homogeneity tends to increase with site size.

In a typical segment, data will be scarce; often, based *only* on segment data, a meaningful characterization cannot be achieved. The working principal for predicting landform soil properties wherever the landform occurs along the route is to utilize all available representative data obtained for the given landform. Combining landform data from different geographic scales and locations is termed "geographical updating." Updating to incorporate additional, new information at a given geographic scale is termed "informational updating." Both forms of updating use Bayesian Techniques, as discussed in the following section.

Geographical updating proceeds down the level of scale in the direction of increasing site specificity. Data for each landform or landform group are: (1) Combined to form a set of landform statistics; and (2) separated by geologic region or physiographic province to form a set of landform/regional statistics. Then, for each landform or landform group, the landform/regional statistics can be used to update the landform statistics to produce posterior, or updated, estimates of the parameters of the property distribution of the landform/region for each local region having occurrences of each particular landform. The updated parameters for each regional landform can be used as prior data for every occurrence of the landform in each segment/site in the particular region to produce, on a segment-by-segment basis, using site-specific data, posterior, updated, best estimates of the parameters of the property distributions for each stratum comprising the soil profile of each segment.

Geographical updating can be accomplished at consecutively greater levels of site-specificity, if data availability and needs warrant it. The principals of updating apply in the same way. With time, as more data become available, each scale can itself be updated to reflect new information.

At the scale of segment/site characterization (where individual strata are identified), updating can be done on the basis of landform and observed soil type (silt, sand, gravelly sand, etc.). For landforms characteristically composed of significantly different soil types, landform prior parameters could be soil type-dependent with updated segment/site-specific estimates conditioned on observed soil types. Updating on the basis of landform and soil type (or any other identifiable factor) can also be done without the landform prior characterized on the basis of soil type.

Each level of characterization has certain uses and limitations, depending on landform property variability and design requirements. Route characterization based on landform identification only may be useful for feasibility studies including route selection. However, unless landform property distributions are not location dependent, it will not support final design for a geotechnically sensitive facility without significant risk of multiple failures or, alternatively, compensatingly high levels of design conservatism—although, for a geotechnically tolerant or high risk facility, it may be adequate. Characterization based on landform and region may support preliminary stages of design for geotechnically sensitive facilities in statistically heterogeneous terrain or may be adequate for final design in more statistically homogeneous terrain and for certain geotechnically tolerant facilities, or both. Final design of geotechnically sensitive facilities in statistically heterogeneous terrain requires more detailed and conservative segment/site-specific route characterization.

In short, the geographic specificity/conservatism in the route characterization required to support a given level of performance/reliability must become greater in proportion to: (1) The statistical heterogeneity of the landforms; and (2) the geotechnical sensitivity of the design. A facility having a route traversing both statistically homogeneous and heterogeneous landforms *and* having a suite of design modes available can be designed such that the geographic specificity/conservatism of the characterization and the design modes are matched with the terrain conditions to achieve the most cost-effective mix meeting project requirements. The matching process requires mile-by-mile design decisions based on multi-technical, economic, and schedule considerations.

Segment-by-Segment Characterization: Soil Profile Developments.— Soil profiles are engineering constructs used to idealize soil stratigraphy and associated material properties within a specific site area or segment length for engineering analysis and design purposes. Segment-by-segment soil profile development conceptually starts with identification of landform profiles, which are then refined as needed using available site-specific data (see Fig. 3). A segment is defined as a reach of alinement

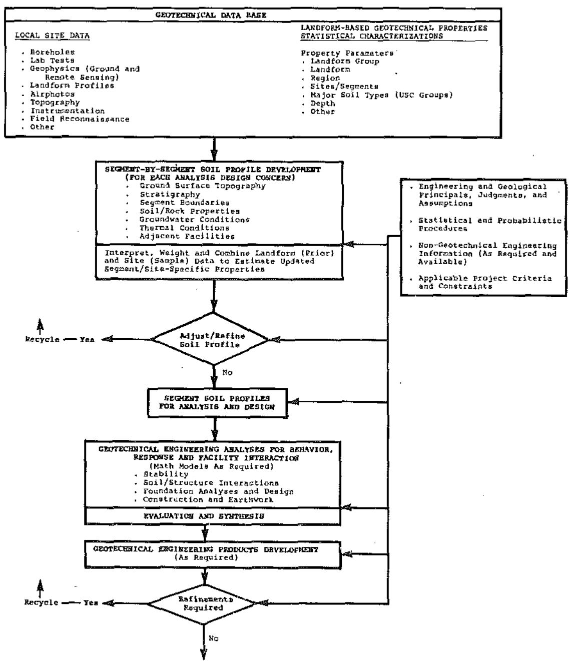


FIG. 3.—Segment-By-Segment Geotechnical Characterization and Analysis: Systematic Approach

which can be adequately characterized for a given (set or single) engineering purpose by a single soil profile. Each point of a route can belong to more than one soil profile (and associated segment) where significantly different characterizations are required for the various specific analysis and design purposes—e.g., within the limits of available data, segment-by-segment soil profiles characterized for slope stability may be different than soil profiles characterized for settlement, frost heave, bearing capacity, or erosion, etc.

Differences between landform profiles and soil profiles are dependent on anticipated engineering behavior. Contiguous landforms may have sufficiently similar engineering characteristics that they can be treated as a single stratum in a soil profile. Or, a single landform may be characterized by more than one stratum because of spatial differences in soil parameters within the occurrence of the landform at that site. With or without site-specific data, landform statistical analyses may show that significant differences in certain properties occur with depth. Such depth dependence should be represented in the soil profile to the extent it affects engineering behavior in a significant and predictable way. Thus, soil profiles can be simplifications or elaborations of landform profiles.

Soil property parameters for each stratum are updated using available sample data and statistically characterized landform prior information. Soil profiles should be discretized or standardized to eliminate any practically significant depth or lateral dependence of soil property parameters so that each stratum is reasonably statistically homogeneous (11). Soil profile development can be an iterative process with successive refinement of the soil profile geometry and parameter estimates, as shown in Fig. 3. The segment/site soil profile is developed so that for each stratum all pertinent geotechnical property statistical parameters are characterized by an updated estimate of their mean and variance.

Development of the site soil profile is dependent on site geology and limited by available site data. Prudent consideration must be given to the impact of very localized geological inhomogeneities which can dominate engineering behavior; such geological details can be spatially complex, interdependent and easy to miss. Conservative compensation for data limitations in soil profile development may be limited to emphasizing potentially critical stratigraphic details observed in borehole logs. Without any site borehole data, soil profiles are limited to inferred landform profiles and a prior knowledge of characteristic landform stratigraphy and associated material properties; interpretation of the geotechnical response of the profile could statistically consider the potential effects of the characteristic stratigraphy and properties. In all cases the potential effect of unanticipated negative conditions should be considered since subsurface conditions are never completely defined in exploration.

Probabilistic Bayesian Updating: Mathematical Procedures

Bayesian updating techniques (10,12,14) can be used to provide a rational and systematic basis for combining landform data obtained from different scales and locations (geographic updating) or additional, subject/site-specific data (informational updating). They require that the

J. Geotech. Engrg. 1984.110:1715-1734.

variability in the property(s) of interest, X (material and geometric, or both), be describable by a known probability density function, PDF[X], having, in general, uncertain statistical parameters, θ, with uncertainty in θ described by PDF[θ]. Subject/site-specific estimates of θ, $\theta(S)$, are made using prior data and available subject/site-specific sample data with Bayes' theorem to make posterior, updated estimates of $\theta(S)$, $\theta(S)''$ as follows: Subject/site-specific parameters $\theta(S)$ are random variables (in the Bayesian sense) having prior distributions PDF[$\theta(S)$]$'$ based on available prior data. Subject/site-specific sample data on property X, $x(S)$, are summarized by a sample likelihood function, $L[\theta(S)/x(S)]$, which gives the relative likelihoods of the uncertain values of $\theta(S)$ given $x(S)$ (1); if $x(S)$ is composed of n random observations of X, following a PDF having parameters $\theta(S)$, then the sample likelihood function becomes:

$$L[\theta(S)/x(S)] = \prod_{j=1,n} \text{PDF}\,[x_j(S)/\theta(S)] \quad \text{(1)}$$

Then, using Bayes' Theorem the posterior, updated PDF of $\theta(S)$, PDF[$\theta(S)$]$''$, is equal to the product of the prior PDF, PDF[$\theta(S)$]$'$, and the sample likelihood function, $L[\theta(S)/x(S)]$, normalized by a constant to insure that PDF[$\theta(S)$]$''$ integrates to unity:

$$\text{PDF}[\theta(S)]'' = \frac{L\,[\theta(S)/x(S)] \cdot \text{PDF}[\theta(S)]'}{\int L\,[\theta(S)/x(S)] \cdot \text{PDF}[\theta(S)]' \cdot d\theta} \quad \text{(2)}$$

If the prior distribution of $\theta(S)$ and the sample-likelihood function are conjugate pairs, posterior distributions of $\theta(S)$ are of the same mathematical form as the prior; then, the mean or expected value, $E[\theta(S)]''$, and variance, Var $[\theta(S)]''$, are simply related to the parameters of the prior distribution. Justification for using conjugate pairs can be based on physical reasoning, empirical evidence, or solely on mathematical convenience and simplicity (unless they are not compatible with available evidence). Where conjugate pairs are not appropriate, $E[\theta(S)]''$ and Var $[\theta(S)]''$ can be obtained from PDF[$(\theta(S)$]$''$ for any PDF[$\theta(S)$]$'$ or $L[\theta(S)/x(S)]$ using Eq. 2.

Table 1 presents the pertinent mathematics (10,12) of two conjugate pair PDF models that are particularly useful for site characterization: a normal probability model and a binomial model. For example, observed PDF[X]s for soil density, moisture content, shear strength parameters, and compressibility parameters tend to follow bell shaped Beta distributions (3,7). The central portions of the Beta distributions can be described by normal (6,7) and lognormal or inverse lognormal distributions (by simple logarithmic transformations the latter two distributions can be transformed into normal distributions). Thus, practical updating of bell shaped distributions of any soil property X (subsequent to suitable transformation if necessary) can be done assuming PDF[X] is normal (defined by the two parameters mean of X, $\bar{X}$, and standard deviation of X, σ) and using the normal PDF model updating equations presented in Table 1. Also, the soil property parameter of interest in grain size characterizations is commonly associated with the proportion, u, compared to a critical value, U, of a soil which is finer than a specified grain

J. Geotech. Engrg. 1984.110:1715-1734.

TABLE 1.—Updating Equations for Soil Property Parameter Estimating

Assumed soil property PDF model and uncertain parameters of model (1)	Prior and posterior PDF of parameters (2)	Updated estimates of mean and variance of parameters for site (3)
NORMAL With uncertain parameters: mean, $\bar{X}$, and standard deviation, σ, of property X $\mathrm{PDF}(X) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2}\left(\frac{x-\bar{X}}{\sigma}\right)^2\right]$	STUDENT, INVERTED-GAMMA-2 $\mathrm{PDF}(\bar{X}) = \left\{\frac{(n-1)^{n/2-1/2}}{\sqrt{\pi}\, s/\sqrt{n}} \cdot \frac{\Gamma[n/2]}{\Gamma\left[\frac{1}{2}n-\frac{1}{2}\right]} \cdot [n-1+n\cdot(\bar{X}-\bar{x})^2/s^2]^{n/2-1/2}\right\}$ $\mathrm{PDF}(\sigma) = \left\{\frac{2\left(\frac{1}{2}n-\frac{1}{2}\right)^{n/2-1/2}}{s\cdot\Gamma\left[\frac{1}{2}n-\frac{1}{2}\right]} \cdot \left(\frac{s^2}{\sigma^2}\right)^{n/2-1/2} \cdot \exp\left[-\frac{1}{2}(n-1)s^2/\sigma^2\right]\right\}$ PDF': $\bar{x} = \bar{x}(Lf)$, $s^2 = s(Lf)^2$, $n = n'$ PDF'': $\bar{x} = \bar{x}''$, $s^2 = s''^2$, $n = n''$ $s''^2 = \frac{[(n-1)\cdot s(S)^2 + (n'-1)\cdot s(Lf)^2 + n\cdot n'[\bar{x}(Lf)-\bar{x}(S)]^2/n+2]}{(n'+n-1)}$	$E[\bar{X}(S)]'' = \bar{x}'' = \frac{n'\cdot\bar{x}(Lf)+n\cdot\bar{x}(S)}{n'+n}$ (I.1) $\mathrm{Var}\,[\bar{X}(S)]'' = \frac{s''^2\cdot(n'+n-1)}{(n'+n)\cdot(n'+n-3)}$ (I.2) $E[\sigma^2(S)]'' = \frac{s''^2\cdot(n'+n-1)}{(n'+n-3)}$ (I.3) $E[\sigma(S)]'' = \frac{\sqrt{s''^2\cdot(n'+n-1)}}{2}\cdot\frac{\Gamma[(n'+n-2)/2]}{\Gamma[(n'+n-1)/2]}$ (I.4) $\mathrm{Var}\,[\sigma(S)]'' = E[\sigma^2(S)]'' - E^2[\sigma(S)]''$ (I.5)
BINOMIAL With uncertain parameter: the probability, Q, property u is below a critical value U $\mathrm{PDF}(Z) = \binom{\eta}{z} Q^z(1-Q)^{\eta-z}$	BETA $\mathrm{PDF}(Q) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\beta)} Q^{\alpha-1}(1-Q)^{\beta-1}$ $\alpha = n'\cdot Q(Lf) + n\cdot Q(S) + 1$ $\beta = n'\cdot[1-Q(Lf)] + n\cdot[1-Q(S)] + 1$	$E[Q(S)]'' = \frac{n'\cdot Q(Lf)+n\cdot Q(S)+1}{n'+n+2}$ (II.1) $\mathrm{Var}\,[Q(S)]'' = E[Q(S)]''\cdot\left[\frac{1-E[Q(S)]''}{n'+n+3}\right]$ (II.2)

Note: $\bar{x}(Lf)$, $s(Lf)^2$ = sample mean and variance of landform (or any) prior; $\bar{x}(S)$, $s(S)^2$ = sample mean and variance of site; $Q(LF)$ = estimated probability $u < U$ based on landform data; $Q(S)$ = estimated probability $u < U$ based on site data; (Lf) identifies information from landform (or any) prior; (S) identifies site-specific information; n' = landform (or any) prior sample weight; n = site sample weight; $n'' = n' + n$ = posterior sample weight.

size (gravel, sand, silt, clay, or any particular grain size fraction). The probability of the number, Z, of fundamentally sized, effectively homogeneous soil volumes composing a site (or stratum) of volume V and having $u < U$ can be modeled by the binomial distribution. The number of effective soil volumes comprising V is the effective number of trials, η, so that Q, the estimated probability that $u < U$ in any effective soil volume can be estimated using the binomial PDF model updating equations presented in Table 1. The more general Eq. 2 can be used if the normal or binomial models in Table 1 are not appropriate. In all cases, proper use of these techniques requires judgment and understanding of their application-specific limitations.

Weighting of Landform Prior Data.—In the mathematics of the updating equations presented in Table 1, prior site information on property X is based on: (1) Aggregating available X data (on a landform basis) into a set of landform sample statistics; and (2) weighting these by a numerical factor n', which represents a prior sample size; n' is proportional to the relevance of the landform statistics to site characterization of property X. These are combined, through updating (using Baye's Theorem), with site-specific sample statistics weighted by the available site sample size n.

This aggregating and weighting approach to using priors is compatible with data bases formed by combining useful results from all available site exploration programs—where the quality, quantity, detail, and geographic extent of the exploration programs may be quite diverse. In general, n' is dependent on site size, landform geological characteristics, and the quality, quantity, and statistical uncertainty of available prior data; establishing suitable values for n' requires both analysis and geotechnical judgment.

As an inspection of the equations in Table 1 suggests, if prior information, measured by n', is adequate, estimates of $\theta(S)$ can be made with any amount of site-specific data (n)—including no data ($n = 0$). However, uncertainty in all estimates, measured by Var $[\theta(S)]''$, decreases with increasing site-specific data (increasing n). In all cases, as n increases the influence and importance of landform prior information on the estimates $\theta(S)''$ decrease.

Bayesian updating can be used for sites where landform prior information must be evaluated on a site-specific basis. Site-specific priors may be necessary where: (1) Available numerical data are inadequate (e.g., because of limited subsurface exploration or testing of previous sites in the landform) to be of practical value in estimating landform prior parameters by mathematical analysis (particularly as concerns variability between sites); or (2) geotechnically, site conditions are qualitatively different from available landform prior data. In practice, evaluating site-specific priors necessarily requires geotechnical knowledge, experience and judgment; it can encompass analysis and synthesis of all available prior data (qualitative and quantitative) with specific relevance to the particular site in question, including geotechnical details of the exploration and testing programs and landform and site geological characteristics.

Certainly, because of the complexities involved, prior data can be usefully evaluated and summarized in various ways, depending on practical needs and constraints. Nevertheless, as Case 2 in the following illustrative example suggests, important engineering implications can be inferred from even very limited data.

As an example of determining suitable values for n' where adequate numerical data are available, consider estimating the site mean of X, $\bar{X}(S)$, where X follows a normal PDF; Eqs. I.1 and I.2 from Table 1 would then be appropriate (in all cases "Lf" denotes landform prior information and "S" denotes site-specific information). To a useful first approximation, the value of n' is a function of α, the estimated ratio of Var $[X(Lf)]$ to the variance of the observed distribution of $E[\bar{X}(S)]$ for all other sites in the subject landform where data are available; n' can be calculated by conservatively assuming: (1) For $n = 0$ (no sample data) Var $[\bar{X}(S)]''$ is equal to an estimate of the variance of $E[\bar{X}(S)]$ for all sites comprising the landform, and Var $[X(S)]$ is equal to the landform data-based estimate of Var $[X(Lf)]$ based on N independent, available samples from the landform; and (2) α, conservatively estimated from available data ($\alpha = 1$ is most conservative), is less than the ratio of Var $[X(S)]$ to Var $[\bar{X}(S)]$. It then follows that n' can be calculated from:

$$\left(\frac{N-3}{N-1}\right)\cdot\frac{(n'+n-1)}{(n'+n)\cdot(n'+n-3)}=\frac{1}{\alpha+n} \quad (3)$$

again, in which

$$\alpha=\frac{\text{Var }[X(Lf)]}{\text{Var }[E[\bar{X}(S)]]}\leq\frac{\text{Var }[X(S)]}{\text{Var }[\bar{X}(S)]} \quad (4)$$

In general, α is a random variable dependent on: (1) Appropriate site

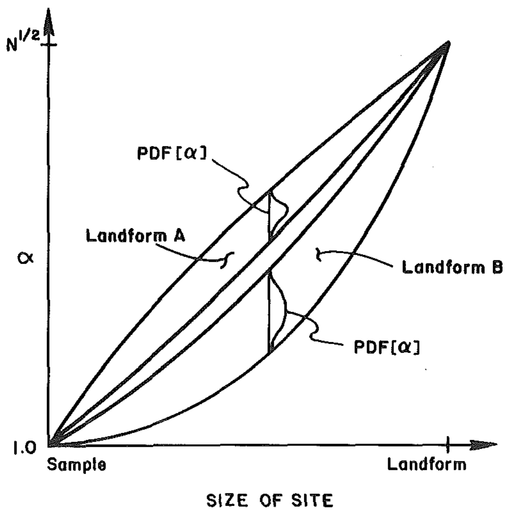


FIG. 4.—Illustration of Factors Affecting α (Curve Shape, Height and PDF[α]) for Two Landforms

size, due to spatial autocorrelation (ranging from 1.0 for sample-sized "sites" to $\sqrt{N}$ for a landform-sized site); (2) on landform geological/statistical characteristics ($\alpha = 1$ reflects greatest site-to-site variability, α being relatively larger for statistically homogeneous landforms); and (3) on the quality, quantity, and statistical uncertainty of available multiple site data, ($\alpha = 1$ reflects least admissible quality, quantity, or greatest uncertainty). These are shown in Fig. 4 (by curve shape, height, and PDF[α]) for two landforms, A and B (having equal sample sizes, N).

As another example, n' for estimating Q, the probability property u is below a critical value U in a given soil volume, could use Eq. II.2 (Table 1) by solving for n' but substituting Var $[Q(S)]$ for Var $[Q(S)]''$ and putting $n = 0$; this yields:

$$n' = \frac{E[Q(S)]'' \cdot (1 - E[Q(S)]'')}{\text{Var}\,[Q(S)]} - 3 \quad (5)$$

Var $[Q(S)]$ could be estimated from data obtained from all past site explorations in the subject landform. The random variable Var $[Q(S)]$ is also landform-dependent, site size-dependent, and data-dependent, as Fig. 5 shows for the two landforms, A and B.

Site data-based estimates are made (utilizing site-specific data only) by neglecting landform prior data and letting $\alpha = n' = 0$. Site data-based estimates are then compared with updated landform (prior) site data estimates. Where comparison indicates data-based and landform/site data estimates are significantly different, the site can be given appropriate attention (more analysis, and, if necessary, more exploration, or changed design); also, in subsequent site analysis the weight of the landform prior could be reduced by decreasing n'. Note that a formal comparison could utilize an estimate of p, the probability that the landform and site soil

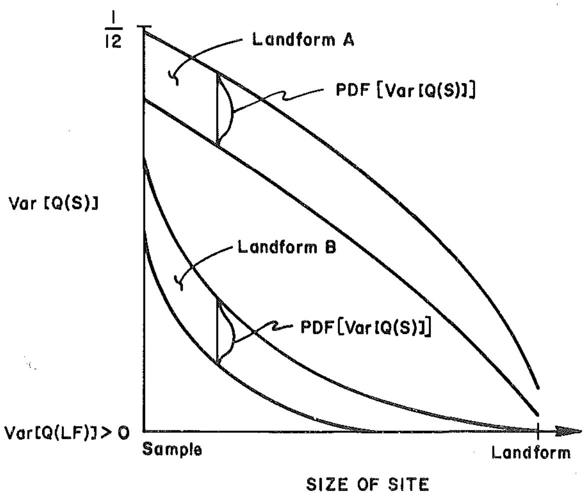


FIG. 5.—Illustration of Factors Affecting Var $[Q(S)]$ **(Curve Shape, Height, and** PDF[Var $[Q(S)]$]) **For Two Landforms**

property parameters $\theta(Lf)$ and $\theta(S)$ differ by more than some geotechnically significant amount Δ ($\Delta \geq 0$). For example, one data-based procedure suitable for means of normal PDF's is to estimate p assuming the difference between $\bar{X}(Lf)$ and $\bar{X}(S)$ is normally distributed and described by Student's t distribution; using the sample statistics defined in Table 1, p can be estimated by

$$p = 2F(t) - 1 \quad (6a)$$

$$t = \frac{|\bar{X}(Lf) - \bar{X}(S)| - \Delta}{\left(\dfrac{s(Lf)^2}{N} + \dfrac{s(S)^2}{n}\right)^{1/2}} \quad (6b)$$

in which $F(t)$ is the cumulative distribution function of the t distribution (9,12,14). A similar procedure could be adapted for $Q(Lf)$ and $Q(S)$ in estimating p. Mathematically, procedures for estimating p for Eq. I.3 and Eq. I.4 (Table 1) should be based on the F distribution (2).

In all cases, developing landform-based priors requires that sampling and testing biases and limitations are considered and that significant skewing affects on parameters are adequately rectified (14). Also, for landforms characteristically composed of significantly different soil types (e.g., heterogeneous tills), landform parameters may be made soil type-dependent and updated site-specific estimates conditioned on observed soil types. Further, updated site-specific estimates for situations where landforms are uncertain and/or occur in complex associations can be made by conditioning estimates on the expected probability of occurrence of the various potential conditions; the sensitivity of such estimates to the uncertainty, particularly the potential extreme conditions, should be properly evaluated.

Generalizing Priors for Updating.—All equations for making updated posterior estimates and associated discussion can be generalized for geographic and informational updating by considering landform (Lf) prior parameters as any appropriate prior (based on geographic collective and/or prior subject/site-specific data) and site-specific (S) sample parameters as any admissible subject/site-specific sample data coming from any sized subject/site-specific geographic subset of Lf. For informational updating, the new prior sample size n' is the posterior sample size from previous geographic updating; and the new sample size is again called n. Informational updating can be "avoided" by reiterating the geographic updating calculation by combining existing and new subject/site-specific data into one composite sample. Updating procedures must not misaccount prior sample sizes in any way that erroneously increases the actual extent or quality of the data base.

Autocorrelation Effects.—Typically, available site data is too sparse to support meaningful probabilistic autocorrelation characterizations of soil and rock strata (although controversial and undemonstrated, it has been suggested (11) that autocorrelation effects may not be site-dependent for similar soil types). The probabilistic updating procedures presented here can partially incorporate autocorrelation effects by including this influence on landform prior and sample statistics. In general, the potential practical importance of autocorrelation effects should be evaluated application-specifically and site-specifically. In particular, the potential influence of autocorrelation on biasing the mean and underestimating the variance of soil property parameters should be considered.

Geotechnical Analysis/Design: Probabilistic Modeling.—Any analysis or design variable V dependent on $\theta(S)$ can be evaluated using the updated parameters $E[(\theta(S)]''$ and Var $[\theta(S)]''$ to provide subject/site-specific probabilistic predictions of geotechnical behavior. V can be modeled as a random variable using the point estimation method or a Taylor series expansion to estimate its mean or expected value, $E[V]$, and variance, Var $[V]$. Assuming a reasonable PDF model, an *estimated* probability of V being below any critical value can be calculated. Whenever appropriate, estimates based on alternative sets of reasonable assumptions or interpretations can be assessed for sensitivity and technical/economic implications. In theory, uncertainty in results can be made as small as desired by obtaining sufficient information on which to base estimates.

Illustrative Application for Arctic Transportation Route

Thawing Slope Stability—Factor of Safety.—Permafrost slopes which are thermally disturbed by facility construction and operation can thaw and become mechanically unstable. The resistance of a thawing slope to mass movement is commonly characterized by a limit equilibrium factor of safety (FS) for the expected critical failure surface A with FS defined as the ratio C/D where C (capacity) is the sum of all forces resisting movement along A, and D (demand) is the sum of all forces contributing to movement along A. In theory, if FS could be evaluated with absolute accuracy, it would give a true measure of the future stability of a slope: $FS \geq 1.0$ would imply certain stability and $FS < 1.0$ would imply certain movement. In practice, FS cannot be evaluated with such accuracy due to uncertainties in the failure mechanism, soil conditions, loading conditions, and thermal response; these uncertainties make FS a random variable.

FS can be evaluated probabilistically using estimates of $E[FS]$, Var $[FS]$, and assuming a reasonable PDF$[FS]$. Here, PDF$[FS]$ is assumed normal in the range $1.0 \leq FS \leq E[FS]$. Given $E[FS]$, Var $[FS]$ and PDF$[FS]$, an estimated probability of instability or "failure," Pf, a random variable defined as $P[FS < 1.0]$, can be calculated as $Pf = P[FS < 1.0/E[FS]$, Var $[FS]$, PDF$[FS]]$. Also useful, the stability margin, U, is defined as $U = (E[FS] - 1)/SD[FS]$; where SD is standard deviation, the square root of the variance.

The previous procedure can be followed for any form of FS generally defined by

$$FS = fs(X_i) + e \quad i = 1, 2, \ldots k \qquad (7)$$

The function $fs(X_i)$ represents the stability model C/D based on the X_i random variables used to model geotechnical, geometric, and loading conditions—e.g., slope geometry, soil shear strength, failure surface geometry, groundwater conditions, excess pore pressures, dead and live loads, stress distribution effects, etc. Variable e accounts for bias and

J. Geotech. Engrg. 1984.110:1715-1734.

random error—reflecting potential inaccuracy and uncertainty in the stability model $fs(X_i)$ itself. Using an effective stress Mohr-Coulomb failure model with attention focused on the uncertain soil strength parameters, $fs(X_i)$ can be written as the sum of two components: $(G \cdot C' + F \cdot \tan \phi')$—where C' and $\tan \phi'$ are the average effective (thawed) soil strength parameters along the critical failure surface; G is a function of all X_i related to cohesion (C') and F is a function of all X_i related to friction ($\tan \phi'$). Using a Taylor series expansion of FS yields [13]

$$E[FS] = E[G] \cdot E[C'] + E[F] \cdot E[\tan \phi'] + \bar{e} \quad \text{(8a)}$$

$$\text{Var}\,[FS] = E[G]^2 \cdot \text{Var}\,[C'] + E[F]^2 \cdot \text{Var}\,[\tan \phi'] + Z \quad \text{(8b)}$$

Z represents in summary form all uncertainty, including random error in the model, other than due to Var $[C']$ and Var $[\tan \phi']$. The expected values and variances of C' and $\tan \phi'$ can be estimated based on regression analysis correlations with standard soil index properties of representative samples recovered during boring and sampling. The regression equations and associated statistical parameters applicable to a given route can be developed from a geotechnically and statistically designed strength testing program using a limited, but adequate, number of generally representative soil samples taken from the route. This approach can form the entire basis for estimating C' and $\tan \phi'$ for a given slope or it can supplement any available site-specific strength data. The model used here is based on average frozen dry density (prior to thawing), $\bar{\gamma}_{df}$, as the predictive soil index property

$$\tan \phi' = \tan [\sin^{-1} \beta \cdot \bar{\gamma}_{df}] \quad \text{(9a)}$$

$$C' = \frac{\rho \cdot \bar{\gamma}_{df}}{\cos \phi'} \quad \text{(9b)}$$

in which ρ and β are data-based, soil type-dependent (updatable) regression parameters. The functional forms of Eq. 9*a*, *b* can be expanded in a Taylor series to calculate $E[\tan \phi']$, Var $[\tan \phi']$ and $E[C']$, Var $[C']$ for substitution into Eq. 8*a*, *b*.

Average Thaw Settlement.—The estimated average thaw settlement potential of a permafrost stratum or strata of average thickness h is TS, where $TS = h \cdot (1 - \bar{\gamma}_{df}/\bar{\gamma}_{dt})_h$; $\bar{\gamma}_{df}$ and $\bar{\gamma}_{dt}$ are stratum frozen and thawed dry densities averaged over the thickness h. A Taylor series expansion estimate of the mean and variance of TS is given in Ref. 13. Confidence intervals on TS for given probability/reliability levels including estimates of the probability of TS being above (or below) any given value can be calculated using $E[TS]$, Var $[TS]$ and an appropriate PDF$[TS]$ model. Also, the coefficient of variation of TS, given by $CV[TS] = SD[TS]/E[TS]$, is a simple, useful measure of uncertainty in TS.

Landform-Based Estimation of Site-Specific Soil Parameters.—As developed here, site-specific estimation of FS and TS is based on prediction of average frozen dry densities using landform-based procedures to augment sparse or nonexistent site-specific subsurface data. In general, the observed PDF$[\gamma_{df}]$ will closely resemble an inverse lognormal distribution, which can be transformed to a normal distribution; but for most practical purposes directly assuming a normal distributing is adequate.

Using from Table 1, Eqs. I.1 and I.2 [with $n' = w$, following (13)], the updated landform/site data estimates of the mean and variance of $\bar{\gamma}_{df}(S)$ for a normal PDF[γ_{df}] are:

$$E[\bar{\gamma}_{df}(S)]'' = \frac{\bar{x}(S) \cdot n + \bar{x}(Lf) \cdot w}{n + w} \quad \text{(10a)}$$

$$\mathrm{Var}\,[\bar{\gamma}_{df}(S)]'' = (n-1) \cdot s(S)^2 + (w-1) \cdot s(Lf)^2 + \frac{\dfrac{n \cdot w \cdot [\bar{x}(Lf) - \bar{x}(S)]^2}{(n+2)}}{(w+n) \cdot (w+n-3)} \quad \text{(10b)}$$

in which n = number of available independent samples from the site; $\bar{x}(S)$, $s(S)^2$ and $\bar{x}(Lf)$, $s(Lf)^2$ are the sample mean and variance of the site and landform γ_{df} data respectively; and w ($w = n'$) is calculated from Eq.

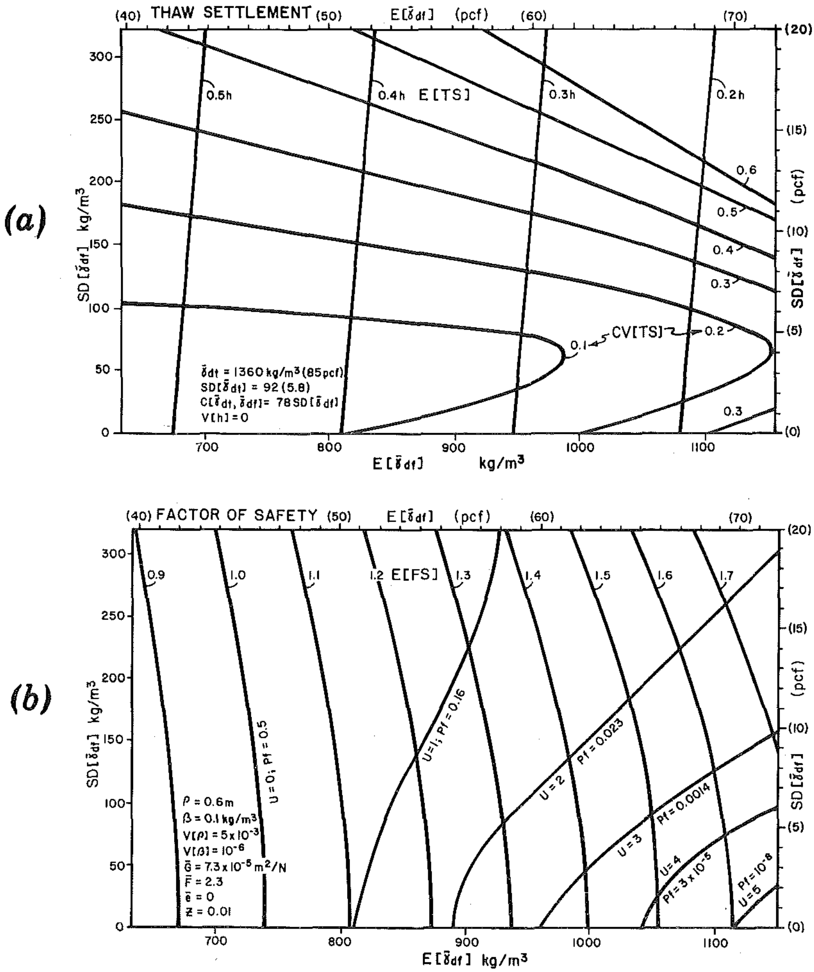
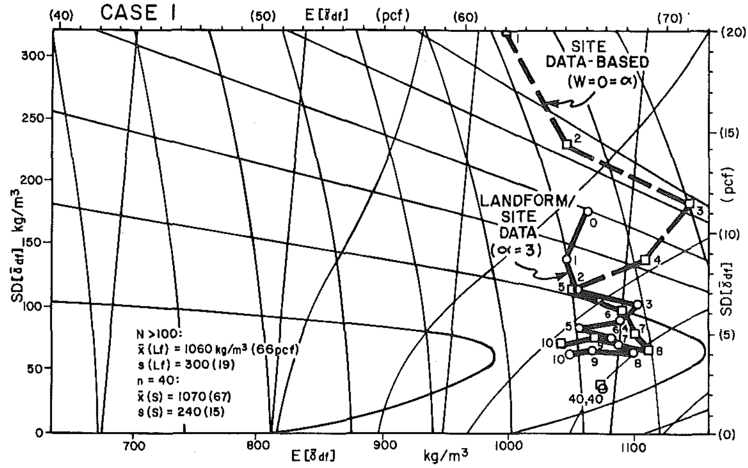


FIG. 6.—Site-Specific Results for Illustrative Application: (*a*) Thaw Settle

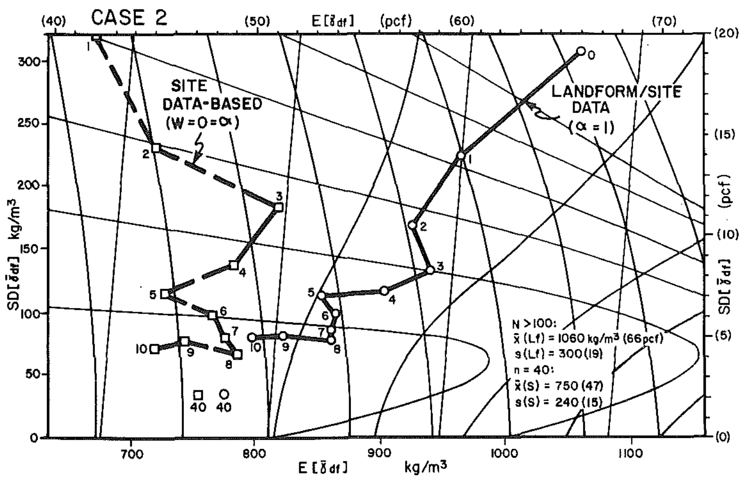
3 where α is the estimated ratio of Var $[\gamma_{df}(Lf)]$ to the variance of the observed distribution of $E[\tilde{\gamma}_{df}(S)]$ for all other sites in the subject land-form.

Site data-based estimates are made (utilizing site-specific data only) by neglecting landform prior data and letting $\alpha = w = 0$. Note that Eq. 10*b* is undefined for site data-based estimates where $n \leq 3$; the assumption that $\tilde{\gamma}_{df}(S)$ is uniformly distributed over a maximum possible range R introduces the least bias and gives Var $[\tilde{\gamma}_{df}(S)] = R^2/12$, centered at $\tilde{x}(S)$, which is at least as conservative as is the estimate of R. In this example R = 1,100 (70 pcf), 800 (50 pcf), and 640 (40 pcf) kg/m³ for n = 1, 2 and 3, respectively. Site data-based estimated are compared with updated estimates using land-form prior data.

Results.—Fig. 6 displays generally representative characteristics of the



(c)



(d)

ment; (*b*) Thaw Plug Stability; (*c*) Case Study 1; (*d*) Case Study 2

preceding equations including specific results for a several-acre permafrost site located in a landform composed of colluvial silt. The site was selected for illustration because it had sufficient soil data to compare predictions based on typically limited data (small n) with predictions based on substantial data (large n—much more soil data than would be available for a typical segment-site along a route).

Results in Fig. 6 are for a critical stratum of frozen silt, a zone at 3–4 m (10–12 ft) depth, plotted as a function of the estimated expected value, $E[\bar{\gamma}_{df}]$, and estimated standard deviation, $SD[\bar{\gamma}_{df}]$, of the (uncertain) average frozen dry density of the stratum, $\bar{\gamma}_{df}(S)$ (any autocorrelation effects are assumed negligible). Fig. 6(a) shows contours of mean thaw settlement, $E[TS]$, and coefficient of variation, $CV[TS]$; Fig. 6(b) shows contours of mean factor of safety, $E[FS]$, stability margin, U, and probability of failure, Pf. Superimposed on these contours, Fig. 6(c) (Case 1) and Fig. 6(d) (Case 2) illustrate sample-by-sample results for two important site situations.

Case 1 shows sample-by-sample results for a site-specific stratum that is similar to the landform as a whole, i.e., $\bar{\gamma}_{df}(S) \simeq \bar{\gamma}_{df}(Lf)$. Case 1 shows actual site data with $\alpha = 3$, representing a relatively statistically homogeneous landform. Both updated landform/site data estimates ($\alpha = 3$) and site data-based estimates ($\alpha = w = 0$) are shown (the numbers next to box and circle symbols on case study results indicate number of samples from the site, n), Case 2 shows actual sample γ_{df} data reduced by 320 kg/m^3 (20 pcf) with $\alpha = 1$, representing a statistically variable landform or one where either little information on other sites is available or it is of poor quality such that a minimum α is considered prudent. Case 2 simulates sample-by-sample results for a site-specific stratum that is substantially less dense than the landform as a whole, i.e., $\bar{\gamma}_{df}(S) << \bar{\gamma}_{df}(Lf)$. A total of $n = 40$ samples (each from one borehole) were obtained from the critical stratum. Sample-by-sample results for the first 10 ($n = 1, 2, \ldots 10$) samples are plotted; results based on all 40 samples are given for comparison.

Case 1 (actual γ_{df} data) shows the efficiency of the updated landform/site data method when the site is in fact similar to the landform: uncertainty in FS and TS is significantly decreased at small n relative to the site data-based estimates; therefore, the site can be qualified as acceptable or unacceptable based on a given Pf criterion with less site-specific data. Case 2 (actual γ_{df} data minus 320 kg/m^3) shows the effect of a "negative surprise" site—one that is in fact much less dense, and therefore less stable, than the landform as a whole. The benefits of the updated landform/site data method for Case 2 sites are derived through comparison with the site data-based predictions: at small n potentially significant differences become obvious—suggesting the need for further exploration, analysis, and/or changed design to achieve the Pf design criterion.

Summary and Conclusions

The landform-based approach outlined here can support a systematic, unified framework for geotechnical characterization and analysis of transportation routes. It can be used to produce segment-by-segment soil profiles and associated properties for engineering analysis and design using available data and analytical capabilities. It is capable of objectively accounting for uncertainty and limitations in available data and geotechnical variability on a site-specific basis. The analytical models can be objectively improved by calibration where sufficient data is available. On a site-by-site basis, the methodology efficiently uses all available data; it can rationally and consistently utilize past experience with relevant geological/geotechnical conditions to augment available site-specific data. It provides a quantitative basis on which the project-specific exploration/characterization/analysis/design process can be optimized, dependent on overall constraints of performance/reliability, cost and time. The addition of new data, through updating, can be incorporated into the process in a clear and documentable way.

The method supports the unification and integration of the planning/exploration/characterization/analysis/design process and products required in a complex project dealing with multiple geotechnical concerns (e.g., slope stability, foundation support, excavatability, etc.) along a variable terrain route (14,15). However, the methodology outlined here cannot be successfully applied in a mechanical, cook book manner—sound interpretive and integrative judgment and technical understanding is unavoidable in successfully developing and utilizing geotechnical evaluations in the overall design process.

Appendix.—References

1. Benjamin, J. R., and Cornell, A. C., *Probability, Statistics, and Decision for Civil Engineers,* McGraw-Hill, New York, N.Y., 1970.
2. Bowker, A. H., and Lieberman, G. J., *Engineering Statistics,* 2nd ed., Prentice-Hall, Englewood Cliffs, N.J., 1972.
3. Harr, M. E., *Mechanics of Particulate Media—A Probabilistic Approach,* McGraw-Hill, New York, N.Y., 1977.
4. Hunter, J. M., et al., "Site and Route Studies," Ch. 5, *Permafrost Engineering Design and Construction,* G. H. Johnston, ed., John Wiley and Sons, New York, N.Y., 1981.
5. Kreig, R. A., and Reger, R. D., "Preconstruction Terrain Evaluation for the Trans-Alaska Pipeline Project," *Geomorphology and Engineering,* Donald R. Coates, ed., Dowden, Hutchinson, and Ross, Stroudsburg, Pa., 1976, pp. 55–76.
6. Lumb, P., "Variability of Natural Soils," *Canadian Geotechnical Journal,* Vol. 3, No. 2, 1966.
7. Lumb, P., "Safety Factors and Probability Distribution of Soil Strength," *Canadian Geotechnical Journal,* Vol. 7, No. 3, 1970.
8. Mollard, J. D., *Landforms and Surface Materials of Canada,* 7th ed., J. D. Mollard and Assoc. Ltd., Regina, Saskatchewan, 1982.
9. Natrella, M. G., *Experimental Statistics,* National Bureau of Standards Handbook 91, 1963.
10. Raiffa, H., and Schlaifer, R., *Applied Statistical Decision Theory,* Harvard University Press, Cambridge, Mass., 1961.
11. Vanmarke, E. H., "Probabilistic Characterization of Soil Profiles," *Proceedings,* Site Characterization & Exploration, ASCE, June, 1978, pp. 199–216.
12. Vita, C. L., "A Landform-Based Probabilistic Methodology for Site Characterization," *Proceedings,* Nineteenth Annual Symposium on Engineering Geology and Soils Engineering, Idaho State University, April, 1982, pp. 339–354.

13. Vita, C. L., "Thaw Plug Stability and Thaw Settlement Evaluation for Arctic Transportation Routes: A Probabilistic Approach," Permafrost: Fourth International Conference, Proceedings, National Academy Press, Washington, D.C., July 1983, pp. 1321–1325.
14. Vita, C. L., "The Systematic Use of Geotechnical Data for Characterization of Transportation Routes," *Civil Engineering Systems*, England, Vol. 1, No. 1, Sept. 1983, pp. 23–32.
15. Vita, C. L., and Rooney, J. W., "A Unified Geotechnical Engineering Systems Approach Applied to Arctic Pipeline Planning and Design," *Proceedings*, Pipelines in Adverse Environments II, ASCE, Nov., 1983, pp. 291–299.
16. Way, D., *Terrain Analysis*, 2nd ed., McGraw-Hill, New York, N.Y., 1978.

J. Geotech. Engrg. 1984.110:1715-1734.