

**INCERTITUDES EN GÉOTECHNIQUE
ET DIMENSIONNEMENT DES OUVRAGES** 25/06/01numéro élaboré sous la direction de
J.-L. FAVRE et J.-P. MAGNANREVUE
FRANÇAISE
DE
GÉOTECHNIQUE
N° 93
4^e trimestre 2000**Sommaire**

Quelques spécificités du problème des incertitudes en géotechnique J.-P. MAGNAN	3
Les différents types d'erreurs et leur prise en compte dans les calculs géotechniques J.-L. FAVRE	11
Utilisation de la statistique descriptive en géotechnique M. CASSAN	21
Une méthode de détermination des valeurs caractéristiques des paramètres géotechniques F. BAGUELIN, J.-B. KOVARIK	35
Deux applications de l'analyse factorielle à des données géotechniques P. POUGET, N. CADOUX, J.-P. MAGNAN	43
Étude et prise en compte de la variabilité spatiale A. BOLLE	55
La méthode des éléments finis stochastiques en géotechnique G. AUVINET, R. MELLAH, F. MASROURI, J.-F. RODRIGUEZ	67
Décision et incomplétude D. BOISSIER	81

La *Revue française de géotechnique* est une publication scientifique trimestrielle parrainée par les Comités français de mécanique des sols, de mécanique des roches, et de géologie de l'ingénieur, qui publie des articles et des notes techniques relevant de ces domaines. Des discussions sur les travaux publiés dans la revue sont également les bienvenues.

La *Revue française de géotechnique* se consacre à l'étude pluridisciplinaire des interactions entre l'activité humaine et le terrain naturel. Elle est donc particulièrement concernée par tout ce qui se rapporte à l'intégration de l'homme dans son environnement, dans une perspective de développement durable, ce qui inclut la prise en compte des risques naturels et anthropiques, ainsi que la fiabilité, la sécurité et la durabilité des ouvrages. Le terrain naturel intervient dans de nombreuses constructions, soit parce qu'il les porte (fondations), les constitue (remblais routiers, barrages, barrières étanches de confinement de déchets, soutènements) ou les contient (ouvrages souterrains, tunnels) ; on y extrait également de nombreuses ressources pour la production d'énergie et de matériaux et on y stocke des déchets divers.

Les terrains naturels sont des milieux complexes, spécifiques et de caractéristiques variables dans l'espace et dans le temps, composés de solides et de fluides qui y circulent ou les imprègnent. L'identification de leurs propriétés, en termes de comportement mécanique et hydraulique, est coûteuse, et donc nécessairement incomplète et incertaine. Les problèmes posés sont variés, et leur résolution engage la responsabilité de l'ingénieur. On peut citer en particulier : la conception, la construction et la maintenance d'ouvrages bâtis sur, dans ou avec le terrain, dans des sites urbains ou extra-urbains ; la stabilité de sites naturels ou construits ; l'étude de la circulation et de la qualité de l'eau souterraine ; l'exploitation des ressources naturelles...

Les instructions aux auteurs sont publiées dans le premier numéro de chaque année, disponibles sur demande, et accessibles sur le site Internet des trois comités (www.geotechnique.org).

Les manuscrits sont à envoyer en trois exemplaires (dont un original) et une disquette contenant le fichier à l'un des rédacteurs en chef :

Pierre DELAGE	Françoise HOMAND	Jean-Paul TISOT
ENPC-CERMES	École de géologie (ENSG)	École de géologie (ENSG)
6-8, av. Blaise-Pascal	BP 40	BP 40
77455 Marne-la-Vallée CEDEX 2	54500 Vandœuvre-lès-Nancy	54500 Vandœuvre-lès-Nancy

Toute proposition de publication est examinée par le Comité de lecture.

REVUE FRANÇAISE DE GÉOTECHNIQUE

Rédacteur en chef : Pierre DELAGE (École nationale des ponts et chaussées)

Co-rédacteurs en chef : Françoise HOMAND, Jean-Paul TISOT (École de géologie de Nancy)

Comité de lecture : Gabriel AUVINET (UNAM, Mexico), Lucien BOURGUET (Hydrogéologue-expert), Bernard CAMBOU (École centrale de Lyon), Roger COJEAN (École des mines de Paris), Emmanuel DETOURNAY (University of Minnesota, USA), Jean-Louis DURVILLE (LCPC), Dominique FOURMAINTRAUX (Total Fina Elf), Alain GUILLOUX (Terrasol), Marc PANET (FC International), Aurèle PARRIAUX (École polytechnique fédérale de Lausanne, Suisse), Pierre VEZOLE (SAE), Gérard VOUILLE (École des mines de Paris)

Revue trimestrielle

Abonnement 2000 (numéros 90 à 93) franco : 710 F

Prix au numéro franco : 198 F (valable également pour les numéros anciens)

La revue est expédiée par avion dans les D.O.M.-T.O.M. et à l'étranger.

Sommaires des numéros anciens sur demande.

Presses de l'École nationale des ponts et chaussées

28, rue des Saints-Pères, 75007 Paris - Tél. : 01 44 58 27 40 – presses.ponts@mail.enpc.fr

Impression : Corlet, Imprimeur, S.A. 14110 Condé-sur-Noireau.

N° d'imprimeur : 52609. Dépôt légal : mai 2001

Les articles publiés dans cette revue n'engagent que la responsabilité de leurs auteurs. Tous droits de reproduction, de traduction et d'adaptation réservés pour tous pays.

Presses de l'École nationale des
Ponts et chaussées

(©) 2000

Commission paritaire n° 60855

ISSN 0181 — 0529

Quelques spécificités du problème des incertitudes en géotechnique

JEAN-PIERRE MAGNAN

Laboratoire central
des ponts et chaussées
58, bd Lefebvre
75732 Paris Cedex 15

Résumé

L'auteur introduit par des exemples les sources usuelles d'incertitudes de la géotechnique : manque de chance dans l'implantation des sondages, erreurs humaines dans la manipulation ou l'interprétation des données, remaniement des éprouvettes de laboratoire... Dans ce contexte, la caractérisation des incertitudes, définies comme l'écart entre le modèle de l'ingénieur et la réalité, est une entreprise complexe, qui ne se réduit pas à l'application de simples analyses statistiques ou probabilistes. Il en est de même pour la théorie de la décision, qui nécessite des schémas d'analyse plus complexes que ceux que l'on avait appliqués dans les années 1960-1970. Néanmoins, ces techniques d'analyse doivent être connues pour éclairer la réflexion des ingénieurs.

Mots-clés : sources d'incertitude, analyse de risque, fondations, erreur.

Some specific aspects of the question of uncertainties in geotechnical engineering

Abstract

The author describes, by means of examples, the usual sources of uncertainty in geotechnical engineering: bad luck when choosing the location of soundings, human errors during data handling or interpretation, remoulding of samples before laboratory testing... In such a context, characterising uncertainties, defined as the distance between the engineer's model and reality, cannot be restricted to the use of elementary statistical or probabilistic techniques of analysis. The situation is the same for the decision theory, which needs more complex models or analysis than those used in the 1960-1970'ies. Nevertheless, these techniques must be known to engineers, in order to help them thinking more efficiently.

Key words : sources of incertainly, risk analysis, foundations, errors.

NDLE : Les discussions sur
cet article sont acceptées
jusqu'au 1^{er} août 2001.

Introduction

Parmi les sciences de l'ingénieur, la géotechnique a une spécificité lourde de conséquences : l'ingénieur chargé de l'aménagement d'un site ou de la conception d'un ouvrage peut spécifier une partie des matériaux qu'il utilise, comme le béton, l'acier, le bois, les matières plastiques... Mais il doit s'accommoder des terrains existant sur le site et, pire encore, commencer par les identifier et les caractériser avant d'entreprendre la conception du projet.

L'identification des sols et des roches s'effectue au moyen de sondages (linéaires), d'essais ponctuels et de mesures géophysiques sur des surfaces ou des volumes. Comme le volume et le budget des reconnaissances sont toujours limités, la connaissance du site d'un projet reste partielle et les données des calculs sont incertaines.

L'intérêt des spécialistes de la géotechnique pour la description statistique des sols et des roches et pour les calculs de probabilités est né de cette incertitude, qui ne traduit pas un état aléatoire de la nature mais le niveau de connaissance ou de méconnaissance que l'on en a.

Le géotechnicien observe la nature à une échelle où il n'y a pas de doute sur la présence et les propriétés des matériaux. Dans chaque volume élémentaire, qui est infiniment grand par rapport aux atomes ou aux rayonnements, il y a un matériau (ou plusieurs matériaux) dont l'état et le comportement sont parfaitement définis et que l'on peut mesurer si on le désire. Mais on ne peut prélever et caractériser tout le sol d'un site. Entre des points, des lignes, des surfaces ou des volumes où l'on fait des mesures, il reste des espaces dont les propriétés doivent être estimées. Dans un volume de sol ou de roche, la nature et les propriétés des matériaux sont donc des fonctions aux valeurs estimées, assimilables aux champs aléatoires de la théorie des probabilités parce qu'en chaque point de nombreuses valeurs sont possibles, même si une seule (inconnue) existe réellement.

Dans ce texte, la probabilité ne décrira pas la nature aléatoire de la matière mais plutôt, comme l'a posé Cornell au début des années 1960, le degré de confiance de l'ingénieur pour le modèle qu'il a construit des massifs de sols et de roches existant sur le site de son étude ou du comportement de l'ouvrage qu'il étudie.

Cette définition implique la nature, qui existe, et un modèle, qui la représente. Tous les ingénieurs font des modèles et chaque modèle a un écart par rapport à la réalité. Les modèles de la géotechnique peuvent être très représentatifs ou très incertains suivant la complexité du site, que l'on ne connaît généralement pas a priori et qu'il faut identifier en même temps que l'on fait les mesures.

Les différentes classes d'erreurs, d'incertitudes et d'aléas que l'on peut rencontrer en géotechnique ainsi que leur prise en compte possible sont détaillés dans l'article de J.-L. Favre (2000). Nous allons les introduire ici à travers quelques exemples. Puis nous examinerons, dans le cas des fondations sur pieux, le niveau de complexité nécessaire des analyses d'incertitudes en géotechnique, qui s'avèrent nettement plus complexes que les premiers travaux des années 1960-1970.

1

Les surprises d'une plaine côtière

La reconnaissance géotechnique d'un terrain de 25 ha (500 m x 500 m) avait été organisée de façon méthodique, avec des sondages au pénétromètre statique et des sondages carottés avec essais de laboratoire (Fig. 1.a). Rien en surface n'indiquait de variabilité particulière des terrains et l'ingénieur chargé de l'étude établit un modèle simple du site, constitué de couches horizontales dont les propriétés étaient la moyenne des résultats des essais de laboratoire. Leur variabilité était faible et ce modèle servit au calcul des tassements du sol sous un remblai général de quelques mètres de hauteur. Ce modèle était bien calé sur les essais œdométriques. Pourtant, il s'avéra complètement faux : le site était traversé, selon l'axe d'implantation des sondages carottés, par une vallée fossile remplie de sols compressibles (Fig. 1.b). Les essais œdométriques représentaient uniquement la vallée. Le reste du site ne comportait que des argiles surconsolidées anciennes, dans lesquelles la vallée avait été creusée.

Cet exemple illustre à la fois un problème d'échantillonnage et un problème d'interprétation des données. L'échantillonnage (le choix de la position des essais) conditionne la représentativité des données. Dans l'exemple présenté, ces données ne reflétaient pas le site entier. Tout géotechnicien est à la merci d'un tel « manque de chance ». En revanche, l'interprétation des données était incorrecte aussi, par inexpérience de l'ingénieur responsable des calculs de tassements : les sondages de pénétration statique permettaient de délimiter clairement la vallée fossile et de guider l'interprétation du site.

Les informations disponibles permettaient donc de définir le site mais pas par une analyse simple des seuls paramètres de compressibilité œdométrique, qui étaient ceux nécessaires au calcul.

L'analyse simultanée d'informations d'origine et de nature diverses est très souvent nécessaire dans les études géotechniques et peut nécessiter des techniques d'analyse complexes, telles les méthodes d'analyse factorielle comme l'analyse en composantes principales qui est présentée dans l'article de Pouget *et al.* (2000).

2

De l'importance du transport des carottes de sol

Sur un autre projet, où l'on voulait construire un barrage à travers la vallée d'un fleuve, les couches superficielles des sols, jusqu'à une profondeur d'une dizaine de mètres, étaient constituées d'argiles molles sensibles. Ces argiles étaient assez résistantes pour porter les quelques mètres de remblai constituant le barrage dans le lit majeur du fleuve. Néanmoins, le projet établi sur la base d'essais de laboratoire réalisés à quelques milliers de kilomètres conclut à la nécessité d'un barrage en palplanches accrochées à une ligne de pieux : le laboratoire avait déclaré que les argiles qu'il avait reçues étaient à l'état liquide et ne pouvaient donc porter aucune charge.

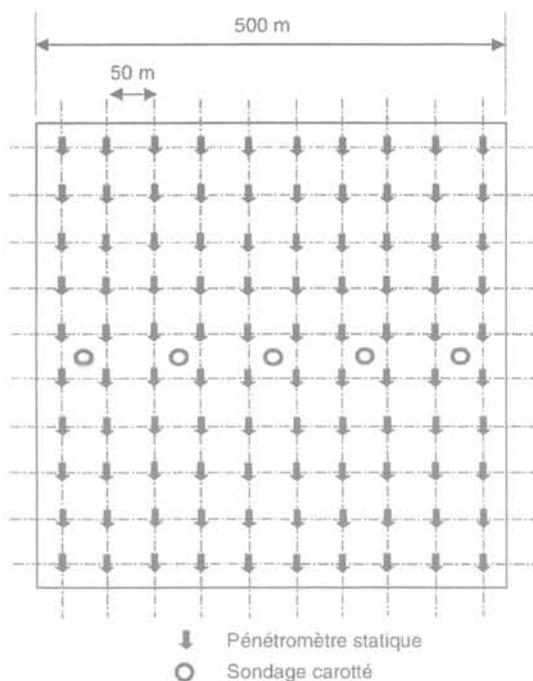


FIG. 1a Implantation des sondages.
Location of soundings.

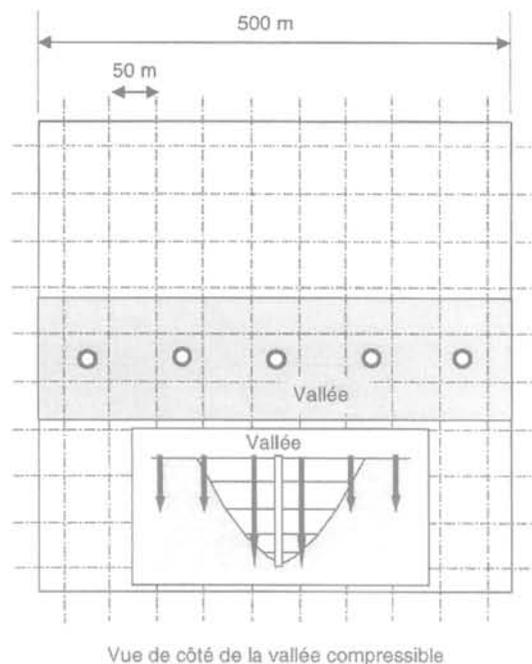


FIG. 1b Position de la vallée compressible.
Location of the soft soil valley.

FIG. 1 Plan schématique du site qui comportait une vallée fossile.

Schematic map of a site with a hidden valley.

Cet exemple montre l'importance de la qualité des mesures : si le résultat d'une mesure est faux (ici par remaniement préalable de l'éprouvette, mais cela peut venir de l'appareil de mesure ou d'une erreur de transmission des résultats), le modèle sera faux, même si tous les résultats des essais sont identiques.

3

Il ne faut pas toujours éliminer les données atypiques

Une des démarches fréquentes des analyses statistiques consiste à éliminer les données qui perturbent l'analyse parce qu'elles sont trop différentes des autres. Dans certains cas, cela fait partie du modèle d'analyse : une pluie catastrophique ou un séisme de magnitude élevée peuvent être traités comme des événements rares à longue période de retour et être éliminés de l'analyse des risques pour une période de durée limitée. Mais, quand il s'agit de propriétés des sols sur un site, il vaut mieux réfléchir soigneusement avant d'éliminer des données, comme l'illustre l'exemple suivant.

Sur le tracé d'une autoroute, la reconnaissance géotechnique avait révélé l'existence de deux couches de sols argileux compressibles, séparées par une couche de sable. La construction fut organisée en deux étapes, séparées de trois mois d'attente pour que la consolidation de l'argile ait le temps de s'effectuer. Lors de la mise en place du remblai de la deuxième phase, une rupture se produisit sur une centaine de mètres de longueur, retardant de plus d'un an l'achèvement des travaux. L'enquête réalisée après la rupture révéla que,

dans la zone du glissement, la couche intermédiaire de sable était remplacée par une couche d'argile sableuse de perméabilité voisine de celle des couches d'argile. L'ingénieur responsable du projet avait bien noté que le sondage réalisé dans cette zone avait donné des résultats un peu différents des autres... Mais au lieu de demander des sondages de contrôle, il avait jugé que c'était une information aberrante et l'avait éliminée de son modèle du site.

Dans les reconnaissances géotechniques, il n'y a pas de raison que la nature soit systématiquement conforme à l'idée simplifiée que l'on se fait d'elle. La ressemblance d'une donnée aux autres n'est pas un critère de tri... Il faut vérifier si ce n'est pas plutôt un indice de la mauvaise qualité du modèle du site, que des données complémentaires pourront permettre d'améliorer.

4

À propos d'erreurs grossières

Lorsque l'on entreprend des analyses statistiques, on admet de façon plus ou moins consciente un certain nombre d'hypothèses, dont la première est qu'il existe des lois de probabilité des objets étudiés. L'image la plus fréquente de ces lois de probabilité est celle d'une fonction dont la valeur moyenne est connue en chaque point mais dont les valeurs réelles varient de façon aléatoire autour de cette valeur moyenne. La possibilité d'estimer les paramètres de la dérive (variation de la valeur moyenne) du paramètre étudié dans l'espace ou en fonction de la profondeur n'est pas une garantie de

la validité du modèle. Ce n'est pas l'analyse statistique qui peut protéger l'ingénieur contre les fautes grossières, par nature non probabilisables, que l'on observe parfois dans les projets.

Un événement survenu lors de la reconnaissance géotechnique du tracé de la voie d'accès à une usine sidérurgique déjà construite a fait croire pendant plus d'un an que la construction de cette voie d'accès devrait être abandonnée: l'origine de cette crise était une simple erreur de numérotation des caisses contenant les carottes des sols prélevés sur le site: des échantillons d'argiles raides prélevés à plus de vingt mètres de profondeur furent échangés avec des échantillons d'argiles molles prélevés en surface. Les essais de compressibilité réalisés sur les carottes donnèrent des caractéristiques très faibles pour les sols prélevés en profondeur et des pressions de préconsolidation très inférieures à la contrainte effective initiale à cette profondeur. L'application mécanique des formules de calcul des tassements montrait que ces couches profondes auraient dû tasser de plusieurs mètres sous le poids des couches supérieures et sous leur propre poids. Il fallut plus d'un an pour que la vraisemblance des données géotechniques fût remise en cause, en s'appuyant sur des règles de cohérence géologique et géotechnique du site sans aucun rapport avec les données analysées.

Ce type d'erreurs, heureusement rare, n'est pas décelable par une simple manipulation des données. Les erreurs signalées dans les exemples précédents (mauvais choix des données à analyser, remaniement du sol avant les essais) sont également très rares et font partie de cette catégorie des erreurs humaines ou grossières que l'on ne sait pas traiter de façon statistique ou probabiliste. Les codes de bonne pratique et recommandations professionnelles tentent de les écarter par une exhortation à un travail de qualité, exécuté par du personnel formé et compétent, voire par la vérification du travail par une autre personne.

5

L'incertitude sur les résultats des calculs dépend de facteurs complexes

Les spécialistes de la géotechnique ne se contentent pas de créer des modèles de sites, ils calculent aussi des ouvrages de formes variées. On comprend facilement que si les données du calcul sont incertaines (aléatoires), le résultat du calcul le sera aussi. La théorie des probabilités dispose pour l'étudier de la notion de fonction aléatoire, qui n'est nullement contraire à la définition de la probabilité que nous avons adoptée: le résultat du calcul est aussi un modèle du comportement réel de l'ouvrage. L'ouvrage a un comportement unique (en général) et la probabilité est celle de l'égalité du calcul (incertain) avec ce comportement réel du sol et de l'ouvrage.

L'évaluation de l'incertitude sur le résultat d'un calcul est compliquée en pratique par l'existence de sources d'incertitudes autres que celles des données (Fig. 2): en général, les modèles de calcul ne sont qu'une approximation de la réalité du fonctionnement mécanique des massifs de sols ou de roches et le modèle géométrique du massif peut être lui-même très

approximatif. Les charges imposées à l'ouvrage par son environnement sont mal connues. Et l'ouvrage finalement obtenu dépend aussi des conditions de sa réalisation. On se trouve donc dans la situation d'estimer des incertitudes sur des résultats numériques de calculs qui ne décrivent que de façon approximative la réalité du site et des travaux.

De telles analyses nécessitent des outils complexes dont la validation est difficile et dont le développement rencontre pour cette raison de nombreux obstacles.

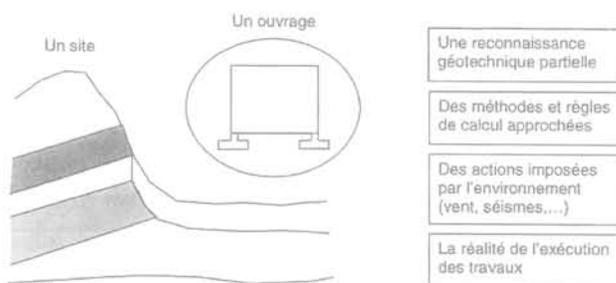


FIG. 2 Les incertitudes des calculs géotechniques
Uncertainly sources in geotechnical design.

6

De l'utilité des corrélations

La rareté des données géotechniques à l'échelle d'un projet courant oblige l'ingénieur à utiliser toute l'information existante, même si elle ne se rapporte qu'indirectement aux phénomènes qui conditionnent la conception et le fonctionnement de l'ouvrage. Les corrélations entre paramètres contribuent à la valorisation de cette information en permettant de contrôler la vraisemblance des résultats des essais, de détecter des anomalies, ou même de fabriquer les données de certains calculs.

L'utilisation de corrélations est facile... les techniques de recherche de corrélations par régression linéaire sont simples et les ouvrages spécialisés en proposent beaucoup. Il faut pourtant être très vigilant dans cette démarche et ne rechercher des corrélations qu'entre des paramètres dont la mécanique des sols ou des roches peut expliquer le lien.

7

Évaluer les risques

L'évaluation des risques est l'objectif affiché de beaucoup de recherches récentes dans tous les secteurs de la société. Le cadre d'analyse est la théorie de la décision, avec des schémas de causes et conséquences, des indices de fiabilité, etc. Les outils sont fournis par les théories statistiques et probabilistes, parmi lesquelles on a observé au cours des dernières années l'émergence de la notion d'ensembles flous pour traiter les phénomènes difficiles à probabiliser. Les schémas d'analyse sont attendus des différentes disciplines concernées, c'est pourquoi il n'est pas éton-

nant que les difficultés pratiques puissent différer d'un domaine à l'autre.

L'évolution des applications des probabilités et statistiques aux études de mécanique des sols, avec l'idée de définir des niveaux de risques associés aux ouvrages, s'avère assez différente des prévisions que l'on a pu faire au cours des années 1970. Le formalisme des méthodes d'analyse de risque et de décision, qui était facile à mettre en œuvre pour des calculs limités à quelques additions ou multiplications, n'a pas suivi l'évolution des méthodes de calcul modernes, mais on s'est aussi rendu compte que l'on manque d'éléments d'évaluation des niveaux de risques réels des ouvrages.

8

L'ouvrage de référence

Pour les ingénieurs et chercheurs qui se sont intéressés aux applications des statistiques et probabilités dans le domaine du génie civil et, plus particulièrement de la géotechnique, au cours des années 70, un ouvrage récent servait de référence, le livre de Benjamin et Cornell (1970) sur les probabilités, les statistiques et la décision pour les ingénieurs civils (Fig. 3).

Dans leur préface, les auteurs écrivaient que : « Ce livre est destiné aux étudiants, praticiens, enseignants et chercheurs en génie civil... (Il s'appuie sur) l'utilisation d'exemples et de problèmes tirés du domaine du génie civil... La raison principale de ce livre (est qu') une nouvelle théorie, encore controversée, se développe autour d'un objectif de prise de décision économique... »

De fait, ce livre avait pour axe central les outils de la

théorie de la décision. Sa lecture était et reste encore revigorante sur les objectifs qu'il est possible de poursuivre. Mais il est utile aussi de regarder les trente années passées et de se demander ce qui a évolué et pourquoi cette idée n'a pas débouché sur des applications pratiques. Nous l'examinerons sur le thème des fondations sur pieux.

À la page 527 de l'ouvrage, on trouve un exemple de problème servant de lien entre le livre de Benjamin et Cornell (1970) et le génie civil : un ingénieur doit sélectionner la longueur de pieux à enfoncer par battage jusqu'à un substratum dont la profondeur est incertaine. Les actions possibles sont de foncer des pieux de 20 m ou de 25 m et les profondeurs possibles du substratum sont de 20 m ou 25 m (Tableau I). Les conséquences des couples « décision-état » sont données dans le tableau I. La solution explique comment prendre la meilleure décision. Une variante du problème consiste à inclure dans l'analyse les résultats d'un essai sonique de recherche du niveau du substratum, puis le battage préalable d'un pieu pour faire disparaître l'incertitude.

TABLEAU Conséquences des actions suivant les états de la nature.

Consequences of actions depending on the states of nature.

Actions États de la nature	Pieu de 20 m	Pieu de 25 m
Rocher à 20 m	Décision correcte, pas de perte	5 m de pieu à couper et enlever, perte de 100 unités
Rocher à 25 m	Équipement de battage à modifier, perte de 400 unités	Décision correcte, pas de perte

L'ouvrage de Benjamin et Cornell définit clairement les fondements philosophiques de leur analyse. Pour eux, l'incertitude naît de la différence entre la réalité et le modèle qui la représente. Le modèle est l'idée que l'ingénieur se fait de la réalité, qu'elle provienne de l'analyse de mesures, de calculs ou d'idées « généralement admises ». On peut pour éclairer les idées citer comme exemples de modèles, le modèle géologique ou géotechnique d'un site, ou le résultat d'un calcul.

Le problème de cette approche est qu'elle ne s'intéresse pas au cœur de l'activité du géotechnicien, qui est la conception et la justification des projets d'ouvrages. Ainsi, dans l'exemple rappelé ci-dessus, le calcul du pieu se limite à « atteindre le substratum rocheux ». C'est un type d'analyse géotechnique pour le moins très élémentaire. C'est probablement aussi le seul type d'analyse que l'on sache vraiment faire, trente ans plus tard.

9

Les questions et compétences de l'an 2000

Les trente années écoulées depuis la publication du livre de Benjamin et Cornell (1970) ont été l'occasion

Probability, Statistics, and Decision for Civil Engineers

Jack R. Benjamin
Professor of Civil Engineering
Stanford University

C. Allin Cornell
Associate Professor of Civil Engineering
Massachusetts Institute of Technology

Probabilités, statistiques et décision pour les ingénieurs civils

J.R. Benjamin, C.A. Cornell

1. Analyse statistique des données
2. Éléments de la théorie des probabilités
3. Modèles probabilistes courants
4. Modèles probabilistes et données observées
5. Théorie élémentaire bayésienne de la décision
6. Analyse de décision des processus aléatoires indépendants

FIG. 3 L'ouvrage de référence des années 70.
The reference book of the 1970'sies.

d'une prise de conscience que l'évaluation des incertitudes sur la conception et la justification des ouvrages était vraiment un problème complexe. On peut le voir sur le même exemple d'un pieu dont on doit justifier la capacité portante en suivant les règles de calcul en vigueur.

Selon les règlements ou normes en vigueur dans la plupart des pays, la justification d'un pieu (Fig. 4) soumis à l'action de charges externes \bar{F} et porté par la

résistance de pointe \bar{Q}_p et le frottement latéral \bar{Q}_s

s'appuie sur deux conditions pour chaque combinaison de charge : une condition d'équilibre de la forme ;

$$\bar{F} < \bar{Q}_s + \bar{Q}_p,$$

et une condition sur les déplacements ou déformations du pieu :

Déplacement < seuil prescrit.

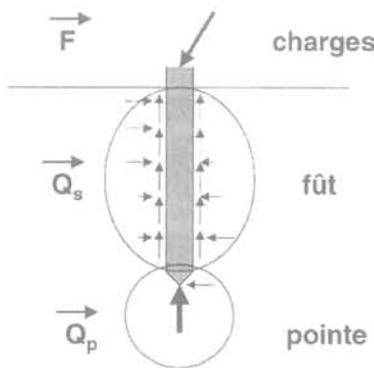


FIG. 4 Charges et résistance du sol sur un pieu.
Loads and soil resistance on a pile.

Les procédures de calcul traitent globalement le problème et donnent les valeurs à comparer pour justifier l'ouvrage. Si l'on veut rechercher par contre les sources d'incertitudes, il faut identifier les modèles qui représentent les différentes réalités qui conditionnent le fonctionnement de l'ouvrage. Ces modèles correspondent aux charges, au comportement du sol, à l'interaction sol-pieu, à l'interaction pieu-structure, à la déformation du pieu, etc. On en vient alors à évaluer les écarts entre la réalité et chaque modèle, ce qui n'est pas facile à réaliser.

Les incertitudes sur les charges peuvent être rapportées aux combinaisons (arbitraires) des charges élémentaires et à ces charges élémentaires : masse des matériaux (géométrie et masse volumique), charges mobiles, reports de charges dus au fonctionnement mécanique de la structure, effets du vent, de la neige, des chocs, des séismes, du frottement négatif...

L'analyse des incertitudes sur les sols et sur les modèles de calcul est traitée différemment selon la méthode de calcul choisie. La méthode pressiométrique, d'usage très courant en France (Fig. 5), utilise des résultats de mesures mécaniques (l'essai pressiométrique) mais aussi des abaques qui expriment l'effet de la nature des sols et des types de pieux, et qui ont été calés sur l'expérience... La fiabilité d'une procédure de calcul dont certaines parties ont été calées sur

l'expérience mais qui dépend tout de même d'essais ponctuels et de qualité inconnue pose des questions difficiles, d'autant que le calcul complet incorpore d'autres phénomènes (charges horizontales, effets de groupe, effet répartiteur de la rigidité de la structure portées par les pieux) qu'il faudrait aussi caractériser.

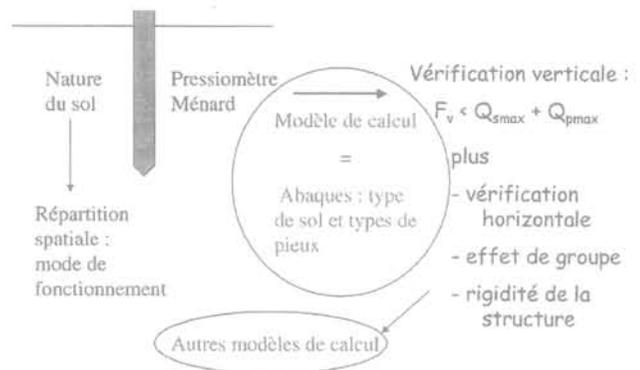


FIG. 5 Modèle d'interaction sol-pieu.
Soil-pile interaction model.

L'utilisation d'autres modèles (méthodes de calcul utilisées au Pays-Bas, calcul en éléments finis...) oblige à refaire complètement l'analyse d'incertitudes.

Cet exemple à peine ébauché illustre à notre avis l'évolution des attentes des ingénieurs géotechniciens vis-à-vis des méthodes d'analyse statistique et probabiliste au cours des trente dernières années : les schémas de calcul des années 1970 étaient suffisamment simples pour permettre le déroulement d'une procédure d'analyse d'incertitude, voire de prise de décision... Mais ils sont trop simples par rapport aux développements modernes des connaissances sur le fonctionnement mécanique des pieux (et, de façon générale, de tous les ouvrages géotechniques). L'étalement global des méthodes de calcul, y compris les méthodes de mesure des paramètres et les méthodes d'exécution des travaux, n'a pratiquement été fait dans aucun pays, à part quelques exceptions comme l'étalement du calcul des pieux au pressiomètre et au pénétromètre statique en France. Cet étalement est indispensable pour évaluer le biais de la procédure de calcul (écart de la valeur moyenne calculée par rapport à la réalité). Sans cette vérification du bon cadrage de la procédure de calcul, il n'est pas possible d'évaluer les incertitudes et les risques associés aux ouvrages. Ce travail est long mais il devra être fait si l'on veut préciser les valeurs réelles des marges de sécurité en géotechnique et éventuellement les réduire un jour de façon volontaire et contrôlée. En l'absence d'un tel travail, l'application des procédures de dimensionnement validées par l'expérience reste le seul garant de la qualité du travail des ingénieurs.

Conclusion

Les études de géotechnique souffrent d'incertitudes dont les origines sont complexes. Nous avons cité l'interpolation ou l'extrapolation des données autour des points de mesure, la fiabilité des résultats de ces mesures, la complexité des problèmes étudiés et le volume réduit des informations disponibles, la caractérisation des résultats des calculs mais aussi l'existence

d'incertitudes ou d'erreurs tellement importantes et inexplicables qu'elles perturbent les schémas classiques de l'analyse statistique ou probabiliste.

L'avenir des procédures d'analyse de risques et de prise de décision, connues sous le nom de théorie bayésienne de la décision, dépend de la possibilité d'inclure dans ces analyses les modèles modernes de fonctionnement des ouvrages. Nous possédons actuellement des règles historiques, partiellement justifiées et codifiées, dont l'utilisation garantit un résultat satisfaisant. Pourra-t-on affiner ces règles pour diminuer en

moyenne la marge de sécurité? La réponse est incertaine : les outils d'analyse dépendent de données expérimentales complexes ou rares... La prévision est difficile dans un contexte de concurrence économique difficile à contrôler.

L'application des techniques d'analyse statistique et probabiliste présente donc un intérêt potentiel dans les différents domaines de la géotechnique et la connaissance de ces méthodes, mais aussi de leurs limites, est un élément important de la formation des ingénieurs de génie civil et de géotechnique.

Bibliographie

Benjamin J.R., Cornell C.A. (1970) – *Probability, statistics and decision for civil engineers*. New York, Mc-Graw-Hill Book Company, 684 p.

Faure J.-L. (2000) – « Les différents types d'erreur et leur prise en compte dans les calculs géotechniques ». *Revue française de géotechnique*, n° 93, p. 11-20.

Pouget P., Cadoux N., Magnan J.P. (2000) – « Deux applications de l'analyse factorielle à des données géotechniques ». *Revue française de géotechnique*, n° 93, p. 43-54.

Les différents types d'erreurs et leur prise en compte dans les calculs géotechniques

J.-L. FAVRE

École centrale de Paris
92295 Châtenay-Malabry

Résumé

Les trois principales sources d'incertitudes sur les propriétés des sols et des roches sont tout d'abord précisées : la variabilité naturelle, les mesures, les modèles. Leur prise en compte possible est indiquée dans le cadre de l'analyse statistique et de la théorie des probabilités. L'attention est particulièrement attirée sur le problème de la réduction de variance, spécifique aux mécanismes de défaillance sur de grands volumes, et sur le poids très pénalisant du manque de mesures comparativement à la variabilité naturelle des terrains.

Mots clés : variabilité naturelle, incertitudes d'échantillonnage, erreurs de modèles, distance d'auto-corrélation, réduction de variance, fiabilité.

The different kinds of errors and their consideration in geotechnical design

Abstract

The three main sources of uncertainties about soils and rocks characteristics are described, namely : the natural variability, measures and models. Their possible consideration is described in the statistical analysis and within the theory of probabilities. Attention is specially drawn on the problem of variance reduction, typical in failure mechanisms linked with great volumes, and on the handicap due to the lack of measures.

Key words : natural variability, uncertainties of samplings, model errors, auto correlation distance, variance reduction, reliability.

NDLE : Les discussions sur cet article sont acceptées jusqu'au 1^{er} août 2001.

Les trois types d'incertitudes pour les sols et les roches

1.1

Le désordre ordonné

Le « désordre » qui affecte les propriétés des sols et des roches n'est pas aléatoire. Il est lié à l'histoire de leur formation qui a ses lois très précises (tectonique, érosion, transport et sédimentation, etc.) et à la nature même de la matière. Mais la complexité de ces lois et des conditions de leur application et la complexité de la matière font qu'il est impossible de connaître parfaitement ces propriétés : on est donc devant un phénomène incertain au sens où on ne peut prévoir de façon certaine ses réalisations, *i.e.* les propriétés en un point d'un massif. La première démarche est de connaître l'essence des phénomènes et pour cela une bonne reconnaissance géologique est toujours indispensable ; elle réduit le caractère incertain de la prévision.

1.1.1

Différents degrés d'incertain en géotechnique

Cet incertain naturel peut avoir divers degrés et donc entraîner différents moyens d'investigation, de prise en compte et modélisation :

- on peut se trouver dans des terrains très erratiques, à fortes hétérogénéités avec des fractures, des zones de faiblesse ou au contraire des points « durs », des arrivées d'eau, etc. Géométrie et propriétés sont très tourmentées. La modélisation est essentiellement déterministe et dichotomique (oui ou non). Elle peut aussi faire appel aux ensembles flous comme aide à la décision. Une telle démarche est illustrée par l'article de Daniel Boissier « Décision et incomplétude » ;

- on peut se trouver dans des terrains peu désordonnés présentant une stratification relativement régulière avec, pour chaque couche, des propriétés relativement régulières aussi. La reconnaissance géotechnique, guidée par le modèle géologique de terrain, prend ici toute son importance pour proposer un modèle mécanique de terrain. Les autres articles se placent plutôt dans ce cas où les analyses statistiques et le calcul des probabilités prennent tout leur sens ;

- enfin, les terrains peuvent avoir été fabriqués par l'homme, qu'il s'agisse de barrages en terre, de remblais, de terre armée, etc. On a affaire alors à des contrôles de fabrication et des tests de réception. Cet aspect n'est pas traité dans ce numéro spécial.

1.1.2

Différentes échelles de l'incertain

Un autre aspect de l'incertain naturel des sols et des roches est qu'il peut se manifester à diverses échelles spatiales, emboîtées, fractales. Ceci est illustré par une étude de l'École des mines de Paris sur le massif des Ardennes à propos des propriétés des roches (Fig. 1) : celles-ci décroissent en fonction de l'échelle d'observation ou du volume concerné, depuis l'échelle atomique,

où la première dislocation introduit une baisse brusque des propriétés ; les échelles suivantes provoquent la poursuite du phénomène tout en l'amortissant de plus en plus jusqu'à l'apparition, à l'échelle microscopique, de la première micro-fissure, et ainsi de suite jusqu'à l'échelle du massif montagneux parcouru de failles.

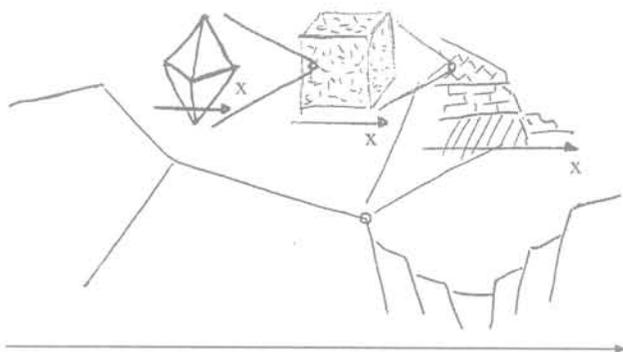
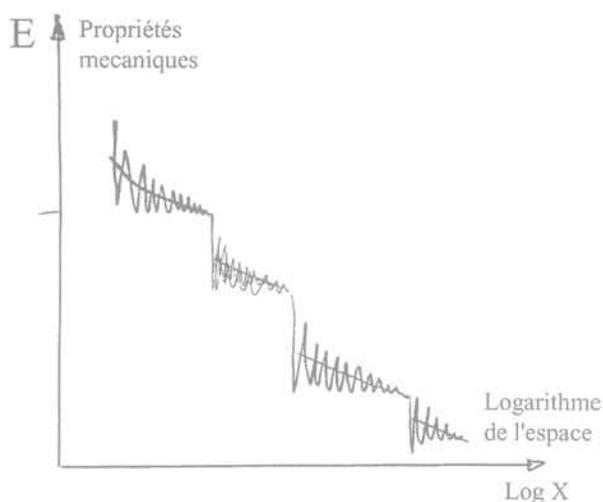


FIG 1 Échelles emboîtées pour les propriétés des roches.
Rock properties overlapped scales.

1.2

Les trois types d'incertitudes

1.2.1

Incertaines active et passive

Mais ce désordre naturel n'est pas la seule incertitude dans la connaissance des propriétés des sols. Une deuxième source vient de notre connaissance imparfaite de la nature : la nature EST. Un phénomène physique incertain, quand il se réalise, se réalise de façon certaine ; il présente une valeur sûre, déterministe, mais nous ne pouvons pas la connaître parfaitement à cause de l'imprécision et de l'imperfection de nos appareils de mesure. Nous commettons une erreur, distance entre une valeur vraie mais inconnue et une valeur donnée par l'appareil de mesure.

La première incertitude liée à la réalisation du phénomène physique est souvent appelée active et la deuxième, liée à la mesure, est souvent appelée passive.

Erreur de modèle

Enfin, il existe pour nous un troisième type d'incertitude, celle liée à l'utilisation de modèles de comportement. Les propriétés des sols et des roches rentrent dans des lois rhéologiques qui représentent plus ou moins bien la réalité de la relation contraintes-déformations. Ainsi on utilise essentiellement le modèle élastique linéaire isotrope-plastique parfait ou fragile, alors que le sol est mieux représenté par un modèle élasto-plastique orthotrope. Quant au comportement réel du sol, personne ne le connaîtra jamais ; on ne fait que l'approcher. On fait donc une erreur de modèle indépendamment de l'erreur que l'on fait sur la mesure des paramètres du modèle choisi.

2

Erreurs et incertitudes liées à la caractérisation des sols

2.1

Les erreurs d'observation et d'enquête

2.1.1

La loi des erreurs

La réalité est vraie, sûre mais inconnue. La mesure en est une estimation. On considérera donc la mesure comme une variable aléatoire. La différence entre la valeur vraie et la mesure est l'erreur de mesure ϵ : c'est aussi une variable aléatoire.

Cette erreur peut avoir une part systématique : $E[\epsilon] \neq 0$. On l'appelle un biais.

Elle a une part aléatoire $\sigma(\epsilon) \neq 0$. C'est la dispersion de l'erreur ou incertitude.

La loi des erreurs est représentée par la fameuse courbe en cloche (Fig. 2). C'est la loi « normale » ou de Laplace-Gauss, appelée ainsi car elle représente le comportement des phénomènes incertains « normaux » qui dépendent d'un grand nombre de petites causes plus ou moins indépendantes et du même ordre de grandeur.

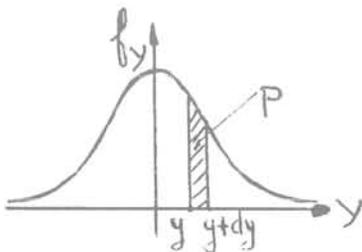


FIG.2 Loi des erreurs.
Errors law.

2.1.2

Erreurs d'observation

On distingue trois types d'erreurs d'observation :

- **Les erreurs de mesure proprement dites, ϵ_{ao}**
Elles sont liées à l'appareil de mesure et à l'opérateur.

Avec les progrès de la métrologie, on a considérablement réduit les erreurs d'imprécision et avec l'acquisition automatique, les erreurs d'opérateur. Encore faut-il avoir fait un étalonnage et un tarage corrects. Mais il reste les erreurs de reproductibilité liées au caractère destructif de la plupart des essais de géotechnique et les erreurs d'opérateur liées à la difficulté de certains essais (au triaxial, au pressiomètre : qualité du trou, etc.).

Enfin, il est bien connu que chaque essai représentant le « coup de main » de l'opérateur, on constate des différences systématiques plus ou moins grandes d'un laboratoire ou d'un bureau de reconnaissance à l'autre.

- **Les erreurs de représentativité, ϵ_{rep}**

Elles proviennent de la transformation de la mesure physique.

On ne mesure guère directement que des longueurs, des masses et des températures et la plupart des appareils sont basés sur des mesures de déplacement. Il faut donc transformer la mesure. On procède alors à un étalonnage de l'appareil et à son tarage. Mais il y a des transformations plus importantes et plus incertaines : section d'un échantillon au triaxial pour la contrainte σ_1 ? Pente initiale d'une courbe de chargement pour le module de Young ? (en fait, il faut parler de modules sécants, la limite élastique sur chemin déviatorique ne dépassant guère 10^{-5}).

- **Les erreurs de l'instant, ϵ_{inst}**

Elles proviennent de la variation de la propriété entre le moment où on la mesure et le moment où le sol est mis en œuvre.

Le cas typique est celui des propriétés du béton. Pour les sols, le remaniement relève de ce type d'erreur, en particulier pour tous les essais mécaniques de laboratoire et pour l'essai pressiométrique.

2.1.3

Les erreurs d'enquête

- **Les erreurs d'enquête proprement dite**

Ce sont les erreurs liées à une mauvaise conduite des reconnaissances.

Les mesures effectuées ne sont pas représentatives du problème : on décrit en figure 3 l'exemple d'un remblai récent mal identifié au droit de chaque appui qui le sollicite, un remblai ancien en biseau pouvant exister sous certains appuis ; la figure 3-b illustre le problème d'une fondation superficielle reconnue à moins de 3 à 5 fois sa largeur, ignorant ainsi une couche faible qui peut piloter les tassements.

- **Les erreurs d'échantillonnage**

Ce ne sont pas, à proprement parler, des erreurs.

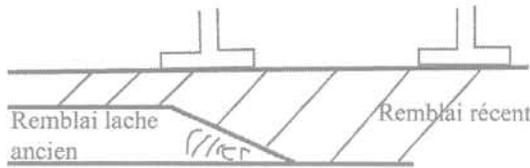
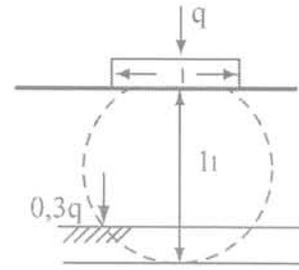


FIG. 3 Exemples de problèmes d'enquête.
Investigation difficultés cases.



Imaginons que chaque mesure est parfaite mais que la propriété varie d'un point à l'autre de l'espace. Cette variabilité naturelle sera modélisée comme une variable aléatoire dont on va estimer les caractéristiques avec les mesures. On a ainsi un échantillon statistique de taille n pour estimer les caractéristiques d'une population infinie. On estime ces caractéristiques par des intervalles de confiance : la moyenne (ou l'écart type) de la propriété, qui est sûre mais inconnue a une probabilité de 95 % par exemple d'être encadrée par deux valeurs dépendant des observations. L'utilisation des intervalles de confiance est illustrée par les articles de Maurice Cassan « Utilisation de la statistique descriptive en géotechnique » et de François Baguelin et Jean-Bernard Kovaric « Une méthode de détermination des paramètres géotechniques » pour le choix des paramètres de projet ou valeurs caractéristiques.

2.1.4

Protection contre les erreurs d'observation et d'enquête et modélisation

La modélisation des erreurs d'observation est naturellement faite par des variables aléatoires. On cherchera à définir en premier lieu l'espérance des erreurs $E[\epsilon_i]$, les biais, et leur variance $\sigma^2(\epsilon_i)$, leur dispersion. La protection contre ces différentes erreurs est claire puisque leurs causes sont claires : maîtriser celles-ci. Une attention particulière doit être apportée à la réalisation des essais : réalisation des trous de sondages pour le prélèvement comme pour le pressiomètre, prélèvement et transport des carottes, conservation des carottes, découpage et mise en place des échantillons.

La modélisation des erreurs d'enquête proprement dite relève de la modélisation des erreurs humaines. Celle-ci n'est pas très avancée, spécialement en géotechnique. La seule parade est de ne jamais négliger le modèle géologique et de faire appel, pour numériser ce modèle, à un vrai géotechnicien.

Les erreurs d'échantillonnage relèvent de la théorie de l'estimation. Elles ne posent aucun problème de modélisation :

- l'article de Jean-Pierre Magnan « Quelques spécificités du problème des incertitudes en géotechnique » illustre le poids que peuvent revêtir celles de ces erreurs qui sont peu ou pas quantifiables ;

- l'article de Pierre Pouget *et al.* « Deux applications de l'analyse factorielle à des données géotechniques » montre tout le parti qu'on peut tirer des analyses factorielles lorsqu'on dispose d'un grand nombre de données, ici pour définir les modèles de terrain.

2.2

La variabilité naturelle

C'est le caractère incertain du phénomène physique.

2.2.1

La modélisation par des fonctions aléatoires de l'espace

Les problèmes mécaniques sont des problèmes sur l'espace : une force est l'intégrale d'une contrainte sur une surface ou dans un massif, un déplacement est l'intégrale d'une déformation dans une direction ou dans un massif. Il s'agit donc d'intégrer des propriétés en tant que fonctions de l'espace. De plus, le sens commun veut qu'on juge les propriétés moyennes d'un volume de sol moins dispersées que ses propriétés ponctuelles : les fortes valeurs sont compensées par les faibles valeurs et plus le volume est grand plus le lissage est fort : on a bien plus de risque avec une semelle étroite qu'avec un radier.

On peut préciser cette idée avec l'exemple d'un essai au scissomètre de chantier :

Soit le profil de C_u , en fonction de la profondeur en trait fin sur la figure 4, enregistré tous les 10 cm avec des ailettes de 10 cm. Si on avait fait l'enregistrement tous les 40 cm avec des ailettes de 40 cm, on aurait obtenu la courbe en trait gras, dont chaque point est la moyenne de 4 points de la précédente. On obtient ainsi un profil moins dispersé. Ceci n'est ni plus ni moins que la règle de la variance de la moyenne de n valeurs, qui veut que celle-ci soit n fois plus petite que la variance de chaque valeur :

$$\text{Variance}[\bar{Z}] = \text{Variance}[Z] / n \quad \text{avec} \quad \bar{Z} = (1/n) \sum_i Z_i$$

ou encore

$$\sigma_{\bar{z}} = \sigma_z / \sqrt{n} \quad (1)$$

La dispersion du deuxième profil est deux fois plus petite que celle du premier profil. Cette question est facilement résolue dans le cas de valeurs discrètes indépendantes par la règle de la moyenne. Elle nécessite la définition de la fonction de covariance ou de la fonction d'auto corrélation des champs aléatoires et de la distance d'auto-corrélation a , distance à partir de laquelle les propriétés deviennent indépendantes, soit 1 m sur la figure 5. On retrouve les mêmes notions

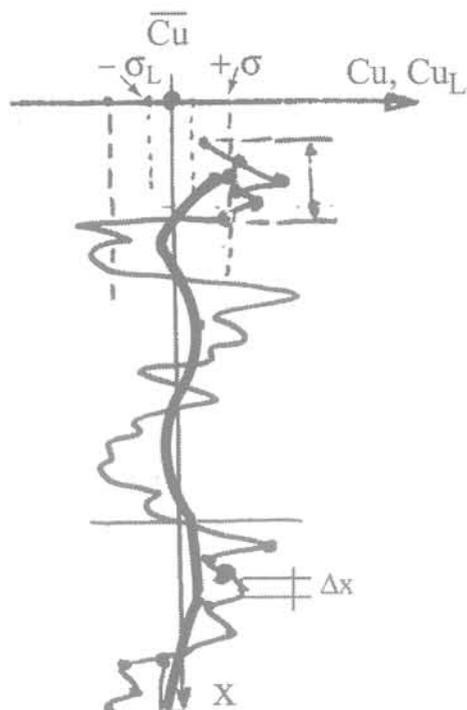


FIG. 4 Profils scissométriques avec deux ailettes différentes.
Two different vane tests profiles.

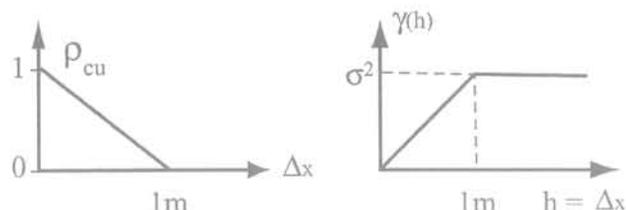


FIG. 5 Correlogramme et variogramme.
Correlogram and variogram.

en géostatistique où le variogramme correspond au correlogramme. Albert Bolle dans son article « étude et prise en compte de la variabilité naturelle » donne différents modèles pour la fonction d'auto corrélation et montre l'interprétation que l'on peut faire des correlogrammes expérimentaux.

2.2.2

Les premiers moments de la fonction aléatoire

Ces premiers moments peuvent être illustrés par le profil d'un sondage pénétrométrique : à chaque niveau (x), tous les 2 cm par exemple, on enregistre une réalisation de la résistance de pointe $R_p(x)$. On peut définir à chaque niveau les caractéristiques de la variable aléatoire « résistance de pointe ». Ces caractéristiques, l'espérance et la variance pour les deux premiers moments, deviennent des fonctions de l'espace $\bar{R}_p(x)$ et $\sigma_{R_p}^2(x)$ (Fig. 6).

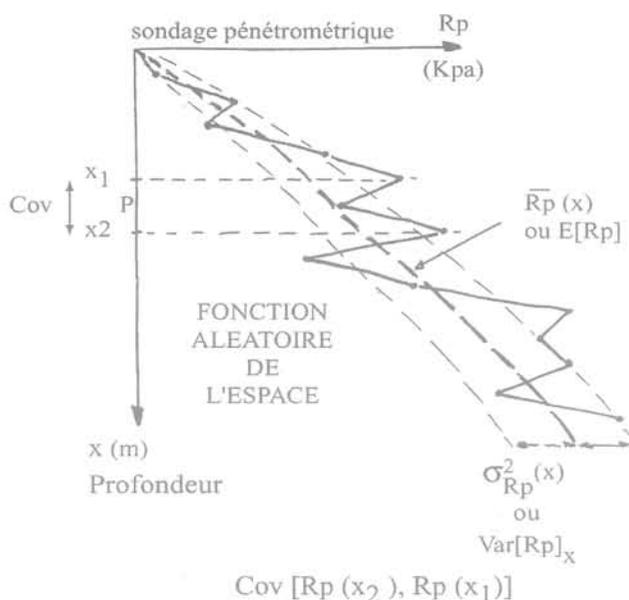


FIG. 6 Résistance de pointe au pénétromètre comme fonction aléatoire de l'espace.
Random field of cone penetration profile.

Pour être complet jusqu'à l'ordre 2, il faut définir la corrélation entre les différents niveaux. C'est la fonction d'auto covariance qui dépend des deux niveaux choisis x_1 et x_2 ou d'un des deux niveaux choisis x et de l'écart entre les deux niveaux Δx .

On cherchera, pour des commodités de modélisation et de traitement, à rendre le phénomène stationnaire, c'est-à-dire ayant les mêmes caractéristiques statistiques en tout point de l'espace. Cette exigence est très difficile à satisfaire et on se contentera de la stationnarité en moyenne d'ordre 2 qui rend seulement les deux premiers moments indépendants de l'espace.

Ainsi :

$$E[R_p^*] = \text{constante}$$

$$\text{Var}[R_p^*] = \text{constante}$$

$$\text{Cov}[R_{p1}^*, R_{p2}^*] = \text{Cov}[\Delta x]$$

$$R_p^* : \text{transformée de } R_p$$

2.3

Prise en compte des erreurs, des incertitudes et de la variabilité naturelle

2.3.1

Les erreurs d'observation

Lorsqu'un paramètre physique entre dans une procédure de dimensionnement d'ouvrage, il devient un paramètre de calcul. Ce paramètre de calcul doit être unique, quel que soit le paramètre physique qui est à son origine : C'est par exemple la valeur naturelle Z_N . Mais on commet sur ce paramètre les erreurs d'observation déjà mentionnées :

- celles de l'appareil et de l'opérateur, e_{op}
- celles de représentation, e_{rep}

– celles de l’instant (ou remaniement pour la géotechnique), e_{inst} .

Ainsi, pour tout paramètre physique mesuré Z_m , on peut écrire :

$$Z_m = Z_N + e_{ao} + e_{rep} + e_{inst} \quad (2)$$

Z_N étant la valeur naturelle ou vraie qui devient le paramètre de calcul

• L’espérance ou le biais

L’espérance de la valeur de calcul est alors :

$$E[Z_N] = E[Z_m] - \{E[e_{ao}] + E[e_{rep}] + E[e_{inst}]\}$$

On ne dispose pas de valeurs pour les biais qu’introduisent, sur la valeur mesurée par rapport à la valeur réelle, les différents facteurs (l’appareil, l’opérateur, les approximations dans la transformation de la mesure, le remaniement).

Le remaniement, la plupart du temps, minimise les valeurs de résistance du sol et semble être le biais le plus important en valeur absolue. Aussi en prenant :

$$E[Z_N] \approx E[Z_m] \quad (3)$$

on se place la plupart du temps du côté de la sécurité.

• La dispersion

La dispersion de la valeur de calcul est donnée par la variance :

$$\text{Var}[Z_m] = \text{Var}[Z_N] + \text{Var}[e_{ao}] + \text{Var}[e_{rep}] + \text{Var}[e_{inst}] + \sum_{i=1 \text{ à } 3} \text{Cov}[Z_N, e_i] + \sum_{i \neq j} \text{Cov}[e_i, e_j]$$

On peut valablement supposer que les erreurs sont indépendantes des valeurs naturelles et qu’elles sont indépendantes entre elles (loi des erreurs). Alors $\text{Cov}[Z_N, e_i] = \text{Cov}[e_i, e_j] = 0$.

Les variances étant des quantités positives, $\text{Var}[e_i] > 0$, alors :

$$\text{Var}[Z_N] \leq \text{Var}[Z_m] \quad (4)$$

Tout comme l’espérance, on ne connaît pas la variance des différentes erreurs. Aussi, en prenant pour dispersion de la valeur de calcul celle des valeurs mesurées, on se place du côté de la sécurité.

• La dépendance d’un point à l’autre de l’espace

Elle est donnée par la covariance entre deux mesures distantes de Δx :

$$\text{Cov}[Z_m(x), Z_m(x + \Delta x)] = \text{Cov}[Z_N(x) + \sum_i e_i(x), Z_N(x + \Delta x) + \sum_i e_i(x + \Delta x)]$$

Les erreurs étant indépendantes entre elles (sauf pour $\Delta x = 0$) et indépendantes de la valeur à mesurer, on a, pour $\Delta x \neq 0$:

$$\text{Cov}[Z_m(x), Z_m(x + \Delta x)] \approx \text{Cov}[Z_N(x), Z_N(x + \Delta x)] \quad (5)$$

A l’heure actuelle, en l’absence de modèles et de grandeurs pour les erreurs d’observation, on prend les valeurs mesurées ce qui, en général, place le géotechnicien du côté de la sécurité.

2.3.2

La variabilité naturelle et les erreurs d’échantillonnage

• Les incertitudes d’échantillonnage

Les incertitudes d’échantillonnage, venant du nombre limité de mesures pour estimer $\mu = E[Z_m]$ et $\sigma = \sqrt{\text{Var}[Z_m]}$, sont prises en compte par des intervalles de confiance qui découlent de la théorie de l’estimation : On prend une forte probabilité $1 - \alpha$ (en général $\alpha = 5\%$) pour que μ et σ tombent dans un intervalle qui dépend du nombre n de mesures faites et des valeurs de la moyenne empirique m et de l’écart type estimé s_e calculés avec ces mesures. Plus n est faible, plus l’intervalle est grand, c’est-à-dire plus l’incertitude est grande. Les articles de Maurice Cassan et de François Baguelin et Jean-Bernard Kovaric utilisent largement cette théorie de l’estimation et des tests d’hypothèses.

• La variabilité naturelle

La variabilité naturelle $Z_N(x)$ est prise en compte par l’indice de sécurité β qui permet théoriquement de calculer la probabilité de ruine. Le fait que cette résistance $Z_N(x)$ soit intégrée dans un volume V concerné par le mécanisme (ou sur une surface Σ ou une ligne L) est pris en compte par un champ de moyenne locale qui introduit une réduction de variance $\gamma(V)$. Cette réduction de variance est liée au coefficient d’auto corrélation de la manière suivante (à 1D) :

$$\gamma(L) = 2/L \cdot \int_0^L (1 - \Delta x / L) \cdot \rho(\Delta x) \cdot d(\Delta x) \quad (6)$$

Soit, avec le modèle triangulaire de corrélation de la figure 5 :

$$\rho(\Delta) = \begin{cases} 1 - |\Delta x| / a, & |\Delta x| \leq a \\ 0 & \text{sinon} \end{cases}$$

$$\gamma(L) = \begin{cases} 1 - L / 3a & L \leq a \\ (a / L)(1 - a / 3L) & L \geq a \end{cases} \quad (7)$$

Si le champ est séparable ($\gamma(V) = \gamma(L_x) \cdot \gamma(L_y) \cdot \gamma(L_z)$), on peut connaître la réduction de variance à 3D de façon très simple en connaissant les distances d’auto corrélation a_x , a_y et a_z et les dimensions du volume sollicité L_x , L_y et L_z .

Illustrons sur un exemple simple comment agit cette réduction de variance et comment on évalue la probabilité de ruine :

Soit une plate-forme offshore, par 100 m de fond, de poids déjaugé $P' = 500\,000$ t, à embase poids de 100 m de côté que l’on veut garantir, grâce à des bèches pénétrant de $D = 10$ m dans le fond marin, contre le ripage

vis-à-vis de forces horizontales (courants, houle, vent, etc.) extrêmes sur sa durée de vie (20 ans): $\bar{H}' = 200\ 000$ t avec un écart type $\sigma_{H'} = 20\ 000$ t. Les caractéristiques du fond marin, à 10 m, sont (moyenne et coefficient de variation) :

$$\begin{aligned} \bar{c}' &= 28 \text{ t/m}^2 & CV_{c'} &= 20\%, \\ \bar{\phi}' &= 10^\circ & CV_{\phi'} &= 13\%, \text{ avec } CV = \sigma_{\phi'} / \bar{\phi}' \end{aligned} \quad (8)$$

$$\begin{aligned} \bar{\gamma}' &= 10 \text{ KN/m}^3 & CV_{\gamma'} &= 7\% \\ \rho_{c\phi} &= -0,6 & COV(c', \phi') &= \rho_{c\phi} \cdot \sigma_{c'} \cdot \sigma_{\phi'} \\ a_v &= a_x = 33 \text{ m}, & a_z &= 1,5 \text{ m} \end{aligned}$$

(les distances d'auto-corrélation sont très différentes verticalement et horizontalement).

En faisant le calcul du coefficient de sécurité F avec les valeurs moyennes, on obtient pour R (résistance au cisaillement, uniquement sur la surface horizontale Σ à 10 m sans compter les effets de bords) :

$$R = \Sigma \bar{c}' + (P' + \Sigma \bar{\gamma}' D) \text{tg} \bar{\phi}' = 31,2 \cdot 10^4 \text{ t} \quad (9)$$

$$S = \bar{H}' = 20 \cdot 10^4 \text{ t}$$

d'où $F = R/S = 1,56$

Théoriquement, la probabilité de ruine est :

$$p_f = \text{Prob}((R - S) = g(Z) \leq 0) \text{ avec } Z = (c', \gamma', \text{tg}\phi', H') \quad (10)$$

$R - S = g(Z)$ est aussi appelé la fonction de performance ou marge de sécurité.

On montre (sous certaines conditions) qu'une valeur approchée de β est :

$$\beta \approx E[g(Z)] / \sigma_{g(Z)} = (\bar{R} - \bar{S}) / (\sigma_R^2 + \sigma_S^2)^{1/2} \text{ (puisque } R \text{ et } S \text{ sont indépendants).}$$

Une valeur approchée (sous certaines conditions) de p_f est :

$$p_f \approx \Phi(-\beta) \quad (11)$$

Φ étant la fonction de répartition de la loi normale.

En écrivant, par pure commodité, $Z = (X_1, X_2, X_3, X_4)$ il vient, en divisant par la surface Σ qui n'est pas incertaine :

$$g(Z) = X_1 + (a + bX_2)X_3 - X'_4 \quad X'_4 = X_4 / \Sigma$$

On peut approximer $g(Z)$ par un développement limité au premier ordre $g_1(Z)$ (linéarisation) en série de Taylor autour des valeurs moyennes :

$$g_1(Z) = g(\bar{Z}) + \Delta X_1 + (a + b\bar{X}_2)\Delta X_3 + b\bar{X}_3\Delta X_2 - \Delta X'_4 \quad (12)$$

On trouve que :

$$E[g_1(Z)] = g(\bar{Z}) = 11,2 \text{ t/m}^2 \quad (13)$$

et que :

$$\begin{aligned} \text{Var}[g_1(Z)] &= \text{Var}[X_1] + (a+b\bar{X}_2)^2 \text{Var}[X_3] + (b\bar{X}_3)^2 \\ &\text{Var}[X_2] + \text{Var}[X'_4] + 2(a+b\bar{X}_2)\text{COV}[X_1, X_3] = 12,96 + 1,44 + \\ &0,02 + 4 - 5,2 = 13,2 \text{ (t/m}^2)^2 \end{aligned} \quad (14)$$

$$\text{soit } \sigma_{g_1(Z)} = 3,6 \text{ t/m}^2$$

Alors, on trouve que $\beta \approx E[g_1(Z)] / \sigma_{g_1(Z)} = 3,1$, ce qui conduit à une probabilité de ruine $p_f \approx 10^{-3}$ inacceptable.

Nous venons de dire qu'il ne faut pas prendre, pour les variances, la valeur ponctuelle mais la valeur réduite: $\text{Var}[X_v] = \text{Var}[X] \cdot \gamma(V)$, X_v étant le champ de moyenne locale sur V .

La réduction de variance pour c' , $\text{tg}\phi'$ et γ' est alors $\Delta(\Sigma) = \gamma(L)^2$:

$$\text{avec } \gamma(L) = (33/100)(1 - 33/300) \text{ (on a } L = 100 \text{ m ; } a = 33 \text{ m)}$$

(γ' serait aussi à intégrer suivant la verticale sur la fenêtre D)

$$\gamma(\Sigma) = 0,086 \quad (15)$$

En appliquant cette réduction aux variances et covariances de c' , $\text{tg}\phi'$ et γ' (mais pas à H') la variance de $g_1(Z)$ devient :

$$\text{Var}[g_1(Z)] = 1,11 + 0,12 + 0,00 + 4 - 0,45 \quad (16)$$

$$\text{soit } \text{Var}[R] = 0,78 \text{ (au lieu de } 9,22)$$

$$\text{et } \text{Var}[S] = 4 \text{ (inchangé)} \quad (17)$$

On trouve $\beta_1 = 5,11$, soit une probabilité de ruine de l'ordre de $2 \cdot 10^{-7}$.

Il est intéressant de voir l'évolution de la contribution respective de la résistance et de la sollicitation (équations 17) ainsi que la contribution de chaque variable X_i , d'une part par sa dispersion propre (valeurs 8) mais encore par le gradient $\delta g(Z) / \delta X_i$ (équations 14 et 16). P. Londe distinguait ainsi les paramètres « lourds ».

Des méthodes similaires ont été développées lorsque l'état limite n'est pas donné par une fonction explicite mais par un calcul aux éléments finis ; c'est ce qu'on appelle la méthode des éléments finis stochastiques qui est présentée dans l'article de Gabriel Auvinet *et al.* « La méthode des éléments finis stochastiques en géotechnique. » Enfin, Albert Bolle propose dans son article « Étude et prise en compte de la variabilité spatiale » une nouvelle technique de simulation fort peu coûteuse en temps calcul et qui évite de passer par les approximations des développements en série de Taylor.

Ces méthodes sont encore loin d'être courantes dans l'ingénierie qui utilise plutôt des méthodes dites « semi-probabilistes » consistant à retenir une valeur caractéristique tirée de l'observation statistique et de procéder à un calcul déterministe avec des coefficients pondérateurs pour tenir compte de différentes sources d'incertitudes supplémentaires.

L'article de Maurice Cassan donne un exemple de choix de valeur caractéristique, utilisant l'analyse sta-

tistique pour la variabilité naturelle et l'échantillonnage et utilisant l'expérience pour la réduction de variance tandis que celui de François Baguelin et Jean-Bernard Kovaric intègre explicitement cette dernière à partir de la dimension de l'ouvrage et de la distance d'auto-corrélation. Cette procédure a été proposée pour l'Euro-codé 7.

3

Erreurs de modèle de comportement et autres erreurs

Nous ne parlons pas explicitement de l'erreur de modèle au cours de ces exposés. Cette notion est très claire théoriquement mais très imprécise pratiquement. Son traitement peut être inclus dans celui des erreurs mal définies ou imprécises grâce aux « sous-ensembles flous » : Daniel Boissier, dans son article « Décision et incomplétude » donne divers exemples d'utilisation des « sous-ensembles flous » pour caractériser les imprécisions sur les paramètres de dimensionnement. Cette procédure peut être explicitement étendue aux biais apportés par les modèles de calcul. On peut en effet montrer (Favre et Touati, 1995) que l'erreur de modèle est un biais pur sans dispersion mais il est impossible de l'isoler complètement de l'erreur d'observation sur ses paramètres.

Parmi les autres erreurs, il y a les approximations d'intégration (calculs aux éléments finis, etc.). De façon plus générale, j'ai présenté au colloque « Risque et Génie civil » (Favre, 2000) l'ensemble des aléas dans les ouvrages géotechniques et leur importance relative, la nature, l'identification et la modélisation de ceux liés au terrain.

4

Poids respectifs de la variabilité naturelle et de l'échantillonnage

On vient de voir que la variabilité naturelle était prise en compte par l'indice de sécurité β calculé à partir des champs de moyenne locale et non à partir des paramètres aléatoires « ponctuels ». Par contre l'incertitude sur les moments (espérance, variance et auto covariance) de ces paramètres est prise en compte avec des intervalles de confiance. Un exemple de leur prise en compte conjointe est donné par l'article de François Baguelin et Jean-Bernard Kovaric mais sans qu'apparaisse leur poids respectif sur la probabilité de ruine de l'ouvrage.

Cette question nous paraît suffisamment importante pour qu'on puisse attirer l'attention sur elle à partir d'un exemple académique (Favre, 1988).

Considérons que toutes les variables de base entrant dans la fonction de performance $g(Z)$ ont fait l'objet du même échantillonnage de taille n .

On cherche une borne inférieure⁽¹⁾ à β au seuil α , c'est-à-dire une valeur β_0 telle que :

$$\text{Prob}(\beta \leq \beta_0) \leq \alpha$$

⁽¹⁾ Plus β est petit, plus la probabilité de ruine est grande.

On vient de voir (§ 2.3.2) qu'une forme approchée de l'indice de sécurité β est :

$$\beta = E[g(Z)]/\sigma_{g(Z)} \quad \text{où } p_f = \Phi(-\beta) \text{ est la probabilité de ruine.}$$

On cherche la borne β_0 sous la forme m_{\min}/s_{\max} .

Soit $m_{\min} - m_{\max}$ l'intervalle de confiance de $E[g(Z)]$ au seuil α

$$\begin{aligned} \text{Prob}(m_{\min} \leq E[g(Z)] \leq m_{\max}) &= 1 - \alpha \\ \text{Prob}(E[g(Z)] \leq m_{\min}) &= \alpha/2 = P(A) \end{aligned} \quad (18)$$

Soit $s_{\min} - s_{\max}$ l'intervalle de confiance de $\sigma_{g(Z)}$ au seuil α

$$\begin{aligned} \text{Prob}(s_{\min} \leq \sigma_{g(Z)} \leq s_{\max}) &= 1 - \alpha \\ \text{Prob}(s_{\max} \leq \sigma_{g(Z)}) &= \alpha/2 = P(B) \end{aligned} \quad (19)$$

Les valeurs de m_{\min} et s_{\max} sont données dans l'article de Maurice Cassan.

La probabilité pour que $\beta = E[g(Z)]/\sigma_{g(Z)}$ soit inférieur à $\beta_0 = m_{\min}/s_{\max}$ est la probabilité de trouver le point de coordonnées $(E[g(Z)]; \sigma_{g(Z)})$ sous la droite $E[g(Z)] = \sigma_{g(Z)}(m_{\min}/s_{\max})$ (Fig. 7-a), soit dans le domaine \bar{D} . Ce domaine est plus petit que le domaine \bar{D}^* (Fig. 7-b).

Le domaine \bar{D}^* est l'union des domaines A et B moins leur intersection, soit :

$$\text{Prob}(\beta \in D) \leq \text{Prob}(\beta \in D^*) = P(A) + P(B) - P(A \cap B) \leq \alpha$$

car $P(A \cap B)$ est positif comme toute mesure de probabilité.

Ainsi,

$$\text{Prob}(\beta \leq \beta_0) \leq \alpha \quad \text{avec } p_0 = \Phi(-\beta_0) \quad (20)$$

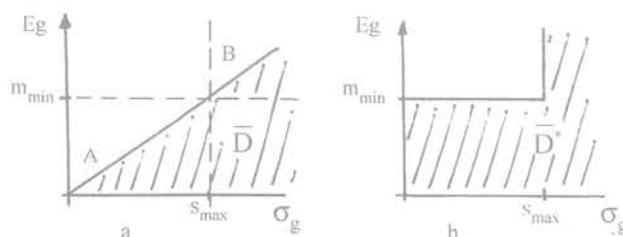


FIG. 7 Domaine de dépassement du seuil β_0 (a) et domaine majorant (b).

Threshold β_0 exceeding domain (a) and raising of this one (b).

Comparons maintenant le poids respectif de la variabilité naturelle et des incertitudes de mesures (cf. tableau ci-après).

Soit par exemple un mur poids sur semelle filante de 2 m de large et 10 m de long. On considère que la charge appliquée au sol est constante et connue (cas favorable). Dans ce cas l'indice de sécurité peut s'écrire

TABEAU Rôle conjugué de la dispersion naturelle et de la reconnaissance.
Natural variability and investigation joint effects.

cv (Z)	20 %	40 %
cv (R) = $\sqrt{\gamma} \cdot cv (Z)$	5 %	10 %
n = ∞ , CV (R) = cv (R)	5 %	10 %
$p_f = p_0$	F = 3, $\beta_0 = 13,3$	F = 3, $\beta_0 = 6,7$
	$p_0 = 0$	$p_0 = 0$
	F = 2, $\beta_0 = 10$	F = 2, $\beta_0 = 5$
n = 8, CV (R) _{max} =	8 %	17 %
	F = 3, $\beta_0 = 8,2$	F = 3, $\beta_0 = 3,92$
	$p_0 = 0$	$p_0 = 5 \cdot 10^{-5}$
Il y a moins de 2 chances sur 10 d'avoir $p_f \geq p_0$	F = 2, $\beta_0 = 6,2$	F = 2, $\beta_0 = 2,94$
	$p_0 = 0$	$p_0 = 1,6 \cdot 10^{-3}$
n = 4, CV (R) _{max} =	12 %	26 %
	F = 3, $\beta_0 = 5,5$	F = 3, $\beta_0 = 2,56$
	$p_0 \leq 10^{-8}$	$p_0 = 5 \cdot 10^{-3}$
Il y a moins de 2 chances sur 10 d'avoir $p_f \geq p_0$	F = 2, $\beta_0 = 4,2$	F = 2, $\beta_0 = 1,92$
	$p_0 = 10^{-6}$	$p_0 = 2,8 \cdot 10^{-2}$

simplement en fonction du coefficient de sécurité F et du coefficient de variation de la capacité portante ultime cv(R) :

$$\beta = (1 - 1/F) / cv(R) \quad (21)$$

Le sol est soit assez homogène (avec un coefficient de variation des propriétés cv(Z) de l'ordre de 20 %), soit peu homogène (cv(Z) de l'ordre de 40 %) (ligne 1 du tableau I).

On fait l'hypothèse que la réduction de variance est importante (cas favorable) et vaut 1/16 ; cela divise le coefficient de variation pour la capacité portante ultime cv(R) par 4 (ligne 2).

a) On suppose qu'on connaît parfaitement le sol, ce qui correspond à un échantillonnage infini (lignes 3).

Alors cv(R) est parfaitement déterminé : son estimateur CV(R), variable aléatoire, est égal à cv(R) (ligne 3-1). Alors $\beta = \beta_0$ et $p_f = p_0$.

Le calcul donne, pour un coefficient de sécurité normal F = 3 (ligne 3-2), un indice de sécurité très grand, que le sol soit homogène ou non, correspondant à des probabilités de ruine très faibles. Pour un coefficient de sécurité faible F = 2 (ligne 3-3), il donne un indice de sécurité moyen dans le cas du sol hétérogène et la probabilité de ruine commence à ne pas être négligeable de 10^{-6} à 10^{-7} . Ceci signifierait que dans le cas d'un sol hétérogène, même si on le connaît très bien, on prend un petit risque avec un coefficient de sécurité faible.

b) Supposons maintenant qu'on connaît imparfaitement le sol. On a fait 8 mesures (lignes 4).

On ne peut estimer avec certitude le coefficient de variation de la capacité portante. Soit $CV(R)_{max}$ la borne supérieure de cv(R) telle que :

$$\text{Prob}\{cv(R) \geq CV(R)_{max}\} = 20 \% \quad (22)$$

avec $\beta_0 = (1 - 1/F) / CV(R)_{max}$.

Pour le sol assez homogène, la borne passe à 8 %. Pour le sol hétérogène, elle passe à 17 % (ligne 4-1).

Pour le sol homogène, la valeur de β_0 reste forte et la probabilité de ruine est non significative. Par contre, pour le sol hétérogène, les valeurs deviennent faibles et les probabilités p_0 deviennent significatives. Mais la probabilité de ruine reste acceptable car on n'a que 2 chances sur 10 d'avoir une probabilité de ruine p_f inférieure à p_0 . On pourrait l'apprécier en écrivant :

$$p_f \approx 2 \cdot 10^{-1} \cdot p_0 \quad (22)$$

Ainsi la probabilité de ruine serait de 10^{-5} avec un coefficient de sécurité normal (ligne 4-2) et de $3 \cdot 10^{-4}$ avec un coefficient faible ce qui est presque excessif (ligne 4-3).

c) Enfin, supposons qu'on connaisse mal le sol. On n'a fait que quatre mesures (lignes 5). La borne $CV(R)_{max}$ devient très grande (ligne 5-1). Pour le sol homogène, que le coefficient de sécurité soit normal ou faible, le risque est très faible. Par contre, pour le sol hétérogène, même avec un coefficient de sécurité normal (ligne 5-2), il est inacceptable (une fondation sur mille aurait des désordres).

Cet exemple n'est qu'illustratif ; il ne donne en aucune façon des valeurs absolues de probabilités de

ruine. Il montre seulement comment le risque évolue en fonction des différents aléas :

– la variabilité naturelle, à elle seule semble ne pas peser (lignes 3). Rappelons tout de même que nous nous sommes mis dans un cas avantageux avec une forte réduction de la variance et une action non aléatoire. Si cette réduction n'avait été que de 1/4, les β_0 auraient doublés et on voit toutes les conséquences sur les p_0 dans les 3 cas (lignes 3, 4 et 5) ;

– on voit l'effet très pénalisant du manque de reconnaissance dès que le sol n'est pas assez homogène. On ne saurait qu'insister sur la nécessité de faire des reconnaissances assez conséquentes dès que le sol est hétérogène ;

– enfin, on retrouve la légitimité de prendre des coefficients de sécurité faibles dans des circonstances favorables (sol assez homogène et bonne reconnaissance).

5

Conclusion

Nous venons de dégager trois grands types d'incertitudes pour l'examen des propriétés des sols, qui tiennent à leur variabilité naturelle, à leurs connaissances et aux modèles que l'on choisit pour ces propriétés.

Les incertitudes de modèle sont des biais, inaccessibles pour le moment sauf par une procédure indirecte basée sur le retour d'expérience d'ouvrages réels (Favre, 1998). L'article de J.-P. Magnan montre l'importance de certains de ces aspects peu quantifiables.

Les articles de M. Cassan et de F. Baguelin et J.B. Kovaric traitent tout d'abord de l'exploitation des mesures et de leur modélisation. Le but recherché à travers la modélisation est de réduire au maximum

l'incertitude sur les paramètres de calcul des ouvrages, que cette incertitude vienne de la variabilité naturelle, de l'échantillonnage ou du lissage par intégration (réduction de variance). Ils indiquent les techniques classiques de modélisation statistique des paramètres fonction ou non de la profondeur et les techniques de tests statistiques. De plus, au cas où l'ingénieur disposerait de la mesure de plusieurs paramètres un nombre non négligeable de fois, on peut utiliser les techniques d'analyses factorielles (Favre, 1983) comme le montre l'article de P. Pouget *et al.*

L'article d'A. Bolle indique comment modéliser les paramètres non plus en fonction de la simple profondeur mais en fonction de tout l'espace par des champs aléatoires. Il dégage la notion d'auto corrélation qui introduit celle de réduction de variance, absolument fondamentale dans les calculs de probabilité de ruine des ouvrages en terre. Il donne une nouvelle méthode de simulation pour ce type de calcul.

G. Auvinet *et al.* traitent ensuite de la prise en compte de cette incertitude dans le dimensionnement des ouvrages en exposant la théorie de la fiabilité et son application au dimensionnement par la méthode des éléments finis stochastiques. Comme nous venons de le préciser, il s'agit de la prise en compte de la variabilité naturelle puisque nous sommes amenés à supposer que $Z_m \approx Z_N$. Ils ne prennent pas en compte les erreurs de modèles vu qu'à l'heure actuelle les propositions dans ce sens commencent à peine à apparaître (Favre *et al.*, 1998). De mon côté, j'ai indiqué plus haut comment juger du poids respectif de la variabilité naturelle et des incertitudes d'échantillonnage.

Enfin, D. Boissier considère l'application possible en matière de prise de décision lorsque les données sont imprécises et incertaines au point de ne pouvoir faire l'objet d'un traitement statistique. Il montre le parti qu'on peut tirer des sous-ensembles flous.

Bibliographie

- Favre J.-L. (1983) – « Analyses statistiques comparées de pénétromètres, scissomètres et pressiomètres auto-foreurs », Symp. Int. Reco. Sols Roches par essais en place, Paris. *Int Assoc Eng Geology*, vol. 2, p. 287-297.
- Favre J.-L. (1988) – Special Lecture « Properties of soil : large uncertainties - a minor aspect of foundation reliability », Symposium on Reliability-Based Design in Civil Engineering, Lausanne, Swiss Federal Institute of Technology, vol. 2, p. 51-66.
- Favre J.-L., Touati K. (1995) – « Probabilistic estimation of the errors in finite element calculation code for shallow foundation », Proc. 7th Int. Conf. Applic. Stat. Prob. soils struct. eng., Paris, A. A. Bakema, vol. 2, p. 1309-1315.
- Favre J.-L., Biarez J., Hachi F. (1998) – « An only model for sand and clay », PROBAMAT 21st Century, PERM, NATO-ASI series, Kluwer academic publishers, vol. 46, p. 465-481.
- Favre J.-L. (1998) – « Errors in geotechnics and their impact on safety », Int. Conf. on Uncertain Structures, Miami 1996. *Computers and Structures*, 67, p. 37-45.
- Favre J.-L. (2000) – « Les incertitudes géologiques et géotechniques », *Risque et Génie civil*, Paris, Presses des Ponts et Chaussées, p. 187-229.

Utilisation de la statistique descriptive en géotechnique

M. CASSAN

S.A. Fondasol Études
BP 767
84035 Avignon Cedex 3

Résumé

Actuellement, le choix des valeurs des caractéristiques géotechniques est relativement subjectif. Il n'en est plus ainsi si l'on soumet les résultats des mesures à une analyse statistique. Cette communication présente quelques méthodes statistiques simples qui conduisent à une évaluation rationnelle de ces caractéristiques et qui prennent en compte les incertitudes et la dispersion des mesures grâce aux notions d'intervalles de confiance pour les caractéristiques indépendantes, et d'ellipses de confiance pour les caractéristiques liées. Mais ces méthodes permettent également d'apprécier l'homogénéité d'un site, par exemple avec le procédé de l'analyse des variances. Nous en avons illustré les approches théoriques par l'examen de quelques cas réels. L'intérêt évident que présentent ces méthodes devrait inciter tous les ingénieurs géotechniciens à les utiliser très largement et à les généraliser dans leurs études.

Mots clés : histogramme, moyenne, variance, ajustement, loi normale, Student, Snedecor, intervalle de confiance, ellipse de confiance, homogénéité.

Descriptive statistics use in geotechnical engineering

Abstract

So far, selecting values for geotechnical data has been rather subjective. One solution is to submit the results to a statistical analysis. This paper presents some simple statistical methods, which lead to a rational evaluation of the data, taking into account the uncertainties and the dispersion of the measures due to confidence interval, for independent data, and confidence ellipses, for linked data. These methods also provide an appreciation of the homogeneity of the studied site, for example by analyzing the variance. These theoretical methods are illustrated with concrete examples.

The evident interest aroused by the statistical methods should motivate the geotechnical engineering community to integrate them in their studies on a casual basis.

Key words : histogram, average, variance, adjustment, normal law, Student, Snedecor, confidence interval, confidence ellipse, homogeneity.

NDLE : Les discussions sur
cet article sont acceptées
jusqu'au 1^{er} août 2001.

Considérations générales

Dans le cas général, la démarche intellectuelle de l'ingénieur, quelle que soit sa spécialité, est essentiellement déterministe et pour l'ingénieur géotechnicien elle s'articule selon un organigramme dont l'aboutissement est un calcul et un dimensionnement, opérations, elles aussi, particulièrement déterministes.

Sur cet organigramme (Fig. 1) on constate qu'à partir des données fournies par les reconnaissances géologiques et géotechniques, le travail de l'ingénieur est d'abord un travail d'analyse qui est relativement facile, puis un travail de synthèse qui est de loin le plus difficile car il doit déboucher d'une part sur une stratigraphie représentative du site étudié et d'autre part sur l'attribution, à chacune des couches définies dans cette stratigraphie, d'une valeur unique des caractéristiques géotechniques qui y ont été mesurées.

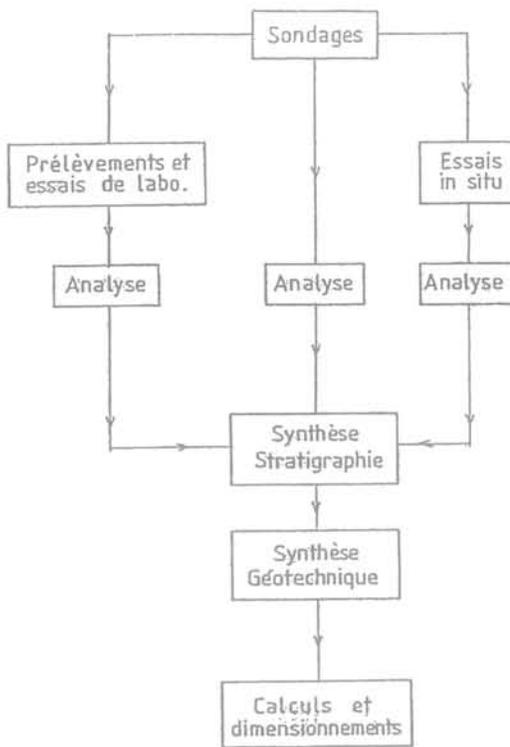


FIG. 1 Organigramme d'une étude géotechnique.
Flow chart of a geotechnical campaign.

C'est ce dernier point qui s'avère très délicat car il se trouve à la limite de l'objectivité et de la subjectivité. Or, en matière scientifique, la subjectivité ne devrait pas exister. Dans ce contexte, après un premier examen critique des données géotechniques, le traitement statistique de ces données apparaît comme un outil extrêmement précieux, susceptible de conférer à la décision de l'ingénieur une plus grande objectivité, l'objectivité totale dans ce domaine étant évidemment très difficile à atteindre sinon impossible.

Mais il faut être très vigilant car, pour que l'outil statistique soit utilisé le plus fréquemment possible et qu'il entre de façon définitive dans le mode de pensée du géotechnicien, il faut qu'il soit simple et maniable. C'est

pourquoi nous estimons qu'il est dangereux de brûler les étapes et d'aborder directement le problème avec des théories trop complexes et des mathématiques trop sophistiquées qui, pour réjouissantes qu'elles soient pour l'esprit, risquent de s'avérer dissuasives et d'éloigner définitivement la statistique de la géotechnique.

La présente communication a pour objet d'exposer comment on peut, en un premier temps, envisager l'utilisation de méthodes statistiques simples pour évaluer les caractéristiques mécaniques à attribuer à chaque couche stratigraphique et éventuellement pour apprécier l'homogénéité d'un site.

Synthèse stratigraphique et synthèse géotechnique

Synthèse stratigraphique

La synthèse stratigraphique consiste à déterminer une stratification telle que, dans chacune des couches lithologiques mises en évidence par les sondages, les variations des caractéristiques géotechniques mesurées puissent être considérées comme « apparemment » indépendantes des variations spatiales. Elle est établie à partir de :

- l'analyse des coupes de forages et des échantillons ;
- l'analyse des diagraphies éventuelles (très important) ;
- l'analyse des résultats des essais *in situ*.

En effet, il ne suffit pas de se fier à la seule stratification lithologique car, dans une couche de nature lithologique ou géologique constante, il peut y avoir des niveaux de caractéristiques géotechniques différentes, comme par exemple des strates d'altération ou des zones superficielles de moindre compacité. Nous donnons plus loin un exemple de cette dernière particularité dans des alluvions fluviales. Des phénomènes analogues peuvent également se produire aux interfaces de sols fins (argiles ou limons par exemple) et de terrains granulaires (sable, graves et galets) dans lesquels la limite entre les deux formations peut s'avérer confuse dans les caisses d'échantillons et où seules les diagraphies et les essais *in situ* permettent de trancher sans ambiguïté.

Synthèse géotechnique

La synthèse stratigraphique ayant permis de délimiter un certain nombre de couches de même nature et apparemment homogènes, le problème consiste à attribuer à une couche donnée une valeur unique de chacune des caractéristiques géotechniques que l'on y a mesurées, ce qui permet de rejoindre le processus déterministe pour la poursuite de l'étude. C'est là qu'intervient le traitement statistique des données, pour lequel on est amené à distinguer deux catégories de caractéristiques géotechniques :

- les caractéristiques indépendantes, directement mesurables ;

– les caractéristiques liées, déterminées simultanément à partir d'un même essai.

Dans la première catégorie, on trouve en particulier :

– les paramètres d'identification tels que la teneur en eau, les limites d'Atterberg, le poids volumique ;

– quelques caractéristiques mécaniques mesurées en laboratoire comme la résistance à la compression simple, la résistance à la traction (essai brésilien) la pression de préconsolidation, l'indice de compression (C_c), le coefficient de consolidation (c_v) ;

– la plupart des essais *in situ*, tels que l'essai pressiométrique avec la pression limite nette et le module de déformation, le scissomètre avec la cohésion non drainée, les termes de pointe et de frottement pénétrométriques, le SPT.

Dans la deuxième catégorie se situent essentiellement la cohésion et l'angle de frottement interne qui peuvent être déterminés soit en laboratoire par essais triaxiaux ou essais de cisaillement à la boîte, soit *in situ* à l'aide d'essais au phicomètre.

3

Méthodes d'estimation statistique des caractéristiques géotechniques

3.1

Caractéristiques indépendantes directement mesurables

3.1.1

Étude de la distribution des mesures

Dans chaque couche apparemment homogène, on considère, dans leur ensemble, toutes les mesures x_i du paramètre étudié, tous forages et toutes profondeurs confondues.

Les théories statistiques nous apprennent qu'il existe des tests permettant de s'assurer du caractère aléatoire d'une série d'observations successives. Je citerai pour mémoire, deux tests non paramétriques assez simples et assez rapides qui sont le test des suites et le test des signes [7]. Mais l'utilisation préalable de ces tests ne paraît pas fondamentale dès lors que l'on cherche à ajuster les observations sur une loi statistique connue, car si cet ajustement est possible c'est que la distribution présente effectivement un caractère aléatoire.

On opère donc de la façon suivante :

a) établissement de l'histogramme des n mesures groupées en k classes ($k \geq 5$) de même « largeur » c'est-à-dire telles que, x_j et x_{j+1} désignant les valeurs de x_i qui limitent la classe j , les différences $[x_{j+1} - x_j]$ soient constantes et égales à $\frac{x_r - x_1}{k}$.

Si n_j désigne la population de la classe j , on peut prendre pour ordonnée de cette classe la valeur n_j ou, comme le font les statisticiens, la fréquence :

$$f_j = \frac{n_j}{\sum n_j} = \frac{n_j}{n}$$

b) calcul de la moyenne empirique \bar{x} de la population, de sa variance empirique s^2 , de sa variance estimée s'^2 et de son coefficient de variation CV :

$$\bar{x} = \frac{1}{n} \sum x_i \quad s'^2 = \frac{\sum (x_i - \bar{x})^2}{n}$$

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} \quad CV = \frac{s}{\bar{x}}$$

La moyenne arithmétique \bar{x} est un estimateur de la vraie valeur M de la moyenne du paramètre étudié, et s^2 l'estimateur de la vraie valeur Σ^2 de sa variance. Le coefficient de variation caractérise la dispersion des mesures et de la distribution de x . C'est un paramètre très important qui peut, presque à lui seul, donner une idée de la qualité d'une reconnaissance ou du caractère particulier du terrain rencontré puisque l'analyse d'un très grand nombre de résultats a permis de définir des plages de valeurs types de ce coefficient de variation pour diverses caractéristiques géotechniques [6 et 14]. Lorsqu'on s'éloigne de ces plages, il faut donc se poser des questions. Mais c'est l'étude de l'histogramme expérimental qui permettra d'aller plus loin dans la démarche statistique.

c) Ajustement de l'histogramme sur une loi de probabilité connue

• Si l'histogramme présente une tendance centrale, c'est-à-dire si l'on voit apparaître une certaine symétrie par rapport à la moyenne ou à la médiane de la distribution, on pourra chercher à ajuster cette distribution sur la loi normale (ou de Laplace- Gauss), ce qui est très facile surtout avec les moyens informatiques actuels.

On utilise pour cela la méthode de la droite de Henry qui consiste à porter en abscisse arithmétique la limite supérieure s_{j+1} de la classe j et en ordonnée gaussienne les fréquences cumulées correspondantes. Rappelons que l'échelle gaussienne est une échelle graduée en $u = F^{-1}(y)$ où $f(u)$ est l'expression de la loi de Laplace-Gauss en coordonnées réduites :

$$y = f(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-\frac{u^2}{2}} du$$

Si les points ainsi obtenus sont alignés, c'est que la distribution suit effectivement la loi de Laplace.

On peut alors vérifier la probabilité que l'on a de se tromper si l'on ne retient pas la normalité de la distribution en utilisant le test du χ^2 pour $n > 30$ (échantillons importants) ou le test de Shapiro-Wilk pour $n < 30$ (petits échantillons) ces deux tests étant bien connus des statisticiens (cf. bibliographie, en particulier 7 et 14).

• Si l'histogramme présente une dissymétrie avec étalement vers la droite, on pourra chercher à l'ajuster sur une loi log-normale dans laquelle c'est le logarithme népérien des mesures qui suit la loi de Laplace-Gauss. On utilise là aussi la méthode de la droite de Henry, mais dans un système de coordonnées gausso-logarithmique.

Les paramètres μ et σ^2 estimateurs respectifs de l'espérance mathématique et de la variance théorique de la loi log normale ajustée, sont liés aux moyenne et variance empiriques par les relations :

$$\mu = \ln \frac{\bar{x}^2}{\sqrt{\bar{x}^2 + s^2}}$$

$$\sigma^2 = \ln \left(1 + \frac{s^2}{\bar{x}^2} \right)$$

• Si ces tentatives d'ajustement s'avèrent infructueuses, on essaiera une loi « bêta » plus complexe. Dans ce cas, il faudra soumettre la loi choisie à l'un des tests d'acceptation évoqués précédemment.

L'estimation de la caractéristique étudiée se fait ensuite à partir de la loi de distribution retenue en utilisant la théorie de l'échantillonnage, comme nous allons le voir.

3.1.2

Estimation à partir de la loi normale

a) Notion d'intervalle de confiance

On admet, dans ce qui suit, que l'ensemble des n mesures réalisées constitue un échantillon unique de taille n prélevé dans une population infinie à distribution **normale** dont la moyenne vraie M et la variance vraie Σ^2 sont inconnues. Toutefois, on en connaît deux estimateurs \bar{x} et s^2 , qui sont eux-mêmes des variables aléatoires alors que M et Σ^2 , qui resteront toujours inconnues, sont de nature parfaitement déterministe. La théorie de l'échantillonnage permet alors de définir pour M et Σ^2 des intervalles de confiance dépendants de \bar{x} , de s^2 et de n , ces intervalles de confiance étant eux-mêmes des variables aléatoires qui ont une certaine probabilité P de recouvrir M et Σ^2 .

Bien que M et Σ^2 ne soient pas aléatoires, l'usage admet, par simplification, que l'on dise que M ou Σ^2 ont la probabilité P d'être compris entre les bornes inférieures I et supérieure S de l'intervalle de confiance et l'on écrira : $\text{Prob} [I \leq M \text{ (ou } \Sigma^2 \leq S)] = P$.

Cette relation correspond à un intervalle bilatéral, mais on peut également avoir des intervalles de confiance unilatéraux tels que :

$$\text{Prob} [M \text{ ou } \Sigma^2 \leq S] = P$$

$$\text{Prob} [M \text{ ou } \Sigma^2 \geq I] = P$$

b) Intervalle de confiance de la moyenne

On démontre que la variable aléatoire \bar{x} a même espérance mathématique, c'est-à-dire même moyenne M , que la population totale infinie, alors que sa variance n'est pas la variance estimée s^2 , mais :

$$s_{\bar{x}}^2 = \frac{s^2}{n}$$

On démontre également que la quantité

$$t = \frac{\bar{x} - M}{s_{\bar{x}}} = \frac{(\bar{x} - M)\sqrt{n}}{s}$$

suit une loi de probabilité appelée *loi de Student* dont la courbe de densité est une courbe symétrique en cloche dans laquelle intervient un paramètre ν appelé *degré de liberté* et qui, pour chaque valeur de ν , donne la valeur t_α ayant la probabilité $\alpha = 1 - P$ d'être dépassée (Fig. 2).

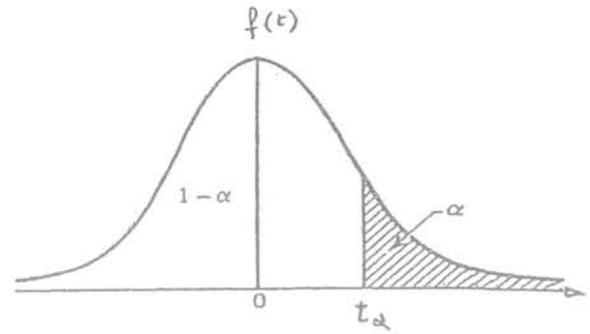


FIG. 2 Loi de Student. Intervalle unilatéral. Student probability density function. Unilateral interval.

Le degré de liberté est égal à la population de l'échantillon diminuée du nombre de paramètre que l'on estime, c'est-à-dire, pour la loi normale, un seul paramètre (la moyenne) d'où $\nu = n - 1$.

Dans le cas de la figure 2 qui correspond à un intervalle de confiance unilatéral, et en se référant à l'expression de t , on a :

$$\text{Prob} \left[M \geq \bar{x} - t_\alpha \frac{s}{\sqrt{n}} \right] = 1 - \alpha = P \quad (1)$$

Dans le cas de l'intervalle de confiance bilatéral (Fig. 3), on écrit :

$$\text{Prob} \left[\bar{x} - t_{\alpha/2} \frac{s}{\sqrt{n}} \leq M \leq \bar{x} + t_{\alpha/2} \frac{s}{\sqrt{n}} \right] = 1 - \alpha = P \quad (2)$$

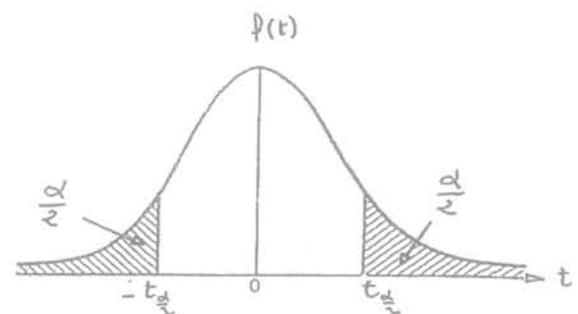


FIG. 3 Loi de Student. Intervalle bilatéral. Student probability density function. Bilateral interval.

L'intervalle de confiance au seuil de probabilité $P = 1 - \alpha$ est donc, pour $n - 1$ degrés de liberté :

$$[I, S] = \bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}} \quad (3)$$

c) Intervalle de confiance de l'écart-type

• Lorsque $n \leq 30$, on démontre que le rapport $(n - 1) \frac{s^2}{\Sigma^2}$ est une variable aléatoire qui suit la loi probabiliste de Pearson, dite loi de χ^2 , et dont le nombre de degrés

de liberté est dans le cas présent égal à $n - 1$, puisqu'on n'estime qu'un seul paramètre.

L'intervalle bilatéral de l'écart-type au niveau de probabilité $P = 1 - \alpha$ est alors (Fig. 4).

$$s \sqrt{\frac{n-1}{\chi^2_{\frac{\alpha}{2}}}} \leq \Sigma \leq s \sqrt{\frac{n-1}{\chi^2_{1-\frac{\alpha}{2}}}} \quad (4)$$

χ^2_{α} étant la valeur de χ^2 ayant la probabilité α d'être dépassée.

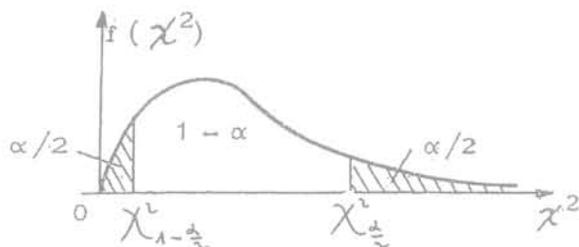


FIG. 4 Loi du χ^2 . Intervalle bilatéral pour $n < 30$. χ^2 probability density function. Bilateral interval for $n < 30$.

Les intervalles de confiance unilatéraux sont de même :

$$0 \leq \Sigma \leq s \sqrt{\frac{n-1}{\chi^2_{1-\alpha}}} \quad (5)$$

$$\Sigma \geq s \sqrt{\frac{n-1}{\chi^2_{\alpha}}} \quad (6)$$

• Lorsque $n > 30$ la distribution de l'écart-type s est à peu près normale. La moyenne vraie reste égale à M et la variance est voisine de $\frac{s^2}{2n}$.

L'expression $\sqrt{2\chi^2} - \sqrt{2n-1}$ suit la loi de Laplace-Gauss réduite d'espérance mathématique 0 et de variance 1.

En désignant par u_{α} la variable normale centrée réduite ayant la probabilité α d'être dépassée, on obtient comme intervalle de confiance bilatéral (Fig. 5):

$$\text{Prob} \left[s \left(1 - \frac{u_{\alpha/2}}{\sqrt{2n}} \right) \leq \Sigma \leq s \left(1 + \frac{u_{\alpha/2}}{\sqrt{2n}} \right) \right] \quad (7)$$

$$[I, S] = s \left(1 \pm \frac{u_{\alpha/2}}{\sqrt{2n}} \right)$$

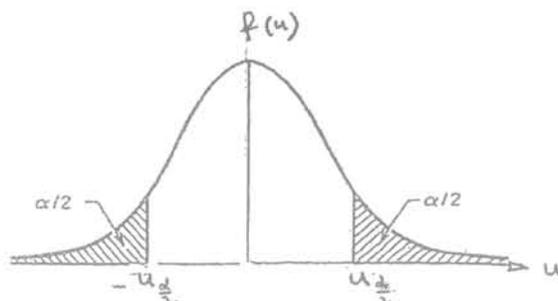


FIG. 5 Loi de Laplace-Gauss. Intervalle bilatéral pour Σ ($n > 30$). Laplace-Gauss probability density function. Bilateral interval for Σ ($n > 30$).

d) Choix de la caractéristique étudiée

À ce stade de l'analyse statistique, la logique voudrait que l'on utilise les résultats précédents pour définir un modèle de sol dans lequel la caractéristique géotechnique étudiée suivrait une loi normale dont les paramètres seraient estimés à partir des intervalles de confiance. Cette estimation devant toujours aller dans le sens de la sécurité, s'il s'agit d'évaluer une caractéristique de résistance, comme par exemple la pression limite nette, la cohésion non drainée ou le nombre N de l'essai de pénétration du carottier (SPT), on pourrait adopter une loi de distribution dont la moyenne M' serait égale à la limite inférieure de son intervalle de confiance et l'écart-type Σ' à la limite supérieure de son intervalle de confiance:

$$M' = \bar{x} - t_{\alpha/2} \frac{s}{\sqrt{n}}$$

$$\Sigma' = s \sqrt{\frac{n-1}{\chi^2_{\alpha/2}}} \quad (n < 30) \quad (8)$$

ou
$$\Sigma' = s \left(1 + \frac{u_{\alpha/2}}{\sqrt{2n}} \right) \quad (n > 30) \quad (9)$$

On pourrait alors adopter comme valeur de la caractéristique géotechnique x étudiée, la valeur x_{β} telle que (Fig. 6) : $\text{Prob}(x \geq x_{\beta}) = P = 1 - \beta$.

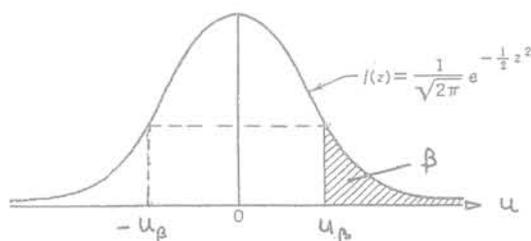


FIG. 6 Loi de Laplace-Gauss. Laplace-Gauss probability density function.

C'est-à-dire en passant à la variable centrée réduite u :

$$\text{Prob}(u > -u_{\beta}) = 1 - \text{Prob}(u > u_{\beta}) = 1 - \int_{u_{\beta}}^{+\infty} f(u) du$$

où $f(u)$ est la fonction de densité de probabilité de la loi normale centrée réduite.

Si l'on prend par exemple $\beta = 5\%$, on a $u_\beta = 1,64$, donc $x_\beta = M' + (-u_\beta) \Sigma' = M' - 1,64 \Sigma'$

C'est à partir de cette valeur que l'on pourrait rejoindre le raisonnement déterministe. Toutefois cette méthode s'avère pénalisante, et conduit à des probabilités de ruine non négligeables de l'ordre de 10^{-2} à 10^{-3} , parce qu'on n'a pas fait intervenir la notion de longueur d'auto-corrélation et la réduction de variance.

Sans chercher à aborder pour l'instant la probabilité de ruine, nous proposons, pour la pratique courante, une méthode beaucoup plus simple qui, depuis vingt ans, nous a donné entière satisfaction et qui consiste à adopter comme caractéristique de calcul la limite inférieure de l'intervalle de confiance de la moyenne :

$$x_c = \bar{x} - t_{\alpha/2} \frac{s}{\sqrt{n}} \quad (10)$$

Lorsque le choix sécurisant n'est pas une estimation par défaut, mais au contraire une estimation par excès de la caractéristique étudiée, comme par exemple un module de déformation permettant de mettre en évidence un tassement différentiel dû à une zone plus raide, on adoptera la limite supérieure de l'intervalle de confiance de la moyenne.

3.1.3

Estimation à partir de la loi log normale

On sait que, dans cette loi, c'est le logarithme de la variable aléatoire qui suit la loi normale. On pourrait donc envisager d'appliquer la méthode précédente, c'est-à-dire d'adopter comme valeur du paramètre étudié celle qui correspond à la limite la plus sécurisante de l'intervalle de confiance de la moyenne de son logarithme.

Nous avons eu l'occasion de procéder ainsi mais les valeurs obtenues étaient extrêmement pénalisantes. C'est pourquoi, maintenant (et cela depuis plusieurs années) nous adoptons simplement la valeur modale x_m de la distribution ajustée (Fig. 7).

Si les paramètres de cette distribution ajustée sont respectivement μ et σ , la variable aléatoire $u = \frac{\ln x - \mu}{\sigma}$ suit la loi normale centrée réduite et on a :

$$x_m = \exp(\mu - \sigma^2) \quad (11)$$

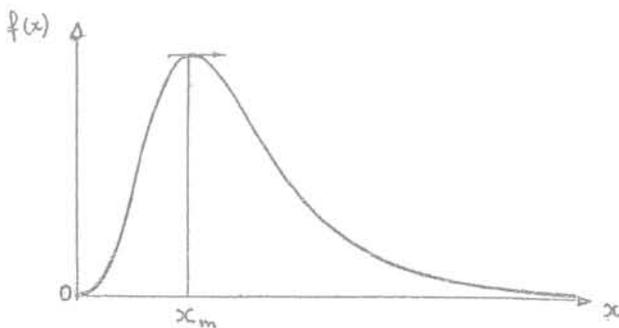


FIG. 7 Loi log-normale.
Log-normal probability density function.

Caractéristiques liées à détermination simultanée

Il s'agit là essentiellement de la détermination de la cohésion effective c' et de l'angle de frottement interne ϕ' en laboratoire, ou d'une évaluation *in situ* au phicomètre de ces paramètres que l'on désigne alors par ϕ_1 et c_1 pour éviter toute confusion.

Dans tous ces essais, on mesure deux variables x et y qui doivent théoriquement être liées par une relation fonctionnelle linéaire, $y = A + Bx$, dans laquelle A et B sont des paramètres fonctions de c' et ϕ' que l'on cherche à déterminer.

La définition des variables x et y est donnée dans le tableau suivant :

Essais	x	y
Triaxial	$\frac{\sigma'_1 + \sigma'_2}{2}$	$\frac{\sigma'_1 - \sigma'_2}{2}$
Cisaillement	σ'	τ'
Phicomètre	σ_1	τ_1

La méthode d'interprétation statistique de ces essais consiste à porter sur un même graphique pour chaque couche de terrain l'ensemble des couples (x, y) tous forages, toutes profondeurs, tous échantillons confondus et toutes éprouvettes, et d'effectuer les calculs suivants :

1) Calcul de la droite des moindres carrés dont le coefficient de régression R doit être supérieur à 0,85. L'équation de cette droite est $\hat{y} = a + bx$, où a et b sont les estimateurs de A et B .

2) Calcul des intervalles de confiance de A et B :

$$[A] = a \pm t_{\frac{\alpha}{2}, n-2} s \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum_1^n (x_i - \bar{x})^2}} \quad (12)$$

$$[B] = b \pm t_{\frac{\alpha}{2}, n-2} \frac{s}{\sqrt{\sum_1^n (x_i - \bar{x})^2}} \quad (13)$$

avec :

$t_{\frac{\alpha}{2}, n-2}$ variable de Student à $n - 2$ degrés de liberté pour un seuil de probabilité $1 - \alpha$.

$$s^2 = \frac{1}{n-2} \sum_1^n (y_i - \hat{y})^2$$

où n est le nombre de couples (x_i, y_i) et $y_i - \hat{y}$ est l'écart des observations par rapport à la droite des moindres carrés, \bar{x} étant la valeur moyenne des x_i .

On peut alors déterminer des estimateurs de A et de B sécurisants, en prenant par exemple les valeurs minimales des intervalles de confiance, ou en admettant comme vraie la valeur de a ou celle de b et en ne prenant en considération que l'intervalle de confiance de l'autre.

Mais la méthode la plus rigoureuse consiste à traiter simultanément A et B en faisant appel à la notion d'ellipse de confiance, c'est-à-dire à l'ellipse qui a la probabilité P de contenir les vraies valeurs de A et B lorsque A est porté en abscisse et B en ordonnée.

En pratique, on fait le changement de variable :

$$X = A - a$$

$$Y = B - b$$

X et Y représentent donc les écarts entre les vraies valeurs de A et B inconnus et leurs estimateurs a et b connus. L'équation de cette ellipse est alors :

$$X^2 + \beta XY + \alpha Y^2 - \gamma = 0 \quad (14)$$

avec :

$$\begin{aligned} \alpha &= \frac{1}{n} \sum_1^n x_i^2 \\ \beta &= \frac{2}{n} \sum_1^n x_i \\ \gamma &= \frac{2s^2 F_p}{n} \end{aligned} \quad (15)$$

F_p est la valeur d'une variable aléatoire F qui suit la loi de Snedecor à $\nu_1 = 2$ et $\nu_2 = n - 2$ degrés de liberté et telle que $\text{Prob}(F < F_p) = P$. Cette valeur est évidemment donnée par des tables numériques.

Par analogie avec le cas des caractéristiques indépendantes, on évaluera A et B en se plaçant sur l'ellipse elle-même en des points que l'on aura jugé à la fois sécurisants et physiquement réalistes.

On en déduit alors les caractéristiques de cisaillement :

$$\text{Essai triaxial} \quad \varphi' = \text{Arc sin } B \quad \text{et} \quad C' = \frac{A}{\cos \varphi'}$$

$$\text{Cisaillement à la boîte} \quad \varphi' = \text{Arctg} B \quad \text{et} \quad C' = A$$

$$\text{Phicomètre} \quad \varphi_1 = \text{Arctg} B \quad \text{et} \quad C_1 = A$$

4

Étude de l'homogénéité d'un site

L'appréciation de l'homogénéité d'un site reste quelque peu subjective mais la statistique descriptive peut lui conférer une certaine objectivité.

Nous présentons ci-après trois méthodes que nous avons eu l'occasion d'utiliser et qui ne sont valables que pour des caractéristiques indépendantes dont la distribution est soit normale soit ramenée à la normale par introduction d'une fonction auxiliaire (loi log-normale, par exemple).

4.1

Méthode approchée

Cette méthode est basée sur la comparaison de deux moyennes et de deux variances et utilise un test d'acceptation.

Considérons donc un site sur lequel on a réalisé r forages dans chacun desquels on dispose de n_i mesures

du paramètre que l'on cherche à évaluer ($1 \leq i \leq r$), ce qui porte le nombre total de mesures sur le site

$$\text{à } N = \sum_1^r n_i \text{ soit alors :}$$

$$\bar{x} = \text{moyenne empirique des } N \text{ mesures}$$

$$s^2 = \text{variance estimée des } N \text{ mesures}$$

$$\bar{x}_i = \text{moyenne empirique des } n_i \text{ mesures réalisées dans le } i\text{-ième forage}$$

$$s_i^2 = \text{Variance estimée des } n_i \text{ mesures réalisées dans le } i\text{-ième forage}$$

4.1.1

Homogénéité des variances

On compare successivement les variances estimées s_i^2 de chaque forage à la variance estimée s^2 de la population totale tous forages confondus.

$$\text{Pour cela, on forme la variable aléatoire } F_i = \frac{s_i^2}{s^2} \text{ qui}$$

suit la loi de Snedecor à $\nu_1 = N - 1$ et $\nu_2 = n_i - 1$ degrés de liberté.

L'hypothèse de l'égalité des variances sera acceptée au niveau de probabilité α si :

$$\frac{1}{F_{\alpha/2, n_i - 1, N - 1}} \leq F_i \leq F_{\alpha/2, n_i - 1, N - 1} \quad (16)$$

ce qui équivaut à accepter l'hypothèse nulle lorsque celle-ci est vraie avec une probabilité $1 - \alpha$, c'est-à-dire que le risque de se tromper en rejetant l'hypothèse est égal à α .

Une autre méthode plus simple mais qui n'est possible que si le nombre de mesures est le même dans tous les forages, est le test de Cochran qui permet de tester simultanément l'égalité de toutes les variances obtenues :

$$s_1^2 = s_2^2 = \dots \dots \dots s_r^2$$

pour cela on forme le rapport :

$$g = \frac{\text{maximum de } s_i^2}{\sum_1^r s_i^2} \quad (17)$$

L'hypothèse d'égalité sera acceptée au seuil de probabilité α si $g \leq g_{\alpha}$, g_{α} étant donné par une table numérique établie pour $\alpha = 0,05$ et $\alpha = 0,01$ et qui est fonction de n_i (nombre de mesures par forages) et r (nombre de forages).

4.1.2

Homogénéité des moyennes

On compare successivement les moyennes \bar{x}_i de chaque forage à la moyenne globale \bar{x} , c'est-à-dire

que l'on cherche à voir si la différence $\bar{x}_i - \bar{x}$ peut être considérée comme nulle avec une probabilité satisfaisante. Le test précédent ayant permis de décider si l'on est en droit d'admettre ou de rejeter l'égalité des

variances, on doit donc considérer ces deux cas, dans chacun desquels les variances restent toujours inconnues.

a) Variances égales

On forme la variable aléatoire

$$t'_i = \frac{\bar{x}_i - \bar{x}}{s \sqrt{\frac{1}{n_i} + \frac{1}{N}}} \quad (18)$$

qui suit la loi de Student à $v_i = n_i + N - 2$ degrés de liberté et dans laquelle s^2 est la variance commune à tous les forages.

Le domaine d'acceptation de l'égalité des moyennes du forage i et de celle de l'ensemble des forages est alors :

$$-t_{\alpha/2, v_i} \leq t'_i \leq t_{\alpha/2, v_i}$$

b) Variances différentes

Sur le plan théorique et purement mathématique il n'est pas possible, dans ce cas, d'utiliser un test paramétrique de décision. Toutefois, selon A.H. Bowker et G.J. Lieberman [2] on peut définir une variable statistique t' qui ne suit qu'approximativement la loi de Student lorsque les moyennes sont égales :

$$t'_i = \frac{\bar{x}_i - \bar{x}}{\sqrt{\frac{s^2}{n_i} + \frac{s_i^2}{N}}} \quad (19)$$

qui suit la loi de Student à v_i degrés de liberté avec :

$$v_i = \frac{\left(\frac{s^2}{n_i}\right)^2 + \left(\frac{s_i^2}{N}\right)^2}{\frac{1}{n_i+1} \left(\frac{s^2}{n_i}\right)^2 + \frac{1}{N+1} \left(\frac{s_i^2}{N}\right)^2} - 2 \quad (20)$$

La valeur numérique de v_i ainsi obtenue n'est en général pas entière ; on adoptera alors l'entier le plus voisin.

Les critères de rejet de l'hypothèse $\bar{x}_i - \bar{x} = 0$ au niveau de probabilité α sont alors les suivants :

$|t'_i| \geq t_{\alpha/2, v}$: les deux moyennes sont significativement distinctes ;

$t'_i \geq t_{\alpha, v}$: les deux moyennes sont distinctes avec $\bar{x}_i > \bar{x}$;

$t'_i \leq -t_{\alpha, v}$: les deux moyennes sont distinctes avec $\bar{x}_i < \bar{x}$.

La probabilité α signifie que l'on a α chance de se tromper si l'on rejette l'hypothèse nulle alors qu'elle est vraie.

Méthode de l'analyse des variances

Cette méthode est plus rigoureuse que la précédente et permet d'apprécier globalement si les mesures effectuées dans tous les forages appartiennent à la même population et si les écarts constatés entre forages résultent du hasard, chaque forage étant considéré comme un échantillon de l'ensemble de la formation. Les hypothèses de base sont : distribution suivant la loi normale et même nombre n de mesures par forages. Le nombre total de mesures dans la formation étudiée est donc $N = nr$.

La méthode consiste alors à évaluer de deux façons différentes la variance s^2 de la population totale dont la moyenne est \bar{x} .

– Évaluation de s^2 à partir des termes relatifs à chaque forage

$$s_1^2 = \frac{1}{r} \sum_{i=1}^{i=r} \left[\frac{1}{n-1} \sum_{j=1}^{j=n} (x_{ij} - \bar{x}_i)^2 \right] \\ = \frac{1}{N-r} \sum_{i=1}^{i=r} \sum_{j=1}^{j=n} (x_{ij} - \bar{x}_i)^2 \quad (21)$$

x_{ij} désignant la j -ième mesure dans le i -ième forage

– Évaluation de s^2 à partir des moyennes de chaque forage :

$$s_2^2 = \frac{n}{r-1} \sum_{i=1}^{i=r} (\bar{x}_i - \bar{x})^2 \quad (22)$$

Le rapport $F = \frac{s_2^2}{s_1^2}$ doit être voisin de 1 et son écart

avec l'unité ne doit pas être significatif. Or F est une variable aléatoire qui suit la loi de Snedecor à $v_1 = r - 1$ et $v_2 = N - r$ degrés de liberté.

La condition d'acceptation de l'hypothèse au seuil de probabilité α est alors : $F \leq F_{\alpha, r-1, N-r}$

S'il n'en est pas ainsi, l'hypothèse doit être rejetée et le terrain peut être considéré comme hétérogène.

Dans la pratique, on présente les résultats sous forme d'un tableau dans lequel la variance est égale à la somme des carrés divisée par le nombre de degrés de liberté.

Sources de variation	Somme des carrés	Degrés de liberté	Variance
Totalité des mesures	$\sum_{i=1}^{i=r} \sum_{j=1}^{j=n} (x_{ij} - \bar{x})^2$	$N - 1$	s^2
En forage	$\sum_{i=1}^{i=r} \sum_{j=1}^{j=n} (x_{ij} - \bar{x}_i)^2$	$N - r$	s_1^2
Entre forages	$n \sum_{i=1}^{i=r} (\bar{x}_i - \bar{x})^2$	$r - 1$	s_2^2

Méthode de Tukey

La méthode de Tukey [2] consiste à chercher si l'intervalle de confiance $\bar{x}_{\max} - \bar{x}_{\min} \pm K$ recouvre zéro avec une probabilité $1 - \alpha$.

S'il est ainsi, tous les couples (\bar{x}_p, \bar{x}_q) , où p et $q \leq r$, sont tels que :

$$\bar{x}_{\max} - \bar{x}_{\min} - K \leq \bar{x}_p - \bar{x}_q \leq \bar{x}_{\max} - \bar{x}_{\min} + K$$

On peut alors admettre que $\bar{x}_p = \bar{x}_q$ avec la probabilité $1 - \alpha$, et que le site est homogène.

S'il n'en est pas ainsi, la différence entre \bar{x}_{\max} et \bar{x}_{\min} doit être considérée comme significative et il faut reprendre le calcul pour d'autres couples de moyennes.

Le facteur k est donné par la relation :

$$k = k_o \sqrt{\frac{S_1^2}{n}}, \quad (23)$$

où k_o est fourni par des tables en fonction de r et de $v = N - r$ et pour $\alpha = 0,05$ et $\alpha = 0,01$ [2].

On voit que cette méthode nécessite l'établissement du tableau d'analyse de la variance.

La méthode de Tukey peut également être utilisée pour préciser les forages qui ont conduit, si tel est le cas, à un rejet de l'hypothèse d'homogénéité par la méthode d'analyse de la variance.

5

Examen de cas réels

5.1

Détermination des caractéristiques géotechniques

5.1.1

TGV Méditerranée (1995-1997). Viaduc sur le Rhône à Avignon

La reconnaissance lithologique a fait apparaître quatre couches principales numérotées de haut en bas de 1 à 4. L'examen des profils pressiométriques a toutefois mis en évidence des différences de compacité dans la couche des alluvions grossières du Rhône selon que l'on se trouve sur les berges, c'est-à-dire à terre, ou dans le lit du fleuve. De plus, dans le lit du fleuve lui-même, on a noté un très important affaiblissement des caractéristiques pressiométriques (E_M et $p_i - p_o$) dans la partie supérieure des alluvions grossières, c'est-à-dire précisément dans la zone qui, avant la stabilisation du lit par les barrages de la Compagnie nationale du Rhône, était le siège d'affouillements importants, pouvant atteindre 8 à 10 mètres de profondeur. Nous avons donc été amenés à diviser cette formation alluvionnaire en trois sous-couches.

Dans le substratum marneux rencontré au-dessous des alluvions, les modules pressiométriques mesurés à terre sont légèrement inférieurs à ceux obtenus dans le lit du Rhône, mais cela pour des raisons purement géologiques. Nous avons donc considéré, dans ce substratum, deux couches géotechniquement différentes. En définitive, alors que la lithologie et la géologie donnaient quatre couches différentes, la synthèse stratigraphique nous a amenés à en définir sept.

La figure 8 montre les histogrammes des modules pressiométriques dans trois de ces couches avec les droites de Henry correspondantes.

Dans les alluvions à terre (Fig. 8a) l'histogramme est *grosso modo* symétrique et les points représentatifs des fréquences cumulées en coordonnées gaussio-arithmétiques sont très correctement alignés, avec un coefficient de régression linéaire de 0,995.

La distribution de E_M est donc normale avec pour la loi ajustée :

$$\bar{x} = 40 \text{ Mpa} \quad s = 19 \text{ Mpa} \quad n = 55$$

Pour un niveau de probabilité $P = 0,95$ soit $\alpha = 1 - P = 0,05$ et $n - 1 = 54$ degrés de liberté, la table de Student donne $t_{\alpha/2} = 2,005$. La limite inférieure de l'intervalle de confiance de la moyenne donné par la formule (10) est alors :

$$I = 40 - 2,005 \frac{19}{\sqrt{55}} = 34,86 \text{ MPa}$$

Nous avons donc retenu pour cette formation : $E_M = 35 \text{ Mpa}$.

Les histogrammes des figures 8b et 8c, concernant respectivement la partie supérieure des alluvions grossières et le substratum situés dans le lit du Rhône, présentent tous deux une asymétrie avec étalement vers la droite.

Les courbes de fréquences cumulées en coordonnées gaussio-logarithmique montrent que les distributions sont log-normales avec :

– alluvions grossières 3a : $\mu = 2,88$ et $\sigma = 0,77$

d'où la valeur modale (formule 11) :

$$X_m = 9,8 \approx 10 \text{ MPa}$$

– substratum : $\mu = 6,11$ et $\sigma = 0,62$

$$\text{d'où la valeur modale : } X_m = 307 \approx 300 \text{ MPa}$$

Nous avons donc adopté :

– alluvions grossières lâches : $E_M = 10 \text{ Mpa}$;

– substratum dans le fleuve : $E_M = 300 \text{ Mpa}$.

En opérant ainsi pour toutes les couches de la synthèse stratigraphique, nous avons obtenu le tableau ci-après qui nous a permis de poursuivre l'étude des fondations de façon déterministe.

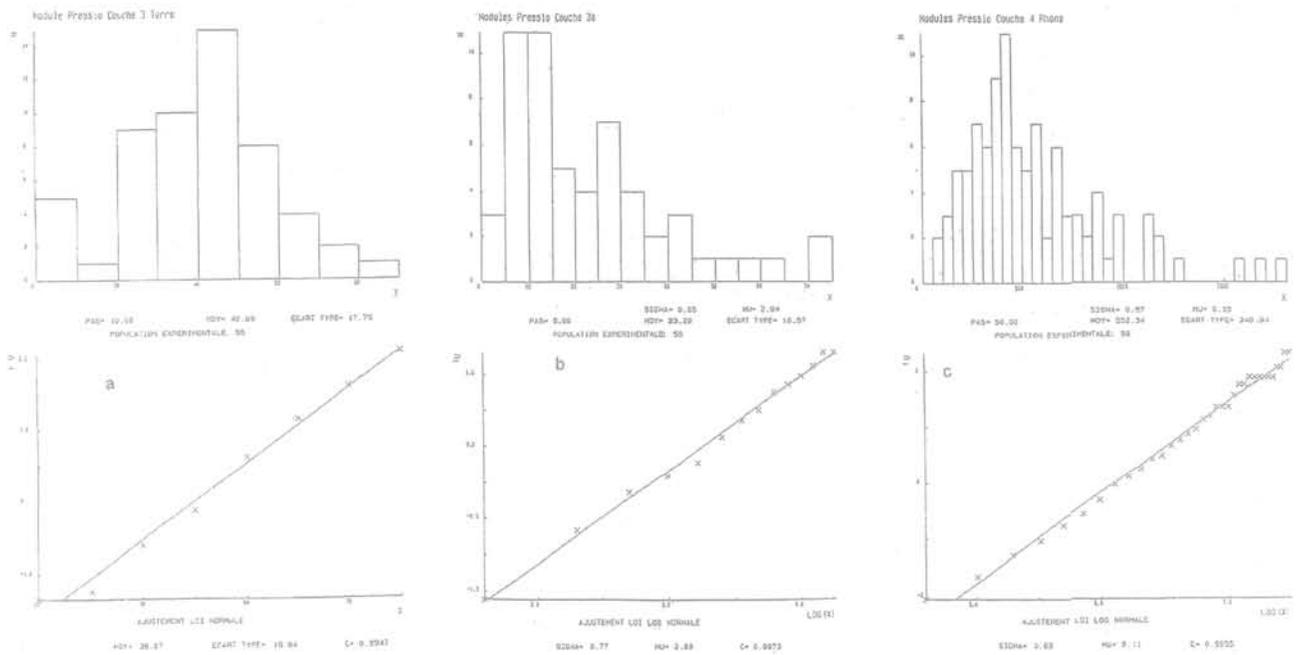


FIG. 8 TGV Méditerranée, viaduc d'Avignon. Distribution des modules pressiométriques.
TGV Méditerranée (High Speed Train Track), bridge in Avignon. Distribution of the pressometer modulus.

Désignation des sols	Nature	E_M (MPa)	$p_1 - p_0$ (MPa)
1	Limons et terrains meubles de couverture	2,2	0,3
2	Sables gris, quelques graviers	11,3	1,0
3	Alluvions grossières à terre	35,0	4,0
3a	Alluvions grossières Rhône	10,0	2,0
3b	Alluvions grossières Rhône	50,0	5,4
4	Substratum Rhône	300,0	7,0
	Substratum terre	220,0	7,0
	Substratum global	270,0	7,0

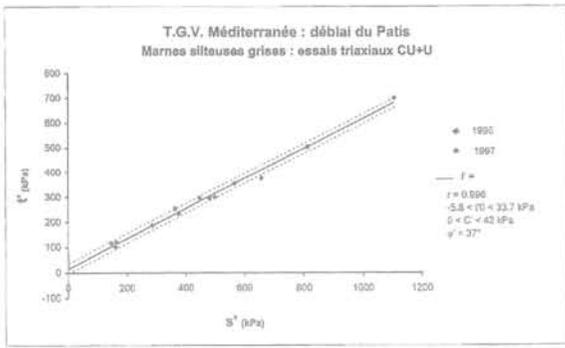
5.1.2

TGV Méditerranée. Déblai du pâti

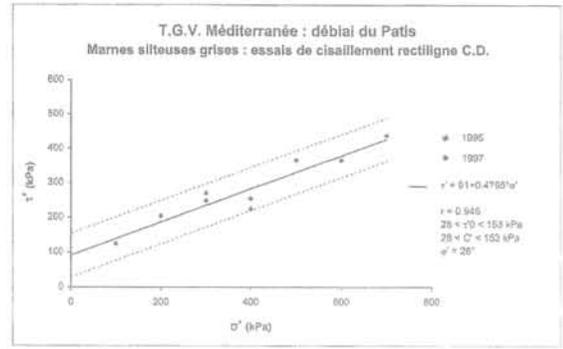
Il s'agit d'une grande tranchée qui entaille une butte de sable compact et de marne silteuse dans le département de la Drôme (plaine de Marsanne). Une première campagne de reconnaissance a été réalisée en 1995 au cours de laquelle, conformément à notre marché, nous avons réalisé dans chacune des deux formations, deux essais triaxiaux CU + U et un essai de cisaillement consolidé drainé à la boîte. C'était vraiment peu pour un tel terrassement !

En 1997, la SNCF nous a demandé une reconnaissance complémentaire avec, à nouveau, deux essais triaxiaux CU + U et un essai à la boîte CD dans chaque couche.

La figure 9 représente les résultats que nous avons obtenus dans les marnes argileuses grises en groupant les deux campagnes avec tous les couples de points (s' , t') ou (σ' , τ'), tous échantillons et toutes profondeurs confondus. Nous remarquons sur le graphe supérieur de la figure (9a) que les points (s' , t') sont remarquablement alignés avec un coefficient de régression de 0,996, ce qui montre l'homogénéité des deux campagnes.



Marnes silteuses grises : essais triaxiaux



Marnes silteuses grises : essais de cisaillement rectiligne

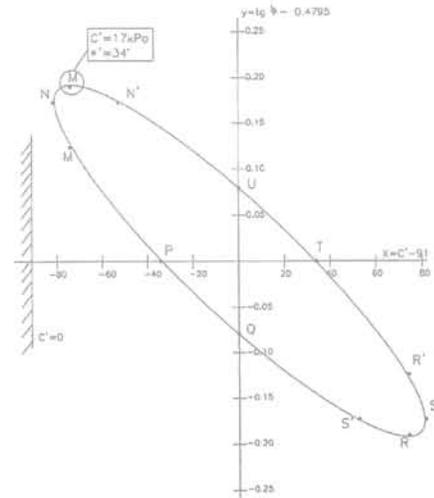
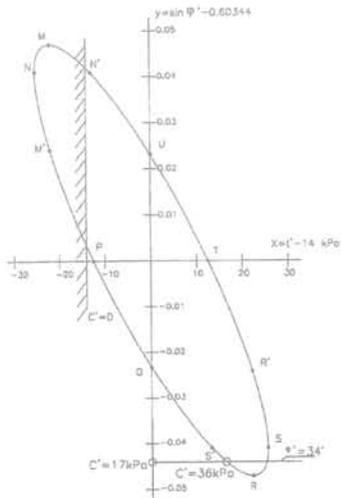


FIG. 9 TGV Méditerranée. Déblai du pâtis. Évaluation de ϕ' et C' . Ellipse de confiance. TGV Méditerranée (High Speed Train Track). Patis Excavation. Evaluation of ϕ' et C' . Confidence ellipse.

Les deux droites en pointillés parallèles à la droite des moindres carrés donnent l'intervalle de confiance à 95 % de l'ordonnée à l'origine de la droite de régression (formule 12). Mais, comme indiqué précédemment, il est plus rigoureux dans ce type d'estimation d'utiliser la notion d'ellipse de confiance et de se placer sur le contour le plus sécurisant de cette ellipse.

La figure 9a (inférieure) montre qu'il est prudent de chercher une valeur de ϕ' et C' vers le bas de la courbe, pour laquelle nous avons adopté un niveau de probabilité de 95 %. Nous pouvons donc prendre : $\phi' = 34^\circ$ et $c' = 36$ kPa pour autant que ce résultat soit compatible avec l'essai de cisaillement à la boîte qui conduit, chose surprenante, à des valeurs de ϕ' significativement plus faibles que celles obtenues dans les essais triaxiaux.

Nous voyons ainsi sur la figure 9b (inférieure) que le seul point qui correspond à $\phi' = 34^\circ$ est le point de contact M de la tangente horizontale supérieure à l'ellipse. En revanche, ce point conduit à $c' = 17$ kPa. Le point qui correspond à $c' = 36$ kPa est voisin de N' et conduit à $\phi' = 33^\circ$ valeur peu différente de 34 degrés.

On peut donc admettre qu'il y a comptabilité (un peu limite certes) entre les deux essais, en ce qui concerne l'angle de frottement interne, mais pas en ce qui concerne la cohésion. Par mesure de prudence, on a adopté $c' = 17$ kPa, mais on a conservé $\phi' = 34^\circ$.

Sur la figure 10, relative aux sables compacts, nous voyons que la linéarité des points représentatifs des essais triaxiaux est également remarquable avec un coefficient de régression linéaire de 0,993, à opposer à

celui des essais de cisaillement à la boîte qui n'est que de 0,843.

S'agissant d'un sable, nous avons résolument adopté une cohésion nulle tout à fait compatible avec les deux types d'essai, et les ellipses de confiance au niveau de probabilité de 95 % montrent qu'il y a également compatibilité des deux types d'essai pour $\phi' = 43^\circ$. Nous avons donc adopté dans les sables $\phi' = 43^\circ$ et $c' = 0$.

5.1.3

Tramway de Strasbourg

Il s'agissait d'évaluer les caractéristiques de résistance au cisaillement ϕ' et c' des alluvions du Rhin, qui ne sont évidemment pas carottables.

La difficulté a été contournée en procédant à des mesures *in situ* au phicomètre, appareil qui permet, à tout niveau, d'appliquer une série de pressions radiales σ_1 sur les parois d'un forage et, pour chacune de ces pressions, d'exercer un effort de traction vertical coaxial au forage jusqu'à la rupture, effort qui correspond, sur la paroi du forage, à une contrainte de cisaillement τ_1 .

Ainsi, à chaque niveau, on dispose d'un certain nombre de couples (5 à 8 environ) de valeurs de σ_1 et τ_1 qui permettent de tracer la droite de Coulomb et d'en déduire des valeurs de ϕ_1 et c_1 peu différentes, en milieu granulaire, de ϕ' et c' , selon l'inventeur de l'appareil, G. Philipponnat.

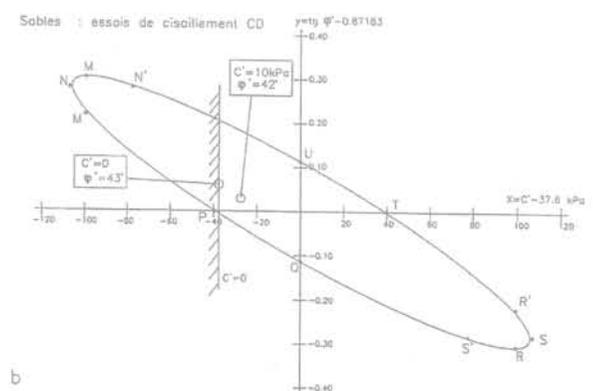
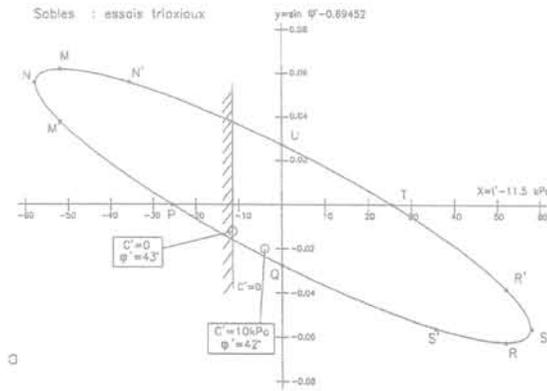
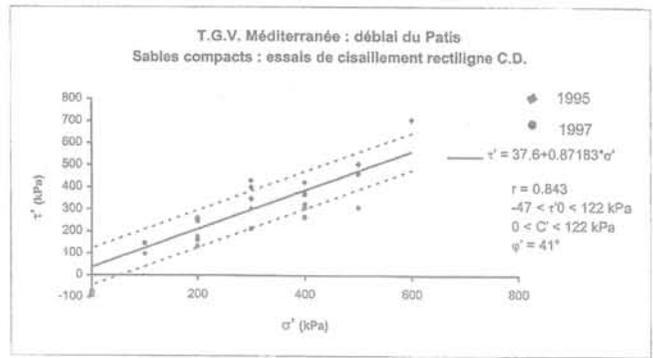
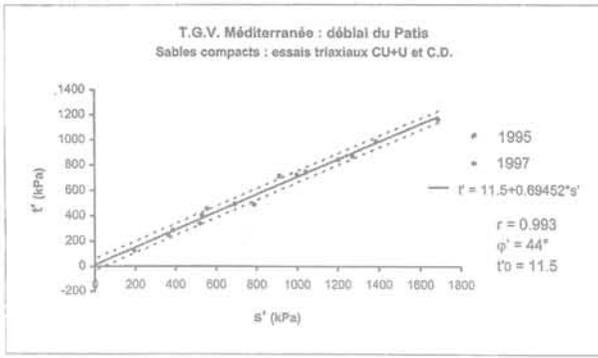


FIG. 10 TGV Méditerranée. Déblai du pâtis. Évaluation de ϕ' et C' . Ellipse de confiance. TGV Méditerranée (High Speed Train Track). Patis Excavation. Evaluation of ϕ' et C' . Confidence ellipse.

Dans le cas présent, on a réalisé de tels essais tous les 1,60 mètre environ jusqu'à 10 mètres de profondeur. Ces essais ont donné des résultats très dispersés avec :

$$23^\circ \leq \phi_i \leq 41^\circ \text{ et } 0 \leq C_i \leq 72 \text{ kPa}$$

Nous avons donc porté sur le même graphique (Fig. 11a) l'ensemble des couples (σ', τ) mesurés tous forages et toutes profondeurs confondues, soit 90 couples. La droite des moindres carrés du nuage de points a conduit à un coefficient de régression de 0,912 ce qui est tout à fait correct et justifie cette approche, et a donné un angle de 34 degrés avec une cohésion de 16 kPa.

Nous n'étions pas d'accord pour attribuer à cette formation essentiellement granulaire une quelconque cohésion, mais le conseil du maître d'œuvre, s'appuyant sur le fait qu'on n'avait pu justifier les déformations des parois du métro de Lyon qu'en faisant intervenir une cohésion dans les alluvions du Rhône, voulait absolument en prendre une dans les alluvions du Rhin, qui ne sont pas plus cohérentes que celles du Rhône.

Or l'examen de l'ellipse de confiance a montré qu'avec une cohésion de 10 kPa, l'intervalle de confiance de ϕ' au niveau de probabilité de 95 % était tel que : $33^\circ \leq \phi \leq 35,5^\circ$

Nous avons donc finalement adopté : $\phi' = 34^\circ$ et $c' = 10$ kPa.

Tout s'est très bien comporté, mais nous avons toutefois clairement précisé, dans notre étude, que la cohésion prise en compte ne constituait pas une liaison entre les particules du sol, comme dans une argile, mais résultait d'une imbrication des grains, et que cette

soit disant cohésion n'était qu'un artifice dû à l'extrapolation de la partie linéaire de la courbe $\tau = f(\sigma')$ pour les faibles valeurs de σ' .

5.2

Homogénéité d'un site

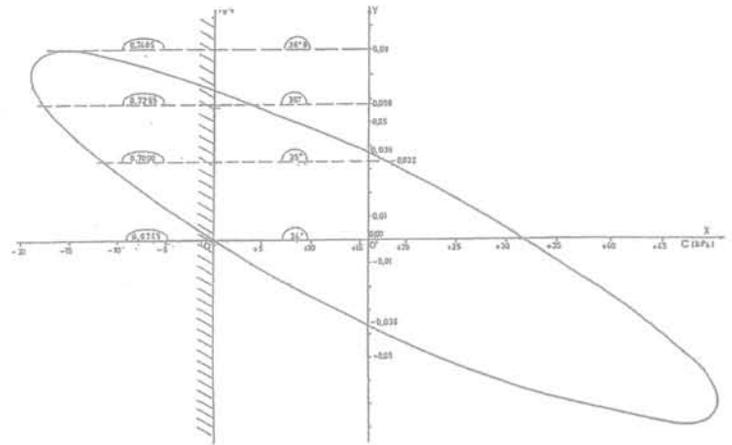
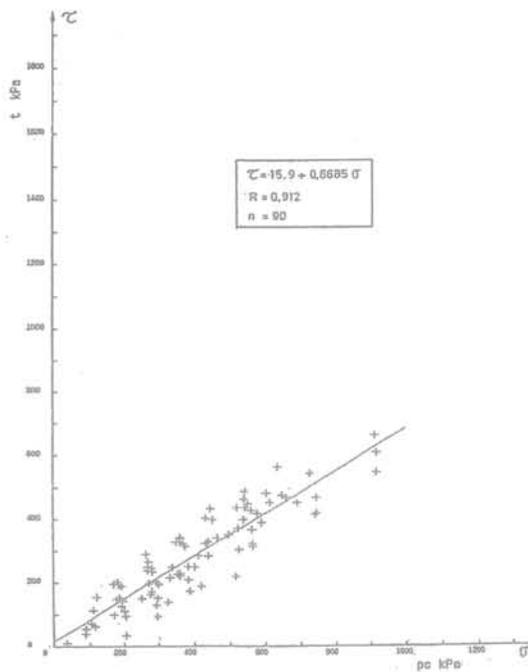
5.2.1

Limons argileux de Cagnes-sur-Mer

Dans ces matériaux relativement cohérents, nous avons réalisé 75 essais au scissomètre qui ont donné des valeurs conventionnelles de la cohésion non drainée et qui ont été répartis sur 5 forages.

Nous avons testé l'homogénéité du site avec la méthode approchée développée en 4.1., en comparant les moyennes empiriques obtenues dans chaque forage à la moyenne empirique globale de l'ensemble des mesures. Avec les mêmes notations, nous avons le tableau suivant :

N° Forages	\bar{x}_i (kPa)	s_i^2 (kPa) ²	n_i	$\frac{s_i^2}{n_i}$
SC1	38	121	26	4,65
SC2	43	64	21	3,05
SC3	49	289	20	14,45
SC4	52	9	3	3,0
SC5	35	121	5	24,2



ESTIMATION GLOBALE

FIG. 11 Tramway de Strasbourg. Évaluation de ϕ_i et C_i au phicomètre. Tramway in Strasbourg. Alluvial deposits of the Rhin River. Évaluation of ϕ_i and C_i with phicometer.

Nous avons par ailleurs pour la population totale :

$$N = 75 \quad \bar{x} = 43 \text{ kPa} \quad s^2 = 169 \quad \frac{s^2}{N} = 2,25$$

À partir des formules (19) et (20), nous obtenons, tous calculs faits pour $\alpha = 0,05$:

N° forage	t'	v	$t_{\alpha/2, v}$	$t_{\alpha, v}$
SC1	-1,896	52	2,009	1,676
SC2	0,00	57	2,000	1,671
SC3	1,466	26	2,056	1,706
SC4	3,928	10	2,228	1,812
SC5	-1,564	5	2,570	2,015

On constate que seul l'écart entre la moyenne de la population totale et la moyenne obtenue dans le forage SC4 est significatif par valeur inférieure, les autres écarts ne l'étant pas.

On peut dire que la zone de SC4 présente une particularité (affaiblissement ou essais défectueux) alors que le reste du site est homogène.

5.2.9

Marnes argileuses de Palavas-les-Flots

Ces marnes ont fait l'objet de prélèvements d'échantillons intacts et d'essais de compression simple

en laboratoire dans le cadre de l'étude du pont des Quatre-Canaux. Ces essais étaient répartis en quatre forages avec trois essais par forage.

Nous avons utilisé ici la méthode d'analyse de la variance.

Les données des mesures nous ont permis de dresser le tableau ci-après :

N° Forage	x_{ij} (MPa)	\bar{x}_i (MPa)	s_i^2 (MPa) ²
SC1	0,68 0,31 0,14	0,377	0,0762
SC4	0,38 0,31 0,38	0,357	0,0016
SC6	0,28 0,48 0,48	0,413	0,0132
SC10	0,35 0,50 0,35	0,400	0,0076

Le tableau d'analyse des variances est alors le suivant :

Source de variation	Somme des carrés	Degrés de liberté	Variances
Totalité des mesures	0,20307	11	$s^2 = 0,01846$
En forages	0,1974	8	$s_1^2 = 0,02468$
Entre forages	0,005534	3	$s_2^2 = 0,00184$

On en déduit : $F = \frac{s_2^2}{s_1^2} = 0,07455$

Au niveau de probabilité $\alpha = 0,05$ on a, d'après la table de Snédécour :

$$F_{\alpha, r-1; N-r} = F_{0,05; 3; 8} = 4,07$$

Comme F est largement inférieur à 4,07, le test est accepté et on peut admettre que le site est homogène.

Appliquons, par curiosité, la méthode de Tukey aux moyennes extrêmes :

on a : $\bar{x}_{\max} = 0,413$ MPa (forage SC6)

$\bar{x}_{\min} = 0,357$ MPa (forage SC4)

Pour $r = 4$, $v = N - r = 8$, $n = 3$ et $\alpha = 0,05$, les tables de Tukey donnent $k_0 = 4,53$ d'où :

$$k = k_0 \sqrt{\frac{s_1^2}{n}} = 4,53 \sqrt{\frac{0,02468}{3}} = 0,41$$

$$\bar{x}_{\max} - \bar{x}_{\min} \pm 0,41 = -0,35/0,47$$

Cet intervalle recouvre largement le zéro, ce qui confirme bien le résultat de l'analyse des variances.

6

Conclusion

Nous avons présenté, dans les développements qui précèdent, quelques-uns des outils que la statistique mathématique met à la disposition de l'ingénieur géotechnicien. Ces outils peuvent paraître, à première vue, d'un emploi compliqué, mais, en fait, leur utilisation,

facilitée par des tables numériques ou des logiciels, est relativement simple, comme le montrent les exemples réels cités.

On a pu voir se dégager tout au long de cet exposé les grandes lignes d'une méthodologie d'analyse qui devrait être appliquée à toute reconnaissance géotechnique. Cette méthodologie comprend trois phases :

– le dépouillement des résultats au sens le plus large du terme avec la classification des paramètres mesurés par nature de terrain, l'établissement des histogrammes et leur ajustement sur une loi de probabilité connue ;

– la détermination des valeurs de calculs à partir des notions d'intervalle de confiance pour les caractéristiques indépendantes et d'ellipse de confiance pour les caractéristiques liées ;

– l'étude de l'homogénéité d'un site par la méthode d'analyse de la variance ou par des méthodes moins sophistiquées utilisant un test d'hypothèse paramétrique double, basé sur la comparaison de deux moyennes empiriques.

Depuis des décennies, dans la grande majorité des cas, les données d'une reconnaissance géotechnique sont traitées par l'ingénieur de façon plus ou moins subjective. Grâce à l'outil statistique, il devient possible, sans faire appel à une mathématique sophistiquée qui pourrait être dissuasive, de conférer à ces choix techniques une certaine objectivité, qui assure la sécurité sans pour autant empiéter sur le jugement et sur l'expérience de l'homme.

Nous ne pouvons donc qu'exhorter les ingénieurs géotechniciens à utiliser systématiquement ces méthodes, qui conduisent à une meilleure appréciation des propriétés d'un sol mais qui permettent également de juger de la qualité d'une expérimentation.

Bibliographie

- Bonitzer J. (1983) – *L'emploi des statistiques*. Paris, Presses des Ponts et Chaussées.
- Bower A.H., Lieberman G.J. (1964) – *Méthodes statistiques de l'ingénieur*. Paris, Dunod.
- Calot G. (1964) – *Cours de statistiques descriptives*. Paris, Dunod.
- Calot G. (1967) – *Cours de calcul des probabilités*. Paris, Dunod.
- Cambefort H. (1971) – *Introduction à la géotechnique*. Paris, Eyrolles.
- Cassan M. (1979) – « Détermination probabiliste des caractéristiques mécaniques des sols ». Conférence ECP, non publiée.
- CEA (1979) – *Statistique appliquée à l'exploitation des mesures*. Paris, Masson (2 tomes).
- Favre J.-L. (1980) – « Milieu continu et milieu discontinu : mesure statistique indirecte des paramètres rhéologiques et approche probabiliste de la sécurité ». Thèse de doctorat d'État en sciences physiques, université Pierre et Marie Curie, Paris.
- Guy J. (1967) – *Éléments de calcul des probabilités et de calcul statistique*. Paris, SEDES.
- Harr M.E. (1981) – *Mécanique des milieux formés de particules*. Presses polytechniques romandes.
- Labrousse C. (1978) – *Statistiques appliquées*. Paris, Dunod (3 tomes).
- Liorzou A. (1976) – *Initiation pratique à la statistique*. Paris, Eyrolles.
- Lipschutz S. (1979) – *Probabilités. Cours et problèmes*. Série Schaum, New York, Mc Graw Hill Inc., 1973, Paris.
- Magnan J.-P. (1982) – *Les méthodes statistiques et probabilistes en mécanique des sols*. Paris, Presses des Ponts et chaussées.
- Moreau M., Mathieu A. (1979) *Statistique appliquée à l'expérimentation*. Paris, Eyrolles.
- Mothes J. (1967) – *Incertitudes et décisions industrielles*. Paris, Dunod.
- Reeb G., Fuchs A. (1967) – *Statistiques commentées*. Paris, Gauthier-Villars.
- Spiegel M.R. – *Théorie et application de la statistique*. Série Schaum, New York-Paris Mc Graw Ill Inc, 1978.

Une méthode de détermination des valeurs caractéristiques des paramètres géotechniques

F. BAGUELIN

Fondaconcept-Fondasol
14, rue Palestro
93500 Pantin

J.-B. KOVARIK

Port autonome de Rouen
34, bd de Boisguilbert
BP 4075
76022 Rouen Cedex

Résumé

En fixant le niveau de risque pour lequel les valeurs caractéristiques des paramètres géotechniques doivent être déterminées, les Eurocodes ouvrent la voie à une approche rationnelle de ces valeurs, basée sur le traitement statistique des données de sol. On présente la méthode pratique qui a été élaborée dans le cadre de la mise au point des *Recommandations pour le calcul aux états-limites des ouvrages en site aquatique*, et qui répond en outre aux exigences de l'Eurocode 7 « Géotechnique » sur la prise en compte de l'étendue de la sollicitation du sol dans la détermination de la valeur caractéristique.

Mots clés : conception, calcul, géotechnique, état-limite, valeur caractéristique, statistiques.

A method for determining the characteristic values of geotechnical parameters

Abstract

In the Eurocodes the level of risk required for the characteristic values of the geotechnical parameters has been specified. Thus the route is opened for a rational approach to the determination of these values, based on a statistical analysis of soil data. In this article, a practical method is presented, which has been developed while drafting the *Recommendations for Limit State Design of Waterways and Harbour Structures*. This method also meets the requirement of Eurocode 7 « Geotechnical design » which calls for taking account of the extent of soil loading in determining a characteristic value.

Key words : design, geotechnics, limite state, characteristic value, statistics.

NDLE : Les discussions sur
cet article sont acceptées
jusqu'au 1^{er} août 2001.

Introduction

La dispersion des propriétés des sols est un problème que doit inévitablement traiter l'ingénieur dans la pratique courante géotechnique. Or, que ce soit dans l'enseignement ou dans la réglementation, cette question est très rarement voire jamais abordée. L'ingénieur est donc laissé à lui-même, et naturellement, si l'on soumet un même cas à différentes personnes, on constate une grande diversité dans le choix des valeurs de calcul à utiliser.

Pour aborder convenablement le problème, il faut distinguer deux questions :

1) *Quel niveau de risque doit-on retenir pour les valeurs « de calcul » ?*

L'approche traditionnelle consistait à travailler avec des coefficients de sécurité globaux, par exemple 3 pour le poinçonnement des fondations superficielles. Mais, on ne savait pas si la valeur « de calcul » de la résistance du sol devait être une valeur moyenne, une valeur prudente, ou une valeur très prudente. Dans les documents actuellement en vigueur pour les fondations des ouvrages, fascicule 62 Titre V du CCTG pour les ouvrages d'art, DTU 13.11 et 13.2 pour les bâtiments, l'approche des coefficients partiels est retenue, mais on constate le même silence sur le choix des valeurs « de calcul ».

2) *Comment déterminer en pratique ces valeurs ?*

En effet, la dispersion des propriétés des sols est causée par des phénomènes physiques complexes ; elle peut être représentée par des lois spécifiques et les outils de traitement statistique habituels sont insuffisants. En outre, le nombre de données est le plus souvent très limité, ce qui accroît la difficulté.

Les Eurocodes apportent une réponse à la première question. Ce peut donc être le point de départ pour élaborer une méthode de traitement statistique des données de sol. C'est ce que nous avons tenté de faire dans le cadre de la mise au point du document *Recommandations pour le calcul aux états-limites des ouvrages en site aquatique* (en abrégé *Recommandations*).

Dans cet article, après avoir précisé les données de départ, nous exposons dans ses grandes lignes la méthode pratique que nous avons retenue dans les *Recommandations* pour répondre à la seconde question. Pour plus de détails on pourra se reporter au document lui-même, diffusé sous forme du CD-Rom ROSA 2000 par le Centre d'études techniques maritimes et fluviales (CETMEF).

Le point de départ

L'objectif, selon les Eurocodes

Les Eurocodes sont encore des normes provisoires, et leur état est encore fluctuant. Cependant les principes généraux n'ont guère varié depuis l'origine, et on peut considérer qu'ils ne seront pas remis en cause dans les versions définitives.

Il existe un Eurocode de base, l'Eurocode n°0, qui donne les principes retenus pour la justification des ouvrages.

D'une manière générale, trois types de paramètres interviennent : 1) les charges ; 2) les propriétés des matériaux, les résistances ou les modules de déformation ; 3) les modèles de calcul. On en détermine d'abord des valeurs dites caractéristiques, qui correspondent à une probabilité maximale de mise en défaut de 5 %.

Ce sont les valeurs caractéristiques qui sont considérées dans les vérifications des états-limites relevant de la catégorie des ELS (états-limites de service). Ces phénomènes préjudiciables (états-limites) correspondent typiquement à l'apparition de défauts gênant l'utilisation de l'ouvrage, mais pouvant être réparés. Ils font en général intervenir les modules de déformation. La probabilité cible d'occurrence de tels états-limites est de l'ordre de 10^{-3} par an.

Pour les vérifications des états-limites relevant de la catégorie des ELU (états-limites ultimes), le risque de rupture est en jeu et ce sont les résistances qui interviennent, sauf cas particuliers (dans ces cas, les modules de déformation doivent être évalués avec une probabilité de mise en défaut de 5 %, et non comme une valeur moyenne). Ce sont les valeurs de calcul qui sont considérées dans les vérifications des ELU, et qui se déduisent des valeurs caractéristiques par l'application de coefficients partiels, majorants pour les charges, minorants pour les résistances. Ces coefficients partiels sont destinés à réduire la probabilité cible d'apparition d'un ELU à une valeur de l'ordre de 10^{-6} par an.

On voit donc que, en général, l'objectif visé pour les valeurs caractéristiques est une certaine prudence : risque de 5 %, soit 1 chance sur 20 de mise en défaut. Ce n'est ni la valeur la plus probable (valeur moyenne), ni une valeur donnant une garantie « absolue » de non franchissement. C'est par la combinaison des différentes valeurs représentatives des paramètres, avec des coefficients partiels, que l'on obtient des probabilités suffisamment basses vis-à-vis du dysfonctionnement ou de la ruine des ouvrages.

L'actuelle version provisoire de l'Eurocode 7, Eurocode dédié aux ouvrages géotechniques, donne les directives suivantes pour la détermination des valeurs caractéristiques des sols et des roches :

La valeur caractéristique d'un paramètre de sol ou de roche doit être choisie comme une estimation prudente de la valeur influençant l'occurrence de l'état limite.

L'étendue de la zone de terrain qui gouverne le comportement d'un ouvrage géotechnique vis-à-vis d'un état limite donné est généralement beaucoup plus grande que celle qui intervient dans un essai sur le sol ou sur la roche et, par conséquent, le paramètre qui contrôle le comportement de l'ouvrage est souvent une valeur moyenne sur une certaine surface ou un certain volume de sol. La valeur caractéristique est une estimation prudente de cette valeur moyenne.

Lorsque des méthodes statistiques sont utilisées, il convient d'établir la valeur caractéristique de telle sorte que la probabilité calculée d'une valeur plus défavorable contrôlant l'occurrence de l'état limite ne dépasse pas 5 %.

La méthode. Cas d'une variable scalaire

3.1

Considérations générales

Pour simplifier le langage, nous parlerons de couche de sol, sachant qu'il peut s'agir en réalité d'une couche ou d'une lentille de sol, ou d'un volume de forme quelconque de sol ou de roche. Et pour commencer, nous supposons que nous avons affaire à une résistance s'exprimant par une variable scalaire unique (exemples : cohésion non drainée c_u , pression limite pressiométrique p_p , résistance pénétrométrique q_p). Le cas où la résistance est caractérisée par deux paramètres liés (cohésion c' et angle de frottement ϕ' , ou variation avec la profondeur z de la pression limite pressiométrique p_p) sera abordé dans la section suivante.

Les données sont N valeurs x_i (x_1 à x_N) du paramètre X . De plus, nous supposons que les N valeurs dont nous disposons sont indépendantes du point de vue statistique. C'est un point dont nous préciserons les conditions un peu plus loin.

On fait l'hypothèse *a priori* que la distribution du paramètre X s'ajuste à une loi normale. Cette loi, classique et justifiée dans la pratique de l'ingénieur par le théorème central limite, est appropriée dans la plupart des cas où l'on a un seul mode, c'est-à-dire une seule valeur dominante. Il est toutefois préférable d'adopter une loi log-normale dans le cas où le coefficient de variation (rapport de l'écart type à la valeur moyenne) est grand (par exemple supérieur à 15 %); ceci permet d'éviter ensuite les valeurs négatives à probabilité non nulle. On travaille alors dans ce cas avec le logarithme du paramètre X en posant $Y = \ln(X)$: Y est alors supposée être une variable aléatoire normale.

La distribution des valeurs observées peut être caractérisée par deux valeurs :

- la moyenne empirique ou observée : $m_x = \sum x_i / N$;
- l'écart type estimé : $s_x = [\sum (x_i - m_x)^2 / (N-1)]^{0.5}$.

La méthode propose de procéder en deux étapes distinctes.

3.2

Le paramètre local

Dans la première étape, on s'intéresse au paramètre local (ou ponctuel), c'est-à-dire à la grandeur résultant des mesures. Sa variabilité ne dépend que de la couche étudiée et ne fait bien sûr pas intervenir l'ouvrage ni le mécanisme de rupture à vérifier. On caractérise sa distribution par deux valeurs particulières, qui sont des valeurs intrinsèques au sol :

- une valeur moyenne inférieure, X_{mi} , définie comme l'estimation par défaut de la moyenne du paramètre local au risque statistique de 25 % (ce qui est conforme aux recommandations de l'Eurocode). Elle se calcule à partir des paramètres de la distribution observée par :

$$X_{mi} = m_x - k_\alpha \cdot s_x$$

On voit que l'objectif de 5 % est rappelé à propos de l'utilisation possible des méthodes statistiques, cette indication ayant valeur générale d'après l'Eurocode 0.

En outre, l'Eurocode 7 associe clairement la notion de valeur caractéristique au mécanisme de ruine (ou de déformation) : *il ne peut donc s'agir d'une valeur intrinsèque au sol*. La probabilité de 5 % intéresse la valeur du paramètre qui entre en compte dans le modèle de calcul ; et l'on comprend qu'un ouvrage de petite taille va être plus vulnérable vis-à-vis de défauts de résistance localisés qu'un grand ouvrage où va s'opérer une certaine moyenne spatiale des variations de résistance.

Le problème ne se réduit donc pas à celui de la détermination du fractile à 5 % de la distribution estimée du paramètre géotechnique dans la couche ou la formation considérée.

2.2

Les conditions préalables

Les méthodes de traitement statistique ne peuvent porter que sur des populations que l'on a clairement identifiées comme homogènes, pour que la variabilité puisse être modélisée par une loi statistique.

Ceci implique que le géotechnicien doit faire auparavant et convenablement son travail habituel de détermination de la structure géologique et géotechnique du site. L'application des méthodes statistiques ne compensera pas les carences éventuelles dans ce domaine. La présence non détectée d'un accident géologique, tel qu'une vallée fossile, ou la méconnaissance de la structure particulière d'une formation, telle qu'une alternance de deux types de sols ou roches (limons et sables dans un site alluvionnaire, ou marne et calcaire dans les flyschs) pourra invalider la reconnaissance, que l'on ait utilisé ou non les méthodes statistiques pour l'interprétation.

Nous supposons également qu'après avoir identifié une formation comme homogène, le géotechnicien a fait son travail habituel de conversion, pour passer de grandeurs mesurées aux grandeurs qui entrent dans le modèle de calcul. Par exemple, les biais liés à une méthode d'essai sont supposés être traités. Ce peut être la correction de plasticité pour l'essai scissométrique, l'interprétation de l'essai Franklin pour la résistance à la compression d'une roche ; ou comme mentionnés dans l'Eurocode 7 : l'effet d'échelle dû à la fissuration ou l'effet de la durée de l'essai par rapport à la durée des charges sur l'ouvrage, etc.

Nous supposons en outre que la dispersion des données relative à la formation étudiée reflète la dispersion spatiale des propriétés correspondantes, et que la dispersion due à la méthode de mesures, soit a pu être isolée et prise en compte par ailleurs (c'est sans doute un vœu pieux dans le cas général), soit est négligée (ce qui va dans le sens de la sécurité).

Ainsi, le problème que nous abordons ici est strictement le traitement de la dispersion spatiale d'un paramètre de sol devant entrer dans un modèle de calcul clairement identifié par ailleurs. L'objectif est d'obtenir une valeur dite caractéristique, au risque de mise en défaut de 5 %.

où k_α est un coefficient d'incertitude statistique, fonction du nombre d'observations N , donné par le tableau I pour le risque statistique $\alpha = 25\%$. Son expression exacte est :

$$k_\alpha = t^{(N-1)\alpha} / \sqrt{N}$$

où $t^{(N-1)\alpha}$ est le fractile à α de la loi de Student à $N - 1$ degrés de liberté.

TABLEAU I k_α [N; $\alpha = 25\%$].

N	2	3	4	5	6	8	10	20	30	100
k_α	0,71	0,47	0,39	0,33	0,30	0,25	0,22	0,15	0,12	0,07

• une *valeur basse*, X_b , définie comme la valeur caractéristique du paramètre local, c'est-à-dire présentant un risque de mise en défaut de 5%. Elle se calcule par :

$$X_b = m_x - k_\alpha \cdot s_x$$

où k_α est un coefficient d'incertitude statistique, fonction du nombre d'observations N , donné par le tableau II.

Son expression exacte est :

$$k_\beta = t^{(N-1)\beta} / \sqrt{N}$$

où $t^{(N-1)\beta}$ est le fractile à β de la loi de Student à $N - 1$ degrés de liberté.

TABLEAU II k_β [N; $\alpha = 5\%$].

N	2	3	4	5	6	8	10	20	30	100
k_β	7,73	3,37	2,63	0,33	2,33	2,18	2,00	1,92	1,76	1,64

3.3

Le paramètre étendu

Dans la deuxième étape, on considère le paramètre étendu, défini comme une moyenne spatiale du paramètre local, sur une surface ou un volume de sol correspondant au mécanisme de rupture à vérifier. La *valeur caractéristique* X_k de la propriété X est définie comme la valeur du paramètre étendu ayant une probabilité de mise en défaut de 5%. On adopte ici une démarche simplifiée dans laquelle la valeur caractéristique (du paramètre étendu) varie entre la valeur basse et la valeur moyenne inférieure du paramètre local X_{mi} et X_b , selon la relation :

$$X_k = X_{mi} - (X_{mi} - X_b) / k_v$$

avec : $k_v = (n_{loc})^{0,5}$

où : k_v est un coefficient de réduction d'écart type ;
et : n_{loc} est le nombre supposé de valeurs indépendantes du paramètre local sur la surface ou dans le volume en cause dans l'état-limite.

Ce nombre $n_{loc}(x, y, z)$ dépend :

- des distances d'auto-corrélation du sol pour la propriété en cause ;
- de l'étendue du volume de sol concerné : n_{loc} est donc le produit de $n_{loc}(x, y, z)$ dans les trois directions.

Lorsque l'étendue de la surface ou du volume sollicité est faible par rapport à la distance d'auto-corrélation, n_{loc} est petit, voire égal à 1 : la valeur caractéristique X_k se rapproche de la valeur basse X_b , voire lui est égale.

À l'inverse, lorsque l'étendue de la surface ou du volume sollicité est grande par rapport à la distance d'auto-corrélation, n_{loc} est grand : la valeur caractéristique X_k se rapproche de la valeur moyenne inférieure X_{mi} , voire lui est égale (phénomène de réduction de variance).

Généralement pour les sols, et c'est le cas en particulier des sols sédimentaires, on peut distinguer une distance d'auto-corrélation horizontale L_H et une distance d'auto-corrélation verticale L_V .

Le tableau III, donné à titre indicatif, présente des ordres de grandeur de valeurs typiques de distances d'auto-corrélation pour la détermination de n_{loc} :

TABLEAU III Valeurs typiques de distances d'auto-corrélation.

Typical values of auto-correlation distances.

Sol	L_H horizontale	L_V verticale
Forte auto-corrélation	15 m	2 m
Auto-corrélation courante	10 m	1 m
Faible auto-corrélation	5 m	0,5 m

La distance L_H étant généralement beaucoup plus élevée que la distance L_V , le nombre n_{loc} sera souvent égal ou proche du nombre de tranches d'épaisseur L_V recoupées par la surface ou le volume concerné. Il convient d'être prudent dans le décompte des valeurs suivant l'horizontale. En effet, si une ligne de glissement a une grande extension horizontale, elle aura tendance à venir se localiser suivant le lit de résistance la plus faible.

On remarquera que les règles de calcul de la pression limite équivalente p_{ie}^* , ou de la résistance pénétrométrique équivalente q_{ce} (fascicule 62 titre V), sont réputées donner des valeurs caractéristiques en partant d'un sondage supposé parfaitement représentatif des variations des propriétés du sol avec la profondeur : il n'est donc pas nécessaire de mettre en œuvre un n_{loc} vertical.

Il convient de revenir maintenant sur la question de l'indépendance statistique des données x_i de départ. En effet, dire qu'il y a auto-corrélation signifie que les valeurs de la variable X à deux emplacements proches ne sont pas statistiquement indépendantes. Plus leur distance respective descend en dessous de la distance d'auto-corrélation, plus la corrélation est forte. Pour que les données de départ puissent être considérées comme statistiquement indépendantes, il convient qu'elles ne se rapportent pas à des emplacements trop proches les uns des autres, mais qu'elles soient convenablement réparties tant dans le sens vertical que dans le sens horizontal, la référence étant les distances d'auto-corrélation L_H et L_V . Si par exemple deux valeurs ont été mesurées au même niveau et à une distance horizontale réduite, alors que les autres valeurs sont bien réparties spatialement, on devra ne compter ces deux valeurs que pour une, avec leur valeur moyenne.

Cas de paramètres multiples

4.1

Considérations générales

Nous nous intéressons à la variation avec la profondeur z d'une propriété mécanique R , résistance (c_v , p_1 ou q_c) ou déformabilité (module E_M ...). On recherche une relation linéaire entre R et z de la forme $R = a + b.z$, éventuellement constante ($R = a$) par la technique des moindres carrés. Le problème consiste alors, connaissant la profondeur z , à estimer l'incertitude sur la résistance R .

Nous nous intéressons aussi aux paramètres de la résistance au cisaillement, cohésion c' et coefficient de frottement $\text{tg}\phi'$ (plutôt que l'angle de frottement ϕ'). Ces deux paramètres doivent être déduits d'une série de N couples (σ_i, τ_i) , résistance au cisaillement τ_i mesurée dans un essai en fonction de la contrainte normale σ_i . Là encore on recherche une relation linéaire entre τ et σ' de la forme $\tau = c' + \text{tg}\phi' \cdot \sigma'$, par la technique des moindres carrés. Le problème est différent parce qu'il consiste à estimer l'incertitude propres aux paramètres de la régression.

La prise en compte des incertitudes statistiques sur la détermination des paramètres a et b d'une régression linéaire permet en effet de déterminer :

- des intervalles de confiance pour la valeur moyenne de la variable R (paramètre de résistance);
- des ellipses de confiance (ellipses de concentration) pour l'évaluation conjointe de la moyenne des paramètres c' et $\text{tg}\phi'$ (résistance au cisaillement).

Pour les propriétés de résistance au cisaillement c' et $\text{tg}\phi'$, l'application d'une méthode statistique sur les paramètres de la régression peut être très complexe. Une démarche pragmatique doit être mise en œuvre : par suite de la corrélation entre ces deux paramètres, on ne peut rigoureusement déterminer de valeurs représentatives que d'une fonction de c' et de $\tan(\phi')$: il resterait donc à considérer la « résistance » $R(c', \tan\phi')$ comme une variable aléatoire dont on calculerait les moments par développement de Taylor, puis les valeurs moyennes inférieures et basses comme dans le cas d'une variable scalaire. En raison de lois physiques (existence des seuils minimum, respectivement 0 et $\text{tg}\phi'_{\text{crit}}$) qui constituent des données supplémentaires par rapport aux données brutes utilisées par la théorie statistique, on choisit de développer une procédure simplifiée en utilisant les facteurs k précédents développés dans le cadre de la variable unique.

4.2

Étude de la régression linéaire

Dans la suite de cet article, on décrit les techniques de régression linéaire avec les variables X , y , a et b , telles que :

- $y = \sigma'$, contrainte normale imposée dans l'essai, ou la profondeur z , considérée dans les deux cas comme variable déterministe;

- $X = \tau$, résistance au cisaillement mesurée dans

l'essai, ou R , paramètre de résistance (pressiométrique...), considérée dans les deux cas comme variable aléatoire;

- $(y_i, X_i) = i^{\text{ème}}$ couple de valeurs mesurées dans les essais;

- a et b , coefficients de la droite de régression (pour l'étude de la résistance au cisaillement, $a = c'$ et $b = \text{tg}\phi'$).

reliées par : $X = a + b.y$, c'est-à-dire : $\tau = c' + \text{tg}\phi' \cdot \sigma'$ ou encore : $R = a + b.z$

Une fois déterminés les paramètres a et b et l'écart type $s_{X|y}$ de la loi conditionnelle $X|y$, on détermine trois droites : D_{mi} , D_b et D_k , moyenne inférieure, basse et caractéristique, analogues aux valeurs X_{mi} , X_b et X_k du cas de la variable scalaire unique.

On peut faire différentes hypothèses sur la variation de l'écart type $s_{X|y}$ avec y .

Nous considérons les cas suivants :

I. Écart type constant : $s_{X|y} = s$

II. Écart type variant linéairement avec y (c'est-à-dire avec la contrainte normale s' ou la profondeur z), cas général : $s_{X|y} = t.g(y)$

avec : $g(y) = c + d.y$, fonction de y donnée *a priori*.

Les paramètres a , b et t sont à estimer en donnant plus de « poids » aux données qui se situent dans les régions de plus faible variance. On choisit de pondérer les écarts quadratiques par des poids w_i en posant : $w_i = 1/g^2(y_i)$.

III. Écart type et résistance proportionnels à la contrainte normale s' ou à la profondeur z , cas particulier de la variation linéaire :

$$X = b.y$$

$$g(y) = b.y$$

Le choix du type de cas à utiliser est laissé à l'appréciation de l'ingénieur. Cependant on trouvera dans les *Recommandations* des indications à ce sujet.

Les valeurs des coefficients a , b , s ou t sont pour les différents cas :

Cas I : s constant

$$b = (\sum x_i y_i - N.m_x.m_y) / (\sum y_i^2 - N.m_y^2)$$

$$a = m_x - b.m_y$$

$$s^2 = \sum (x_i - a - y_i.b)^2 / (N-2)$$

Cas II : cas général linéaire

$$b = \frac{\sum w_i (\sum w_i x_i y_i) - (\sum w_i x_i)(\sum w_i y_i)}{\sum w_i (\sum w_i y_i^2) - (\sum w_i y_i)^2}$$

$$a = \frac{\sum w_i x_i - b \sum w_i y_i}{\sum w_i}$$

$$t^2 = \frac{\sum w_i (x_i - a - b y_i)^2}{N-2}$$

Cas III : cas particulier linéaire

$$b = \frac{\sum x_i \cdot y_i}{\sum y_i^2}$$

$a = 0$ (par définition)

$$t^2 = \frac{\sum w_i (x_i - b \cdot y_i)^2}{N - 1}$$

Une application intéressante du cas général II est de chercher à ajuster les coefficients c et d sur les coefficients a et b de la régression, ce qui revient à dire qu'on considère que la droite de régression D et les droites dérivées, D_{mi} , D_b et D_k forment un faisceau dont le sommet est situé sur l'axe des y (c'est-à-dire σ' ou z). On procède par itérations successives. En prenant comme premières valeurs de c et d les valeurs a_1 et b_1 issues de la régression avec $s_{x|y}$ constant, il suffit en général de deux ou trois itérations pour obtenir une convergence satisfaisante. Cette approche revient à faire l'hypothèse que le coefficient de variation de la loi conditionnelle $X|y$ est constant avec y .

4.3

Détermination des droites représentatives

Nous proposons de déterminer les paramètres des droites représentatives, D_{mi} et D_b , de la manière suivante pour chacun des cas d'application.

On note :

$k_\alpha = k_\alpha(N; \alpha = 25\%)$ valeurs du tableau I
 $k_\beta = k_\beta(N; \beta = 5\%)$ valeurs du tableau II

Cas I :

$a_{mi} = a - k_\alpha \cdot s$ $b_{mi} = b$
 $a_b = a - k_\beta \cdot s$ $b_b = b$

Cas II avec $a \equiv c, b \equiv d$:

$a_{mi} = a \cdot [1 - k_\alpha \cdot t]$ $b_{mi} = b \cdot [1 - k_\alpha \cdot t]$
 $a_b = a \cdot [1 - k_\beta \cdot t]$ $b_b = b \cdot [1 - k_\beta \cdot t]$

Cas III :

$a_{mi} = 0$ $b_{mi} = b \cdot [1 - k_\alpha \cdot t]$
 $a_b = 0$ $b_b = b \cdot [1 - k_\beta \cdot t]$

Les figures 1 à 3 illustrent les droites ainsi déterminées.

La détermination de la droite caractéristique D_k se fait à partir du coefficient de réduction d'écart type k_v de manière analogue à celle de la variable scalaire unique :

$$a_k = a_{mi} - (a_{mi} - a_b) / k_v$$

$$b_k = b_{mi} - (b_{mi} - b_b) / k_v$$

Par la procédure ainsi proposée nous avons cherché à garantir que les résistances utilisées dans le modèle de calcul ne seront pas mises en défaut au risque de 5%, conformément à l'objectif général des valeurs caractéristiques. Dans le cas particulier de l'écart type constant, la procédure revient à linéariser la limite basse d'un intervalle de confiance de X ($= \tau$ ou R) qui, considérée comme fonction de y ($= \sigma'$ ou z), est théoriquement hyperbolique.

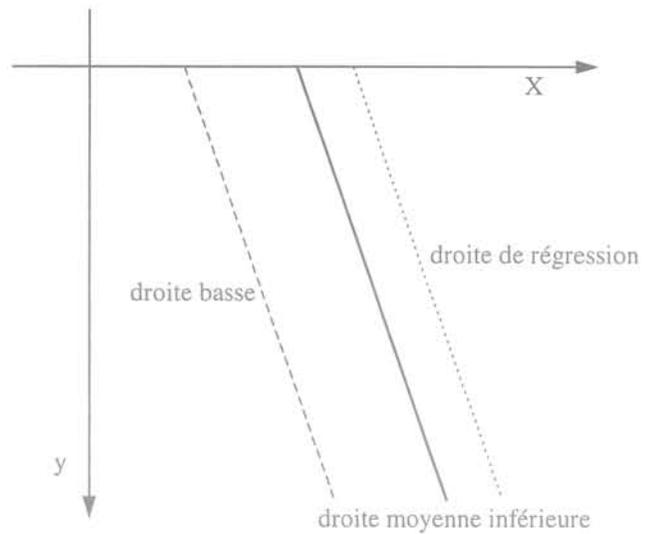


FIG. 1 (Cas I) : Droite de régression, droites moyenne inférieure et basse. Regression line, low average and lower lines.

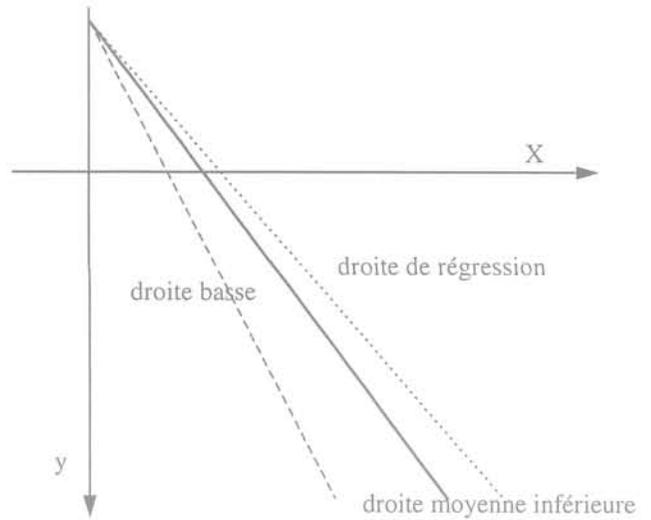


FIG. 2 (Cas II avec $a \equiv c, b \equiv d$) : Droite de régression. Droites moyenne inférieure et basse. Regression line. Low average and lower lines.

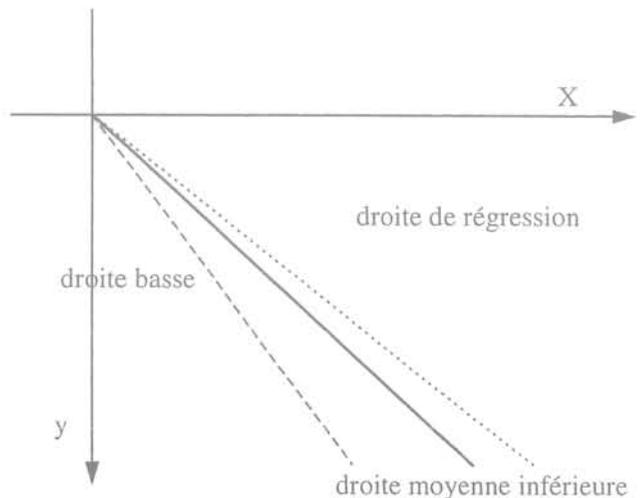


FIG. 3 (Cas III) : Droite de régression. Droites moyenne inférieure et basse. Regression line. Low average and lower lines.

Conclusion

La méthode que nous avons élaborée dans le cadre des *Recommandations* est une proposition permettant d'appliquer les principes de l'Eurocode 7 « Géotechnique ». En prévoyant une démarche en deux étapes, elle permet de distinguer clairement ce qui relève uniquement du sol et ce qui est lié à l'ouvrage.

On trouvera dans les *Recommandations* des conseils d'application à divers paramètres tels que la résistance au cisaillement des sols, la pression limite pressiométrique ou la résistance de pointe pénétrométrique. Cette démarche peut d'ailleurs être suivie dans son esprit, que l'on fasse appel ou non aux méthodes de traitement statistique pour déterminer les valeurs représentatives de la résistance caractérisant le sol : valeurs moyenne inférieure et basse, ou prenant en compte en outre la sollicitation de l'ouvrage : valeur caractéristique.

La méthode de traitement statistique proposée respecte l'objectif du risque de 5 % indiqué par les Eurocodes pour les valeurs caractéristiques des résistances. Par rapport à la théorie statistique et probabiliste, elle comporte un certain nombre de simplifications, notamment concernant la réduction d'écart type, qui semble adaptée à la pratique géotechnique.

Certains aspects doivent faire l'objet de développements ultérieurs. Ainsi en est-il par exemple de la dispersion liée aux essais, négligée dans la méthode, de la réduction de variance et de l'utilisation des distances d'auto-corrélation en liaison avec les différents types de mécanismes de ruine, du traitement de multicouches.

L'aspect quantitatif ne manquera pas de susciter des critiques, en raison notamment du petit nombre de données généralement disponibles dans la pratique courante géotechnique. La méthode illustre en tout cas clairement le lien entre la quantité d'informations et la réduction des incertitudes et elle peut aider à faire prendre conscience de la nécessité d'étoffer les connaissances géotechniques.

Nous pensons qu'il est préférable de traiter le problème de la dispersion des données géotechnique en référence à une méthode explicite, en l'adaptant si nécessaire aux particularités du cas à traiter, plutôt que de laisser à chacun le soin de régler le problème selon ses habitudes.

Nous espérons donc que la méthode proposée constituera un premier pas pour faire évoluer la situation actuelle et susciter chez les spécialistes en statistiques un effort d'adaptation des outils théoriques aux besoins des ingénieurs géotechniciens.

REMERCIEMENTS

Cette étude a été financée par le ministère de l'Équipement, du Logement et des Transports dans le cadre du programme de développement des méthodes semi-probabilistes aux états-limites pour les ouvrages en site aquatique.

Les auteurs remercient MM. les Professeurs Vidal Cohen et Jean-Louis Favre pour l'aide apportée à la mise au point de cette méthode.

Bibliographie

Eurocode 0, EN 1990, draft juin 2000.
 ENV 1991-1, AFNOR.
 Eurocode 7, EN 1997, draft PT version « e », 2000.
 ENV 1997-1, AFNOR.
 Ang A.H.S., Tang W.H. (1975) – *Probability concepts in engineering planning and*

design. John Wiley & Sons.
 CETMEF (2000) – *Recommandations pour le calcul aux états-limites des ouvrages en site aquatique*. CD Rom ROSA 2000.
 Magnan J.-P. (1982) – *Les méthodes statistiques et probabilistes en mécanique*

des sols. Presses des Ponts et Chaussées.
 Linglin S. (1998) – « Élaboration d'une méthode statistique pour la détermination des valeurs caractéristiques des propriétés des sols ». Rapport de stage. CETMEF, Compiègne.

Deux applications de l'analyse factorielle à des données géotechniques

**P. POUGET
N. CADOUX***

Laboratoire régional
des ponts et chaussées
8-10, rue Bernard-Palissy
BP 11
63014 Clermont-Ferrand
* Stagiaire CUST-3^e A

J.-P. MAGNAN

Laboratoire central
des ponts et chaussées
58, bd Lefebvre
75732 Paris Cedex 15

Résumé

La méthode d'analyse factorielle a été appliquée à l'analyse des données géotechniques provenant de deux sites : une zone de sols argileux compressibles où un grand nombre d'essais œdométriques a été réalisé dans un sondage et une zone de sols surconsolidés et peu déformables dans laquelle se trouve une vallée fossile remplie de sédiments argileux mous. Le premier site a été analysé pour trouver les relations existant entre les paramètres et définir les frontières des couches. Le second site a été analysé pour retrouver la géométrie de la vallée fossile à partir de mesures au pénétromètre statique. Ces deux exemples illustrent les possibilités de l'analyse factorielle, qui a fourni le résultat cherché dans les deux cas. Ils ont aussi mis en évidence quelques sources d'erreurs possibles.

Mots clés : analyse statistique, analyse factorielle, site compressible, vallée fossile, exemples.

Two applications of factor analysis to geotechnical data

Abstract

The factor analysis method has been applied to the analysis of geotechnical data from two sites : a zone of compressible clayey soils where a large number of oedometer compressibility tests were made and a zone of overconsolidated stiff soils where a hidden valley filled up with soft clayey sediments was found. The first site was analysed with the intention to find relationships between the soil parameters and to define the layer boundaries. The second site was analysed in order to find the geometry of the hidden valley from cone penetration tests. The two examples show the possibilities offered by factor analysis, which gave the intended results. They displayed possible sources of errors, too.

Key words : statistical analysis, factor analysis, compressible site, hidden valley, examples.

NDLE : Les discussions sur cet article sont acceptées jusqu'au 1^{er} août 2001.

Introduction

Malgré les nombreux travaux qui leur ont été consacrés, les méthodes d'analyse de données restent pour les géotechniciens des outils mal connus et utilisés de façon très marginale, sous forme notamment de calculs statistiques élémentaires tels que les calculs de moyenne et de variance et les études de régression linéaire, où la statistique porte sur la variabilité des données autour d'une droite définie par la méthode des moindres carrés.

L'analyse factorielle, utilisée dès les années 1970 pour des applications de géotechnique (Favre, 1980), offre des possibilités complémentaires pour déterminer les relations existant à l'intérieur d'un ensemble de paramètres. Les deux études présentées dans cet article ont été réalisées sur deux sites où de très nombreuses données étaient disponibles :

– la première a été effectuée sur le site expérimental de Cubzac-les-Ponts (sols argileux compressibles), déjà analysé par Baghery et Magnan (1983) par des méthodes d'analyse de corrélations et de régression linéaire. Il s'agissait de voir si l'utilisation de l'analyse factorielle, sous la forme de l'analyse en composantes principales, apportait des résultats différents des analyses déjà effectuées ;

– la seconde étude portait sur la recherche automatique d'ensembles de terrains homogènes dans la reconnaissance géotechnique d'un site. Cette seconde application a été réalisée sur les données d'une étude géotechnique exécutée quelques années plus tôt et qui avait mis en évidence l'existence d'une vallée fossile remplie de sols compressibles dans un environnement de sols surconsolidés et résistants.

Ces deux applications sont décrites successivement dans cet article, après un rappel succinct des principes de l'analyse factorielle de données.

La méthode d'analyse factorielle (rappels)

L'analyse factorielle cherche à représenter la variabilité d'un ensemble de n variables aléatoires $\{X_i\}_{i=1, n}$ par un plus petit nombre p d'autres variables $\{Y_j\}_{j=1, p}$ appelées catégories ou facteurs. Ce travail se fait à partir de m observations des variables X_i , soit $\{x_{ik}\}_{i=1, n; k=1, m}$.

Pour atteindre cet objectif, on construit par étapes une suite de variables Y_j qui représentent chacune l'essentiel de la variabilité restante des données. Ce travail peut être représenté comme la recherche des axes principaux des m points observés dans l'espace à n dimensions d'axes $\{X_1, X_2, \dots, X_n\}$. Pour $n = 3$, cette procédure est représentée sur la figure 1 :

– la première phase est la recherche de l'axe Y_1 , qui passe par le centre de gravité G des points et minimise la somme des carrés des distances des points $\{x_{ij}\}$ à ce premier axe ;

– la deuxième phase est la recherche de l'axe principal d'inertie Y_2 dans le plan passant par G et perpendiculaire à l'axe Y_1 . La procédure est la même que pour Y_1 : on cherche l'orientation de l'axe Y_2 qui minimise la somme des carrés des distances des points au plan $Y_1 Y_2$;

– le troisième axe est perpendiculaire aux deux premiers.

Quand il existe plus de trois variables X_i , la procédure est la même avec un axe principal Y_1 et des axes successifs définis par minimisation par rapport à des hyperplans.

L'expression mathématique de la procédure de recherche des axes $\{Y_j\}$ est décrite dans divers ouvrages (Bertier et Bourroche, 1977 ; Dagnelie, 1977 ; Lebart et al., 1982). Elle est brièvement rappelée ci-après.

La relation cherchée entre les n variables $\{X_i\}$ et les p variables $\{Y_j\}$ peut s'écrire sous la forme générale :

$$\begin{pmatrix} Y_1 \\ \dots \\ Y_p \end{pmatrix} = \begin{pmatrix} u_{11} & \dots & u_{1n} \\ \dots & u_{ji} & \dots \\ u_{p1} & \dots & u_{pn} \end{pmatrix} \begin{pmatrix} X_1 \\ \dots \\ X_n \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \dots \\ \epsilon_p \end{pmatrix}$$

$$\{Y\} = \{U\} \{X\} + \{\epsilon\}$$

où $\{\epsilon\}$ est la matrice des résidus (écarts entre les valeurs réelles et les valeurs calculées) de la relation. La matrice $\{U\}$, pour sa part, est la matrice de passage du repère orthonormé des variables $\{X_i\}$ au repère orthonormé des nouvelles variables $\{Y_j\}$. Cette matrice contient les coordonnées en ligne des vecteurs unitaires des axes $\{Y_j\}$ par rapport à la base des vecteurs unitaires des axes $\{X_i\}$. Compte tenu de son origine, la matrice $\{U\}$ a les propriétés suivantes :

$$\sum_i u_{ji}^2 = 1 \quad (\text{vecteurs unitaires})$$

et

$$\sum_i u_{ji} u_{ki} = 0 \text{ pour } j \neq k \quad (\text{vecteurs orthogonaux par deux}).$$

Par conséquent la matrice $\{U\}$ est telle que $\{U\} \{U\}^t = \{1\}$ (matrice unité) et $\{U\}^t$ est l'inverse de $\{U\}$.

Les points P_k observés (m valeurs des $\{X_i\}$, soit $\{x_{ik}\}$) ont pour coordonnées dans le nouveau repère des $\{Y_j\}$:

$$\{y_{jk}\} = \{u_{ji}\} \{x_{ik}\}.$$

Dans ces matrices, les coordonnées $\{x_{ik}\}$ des points P_k sont rangées en colonnes et les $\{y_{jk}\}$ correspondants en lignes.

Si l'origine des axes est placée au centre de gravité G des points P_k , la distance entre le point G et la projection H_{k1} d'un point P_k sur l'axe portant la variable Y_1 (Fig. 2) vaut :

$$\{y_{1k}\} = \{u_{1i}\} \{x_{ik}\}.$$

Le carré de cette distance est égal à :

$$\begin{aligned} |GH_{k1}|^2 &= \{u_{1i}\} \{x_{ik}\} \{x_{ik}\}^t \{u_{1i}\}^t \\ &= \sum_i u_{1i}^2 x_{ik}^2 \end{aligned}$$

La somme Σ^2 des carrés des distances GH_{k1} pour tous les points P_k est donnée par le produit de matrices

$$\Sigma^2 = \{y_{1k}\} \{y_{1k}\}^t = \{u_{1i}\} \{x_{ik}\} \{x_{ik}\}^t \{u_{1i}\}^t.$$

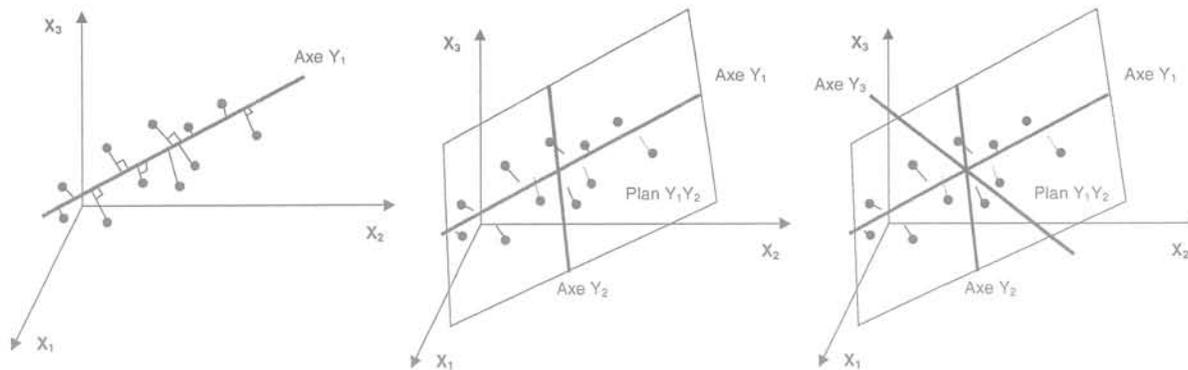


FIG. 1 Procédure de recherche des axes principaux d'inertie.
Procedure for finding the principal inertia axes.

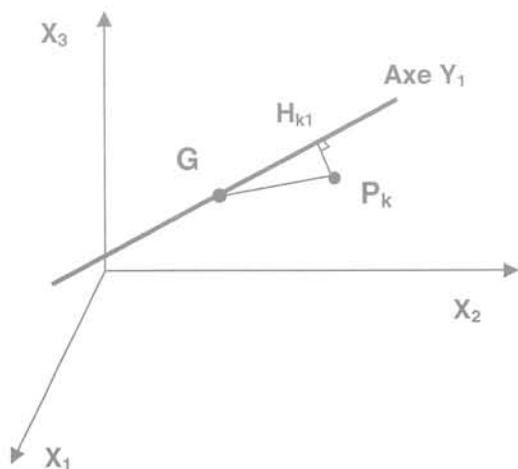


FIG. 2 Distance du point P_k à l'axe Y_1 .
Distance from point P_k to axis Y_1 .

La somme des carrés des distances des points P_k à l'axe GY_1 est égale à :

$$\sum_k H_{k1} P_k^2 = \sum_k G P_k^2 - \sum_k G H_{k1}^2.$$

Comme $\sum_k G P_k^2$ est constant pour un ensemble de points P_k donné, minimiser $\sum_k H_{k1} P_k^2$ revient à maximiser $\sum_k G H_{k1}^2 = \Sigma^2$.

Pour maximiser Σ^2 par le choix optimal de l'axe GY_1 , on procède de la façon suivante.

On recherche le vecteur unitaire de GY_1 de composantes $\{u_{ij}\}$ sur la base orthonormée des $\{X_i\}$ qui maximise le produit $\{u_{ij}\} \{x_{ik}\} \{x_{ik}\}^t \{u_{ij}\}^t$.

La matrice $\{A\} = \{x_{ik}\} \{x_{ik}\}^t$ est symétrique et définie positive, donc diagonalisable, avec des valeurs propres positives. Soit $\{B\}$ la matrice des coordonnées en colonnes des vecteurs propres unitaires de $\{A\}$ dans le repère des $\{X_i\}$ et $\{\Lambda\}$ la matrice de ses valeurs propres :

$$\{B\} = \{b_{\alpha i}\} \quad \{\Lambda\} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_\alpha & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}.$$

On peut écrire :

$$\{A\} \{B\} = \{B\} \{\Lambda\}$$

d'où :

$$\{A\} = \{B\} \{\Lambda\} \{B\}^t \text{ car } \{B\} \{B\}^t = \{1\}.$$

Si l'on écrit les coordonnées des vecteurs propres unitaires $\{b_{\alpha i}\}$ dans leur repère (qui est celui des Y_j cherchés), on obtient une nouvelle matrice :

$$\{C\} = \{c_{\alpha i}\} = \{B\}^t \{U\}^t,$$

et l'on peut écrire que

$$\{U\} \{X\} \{X\}^t \{U\}^t = \{U\} \{B\} \{\Lambda\} \{B\}^t \{U\}^t = \{C\}^t \{\Lambda\} \{C\}.$$

La condition $\{U\}^t \{U\} = \{1\}$ devient alors

$$\{C\}^t \{C\} = \{U\} \{B\} \{B\}^t \{U\}^t = \{U\} \{U\}^t = \{1\}.$$

La maximisation de $\{U\} \{X\} \{X\}^t \{U\}^t$ pour la première direction propre correspond donc à celle de l'expression :

$$\sum_\alpha \lambda_\alpha c_{\alpha 1}^2$$

sous la condition $\sum_\alpha c_{\alpha 1}^2 = 1$.

En effet, les $c_{\alpha 1}$ sont les coordonnées des vecteurs propres unitaires sur la première direction propre. On peut remplir les conditions imposées en choisissant la direction propre associée à la plus grande valeur propre λ_α . On a ainsi déterminé l'orientation du premier facteur Y_1 .

La procédure se continue en utilisant successivement les valeurs propres et directions propres dans l'ordre décroissant des valeurs propres.

Si l'on détermine également les valeurs propres μ_α de la matrice $X'X$ et les vecteurs propres unitaires v_α correspondants, on démontre que :

$$\lambda_\alpha = \mu_\alpha$$

$$X v_\alpha = \sqrt{\lambda_\alpha} u_\alpha,$$

d'où

$$X = \sum_{\alpha=1}^{p} \sqrt{\lambda_\alpha} v_\alpha u_\alpha^t.$$

Le taux d'inertie ou part de variance expliquée

$$\tau_p = \sum_{\alpha=1}^p \lambda_\alpha / \sum_{\alpha=1}^n \lambda_\alpha$$

permet d'apprécier la qualité de la représentation des X_i par les Y_j .

3

Applications, généralités

Les analyses en composantes principales présentées dans cet article ont été réalisées au moyen du logiciel STATGRAPHICS. Le calcul fournit :

- un vecteur contenant les valeurs propres, qui expriment la variance de chaque composante principale ;
- une matrice contenant les vecteurs propres, qui sont les pondérations à appliquer aux variables pour obtenir les facteurs ;
- une matrice contenant les corrélations entre les variables de départ et les facteurs (utile à l'interprétation des facteurs).

La projection des individus et des caractères est réalisée sur l'espace principal considéré. En général, comme nous avons pu le vérifier, le plan défini par les deux premiers axes factoriels comprend un pourcentage important de l'information.

L'interprétation des résultats et de leurs représentations graphiques conduit aux informations suivantes :

- l'origine des axes G est le centre de gravité du nuage : il représente la tendance moyenne des individus. Les individus situés loin du centre présentent un écart important par rapport à cette moyenne et influencent beaucoup les statistiques ;
- l'inertie d'un axe, représentée par la valeur propre associée à son vecteur directeur, témoigne de la représentativité de cet axe par rapport au nuage des points : plus l'inertie du plan est forte, plus la projection du nuage sur ce plan est réaliste avec un minimum de perte d'information ;
- les coordonnées des points sur les axes factoriels sont les coefficients de corrélation des individus qu'ils représentent avec ces axes. La valeur du produit scalaire des vecteurs joignant deux points à l'origine est le coefficient de corrélation entre les deux variables correspondantes : des vecteurs voisins indiquent des variables fortement corrélées ;
- la proximité, dans le plan, d'individus et de variables n'a pas de signification. Un lien entre ces paramètres nécessite une interprétation de la signification des facteurs.

4

Analyse de la compressibilité de l'argile molle de Cubzac

Lors de l'étude géotechnique des sols du site expérimental des Laboratoires des Ponts et Chaussées à Cubzac-les-Ponts, un sondage carotté de 20 cm de diamètre et 9 m de longueur a été entièrement consacré à l'étude des variations des résultats des essais œdométriques en fonction de la profondeur (Sondage A44, Remblai D, Haghgou, 1983) : on dispose ainsi de 168 résultats d'essais œdométriques, à raison de deux essais au même niveau tous les 6,5 cm de profondeur, avec des interruptions périodiques dues au carottage (remaniement des extrémités des carottes).

L'étude générale du site a montré que la coupe des sols dans la zone du sondage A44 est la suivante :

- de 0 à 1,3 m, couche superficielle d'argile molle surconsolidée, correspondant aux fluctuations annuelles de la nappe phréatique ;
- de 1,3 à 2,5 m, vase limono-argileuse moyennement organique, très compressible et légèrement surconsolidée ;
- de 2,5 à 9 m, vase argilo-limoneuse faiblement organique, très compressible et légèrement surconsolidée, plus dense de 2,5 à 5,2 m ($w = 65\%$, $\gamma = 16 \text{ kN/m}^3$) que de 5,2 à 9 m ($w = 80\%$, $\gamma = 14,8 \text{ kN/m}^3$).

Les données géotechniques du site expérimental de Cubzac-les-Ponts ont été analysées par Baghery et Magnan (1983) et par Haghgou (1983), en utilisant les méthodes classiques de la statistique et notamment l'analyse de régression linéaire. L'étude présentée ici cherchait à savoir si l'analyse factorielle pouvait apporter des informations complémentaires par rapport à l'analyse de régression linéaire.

Le travail d'analyse en composantes principales a été effectué sur un tableau comportant 168 lignes et 13 colonnes :

- colonne 1 : i , numéro de l'éprouvette ;
- colonne 2 : z , profondeur (m) ;
- colonne 3 : w , teneur en eau naturelle (%) ;
- colonne 4 : γ , poids volumique (kN/m^3) ;
- colonne 5 : γ_{dr} , poids volumique du sol sec (kN/m^3) ;
- colonne 6 : e_o , indice des vides en place ;
- colonne 7 : C_{sv} , indice de gonflement ;
- colonne 8 : C_{cr} , indice de compression ;
- colonne 9 : σ'_p , pression de préconsolidation (kPa) ;
- colonne 10 : c_v , coefficient de consolidation verticale (m^2/s) ;
- colonne 11 : k_v , coefficient de perméabilité verticale (m/s) ;
- colonne 12 : e_k , paramètre de la relation entre e et k_v ($e = e_k + C_k \lg k_v$) ;
- colonne 13 : C_{kr} , paramètre de la relation entre e et k_v .

L'analyse en composantes principales réalisée sur l'ensemble des données a montré que les deux premiers axes factoriels représentaient toujours de 70 à 100 % de l'information initiale. L'exposé des résultats sera pour cette raison limité à ces deux axes.

4.1

Analyse au niveau des paramètres

L'analyse réalisée sur l'ensemble des mesures entre 1,085 m et 9,075 m de profondeur montre un plan factoriel (Fig. 3) dont l'inertie totale est de 72,6 % (59,5 % pour le premier axe, 13,1 % pour le deuxième). Sur ce plan sont représentés simultanément les points correspondant aux individus et les vecteurs correspondant aux variables. Quatre groupes de vecteurs rapprochés les uns des autres peuvent être observés. Ils sont associés respectivement aux variables suivantes :

- γ et γ_{dr} ,
- w , e_o , C_{sv} , e_k et C_{kr} ,
- c_v et k_v ,
- z et σ'_p .

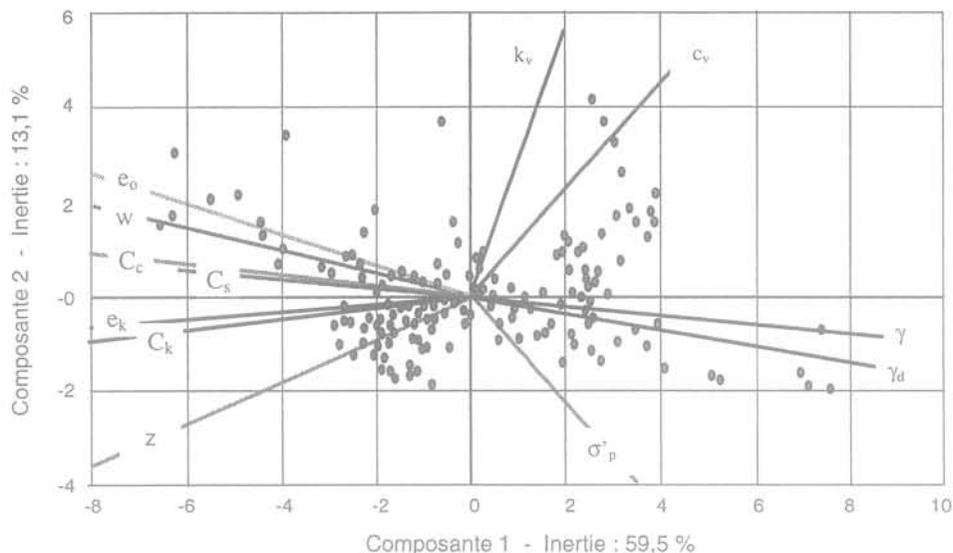


FIG. 3 Plan factoriel 1-2 de l'analyse en composantes principales (ensemble des données).
Plot of factors 1 and 2 from factor analysis (all data).

Le rapprochement des vecteurs ou leur opposition montrent des relations marquées entre les paramètres, correspondant à des corrélations positives dans le cas de deux vecteurs proches l'un de l'autre et à des corrélations négatives en cas de vecteurs opposés.

Dans les quatre groupes ainsi définis, on retrouve des paramètres présentant des relations bien connues, comme par exemple, pour les vecteurs proches, γ et γ_d , avec la relation :

$$\gamma = \gamma_d (1 + w)$$

qui fait intervenir w , dont le vecteur représentatif est pratiquement opposé à celui de γ_d .

Pour d'autres paramètres, la relation paraît simplement logique : ainsi, l'indice de compression C_c matérialise la « capacité du sol à être comprimé », qui augmente avec l'indice des vides e_0 . L'indice des vides varie comme la teneur en eau, et en sens inverse du poids volumique...

Les informations tirées de l'analyse en composantes principales doivent tenir compte des connaissances en mécanique des sols dans la manière de s'intéresser à certains paramètres. Ainsi, on peut observer que les variables w et e_0 sont symétriques des variables γ et γ_d avec une représentativité de l'axe principal de 92,7 % (Fig. 4). Ces résultats vont dans le sens des corrélations existant entre ces paramètres (tableau I).

L'interprétation de la position des vecteurs représentant les variables dans les plans factoriels permet donc de mettre en évidence des relations entre les caractéristiques géotechniques, qui découlent de la formulation des propriétés des sols ou de phénomènes liés à leur formation. Toutefois, pour des types de sols aussi bien étudiés que les argiles molles, cette analyse ne fait que retrouver des relations et corrélations bien connues en mécanique des sols.

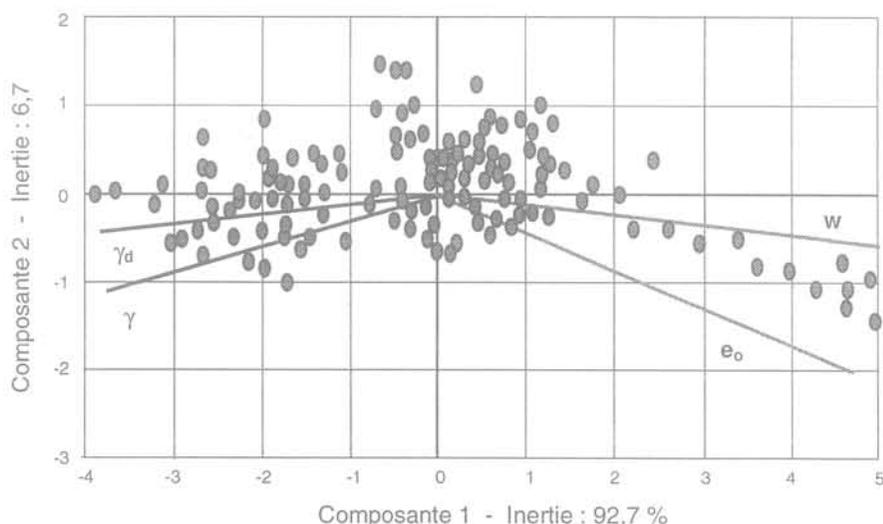


FIG. 4 Plan factoriel 1-2 de l'analyse en composantes principales des paramètres w , e_0 , γ , γ_d .
Plot of factors 1 and 2 from factor analysis (w , e_0 , γ , γ_d).

TABEAU I Coefficients de corrélation linéaire entre les paramètres w , e_o , γ et γ_d .
Correlation coefficients of soil parameters w , e_o , γ et γ_d .

	w	e_o	γ	γ_d
w	1	0,95	-0,90	-0,96
e_o		1	-0,76	-0,86
γ			1	0,97
γ_d				1

4.2

Analyse au niveau des individus

La représentation graphique de l'analyse en composantes principales peut être complétée par la numérotation des différents individus analysés. On peut ainsi définir des groupes représentant des individus proches les uns des autres. La détermination de ces groupes se fait « à la main » sur le graphique et la question de la séparation des zones reste généralement marginale (quelques points des zones de transition peuvent prêter à ambiguïté). La figure 5 montre le plan factoriel de l'ensemble des données. La numérotation des individus permet de définir quatre groupes d'individus, qui correspondent à des couches superposées :

- couche 1 (point 1 à 7) : de 1,085 à 1,255 m ;
- couche 2 (points 15 à 25) : de 1,76 à 2,37 m ;
- couche 3 (points 33 à 84) : de 2,63 à 5,015 m ;

- couche 4 (points 85 à 168) : de 5,28 à 9,075 m.

On note le nombre relativement faible des points non comptabilisés dans la définition de ces couches : l'imprécision correspondante de la délimitation des couches est de l'ordre de 30 cm.

L'analyse réalisée sur des plans factoriels limités à certains paramètres conduit à la définition de groupes d'amplitudes très proches de celles obtenues au niveau de l'ensemble des données. A titre d'exemple, le plan factoriel de l'analyse en composantes principales utilisant les paramètres w , e_o , γ , γ_d (Fig. 6) montre une répartition des points différente de celle obtenue sur la figure 4 mais l'ordre et l'épaisseur des couches qui en résultent sont pratiquement identiques. On note que le plan factoriel obtenu dans ce cas possède une inertie de 99,4 %, ce qui est dû au fait que, pour un sol saturé, il n'y a que deux paramètres indépendants parmi les quatre : en effet, il existe deux relations entre w , e_o , γ , γ_d :

$$\gamma = \gamma_d (1 + w)$$

$$e_o = \frac{\gamma - \gamma_d}{\gamma_w + \gamma_d - \gamma}$$

La figure 7 montre le plan factoriel 1-2 obtenu pour une autre combinaison de variables : z , w , C_c et σ'_p . Cette fois encore la disposition des points sur le graphique est différente, mais les couches obtenues sont les mêmes que dans les autres diagrammes.

Comme l'indique le tableau II, la prise en compte de tout ou partie des paramètres mesurés conduit à définir quatre couches de sol quasiment identiques à quelques centimètres près, c'est-à-dire à un ou deux niveaux de mesure près. Ces couches sont très proches de celles définies par la coupe géologique du sondage.

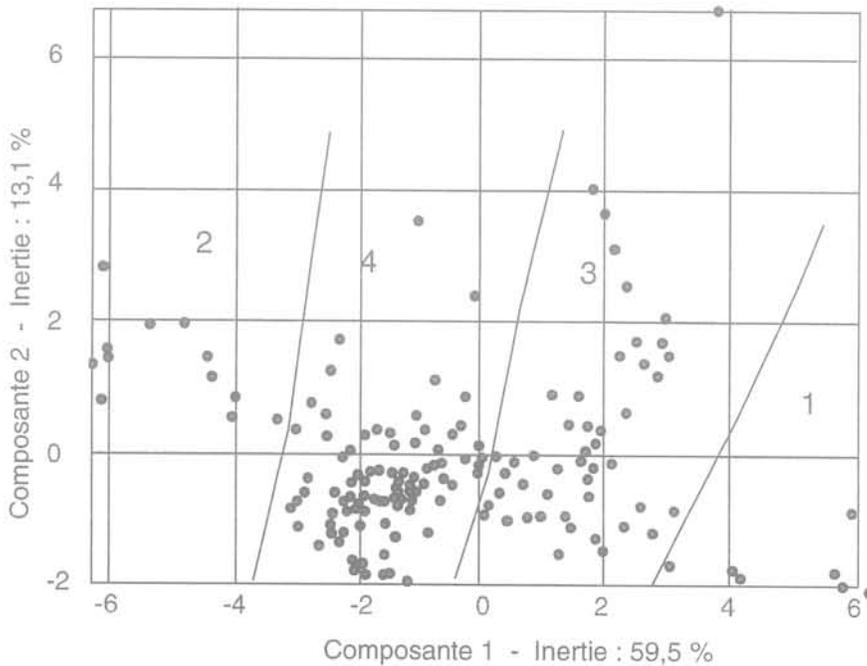


FIG. 5 Plan factoriel 1-2 de l'analyse en composantes principales avec l'ensemble des paramètres, pour identification des individus.
Plot of factors 1 and 2 from factor analysis, including all soil parameters, for analysis at the level of individuals.

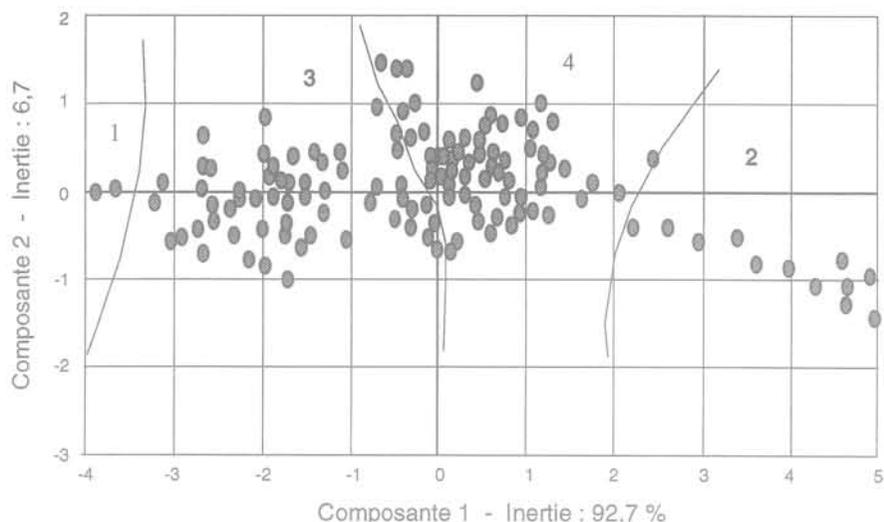


FIG. 6 Plan factoriel 1-2 de l'analyse en composantes principales avec les paramètres w , e_v , γ et γ_d .
Plot of factors 1 and 2 from factor analysis (soil parameters w , e_v , γ and γ_d).

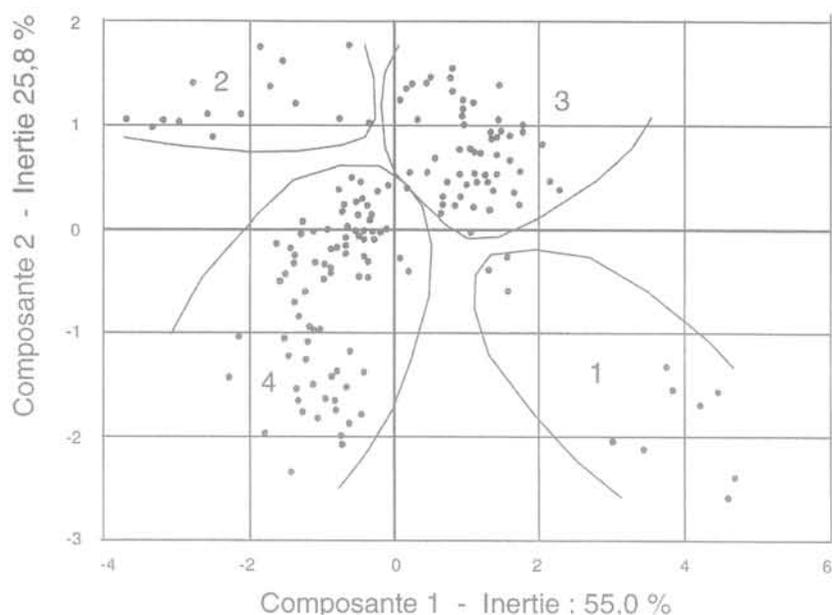


FIG. 7 Plan factoriel 1-2 de l'analyse en composantes principales avec les paramètres z , w , C_c et σ'_p .
Plot of factors 1 and 2 from factor analysis (soil parameters z , w , C_c and σ'_p).

On observe dans le tableau III que la représentativité des deux premiers axes factoriels est d'autant meilleure que le nombre de variables indépendantes (non liées par une relation mathématique) est faible. Les exceptions correspondent à des groupes de paramètres fortement corrélés (indice de compression et indice des vides, par exemple).

On remarque que, si l'analyse en composantes principales peut être utilisée pour déterminer les différentes couches de sol, elle ne permet pas de leur attribuer des

caractéristiques géotechniques. Une étude statistique au niveau de chaque couche doit donc compléter les résultats. Cette étude statistique, menée sur l'ensemble des paramètres de chacune des couches mises en évidence, montre des coefficients de variation des paramètres par couche légèrement inférieurs ou en tout cas du même ordre de grandeur que ceux mesurés sur l'ensemble des paramètres (une seule couche). A titre d'exemple, le tableau III donne les valeurs relatives au poids volumique γ du sol.

TABEAU II Zones déterminées par différentes analyses en composantes principales.
Zones of similar soils (layers) obtained from various factor analyses.

z	Variables intégrées dans l'analyse											Variables non liées	Inertie du plan factoriel 1-2	Définition des couches				
	w	γ	γ_d	e_o	C_s	C_c	σ_p	c_v	k_v	e_k	C_k			Zone 1 (m)	Zone 2 (m)	Zone 3 (m)	Zone 4 (m)	
x	x	x	x	x	x	x	x	x	x	x	x	x	8	72,16 %	1,085 à 1,255	1,76 à 2,37	2,63 à 5,015	5,28 à 9,075
	x	x	x	x									2	99,4 %	1,085 à 1,255	1,76 à 2,435	2,63 à 5,34	5,40 à 9,075
	x	x	x	x	x	x							4	93,8 %	1,085 à 1,385	1,76 à 2,435	2,63 à 5,015	5,28 à 9,075
x	x					x	x						4	80,8 %	1,085 à 1,385	1,76 à 2,435	2,50 à 5,015	5,28 à 9,075
x						x	x						3	82,2 %	1,085 à 1,385	1,76 à 2,435	2,50 à 5,015	5,28 à 9,075
	x						x	x					3	98,3 %	1,085 à 1,385	1,76 à 2,37	2,435 à 5,34	5,40 à 9,075
						x	x	x	x	x	x		6	73,4 %	1,085 à 1,385	1,76 à 2,435	2,50 à 5,015	5,28 à 9,075
x	x		x					x	x	x	x		6	79,3 %	1,085 à 1,255	1,76 à 2,37	2,50 à 5,28	5,34 à 9,075

TABEAU III Coefficients de variation du poids volumique par couche et global.
Coefficients of variation of unit weight for each layer and for the entire ground.

Couche	Coefficient de variation
1	6,64 %
2	2,82 %
3	3,25 %
4	2,25 %
Tout le sol	5,65 %

5

Application au zonage géotechnique

5.1

Introduction

Dans le cadre de l'étude géotechnique du site destiné à la construction d'un parking pour avions dans un aéroport, un grand nombre de sondages au pénétromètre statique a été réalisé dans une zone d'environ 1 km de longueur et 0,6 km de largeur (Fig. 8).

L'étude géotechnique a montré qu'il existe sur ce site une vallée fossile remplie de sédiments de faible résistance (sables et argiles) et dont la profondeur dépasse 20 m. La position de la vallée a été déduite de l'analyse manuelle des courbes de variation de la résistance de cône q_c en fonction de la profondeur. Les sols peu résistants ont été définis comme ayant une résis-

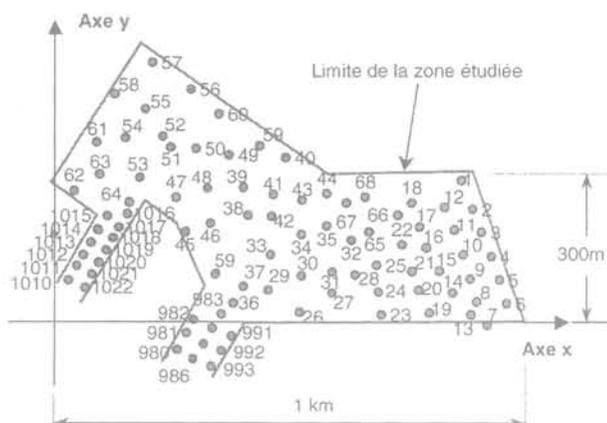


FIG. 8 Disposition des sondages pénétrométriques sur le site du projet.
Layout of the penetration tests (CPT) at the construction site.

tance de cône inférieure à 2 MPa. La vallée est représentée sur la figure 9.

Les formations rencontrées à partir de la surface du sol sont successivement :

- de la terre végétale ou du remblai sur une épaisseur de 0,3 à 1,5 mètre ;
- de l'argile grise dont l'épaisseur peut atteindre 22 mètres. On peut rencontrer dans cette formation des argiles vasardes d'épaisseur variable ;
- de l'argile jaunâtre très consistante à dure, qui constitue le substratum pour le projet étudié.

La couche d'argile grise est absente au dehors de la vallée et a une épaisseur croissante quand on passe du bord au centre de la vallée. Des intercalations de sables et graviers sont rencontrées assez rarement jusqu'à une profondeur de 16 mètres ; elles sont plus nombreuses entre 16 et 22 mètres de profondeur. Ces couches ont une épaisseur et une extension latérale très irrégulières.

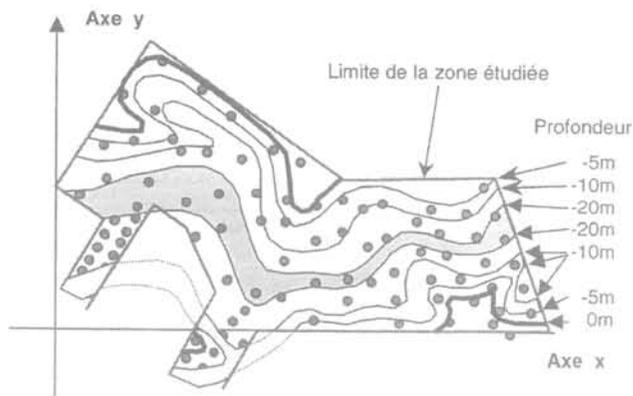


FIG. 9 Position de la vallée fossile sur le site étudié.
Location of the hidden valley across the site.

La configuration de ce site a paru favorable pour essayer de délimiter l'étendue de la vallée fossile par une technique automatique d'analyse de données, en appliquant les outils de l'analyse factorielle.

A cette fin, comme cela avait été fait dans l'analyse manuelle, les pénétrogrammes (variations de la résistance de cône q_c et du frottement latéral f_s en fonction de la profondeur) ont été simplifiés pour être enregistrés dans un tableau informatique. Les quatre-vingt-douze sondages ont été enregistrés à raison d'une valeur moyenne de q_c et f_s tous les 50 cm. La figure 10 montre le pénétrogramme du sondage 25.

Le tableau de données comprend aussi cinq autres paramètres :

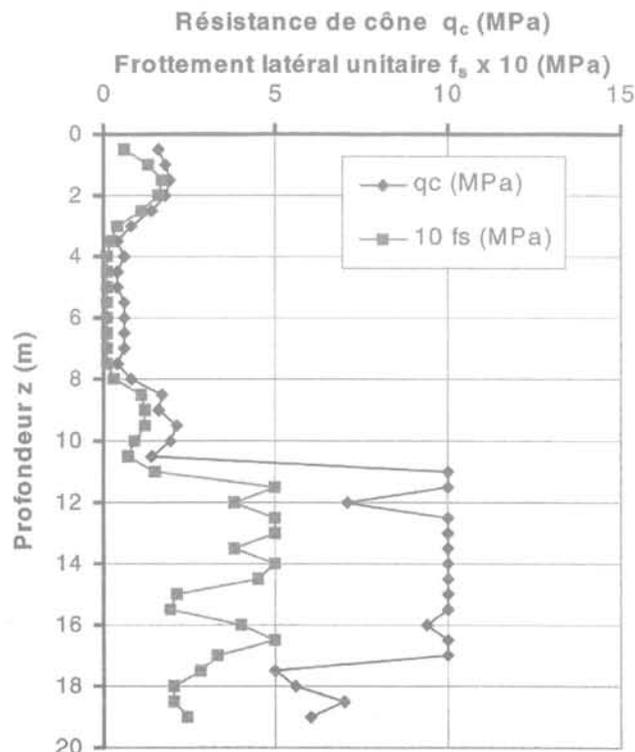


FIG. 10 Pénétrogramme du sondage 25.
Penetrogramme of CPT sounding Nr. 25.

- le numéro de l'enregistrement (les valeurs sont enregistrées dans l'ordre des profondeurs croissantes et par ordre croissant des numéros des sondages) ;
- le numéro du sondage ;
- l'abscisse du point de sondage dans le repère choisi ;
- l'ordonnée du point de sondage ;
- la profondeur de la mesure.

5.2

Étude dans la direction verticale

Chaque sondage a d'abord été analysé séparément. Dans chacun des plans factoriels ainsi élaborés (où chaque individu apparaît par sa numérotation croissante avec la profondeur), on observe des zones où ces individus sont proches les uns des autres : les familles ainsi constituées définissent des couches. Un exemple de résultat (sondage 25) est donné sur la figure 11, où l'on peut observer l'existence de quatre couches.

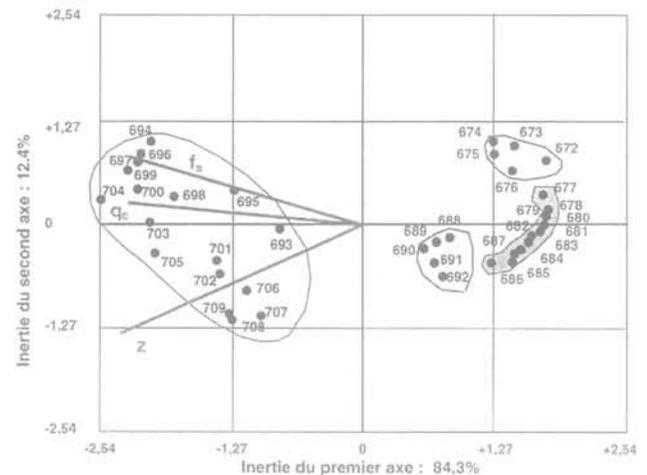


FIG. 11 Biplot factoriel de l'analyse en composantes principales des données du sondage 25 (axes d'inertie 1 et 2). Mesures 672 à 709.
Plot of the first factors of data from CPT sounding Nr. 25 (inertia axes 1 and 2). Data 672 to 709.

Trois à sept (quatre à cinq le plus souvent) zones ou couches ont ainsi été déterminées par sondage. Chacune a fait l'objet d'une analyse statistique avec calcul de la moyenne, de la variance et du coefficient de variation des variables q_c et f_s . On observe une très nette diminution du coefficient de variation par rapport au coefficient de variation mesuré sur toute la hauteur du sondage (tableau IV), ce qui confirme l'homogénéité des couches ainsi définies. On constate toutefois que ce phénomène est moins net dans les couches inférieures des sondages, où des intercalations de bancs de sable se font sentir.

Sondage	Zone	Numéros des individus	Profondeur z(m)	q_c moyenne (MPa)	q_c σ^2 (MPa ²)	q_c CV (%)	f_s moyenne (kPa)	f_s σ^2 (kPa) ²	f_s CV (%)
30	1	809 à 811	1 à 2	2,53	0,173	16,4	120	100	8,3
	2	812 à 823	2,5 à 8	1	0,075	27,3	37	300	49,8
	3	824 à 841	8,5 à 17	1,22	0,153	31,9	54	700	49,7
	4	842 à 846	17,5 à 19,5	4,34	7,448	62,9	220	24 800	71,6
	Global	809 à 846		1,88	3,855	104,6	86	10 800	120,7
65	1	2013 à 2017	0,5 à 2,5	2,9	0,755	30	192	6 600	42,4
	2	2018 à 2025	3 à 6,5	0,93	0,054	25	34	100	22
	3	2026 à 2032	7 à 10	2,5	0,273	20,9	161	800	17,3
	4	2033 à 2044	10,5 à 16	2,08	0,68	39,7	102	1 600	39,8
	5	2045 à 2052	16,5 à 20	7,35	7,003	36	346	8 000	25,9
	Global	2013 à 2052		3,08	6,646	83,8	159	14 300	75,3

5.3

Étude dans un plan horizontal

La technique d'analyse en composantes principales a été utilisée pour analyser les résultats des essais pénétrométriques dans des plans horizontaux choisis à différentes profondeurs (le fichier de données a été conçu de manière à comporter l'altitude de chaque niveau de mesure). Pour les sondages dont la longueur est inférieure à 20 mètres, les données ont été complétées jusqu'à cette profondeur par une valeur correspondant à la dernière valeur mesurée (ce sont des sondages dans des argiles raides ne posant pas de problèmes géotechniques).

L'analyse a été réalisée sur la résistance de cône q_c (qui est *a priori* le paramètre mesuré le plus représentatif pour les sols compressibles) et a consisté à établir des cartes de courbes isométriques de ce paramètre suivant la profondeur. L'intérêt de l'analyse en composantes principales est qu'elle permet des représentations factorielles où l'on peut analyser des ressemblances entre individus et les classer. Dans le cas présent, on analyse seulement deux variables q_c et f_s et toute la variabilité est représentée dans le plan factoriel 1-2. Le travail se décompose comme suit :

- à une profondeur donnée, traçage du plan factoriel calculé à l'aide des paramètres mesurés. Décomposition du plan en dix zones ou classes d'appartenance ;
- sortie du même plan factoriel en faisant apparaître cette fois le numéro de la mesure. Par comparaison avec la figure précédente, on sait alors à quelle classe appartient chaque sondage ;
- traçage, sur le plan de situation des sondages, des courbes correspondant aux limites de classes. Ce traçage s'effectue « à la main » avec interpolation entre les points du graphe. Des formes de couches insolites ont parfois été trouvées par cette démarche (liées à l'imprécision du système, d'une part, et à la présence de

niveaux de sables et graviers, d'autre part), mais, globalement, les indéterminations ont été très réduites et les cartes ainsi obtenues prêtent peu à discussions.

On dispose ainsi d'une carte pour chacune des profondeurs prises en compte, soit tous les 2,5 mètres. Les figures 12 et 13 montrent les courbes isométriques obtenues aux profondeurs de 10 et 20 mètres (la profondeur est donnée par rapport à l'altitude de référence, qui est celle de l'extrémité supérieure du sondage le plus élevé). L'analyse des courbes permet de préciser la situation et la forme de la vallée fossile ; elle donne une information sur la forme du fond de celle-ci (forme se rapprochant du V plutôt que du U).

La figure 14 compare la géométrie de la vallée à 20 m de profondeur (zone de résistances de cône q_c inférieures à 2 MPa) dans l'analyse qui vient d'être présentée et dans l'analyse manuelle effectuée lors de l'étude du site. On observe que la concordance est

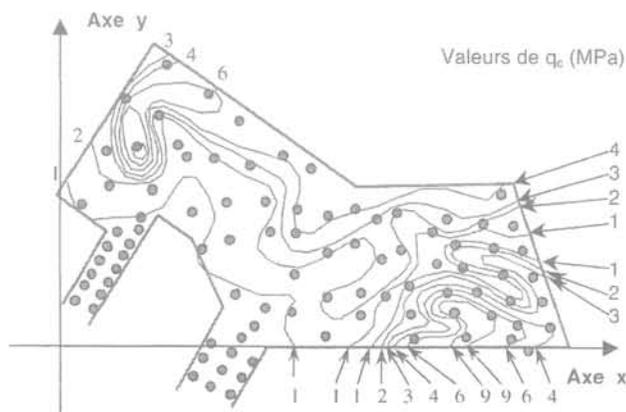


FIG. 12

Courbes isométriques de résistance de cône q_c à la profondeur de 10 m.
Lines of equal cone resistance q_c at 10 m depth.

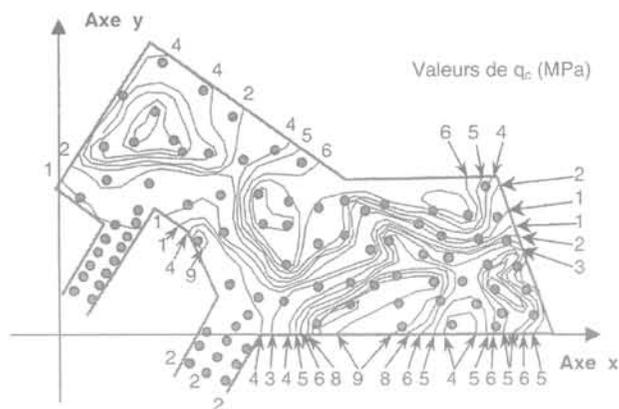


FIG. 13 Courbes isométriques de résistance de cône à la profondeur de 20 m. Lines of equal cone resistance q_c at 20 m depth.

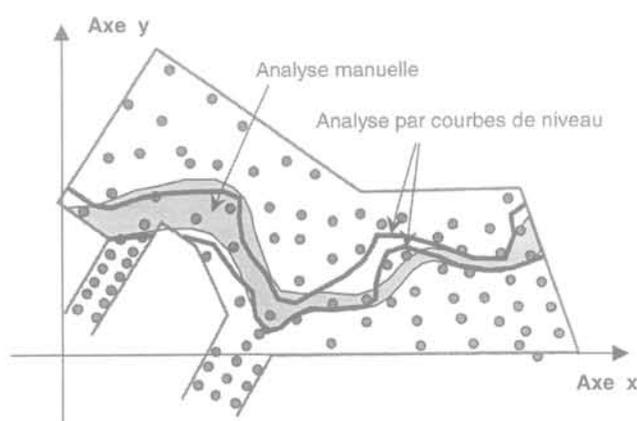


FIG. 14 Comparaison des formes de la vallée à 20 m de profondeur. Comparison of the shapes of the valley at 20 m depth.

assez bonne. La principale différence est située au centre de la figure où la vallée est tracée de part et d'autre de deux sondages dans les deux analyses, ce qui est possible si l'on limite l'analyse à ce seul niveau.

La figure 15 compare les vallées à la profondeur de 10 m. Cette fois encore on retrouve une allure compa-

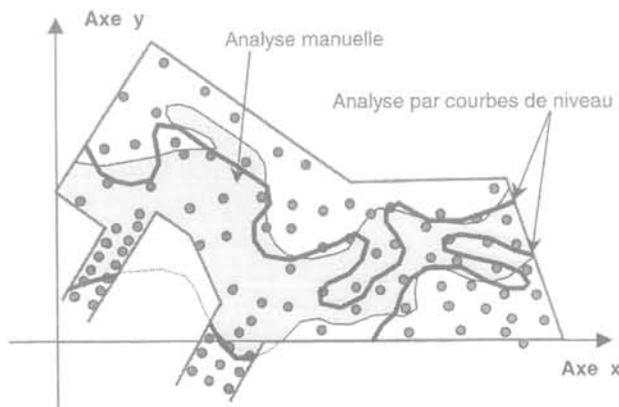


FIG. 15 Comparaison des formes de la vallée à 10 m de profondeur. Comparison of the shapes of the valley at 10 m depth.

table des zones de résistance de cône faibles ($q_c < 2$ MPa), avec des divergences d'appréciation sur quelques sondages.

Ces différences proviennent principalement des choix faits dans l'analyse manuelle globale, où la contrainte d'un rétrécissement progressif de la vallée en profondeur a conduit à éliminer des valeurs localement plus fortes des q_c , afin d'obtenir un résultat logique. Les analyses par niveau, quand elles ne sont pas corrigées par une analyse « tridimensionnelle » tenant compte des règles de la géomorphologie, peuvent en effet donner des résultats incohérents, comme on le voit dans ce cas sur la figure 16.

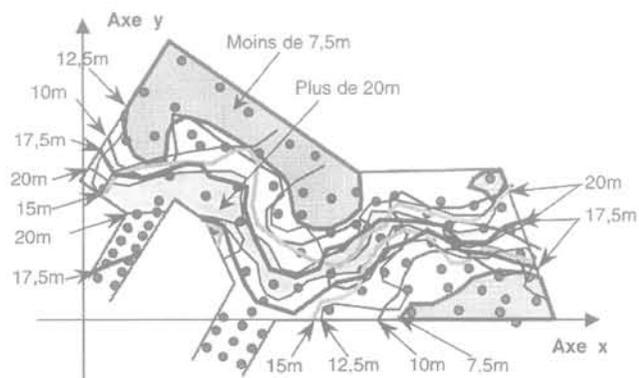


FIG. 16 Comparaison des résultats des analyses par plans horizontaux. Comparison of the results of the analyses made on horizontal planes.

6

Conclusion

Les deux applications de l'analyse factorielle (analyse en composantes principales) présentées dans cet article ont montré que cette technique d'analyse de données peut être appliquée avec succès à des données géotechniques. Pour l'étude des propriétés déduites d'une série importante d'essais œdométriques réalisés dans le même sondage, les relations connues pour le comportement des argiles ont été retrouvées, mais sans qu'apparaissent des idées de relations nouvelles. Pour la recherche d'une ancienne vallée remblayée de sédiments plus compressibles dans une zone de sols plus raides, la démarche de l'analyse factorielle a permis de donner un caractère plus systématique à la recherche des frontières de cette vallée, mais elle doit être contrôlée pour assurer la cohérence du modèle du site et complétée pour déterminer les propriétés représentatives des différents sols identifiés. Le bilan de ces études exploratoires est donc satisfaisant pour la mise en œuvre et le bon fonctionnement du système d'analyse, qui utilise un logiciel spécialisé accessible à tous les utilisateurs. Mais les conclusions tirées des résultats bruts de l'analyse factorielle doivent être soigneusement pesées et confrontées aux connaissances disponibles pour éviter des interprétations erronées.

Bibliographie

- Bagheri S., Magnan J.-P. (1983) – « Analyse probabiliste de la stabilité et des tassements des remblais du site expérimental de Cubzac-les-Ponts ». Paris, Laboratoire central des ponts et chaussées, Rapport de recherche LPC, n° 122, 71 p.
- Bertier P., Bourroche J.-M. (1977) – *Analyse de données multidimensionnelles*. Presses universitaires de France, 2^e éd.
- Boulefkhad T. (1986) – « Influence de la variabilité des propriétés des sols sur le calcul des tassements des fondations superficielles ». Thèse de docteur ingénieur en génie civil présentée à l'université de Clermont II UFR de Recherche scientifique et technique.
- Dagnelie P. (1977) – *Analyse statistique à plusieurs variables*. Presses agronomiques de Gembloux, diffusion Vander-Oyez réédition.
- Favre J.-L. (1980) – « Milieu continu et milieu discontinu : mesure statistique indirecte des paramètres rhéologiques et approche probabiliste de la sécurité ». Thèse de doctorat d'État en sciences physiques, université Pierre-et-Marie-Curie, Paris.
- Haghighou M. (1983) – « Étude du comportement d'un remblai expérimental sur sol compressible ». Thèse de docteur ingénieur présentée à l'université Pierre-et-Marie-Curie, Paris VI.
- Lebart L., Morineau A., Fenelon J.-P. (1982) – *Traitement des données statistiques. Méthodes et programmes*. Paris, Dunod, 516 p.

Étude et prise en compte de la variabilité spatiale

A. BOLLE

Université de Liège
(Belgique)
Chemin des Chevreuils
1 - B-4000 Liège

Résumé

Après un bref rappel de la nature et de l'origine de la variabilité des caractéristiques géotechniques des sols et des massifs rocheux, le choix d'un modèle est proposé. La variabilité est représentée par l'évolution spatiale de la moyenne et par la dispersion autour de cette moyenne. Cette dispersion manifeste généralement une structure plus ou moins régulière en liaison avec la périodicité des phénomènes géologiques ou anthropiques, caractérisée par l'auto-corrélation. Les effets de cette variabilité et de sa structure d'auto-corrélation peuvent être analysés par des méthodes de calcul probabiliste adaptées, autorisant l'utilisation de grands nombres de variables aléatoires. Quelques exemples d'application illustrent les propositions formulées.

Mots-clés : sols, variabilité spatiale, auto-corrélation, techniques de perturbation.

Investigation and allowance for the spatial variability

Abstract

After a short recall about the nature and the origin of the variability of the geotechnical characteristics in soils and rock masses, a model is suggested. The variability is modelled by the spatial trend and by the scattering around that mean value. Generally the dispersion exhibits a more or less regular structure, related to the periodicity of the geological phenomena, characterised by the autocorrelation. The effects of that variability and of its autocorrelation structure can be analysed by suited probabilistic computational methods, allowing to use large numbers of random variables. Some case studies are proposed as examples.

Key words: soils, spatial variability, autocorrelation, perturbation techniques.

NDLE : Les discussions sur cet article sont acceptées jusqu'au 1^{er} août 2001.

Introduction

Le présent exposé veut s'adresser à des praticiens, et non à des mathématiciens. Dans un souci de clarté, de nombreuses simplifications parfois assez radicales ont été apportées aux hypothèses et aux développements mathématiques.

Les quelques notions de base servant à modéliser la variabilité spatiale des paramètres géotechniques des sols et des massifs rocheux sont brièvement rappelées ci-dessous.

Ainsi, notre approche de la variabilité spatiale supposera d'abord que les valeurs mesurées ne reflètent que la seule variabilité naturelle des propriétés du matériau, en négligeant les autres sources de dispersion, en particulier les erreurs de mesure.

Ensuite, notre échelle d'observation de la variabilité correspondra à celle qui concerne le constructeur. Elle sera comprise entre, d'une part, les dimensions de la couche de terrain considérée comme « homogène en grand » et, d'autre part, la dimension du volume d'échantillonnage de la grandeur mesurée. Les dimensions de ce volume d'échantillonnage peuvent aller de quelques centimètres pour des mesures en laboratoire, à plusieurs décimètres, voire plusieurs mètres, pour les techniques d'investigation *in situ* (pressiomètre, pénétromètre).

La modélisation utilisée repose sur un choix arbitraire, à savoir celui des champs aléatoires limités aux moments statistiques d'ordre deux. En termes simples, les fluctuations des paramètres dans l'espace (et éventuellement dans le temps) seront décrites par leurs moyennes, leurs variances et leur auto-corrélation.

Ce choix délibéré de la notion d'auto-corrélation permet de décrire par des modèles d'une écriture simple, comportant très peu de paramètres, divers types de grandeurs géotechniques mesurées sur terrain. Il permet également, en profitant de certaines propriétés mathématiques de la notion de corrélation, de proposer des techniques de simulation simples et efficaces.

2

Les champs aléatoires

Il existe d'autres possibilités de modélisation, comme l'approche de Matheron (1965) basée sur le concept de fonctions aléatoires, qui se concrétise par la représentation du variogramme, mais elles ne seront pas développées ici. On utilisera exclusivement les champs aléatoires homogènes en se limitant à l'information statistique de second ordre, selon Vanmarcke *et al.* (1986).

Les grandeurs mesurées, par exemple la valeur prise par un paramètre géotechnique caractérisant le comportement d'une couche de sol, sont alors représentées par :

- une fonction décrivant l'évolution, sur le domaine considéré, des moyennes et variances locales, c'est-à-dire estimées sur un petit sous-domaine ;
- une fonction d'auto-corrélation, à une ou plusieurs dimensions.

La figure 1 présente l'exemple d'un paramètre géotechnique mesuré en fonction de la profondeur. La pre-

mière composante du modèle va prendre en compte l'évolution moyenne avec la profondeur, ou dérive, qui peut être ici assimilée à une droite.

En soustrayant cette dérive, il est alors possible de ramener les valeurs mesurées à un champ aléatoire stationnaire, c'est-à-dire un ensemble de valeurs aléatoires dont la moyenne et la variance locales sont des constantes.

On obtient ainsi une nouvelle variable dite stationnaire, dont la moyenne est constante sur le domaine et souvent choisie égale à zéro. Cette variable sera ensuite caractérisée par sa variance, mesurant l'importance de la dispersion autour de la moyenne, et par la structure spatiale (ou éventuellement spatio-temporelle) de cette variance, décrite par l'auto-corrélation.

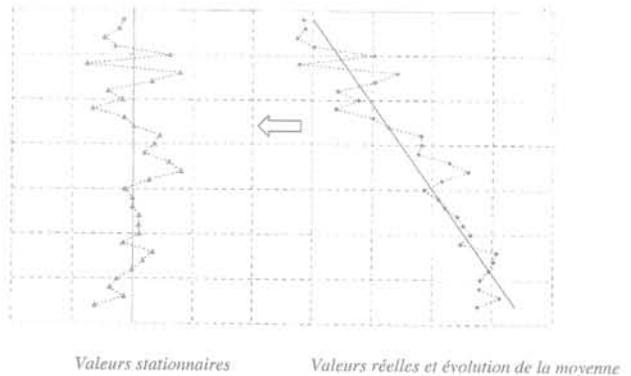


FIG. 1 Exemple de stationnarisation de mesures géotechniques.

Example of stationnarisation of geotechnical data.

3

La fonction d'auto-corrélation

3.1

Définition

La fonction d'auto-corrélation permet de décrire la structure de la dispersion qui subsiste après l'élimination de la dérive. On peut notamment observer des oscillations plus ou moins régulières liées à la périodicité des processus géologiques, combinées à des variations plus erratiques.

En se référant à la figure 1, on désignera par $W(Z)$ la variable rendue stationnaire, fonction de la profondeur Z . Sa dispersion résiduelle est mesurée par $\sigma^2(W)$, la variance, supposée constante.

Sur le diagramme, on peut constater que deux mesures réalisées à des profondeurs voisines ont tendance à présenter de valeurs proches, tantôt inférieures, tantôt supérieures à la moyenne. Par contre, des mesures plus éloignées évoluent de manière plus indépendante. Ceci peut se traduire par la notion d'auto-corrélation, qui exprime la corrélation entre les valeurs $W(Z)$ observées et ces mêmes valeurs $W(Z)$ décalées d'une distance notée ξ .

La valeur de l'auto-corrélation est obtenue par une formule statistique similaire à celle utilisée pour esti-

mer la corrélation entre deux grandeurs différentes. Elle a la forme d'un coefficient de corrélation :

$$\rho(\xi) = \frac{\text{cov}(W(Z), W(Z + \xi))}{\sigma^2(W)}$$

$$\equiv \frac{1}{\sigma^2(W)} \left\{ \frac{1}{N-1} \sum_{i=1}^N (W(Z_i) \times W(Z_i + \xi)) \right\} \quad (1)$$

Si l'on fait varier la valeur du décalage ξ , on obtient différentes valeurs de la corrélation qui forment le diagramme expérimental d'auto-corrélation, ou auto-corrélogramme, tel celui représenté sur la figure 2. Par définition, on a toujours $\rho(0) = 1$.

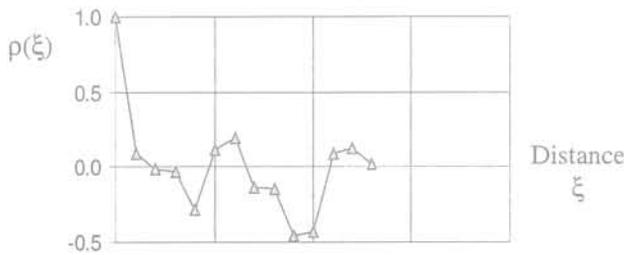


FIG. 2 Auto-corrélation des données de la figure 1.
Autocorrelation of the data from figure 1.

3.2

Constataions experimentales

Deux exemples d'auto-corrélogrammes, correspondant à des mesures réelles, vont illustrer certaines caractéristiques générales observées dans les sols et les massifs rocheux naturels.

À l'occasion d'un programme de recherche sur le comportement des pieux, De Beer *et al.* (1977) ont réalisé un grand nombre d'essais géotechniques sur un site comportant une épaisse couche d'argile assez homogène. La figure 3 est l'auto-corrélogramme moyen des 20 essais de pénétration au cône réalisés à faible distance l'un de l'autre dans cette couche homogène, obtenu par Maertens (1990).

À l'occasion de l'étude d'une fondation de grand pont en arc dans un massif rocheux très perturbé, Bolle (1988, 1994) et Bolle *et al.* (1989) ont analysé la variabilité des 191 valeurs du module pressiométrique mesurées dans une série de forages verticaux et inclinés réalisés à l'emplacement d'une culée. Devant l'impossibilité de dégager une représentation stratigraphique cohérente, le massif rocheux a été assimilé à un corps isotrope très hétérogène, et son auto-corrélation, représentée sur la figure 4, a été étudiée sur le volume en utilisant une notion isotrope de la distance ξ .

En se basant sur ces deux exemples, on peut caractériser l'allure habituellement observée pour l'auto-corrélation de propriétés géotechniques, à savoir : une décroissance monotone à partir de l'origine, suivie par des oscillations plus ou moins régulières autour de zéro.

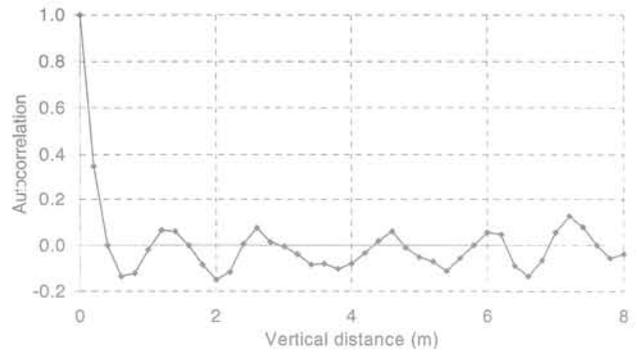


FIG. 3 Auto-corrélogramme vertical (1D) moyen de 20 essais CPT.
Mean vertical autocorrelogram (1D) of 20 CPT tests.

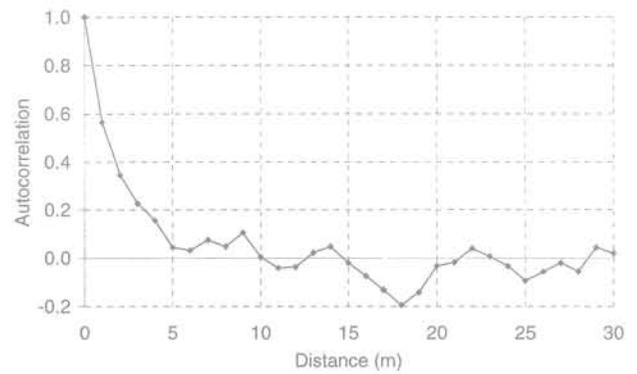


FIG. 4 Auto-corrélogramme isotrope 3D de 191 modules pressiométriques.
Isotropic 3D autocorrelogram of 191 pressuremeter moduli.

3.3

Critique des modèles classiques

La plupart des modèles proposés dans la littérature pour représenter ces fonctions d'auto-corrélation expérimentales se contentent de représenter la décroissance au voisinage de l'origine. La fonction $\rho(\xi)$ passe ainsi de l'unité à zéro de manière monotone, puis reste nulle (Fig. 5). Les oscillations, toujours présentes de manière significative dans les auto-corrélogrammes expérimentaux, n'apparaissent pas dans ces modèles classiques.

Or, on peut montrer que ces oscillations sont une conséquence directe du caractère stratifié, lité, à motifs répétitifs, des formations naturelles, et qu'elles en constituent donc une composante essentielle. Un milieu

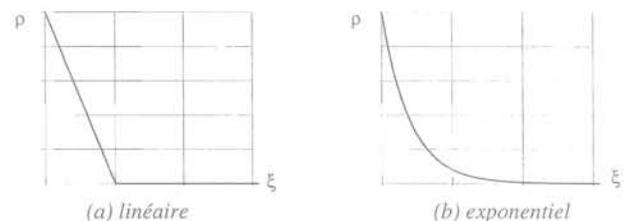


FIG. 5 Deux modèles classiques d'auto-corrélation.
Two classical autocorrelation models.

Le spectre de puissance

4.1

Définition

Par une transformation du problème, familière en électricité et en traitement du signal, on va faire apparaître de manière explicite les divers paramètres géométriques et leur importance relative.

Vanmarcke *et al.* (1986) rappellent que la fonction d'auto-corrélation $\rho(\xi)$ peut être exprimée de manière équivalente sous la forme d'un spectre de puissance $s(\kappa)$, à l'aide de la paire de transformations de Wiener-Khinchine :

$$\rho(\xi) = \int_{-\infty}^{\infty} s(\kappa) \cdot e^{i\kappa\xi} \cdot d\kappa \Leftrightarrow s(\kappa) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \rho(\xi) \cdot e^{i\kappa\xi} \cdot d\xi \quad (2)$$

dans lesquelles la variable κ a la dimension de l'inverse d'une longueur.

Ces fonctions sont paires en κ et ξ , et la fonction de densité spectrale peut être définie de manière plus simple par $g(\kappa) = 2 \cdot s(\kappa)$. La transformation se réduit alors à une paire de transformées de Fourier en cosinus, avec les deux variables κ et ξ positives :

$$\rho(\xi) = \int_0^{\infty} g(\kappa) \cdot \cos(\kappa \cdot \xi) \cdot d\kappa \Leftrightarrow g(\kappa) = \frac{2}{\pi} \int_0^{\infty} \rho(\xi) \cdot \cos(\kappa \cdot \xi) \cdot d\xi \quad (3)$$

Les modèles classiques simples déjà illustrés plus haut, dans lesquels $\rho(\xi)$ reste positif, sont définis par un seul paramètre. La figure 8 présente côte à côte les auto-corrélogrammes et les spectres de puissance correspondants.

On remarque que, pour chacun de ces exemples, le spectre de puissance présente un maximum en $\kappa = 0$, suivi d'une décroissance. Compte tenu de la signification de la variable κ , inverse d'une longueur, ce maximum correspondrait ainsi à un phénomène de longueur (ou de portée) infinie, ce qui n'est guère compatible avec la réalité physique d'une couche de sol ou de roche, nécessairement de dimensions limitées.

stratifié présente nécessairement ce comportement oscillant au-delà de la partie décroissante initiale.

Imaginons une couche de sol idéalisée formée d'une alternance régulière de deux matériaux similaires, mais de caractéristiques moyennes légèrement différentes (Fig. 6). On ajoute ensuite de la dispersion à ce modèle idéal (irrégularités dans l'alternance des sous-couches, variations erratiques et erreurs de mesure), de manière à se rapprocher d'une situation naturelle plausible. Dans une approche d'ingénieur, un tel sol sera généralement considéré comme une seule couche homogène « en grand ».

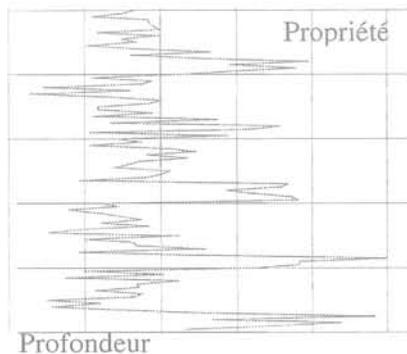


FIG. 6 Modèle idéalisé de massif stratifié.
Idealised model of a stratified soil mass.

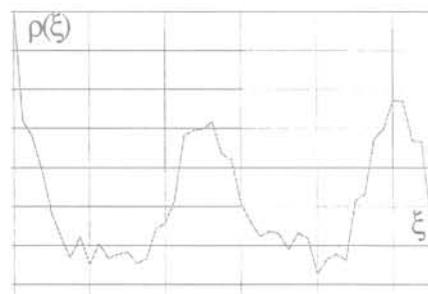
On peut montrer (Bolle, 1990) que l'auto-corrélation d'un tel milieu ne dépend pas de la dispersion des caractéristiques, mais qu'elle est fonction uniquement de la géométrie de ce massif zoné. Les auto-corrélogrammes obtenus se rapprochent des courbes expérimentales habituellement observées (Fig. 7), avec notamment de larges oscillations assez régulières autour de zéro.

La modélisation de l'auto-corrélation ne peut donc pas, comme les modèles classiques évoqués plus haut, ignorer cet aspect. Le nombre de paramètres intervenant dans la définition de telles fonctions devra toutefois rester très limité, car la quantité d'information réellement significative à ce niveau de l'analyse est souvent très réduite.

On vient de voir que les paramètres de l'auto-corrélation sont de nature géométrique. En conséquence, outre les longueurs qui caractérisent les alternances internes, une autre longueur devrait de toute évidence jouer un rôle important, à savoir la dimension du volume d'échantillonnage de la grandeur mesurée.



(a) Variation de la propriété avec la profondeur



(b) Auto-corrélogramme

FIG. 7 Auto-corrélation d'un massif stratifié idéalisé.
Autocorrelation of an idealised stratified soil mass.

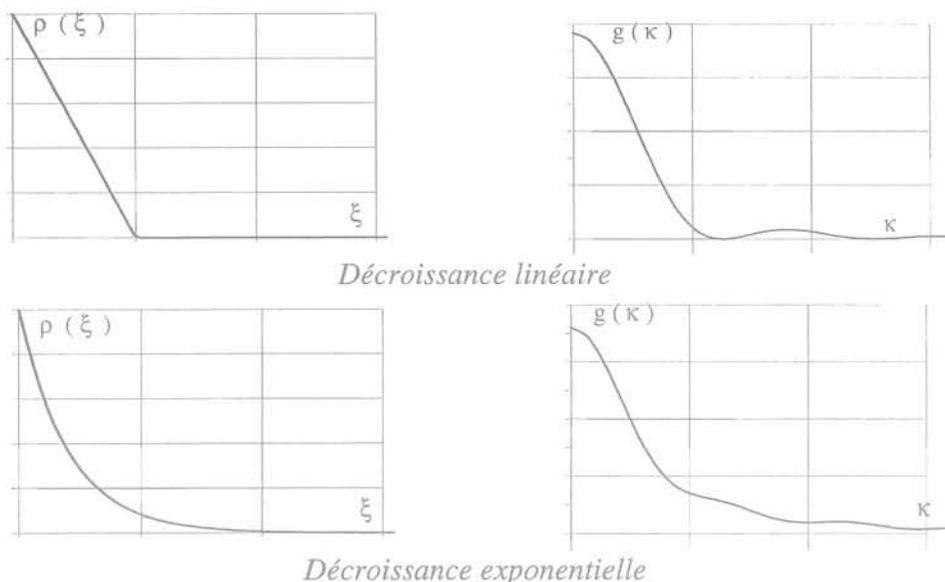


FIG. 8 Fonctions d'auto-corrélation (à gauche), et leur spectre de puissance (à droite). Autocorrelation functions (left) and their power spectra (right).

4.2

Modèles d'auto-corrélation avec oscillations

Un des rares modèles de ce type a été proposé par Baker et Zeitoun (1987), en utilisant le Principe d'entropie (de l'information) maximum. Ils ont appliqué simplement un filtre « passe bas » à un bruit blanc gaussien. En d'autres termes, ils ont supposé que les variations des propriétés du milieu étaient purement erratiques, en limitant leur portée à une longueur minimum que l'on pourrait mettre en relation avec la dimension du volume d'échantillonnage de la grandeur mesurée.

Ce modèle extrêmement simple, à un seul paramètre, donne un résultat assez proche des auto-corré-

logrammes expérimentaux, avec des incursions de p dans la partie négative (Fig. 9).

L'auto-corrélogramme du massif stratifié idéalisé présenté sur la figure 7 peut aussi être transformé en spectre de puissance (Fig. 10).

On remarque ici que le maximum du spectre de puissance ne se présente plus à l'origine, mais en une valeur de κ qui correspond à la principale longueur d'onde des oscillations de ρ , c'est-à-dire au « pas » des sous-couches du massif. On peut encore noter que, en toute rigueur, le spectre de puissance devrait rester strictement non-négatif, mais les petites excursions dans la partie négative résultent des approximations numériques lors des transformations.

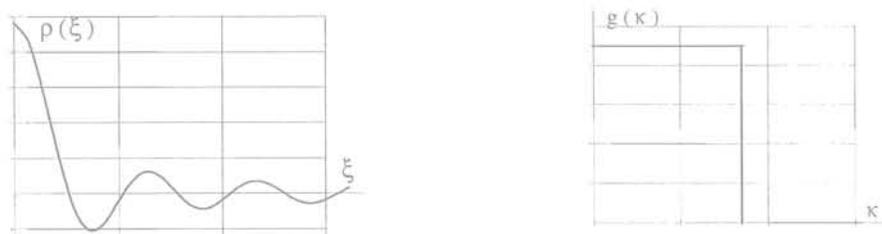


FIG. 9 Fonction d'auto-corrélation (à gauche) d'un bruit blanc gaussien filtré (à droite). Autocorrelation function (left) of a filtered gaussian white noise (right)

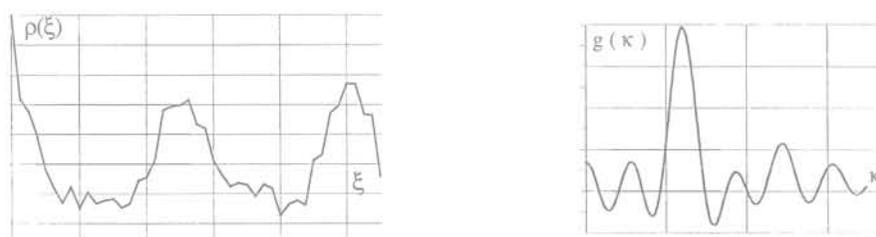


FIG. 10 Auto-corrélation d'un massif stratifié (à gauche) et spectre correspondant (à droite). Autocorrelation function of a stratified soil mass (left) and associated spectrum (right).

Construction d'un modèle d'auto-corrélation

5.1

Proposition pour un modèle d'auto-corrélation

À l'exception du modèle de Baker et Zeitoun (1987), les modèles à un seul paramètre semblent peu aptes à décrire correctement la structure d'auto-corrélation des propriétés observées dans les sols et les massifs rocheux. Un second paramètre devra donc être introduit.

La construction est réalisée en choisissant un spectre de puissance, puis en le transformant en fonction d'auto-corrélation.

Plusieurs modèles pourraient correspondre aux principes évoqués plus haut. A titre d'exemple, on peut choisir un spectre de puissance en forme d'exponentielle décroissante caractérisée par un paramètre λ , bornée à gauche par une longueur d'onde L :

$$g(\kappa) = S \cdot \exp\left(-\frac{\kappa}{\lambda}\right) \quad \forall \kappa \geq \frac{2 \cdot \pi}{L} \quad (4)$$

Ce spectre correspond à une fonction d'auto-corrélation avec oscillations :

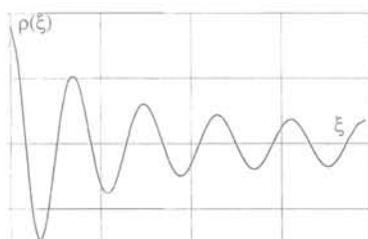
$$\rho(\xi) = \frac{1}{1 + (\xi \cdot \lambda)^2} \cdot \left(\cos \frac{2 \cdot \pi \cdot \xi}{L} - \xi \cdot \lambda \cdot \sin \frac{2 \cdot \pi \cdot \xi}{L} \right) \quad (5)$$

Cette fonction d'auto-corrélation et le spectre correspondant sont représentés côte à côte sur la figure 11.

5.2

Ajustement sur des données expérimentales

Les paramètres L et λ peuvent être obtenus par ajustement sur les auto-corrélogrammes expérimentaux $\rho(\xi)$, en utilisant la méthode des moindres carrés. Dans les exemples qui suivent, les points $\rho(\xi_i)$ sont pondérés par le nombre N' de paires de données expérimentales, comme défini par la relation (1).



La figure 12 montre les résultats de l'ajustement sur les deux exemples de résultats expérimentaux déjà présentés sur les figures 3 et 4. On peut y remarquer un certain nombre de caractéristiques qui ne sont pas de simples coïncidences.

Pour les essais au pénétromètre, le paramètre L peut être mis en relation avec une alternance assez régulière d'autres propriétés, d'un pas d'environ 1,50 m, qui avait été clairement identifiée par des analyses stratigraphiques. Dans les deux exemples, le second paramètre pourrait correspondre à une dimension caractéristique du volume d'échantillonnage.

Enfin, pour le cas des modules pressiométriques du massif rocheux, le paramètre $L = 25,6$ m correspond à un phénomène principal, de grande longueur d'onde. On pourrait aisément y superposer un second phénomène de plus courte longueur d'onde (4 à 5 mètres).

Ce travail de développement d'un modèle reste à améliorer, mais on peut déjà constater sur ces deux exemples une concordance très satisfaisante avec les observations expérimentales. La forme du spectre pourrait être remplacée par une courbe plus douce, mais toujours caractérisée par deux paramètres équivalents : (1) une valeur centrale, ou un pic, et (2) une dispersion autour de ce pic, ce qui conduirait encore à des auto-corrélogrammes avec oscillations amorties.

5.3

Application à deux dimensions

La notion d'auto-corrélation peut être définie, comme dans l'équation (1), pour une distance ξ mesurée selon une seule direction (1D), généralement la profondeur, pour une direction quelconque (isotrope), ou pour 2, voire 3 directions (2D et 3D).

Dans la pratique, le nombre de mesures disponibles limite souvent les approches à 2 ou 3 dimensions. Dans ce cas, pour obtenir un nombre significatif de paires de points de mesure pour une direction donnée, on est amené à relâcher un peu les conditions, par exemple en remplaçant les points de mesure par un volume de mesure centré sur le point.

Les résultats correspondant à 191 modules pressiométriques mesurés dans un massif rocheux très hétérogène ont déjà été présentés sur la figure 4 sous la forme d'un auto-corrélogramme isotrope. Les mêmes données ont aussi été analysées en 2 dimensions dans le plan horizontal. Chaque auto-corrélogramme expérimental, correspondant à une direction donnée, a été ajusté sur le modèle en exponentielle tronquée, donnant ainsi une paire de valeurs L et $1/\lambda$ pour chaque direction. Ces deux paramètres sont reportés, en coor-

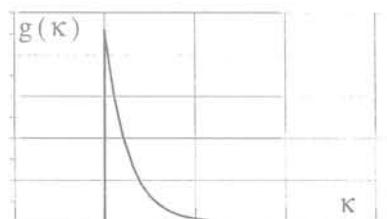
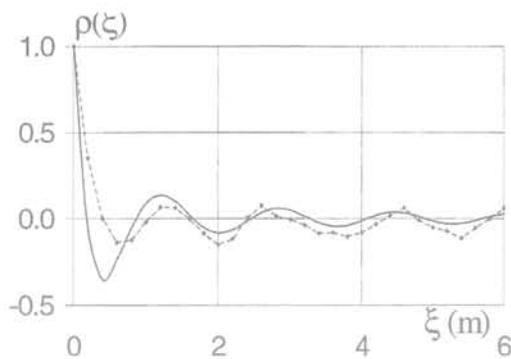


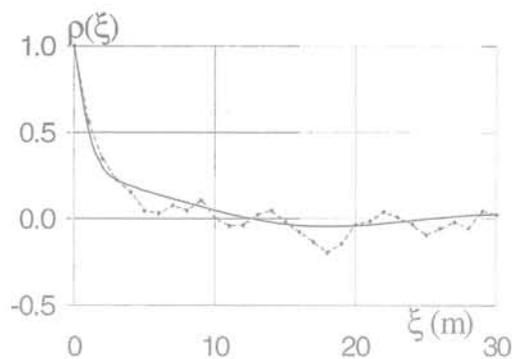
FIG. 11 Fonction d'auto-corrélation (à gauche) obtenue par la transformation d'un spectre en exponentielle tronquée (à droite).

Autocorrelation function (left) obtained from the truncated exponential spectrum (right).



(a) Pénétrètres CPT(b)

$L = 1,63 \text{ m}$ et $1/\lambda = 0,17 \text{ m}$



(b) Modules pressiométriques

$L = 25,6 \text{ m}$ et $1/\lambda = -0,84 \text{ m}$

FIG. 12 Exemples d'ajustement sur des auto-corrélogrammes expérimentaux.
Examples of fitting to experimental autocorrelograms.

données polaires, sur la figure 13, manifestant clairement une allure elliptique.

Une ellipse a ensuite pu être ajustée sur ces points, caractérisée par trois paramètres : la longueur des deux axes et un angle α donnant la direction du grand axe par rapport au nord :

$$\left(\frac{r \cdot \cos(\theta + \alpha)}{L_1} \right)^2 + \left(\frac{r \cdot \sin(\theta + \alpha)}{L_2} \right)^2 = 1 \quad (6)$$

$$\left(\lambda_1 \cdot r \cdot \cos(\theta + \alpha) \right)^2 + \left(\lambda_2 \cdot r \cdot \sin(\theta + \alpha) \right)^2 = 1 \quad (7)$$

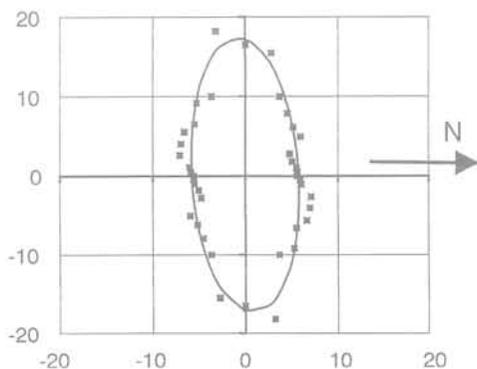
Les orientations des axes principaux des deux courbes coïncident pratiquement, de même que l'allongement relatif des ellipses (rapport de grand axe au petit axe). De toute évidence, ces « directions principales » traduisent une certaine structure interne, cachée, du massif rocheux considéré, inaccessible à une analyse classique, mais apparaissant clairement au travers de la procédure proposée.

6

Conclusions relatives à l'étude de la variabilité spatiale

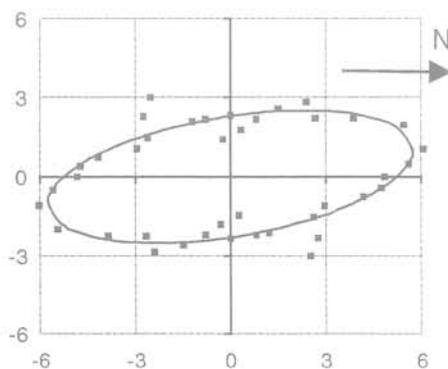
Les modèles classiques d'auto-corrélation, appliqués aux propriétés des sols et des massifs rocheux, ne s'adressent qu'à la première partie de la courbe, correspondant à la « persistance » ou la « portée » des propriétés. Mais les auto-corrélogrammes expérimentaux habituellement observés comportent toujours, après la décroissance monotone proche de l'origine, des oscillations plus ou moins régulières autour de zéro, qui résultent du caractère périodique lié à l'alternance des propriétés.

Le modèle proposé s'ajuste de manière assez satisfaisante aux données expérimentales. Ce modèle, qui pourrait encore subir diverses améliorations, utilise seulement deux paramètres : (1) une longueur d'onde principale et (2) une « distance d'amortissement », ces paramètres ayant tous deux la dimension d'une longueur.



(a) Paramètre L (m)

$L_1 = 17,4$, $L_2 = 5,8$, $\alpha = -87^\circ$



(b) Paramètre $1/\lambda$ (m)

$1/\lambda_1 = -5,8$, $1/\lambda_2 = -2,3$, $\alpha = 12^\circ$

FIG. 13 Paramètres des auto-corrélogrammes 2D : valeurs expérimentales et ellipses ajustées.
Parameters of 2D autocorrelograms : experimental data and fitted ellipses.

Les analyses peuvent être menées à une seule dimension : la verticale pour les couches sédimentaires ou une direction quelconque (isotrope), pour les massifs rocheux à structure complexe indéterminée. Elles peuvent aussi, lorsque le volume d'information disponible est suffisant, être menées à deux ou à trois dimensions, permettant ainsi de mettre en évidence l'anisotropie de la structure d'auto-corrélation. Les paramètres peuvent alors apparaître sous la forme d'ellipses ou d'ellipsoïdes de paramètres principaux, avec une signification proche des notions de contraintes ou déformations principales connues en théorie de l'élasticité.

7

Prise en compte de la variabilité spatiale dans les applications

7.1

Position du problème

Dans une couche de sol (ou un massif rocheux) homogène « en grand », une propriété géotechnique mesurée en différents points pourra être assimilée à un champ aléatoire, et représentée par les composantes suivantes :

- la dérive qui est une fonction décrivant l'évolution, sur le domaine considéré, des moyennes et variances locales, c'est-à-dire estimées sur un petit sous-domaine ; cette fonction comportera éventuellement une transformation de la variable de base ; elle prendra une forme simple adaptée à chaque cas particulier (linéaire, quadratique ou autre) ;
- après déduction de cette fonction, les données expérimentales forment un champ aléatoire homogène (jusqu'au second moment statistique), ou stationnaire ; la moyenne est souvent choisie égale à zéro au cours de l'étape (1), et la variance, rendue constante sur le domaine, reste le paramètre unique ;
- la structure de la variabilité spatiale résiduelle est représentée par la fonction d'auto-corrélation, à une ou plusieurs dimensions, qui peut être décrite par un ou deux paramètres ; le caractère oscillant de cette fonction constitue une caractéristique fondamentale observée sur les matériaux naturels.

La prise en compte d'une telle propriété géotechnique dans une application requiert la simulation de la valeur prise en tout point par cette propriété. Mais, les données expérimentales disponibles sont toujours trop peu nombreuses, et il faut donc « reconstruire » ces valeurs sur le domaine. Des techniques optimisées d'interpolation existent, comme le krigeage (Krige, 1962). Ces techniques peuvent fournir en tout point une estimation de la grandeur, sous la forme d'une valeur probable et de l'incertitude qui y est attachée.

La simulation par la méthode des champs aléatoires permet de conserver le caractère incertain, aléatoire, des valeurs possibles de la grandeur en tout point du domaine étudié. Ces valeurs sont reconstruites comme la somme de deux composantes distinctes :

– une fonction déterministe décrivant la variation, sur le domaine considéré, de la *moyenne* (et éventuellement de la variance) ;

– une variation aléatoire de la propriété autour de cette moyenne, présentant une certaine amplitude, la variance, et une structure d'auto-corrélation données.

Cette seconde composante, de nature aléatoire, incertaine, sera introduite dans les applications et elle produira en finale une incertitude sur les résultats obtenus.

Il faut souligner que la simulation de la variabilité spatiale, comme proposée ci-dessus, entraîne systématiquement l'utilisation d'un grand nombre de variables aléatoires, qui sont simplement les valeurs, affectées d'incertitude, de la grandeur considérée en tous les points (ou volumes élémentaires) choisis pour représenter le domaine.

7.2

Techniques de simulation existantes

Plusieurs méthodes stochastiques, c'est-à-dire utilisant le hasard, ont été proposées pour simuler ces variations spatiales, par exemple pour fournir les données à introduire dans des calculs par éléments finis. Par exemple, Zeldin et Spanos (1995) améliorent les approches spectrales ou ARMA par le développement des ondelettes (wavelets).

La simulation pure du hasard, connue sous le nom de méthode de Monte-Carlo, reste la méthode de référence, mais le volume de calcul nécessaire pour réellement converger vers la solution peut atteindre une taille astronomique.

Des méthodes semi-analytiques, de premier ou de second ordre, ont aussi été développées, comme les éléments finis stochastiques (Cambou, 1975, Vanmarcke *et al.*, 1986), en se basant sur des développements en séries de Taylor. Mais ces méthodes présentent diverses limitations et elles nécessitent la recherche (souvent réalisée par voie numérique) des dérivées partielles par rapport à chaque variable aléatoire.

Enfin, plusieurs méthodes d'estimation ponctuelle ont été proposées, inspirées de la technique imaginée par Rosenblueth (1975), mais elles sont très rapidement limitées (Bolle, 1988).

7.3

Technique de simulation proposée

Pour faire face aux difficultés et aux limitations inhérentes à ces méthodes de simulation, une technique de calcul originale a été mise au point (Bolle, 1988, 1994) pour simuler divers processus stochastiques avec de grands nombres de variables aléatoires. Cette méthode est plus performante que la simulation de Monte-Carlo, car le nombre de calculs nécessaires est très fortement réduit. Elle ne peut toutefois s'appliquer, en toute rigueur, qu'à des fonctions « gentiment » continues.

La méthode proposée utilise uniquement, de manière interne, des variables aléatoires non corrélées à distribution gaussienne. Dès lors, la modélisation des variables spatiales devra subir une transformation permettant de l'exprimer sous la forme d'un ensemble de

variables aléatoires indépendantes (au sens de non corrélées) à distribution symétrique. Cette seconde condition est en général automatiquement respectée si l'analyse de la variabilité spatiale est limitée aux moments statistiques de second ordre.

8

Méthode de calcul probabiliste proposée

L'idée de base consiste à rechercher des estimateurs des premiers moments statistiques de la fonction étudiée, comme dans les méthodes basées sur le développement en séries de Taylor. Ces moments statistiques sont estimés par l'addition des effets dispersants, des incertitudes, attribuables à la variabilité de chaque variable aléatoire indépendante.

Plus précisément, les premiers moments statistiques de la distribution sont estimés en ajoutant à un terme central la somme des termes perturbateurs liés à chaque variable aléatoire indépendante. Le terme central représente la valeur non aléatoire de la fonction, c'est-à-dire la valeur obtenue lorsque toutes les variables sont considérées comme déterministes, ou de variance nulle.

Le calcul des perturbations, définies comme les effets générés par la dispersion de chaque variable, utilise une variante optimisée de la méthode des estimations ponctuelles de Rosenblueth (1975), avec un schéma de discrétisation plus élaboré, comportant trois points au lieu de deux. Une distribution continue est ainsi remplacée par une distribution discrète en respectant jusqu'au moment statistique de quatrième ordre inclus.

8.1

Transformation du champ aléatoire en variables indépendantes

En vue de la résolution d'un problème concret, les propriétés géotechniques connues sous la forme d'un champ aléatoire doivent être exprimées en une série de points (ou de volumes élémentaires) de l'espace. Un exemple habituel est un calcul par éléments finis, qui demande de connaître les propriétés des matériaux en chaque élément, voire en chaque nœud.

On désignera par $X = (x_1, x_2, \dots, x_N)$ les valeurs de ces propriétés. Elles sont caractérisées par les moments statistiques d'ordre 1 et 2 de leur distribution couplée, soit :

- la moyenne des distributions $M = (\bar{x}_1 \dots \bar{x}_N)$;
- les écarts types des distributions, donnés par la

matrice diagonale $S = \begin{bmatrix} \sigma_{x_1} & & \emptyset \\ & \ddots & \\ \emptyset & & \sigma_{x_N} \end{bmatrix}$;

- la matrice de corrélation $R(X)$, qui comporte une diagonale unitaire et des termes symétriques exprimant l'auto-corrélation spatiale de la grandeur considérée ; ces termes sont obtenus en calculant, avec le modèle proposé, la valeur de $\rho(\xi)$ pour chaque couple de points de l'espace.

Des variables réduites normées et centrées formant un vecteur T , parallèle à X , de moyenne nulle et de variance unitaire, sont définies par la transformation linéaire réversible :

$$X = M + S.T \Leftrightarrow T = S^{-1}.(X - M) \quad (8)$$

avec $S^{-1} = \begin{bmatrix} 1/\sigma_{x_1} & & \emptyset \\ & \ddots & \\ \emptyset & & 1/\sigma_{x_N} \end{bmatrix}$

La matrice de corrélation de $R(T) \equiv R(X)$ peut être transformée par une opération de diagonalisation, selon Rosenblatt (1952), en les éléments suivants :

- une matrice diagonale formée des valeurs propres de $R(T) \equiv R(X)$, que l'on peut regrouper sous la forme d'un vecteur $L = \text{Diag}(\lambda_1, \lambda_2, \dots, \lambda_N)$;
- la matrice (normée) A des vecteurs propres correspondant à ces valeurs propres.

On peut alors définir de nouvelles variables aléatoires notées $U = (u_1, u_2, \dots, u_N)$ données par :

$$U = A.T = A.S^{-1} (X - M) \quad (9)$$

On peut montrer que ces nouvelles variables aléatoires sont indépendantes, de moyenne nulle, avec $\bar{U} = (0 \ 0 \ \dots \ 0)$, et de variances $\sigma^2 = L = (\lambda_1, \lambda_2, \dots, \lambda_N)$.

La transformation de X , vecteur des variables réelles exprimées en unités physiques, en variables non corrélées U , de moyenne nulle et de variances L , est réversible :

$$U = A.S^{-1}.(X - M) \Leftrightarrow X = M + S.A^T.U \quad (10)$$

en remarquant que, par suite des propriétés de A , on a $A^{-1} = A^T$.

On peut résumer ces transformations, en langage plus concret, en une mise à l'échelle des grandeurs physiques, puis en une rotation (dans un espace à N dimensions) en tout point assimilable à la recherche des contraintes principales dans un corps élastique.

Bien entendu, il est possible d'inclure dans cette transformation toute autre variable aléatoire intervenant dans le problème, à condition de lui donner une distribution proche de la normale, et d'exprimer son éventuelle corrélation avec les autres variables.

8.2

Estimation ponctuelle à trois points

Chacune des variables aléatoires indépendantes U_i , à distribution normale de moyenne nulle et de variance λ_i , peut être représentée de manière équivalente (jusqu'au quatrième moment statistique inclus) par la distribution discrète en trois points représentée sur la figure 14. Les caractéristiques de cette distribution discrète s'écrivent simplement :

$$u_i = \sqrt{3\lambda_i} \quad P_1 = \frac{1}{6} \quad P_0 = \frac{2}{3} \quad (11)$$

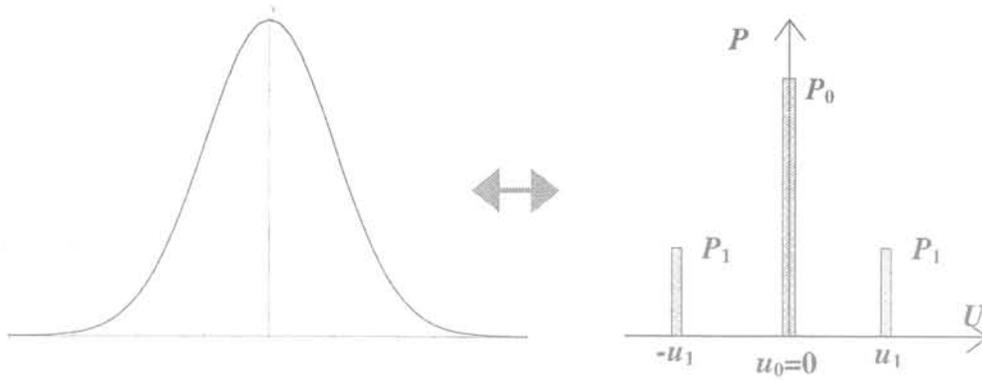


FIG. 14 Discretisation de Rosenblueth modifiée, en 3 points.
Modified Rosenblueth's discretization with 3 points.

8.3

Développement de la méthode de simulation

La fonction $Y(\mathbf{X})$ représente le phénomène à simuler en faisant appel, par exemple, à un code de calcul par éléments finis. Le résultat du calcul sera la distribution de $Y(\mathbf{X})$ exprimée sous la forme des premiers moments statistiques.

Le terme central Y_c de la distribution de $Y(\mathbf{X})$ est la valeur de cette fonction calculée au point central (moyenne) de chaque variable aléatoire \mathbf{X} :

$$Y_c = Y(\bar{\mathbf{X}}) = Y(\mathbf{M}) \quad (12)$$

Il s'agit du moment statistique d'ordre 1, non perturbé, de la fonction $Y(\mathbf{X})$. De même, la valeur non perturbée des moments statistiques d'ordre supérieur vaudra respectivement :

- variance (ordre 2) = 0 ;
- asymétrie (ordre 3) = 0 ;
- aplatissement (ordre 4) = 3, c'est-à-dire la valeur correspondant à une distribution normale.

Les perturbations sont estimées par rapport à ces tendances centrales (valeurs non perturbées), à l'aide des valeurs de la fonction $Y(\mathbf{X})$ calculées aux deux points latéraux de la discrétisation de la distribution de chaque variable aléatoire, que l'on appellera les valeurs perturbées de la fonction. On désigne les valeurs de la fonction $Y(\mathbf{X})$ calculées en ces deux points, pour la variable U d'indice i , par la notation :

$$\begin{cases} Y_i^+ = Y(\mathbf{M} + \mathbf{S} \cdot \mathbf{A}^T \cdot (0, \dots, +u_i, \dots, 0)) \\ Y_i^- = Y(\mathbf{M} + \mathbf{S} \cdot \mathbf{A}^T \cdot (0, \dots, -u_i, \dots, 0)) \end{cases} \quad (13)$$

La distribution discrète de la fonction $Y(\mathbf{X})$, dans laquelle une seule variable U d'indice i est considérée comme aléatoire, se présente alors sous la forme représentée sur la figure 15.

Les notations suivantes simplifieront les écritures :

$$\delta_i^+ = Y_i^+ - y_i = \frac{1}{6} \cdot (5 \cdot d_i^+ - d_i^-) \quad \text{et}$$

$$\delta_i^- = Y_i^- - y_i = \frac{1}{6} \cdot (-d_i^+ + 5d_i^-) \quad (14)$$

$$\varepsilon_i = y_i - Y_c = \frac{1}{6} \cdot (d_i^+ + d_i^-) \quad (15)$$

$$\eta_i = \frac{d_i^+ \cdot d_i^-}{36}$$

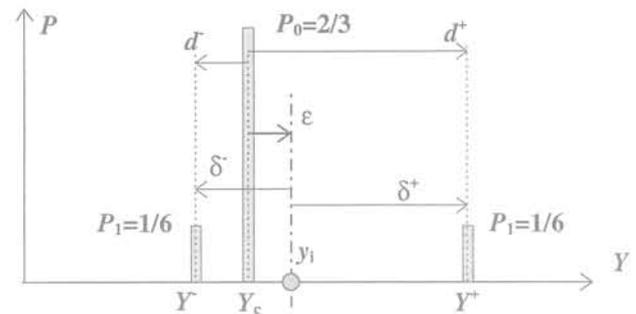


FIG. 15 Calcul des termes perturbateurs de $Y(\mathbf{X})$ sous l'effet de la variance de U_i .
Calculation of the perturbing terms of $Y(\mathbf{X})$ due to the variance of U_i .

L'estimateur de la moyenne de $Y(\mathbf{X})$, soit y_i , peut alors être aisément exprimé par :

$$y_i \equiv Y_c + \varepsilon_i \quad (16)$$

Pour les moments d'ordre supérieur à un, on adoptera comme principe que la seule variable U d'indice i est aléatoire et, par conséquent, que la distribution discrète de la fonction $Y(\mathbf{X})$ est complètement représentée par ce schéma en trois points. Moyennant cette hypothèse, on considère que y_i est une moyenne locale de la distribution, et que les δ_i^\pm sont des perturbations locales nettes de cette moyenne.

Les moments perturbateurs d'ordre k seront obtenus par la formule générale suivante :

$$\begin{aligned} d\mu_{k,i} &\equiv \sum_j P_j \cdot (y - y_i)^k \\ &= P_0 \cdot (-\varepsilon_i)^k + P_1 \cdot \left((\delta_i^+)^k + (\delta_i^-)^k \right) \end{aligned} \quad (17)$$

à partir de laquelle on peut obtenir les diverses expressions correspondant à la variance, la dissymétrie et l'aplatissement.

Ensuite, compte tenu de l'indépendance des variables de type **U**, on pose que la perturbation totale des différents moments statistiques de **Y(X)** peut être obtenue par une simple **sommation** des effets des perturbations dues à chaque variable.

L'estimation de la moyenne de la distribution de la fonction **Y(X)** sera ainsi donnée par :

$$\bar{Y} \equiv Y_c + \frac{1}{6} \sum_{i=1}^N (d_i^+ + d_i^-) = Y_c + \sum_{i=1}^N \varepsilon_i \quad (18)$$

La variance de la distribution de la fonction **Y(X)** sera estimée par une expression similaire :

$$\sigma^2(Y) \equiv \sum_{i=1}^N d\mu_{2,i} = \sum_{i=1}^N (5 \cdot \varepsilon_i^2 - 12 \cdot \eta_i) \quad (19)$$

La dissymétrie et l'aplatissement de la fonction **Y(X)** peuvent être estimés par des développements conduisant à des expressions similaires.

8.4

Calcul simultané de plusieurs fonctions et estimation des corrélations

Dans la première partie des développements, pour ne pas alourdir l'écriture, on n'a considéré qu'une seule fonction scalaire **Y(X)** sans indice. Mais on a plus souvent affaire à plusieurs fonctions simultanées **Y(X) = (Y₁(X), Y₂(X), ..., Y_M(X))**. Il suffit de penser, par exemple, aux résultats fournis par la méthode des éléments finis, qui sont obtenus à chaque phase de calcul en tous les nœuds (ou points d'intégration) du réseau.

Les moments statistiques croisés d'ordre deux (covariances) de tous les couples de fonctions **{Y_j(X), Y_k(X)}** peuvent aussi être calculés par la même méthode. Pour chacune de ces fonctions, on peut encore calculer le terme central de leur distribution, **Y_{j,c}** et **Y_{k,c}**, qui sont la valeur de la fonction calculée au point central de chaque variable aléatoire. Il s'agit des moyennes non perturbées des fonctions :

$$Y_{j,c} = Y_j(\bar{X}) = Y_j(\mathbf{M}) \text{ et } Y_{k,c} = Y_k(\bar{X}) = Y_k(\mathbf{M}) \quad (20)$$

On désignera les valeurs perturbées des fonctions **Y_{j,c}** et **Y_{k,c}** calculées aux deux points latéraux de la discrétisation de la distribution de la variable **U** d'indice **i** par la notation :

$$\begin{cases} Y_{j,i}^+ = Y_j(\mathbf{M} + \mathbf{S} \cdot \mathbf{A}^T \cdot \{0, \dots, \pm u_i, \dots, 0\}) \\ Y_{k,i}^+ = Y_k(\mathbf{M} + \mathbf{S} \cdot \mathbf{A}^T \cdot \{0, \dots, \pm u_i, \dots, 0\}) \end{cases} \quad (21)$$

Par des développements similaires, qui ne seront pas exposés ici, on peut ensuite obtenir une estimation de la covariance, puis de la corrélation entre chaque couple de variables résultats **Y_j(X)** et **Y_k(X)** :

$$\begin{aligned} \text{cov}(Y_j, Y_k) &\equiv \sum_{i=1}^N d\mu_{jk,i} \\ &= \sum_{i=1}^N \left(\frac{d_{j,i}^+ \cdot d_{k,i}^+ - d_{j,i}^- \cdot d_{k,i}^-}{6} - \varepsilon_{j,i} \cdot \varepsilon_{k,i} \right) \end{aligned} \quad (22)$$

$$\rho(Y_j, Y_k) = \frac{\text{cov}(Y_j, Y_k)}{\sigma(Y_j) \cdot \sigma(Y_k)} \quad (23)$$

Dans la conduite pratique des calculs, il suffit d'ajouter aux opérations relatives à chacune des fonctions **Y(X)** considérées isolément, une sommation relative à chaque couple **(j, k)**.

8.5

Remarques sur les volumes de calcul

L'amélioration apportée par une telle méthode par comparaison avec les autres méthodes de simulation réside d'abord dans le nombre de simulations de **Y(X)**, qui est réduit à **2.N + 1** pour **N** variables aléatoires. Ceci permet l'approche de problèmes faisant intervenir de grands nombres de variables aléatoires. On peut remarquer que ce nombre de simulations correspond précisément au nombre idéal permettant une utilisation optimale de l'information de départ.

Dans un problème à **N** variables aléatoires tel que nous l'avons défini, on dispose de **2.N** données qui sont les deux premiers moments statistiques de chaque variable. L'indépendance entre ces variables est assurée par la technique d'estimation elle-même, qui utilise des variables aléatoires rendues non corrélées. Au calcul déterministe habituel, on ajoute donc strictement le nombre de simulations qui correspond au nombre d'informations élémentaires présentes dans les données.

On peut considérer qu'il n'y a de la sorte aucune perte ou dilution d'information dans le processus ou, en d'autres termes, aucune augmentation de l'entropie de l'information.

De surcroît, le nombre de simulations nécessaires, estimé a priori à **2.N + 1**, peut souvent être fortement réduit. En effet, lors de la transformation des variables **X** en variables non corrélées de type **U**, les valeurs propres sont obtenues en ordre décroissant, et elles deviennent souvent négligeables au-delà d'un certain rang, soit **n < N**. Or, ces valeurs propres représentent la variance des variables de type **U**. Si cette variance devient négligeable, le caractère aléatoire des variables correspondantes disparaît, et elles se réduisent à des grandeurs déterministes, fixes, égales à zéro.

Dans l'application de la méthode, les variables **U** d'indice **i > n** produiront des perturbations nulles ou négligeables. Les simulations correspondantes sont inutiles, et le nombre total de simulations nécessaires est ramené à **2.n + 1**. Dans les cas habituels, **n** dépasse rarement quelques dizaines.

Justification et domaine d'application de la méthode

La justification de base de la méthode se fonde sur l'hypothèse que les effets du caractère aléatoire de chaque variable aléatoire indépendante sont aussi indépendants, et qu'ils peuvent dès lors être additionnés. La justification de cette hypothèse, et en particulier l'additivité des effets de chaque variable aléatoire indépendante, n'a toutefois pas encore été étudiée in extenso par des principes mathématiques rigoureux.

Mais des considérations sur la bonne utilisation de l'information, basées sur le concept d'entropie de l'information, permettent intuitivement de fournir une certaine justification logique au choix de cette technique.

Enfin, de nombreuses applications ont été réalisées sur des fonctions simples, mais fortement non linéaires, et comparées de manière très satisfaisante avec les solutions théoriques. Des cas pratiques plus complexes

ont été comparés aux résultats fournis par d'autres méthodes de simulation, dont la méthode de référence de Monte-Carlo.

La seule réserve à l'application de la méthode concerne la continuité des fonctions analysées. Compte tenu de la manière de calculer les perturbations, qui fait intervenir un calcul des différentielles d'ordre 1 et 2, il faut que les dérivées estimées par ces différentielles existent sur le domaine balayé par la fonction.

Un contre-exemple de cette nature a été mis en évidence par Fischbach (1995) dans une application consacrée à la stabilité des pentes dans un milieu aléatoire stratifié. Au cours de la simulation, on a parfois constaté que la ligne de glissement (cercle) critique, qui passait généralement par le pied du talus, se trouvait assez brutalement « attirée » par une couche anormalement moins résistante. Ce saut brutal n'est pas compatible avec les hypothèses de la méthode, et ce type de situation nécessite alors le recours à une technique plus générale comme Monte-Carlo.

Bibliographie

- Baker R., Zeitoun D.G. (1987) – « Soil variability and the maximum entropy principle ». *ICASP 5*, Vancouver, p. 642-649.
- Bolle A. (1988) – « Approche probabiliste en mécanique des sols avec prise en compte de la variabilité spatiale ». Thèse n° 743, EPF Lausanne.
- Bolle A., Bonnechère F., Crémer J.-M. (1989) – « Probabilistic approach of the settlement of an arch bridge ». *XIII ICSMFE*, Rio de Janeiro.
- Bolle A. (1990) – « Méthodes statistiques en infrastructures ». Notes de cours à l'université de Liège (non publié).
- Bolle A. (1994) – « How to manage the spatial variability of natural soils. *Probabilities and Materials*. D. Breyssse, ed., p. 505-516.
- Cambou B. (1975) – « Application of first-order uncertainty analysis in the finite element method in linear elasticity ». *Proceedings, Second International Conference on Application of Statistics and Probability in Soil and Structure Engineering*, Aachen, p. 67-87.
- De Beer E., Lousberg E., Wallays M., Carpentier R., De Jaeger J., Paquay J. (1977) – « Bearing capacity of displacement piles in stiff fissured clays ». *Verslag over navorsingen n° 39 van het IWONL*, Brussels.
- Fischbach S. (1995) – « Étude de la stabilité des talus dans un cadre probabiliste ». Travail de diplôme, Université de Liège.
- Krige D.G. (1962) – « Economic aspects of stoping through unpayable ore ». *Journal South African Institute of Mining and Metallurgy*.
- Maertens J. (1990) – « Aanwending van de statistiek voor de verwerking van grondmechanische karakteristieken ». *Geopropa 90* (ed. Belgian Group of the ISSMFE), Brussels.
- Matheron G. (1965) – « Les variables régionalisées et leur estimation. Une application de la théorie des fonctions aléatoires aux sciences de la nature. Paris, Masson et Cie.
- Rosenblatt M. (1952) – « Remarks on a multivariate transformation ». *Ann. Math. Stat.*, vol. 13, p. 470-472.
- Rosenblueth E. (1975) – « Point estimates for probability moments ». *Proc. Nat. Acad. Sci. USA*, vol. 72, n° 10.
- Vanmarcke E., Shinozuka M., Nagakiri S., Schuëller G.I., Grigoriu M. (1986) – « Random fields and stochastic finite elements ». *Structural Safety* 3, p. 143-166.
- Zeldin B., Spanos P. (1995) – « Random field simulation using wavelet bases ». *ICASP 7*, Paris, p. 1275-1283.

La méthode des éléments finis stochastiques en géotechnique

G. AUVINET

Instituto de Ingeniería
UNAM
Cd. Universitaria
Apdo Postal 70-472
Coyoacán 04510
Mexico

**R. MELLAH
F. MASROURI**

Laboratoire Environnement
Géomécanique et Ouvrages
École Nationale Supérieure
de Géologie, INPL, Nancy

J. F. RODRIGUEZ

Instituto de Ingeniería
UNAM, Mexico

Résumé

Cet article présente différentes techniques permettant d'appliquer la méthode des éléments finis stochastiques (MEFS) aux ouvrages géotechniques. La MEFS a pour but d'évaluer la marge d'erreur sur les résultats des modélisations numériques du comportement de ces ouvrages, compte tenu des diverses incertitudes qui affectent ces modélisations et, en particulier, de celles qui portent sur les paramètres du sol. On montre comment la méthode des éléments finis peut être couplée à des techniques probabilistes telles que la méthode de Monte-Carlo ou les diverses techniques de perturbation, y compris les méthodes du « premier ordre-seconds moments » et la méthode des estimations ponctuelles. Les avantages et les inconvénients de ces diverses techniques sont examinés. L'application de la MEFS est brièvement illustrée par des analyses d'incertitude sur le champ de contraintes dans une structure simple et sur le champ des déplacements dans un remblai. Les résultats de ces analyses et d'autres similaires qui seront présentées dans de prochains articles montrent que la MEFS, malgré ses limites actuelles, permet aux ingénieurs géotechniciens d'introduire un degré de réalisme supplémentaire bien nécessaire dans les modélisations aux éléments finis.

Mots-clés : géotechnique, incertitude, méthode des éléments finis stochastiques, sols, méthode de Monte Carlo, techniques de perturbations.

The stochastic finite element method in geotechnical engineering

Abstract

This paper presents several techniques making it possible to apply the Stochastic Finite Element Method (SFEM) to geotechnical structures. SFEM can be used for assessing the error margin on the results of numerical modeling of such structures, taking into account the main uncertainty sources, including those concerning soil parameters. As shown in the paper, the Finite Element Method can be coupled with probabilistic techniques such as the Monte-Carlo Method and different perturbation techniques including the First Order-Second Moment method and the Point Estimate method. The advantages and limitations of these different techniques are discussed. The application of the SFEM is briefly illustrated by analyses of uncertainty on the stress field within a simple structure and on the displacement field within an embankment. The results of these analyses and those of similar ones that will be presented in future papers show that, in spite of its current limitations, the SFEM can help geotechnical engineers to introduce a quite necessary additional degree of realism in Finite Element modeling.

Key words : geotechnical engineering, uncertainty, stochastic finite element method, soils, Monte Carlo method, perturbation techniques.

NDLE : Les discussions sur cet article sont acceptées jusqu'au 1^{er} août 2001.

Introduction

L'application de l'informatique et de méthodes numériques telles que les différences finies ou les éléments finis a connu un large développement dans le domaine de la modélisation des déplacements, déformations et contraintes dans les sols et les ouvrages en terre.

Cependant, les sources d'incertitude dans les analyses aux éléments finis en géotechnique sont nombreuses. Le choix des paramètres des lois de comportement à introduire dans le calcul est particulièrement délicat. Les propriétés des sols naturels ou compactés présentent généralement des variations spatiales importantes et les techniques directes ou indirectes utilisées pour les déterminer sont souvent peu fiables. Aux incertitudes sur les propriétés, s'ajoutent celles concernant les sollicitations, les conditions aux limites et la méthode de calcul elle-même. On s'accorde en général à reconnaître que les incertitudes les plus graves sont celles induites par une mauvaise connaissance des propriétés des sols.

Les techniques statistiques et probabilistes permettant d'évaluer l'incertitude sur le comportement des sols et des ouvrages liée aux diverses incertitudes citées ci-dessus dans les analyses aux éléments finis portent globalement le nom de « méthode des éléments finis stochastiques » (MEFS). Comme on le montre dans ce qui suit, les différentes techniques disponibles pour la mise en œuvre de la MEFS ont leur domaine d'application, leurs limitations et leurs mérites respectifs. Les exemples présentés montrent par ailleurs que les résultats typiques obtenus par la MEFS sont souvent d'interprétation délicate.

Les premiers travaux sur l'application de la MEFS à la géotechnique ont été exposés par Cambou et Auvinet (1974, 1975, 1977). Un logiciel a été développé pour l'analyse stochastique des déplacements et des contraintes dans le sol au cours des excavations réalisées pour la construction du métro de Mexico. Une méthodologie analogue a été utilisée par Baecher *et al.* (1981) pour l'étude des incertitudes sur les tassements des fondations. Un logiciel de MFES a été développé par Magnan (1987) (Code de calcul PROBEF). Magnan *et al.* (1995) ont par la suite appliqué la MEFS à l'analyse de la consolidation uni- et bidimensionnelle des sols sous une semelle isolée. Orlandi (1996) et Bouayed (1997) ont abordé l'analyse stochastique des barrages en terre en utilisant la méthode FOSM en élasticité linéaire et non-linéaire mais en se limitant dans ce dernier cas aux incertitudes sur les déplacements. Mellah (1999) a présenté une évaluation de l'applicabilité de la méthode au cas des analyses par éléments finis en élasto-plasticité.

Dans le domaine des écoulements en milieu poreux obéissant à la loi de Darcy, Griffith *et al.* (1993) ont étudié l'influence de la variabilité spatiale de la perméabilité sur l'écoulement de l'eau sous un barrage par simulation de champs stochastiques. Par ailleurs, Bencheikh (1994) et López Acosta (1999) ont appliqué la méthode des éléments finis stochastiques à l'étude des écoulements souterrains dans des milieux à perméabilité incertaine décrite au moyen de variables aléatoires.

Les incertitudes dans les analyses aux éléments finis en géotechnique

2.1

Les principales sources d'incertitude

Si, pour les structures, la diversité des matériaux est relativement réduite et leurs propriétés sont maintenant assez bien connues, à l'opposé, la mécanique des sols doit s'accommoder de matériaux que la nature a dotés de caractéristiques complexes et variables dans l'espace et dans le temps et dont la mesure est délicate. De ce fait, les paramètres mécaniques que l'on doit introduire dans les calculs en géotechnique, et en particulier dans ceux réalisés par la méthode des éléments finis, sont souvent mal connus. On doit ajouter à cela les incertitudes sur les sollicitations et les conditions aux limites ainsi que l'erreur que peuvent introduire les hypothèses et approximations du modèle mécanique utilisé. Les incertitudes associées aux erreurs humaines peuvent également être extrêmement importantes mais elles échappent en général aux modélisations probabilistes ou autres.

2.1.1

Incertitudes sur les paramètres mécaniques des sols

La connaissance des lois de comportement des sols naturels ou compactés joue un rôle fondamental dans les calculs de la géotechnique et en particulier dans les analyses aux éléments finis effectuées en vue de la détermination des champs de déplacements, de contraintes et de déformations au sein des ouvrages et de leur environnement. Les paramètres de la loi de comportement (élastique linéaire ou non-linéaire, élastoplastique, viscoélastique) sont souvent estimés à partir de l'expérience, de corrélations avec les propriétés physiques et, dans le meilleur des cas, à partir d'un nombre limité d'essais en place ou au laboratoire. Il est clair que l'on doit distinguer deux types d'incertitudes : celles qui sont associées aux variations spatiales et celles qui sont dues à la mesure ou à l'estimation des paramètres d'un élément de sol donné.

• Variabilité spatiale

Par nature, les sols sont des matériaux hétérogènes, leurs propriétés mécaniques et physiques présentent une variation souvent considérable d'un point à l'autre du milieu géotechnique étudié. La connaissance de la géologie du site ou du procédé de construction de l'ouvrage permet en général de définir des sous-domaines à caractéristiques à peu près homogènes. Il peut ne s'agir toutefois que d'une homogénéité statistique dissimulant des variations spatiales souvent très significatives. Il est donc nécessaire de reconnaître que les propriétés des sols sont des fonctions aléatoires spatiales.

• Incertitudes sur les mesures et estimations des paramètres

Les lois de comportement courantes sont des modèles qui donnent du comportement réel du sol une

description plus ou moins satisfaisante. Dans la plupart des cas, les courbes expérimentales décrivant le comportement du sol au cours d'essais de laboratoire ou en place sont ajustées à celles prévues par modèle choisi, afin de déterminer les paramètres mécaniques du sol considéré. L'écart observé entre le comportement réel du sol et la réponse du modèle choisi est dû principalement :

- aux erreurs aléatoires commises lors des essais réalisés ;
- aux erreurs systématiques dues à un biais dans la mesure, souvent inhérent à l'essai réalisé (on observe par exemple des différences systématiques entre la résistance au cisaillement mesurée au scissomètre et au triaxial) ou lié au remaniement des échantillons et autres facteurs similaires.

Un biais peut également être introduit lorsque la loi de comportement ajustée est choisie de façon trop arbitraire et surtout quand les paramètres du sol sont estimés de façon subjective (jugement d'expert) ou à partir de corrélations statistiques entre propriétés physiques et mécaniques. Ce biais, le plus souvent mal connu, donc aléatoire, peut être additif (résultat de l'essai ou de l'estimation valide à une constante aléatoire près) ou, le plus souvent, multiplicatif (erreur systématique aléatoire proportionnelle au résultat de la mesure ou de l'estimation).

2.1.2

Incertitudes sur les sollicitations et conditions aux limites

Les forces extérieures et les forces de volume à prendre en compte dans les analyses géotechniques réalisées par la méthode des éléments finis, sont souvent elles aussi mal connues. Il en est ainsi des forces tectoniques ou des contraintes transmises au sol par les fondations, fréquemment estimées à partir de descentes de charges approximatives. Les conditions aux limites elles-mêmes ne peuvent souvent être représentées que de façon approchée. Les incertitudes ainsi introduites sont souvent difficiles à évaluer. Si l'on dispose des éléments nécessaires, ce type d'incertitude peut être représenté par un vecteur de variables aléatoires susceptibles d'être prises en compte dans les analyses stochastiques (Cornell, 1971).

2.1.3

Incertitudes sur la méthode de calcul

Il existe des méthodes théoriques pour déterminer l'erreur implicite dans les analyses aux éléments finis *a priori* et *a posteriori* (Mestat et Prat, 1999). On trouve par ailleurs dans la littérature, mais relativement dispersées, diverses études de sensibilité sur les différentes caractéristiques d'une modélisation numérique (maillages, méthode d'intégration, critère de plasticité...). En général, ces études ne sont pas réalisées de manière systématique et ne se prêtent donc guère au traitement probabiliste et statistique, qui permettrait une étude des facteurs susceptibles d'introduire un biais dans les calculs aux éléments finis. L'approche de ce problème est souvent empirique et passe par la validation des résultats sur essais, ouvrages types, et ouvrages réels. Si les éléments nécessaires sont réunis, il est toutefois possible de représenter l'erreur sur la méthode de calcul au moyen d'un terme aléatoire d'erreur multiplicatif ou additif (Cornell, 1971).

2.1.4

Conclusion

Nous avons souligné l'importance que peuvent revêtir les incertitudes sur les propriétés du sol, mais aussi sur les sollicitations, les conditions aux limites et sur la validité du modèle lui-même dans les analyses aux éléments finis. Il semble qu'en géotechnique, le plus souvent, l'incertitude sur les paramètres mécaniques des matériaux considérés soit la plus importante. Dans la suite, nous nous intéresserons donc plus particulièrement à la modélisation de cette incertitude en vue de sa prise en compte dans les analyses aux éléments finis.

2.2

Modélisation de l'incertitude sur les propriétés des sols

2.2.1

Champ stochastique

Soit $V(X)$, la valeur prise par une propriété du sol donnée au point X d'un domaine R^p ($p = 1, 2, 3$). En chaque point, cette valeur, en général non ou mal connue, peut être considérée comme une variable aléatoire. L'ensemble des variables aléatoires du domaine constitue une *champ stochastique* qui peut être décrit au moyen des éléments suivants :

- Espérance mathématique : $\mu_{V(X)} = E[V(X)]$ (1)
- Variance : $\sigma_{V(X)}^2 = \text{Var}[V(X)]$ (2)
- Auto-corrélation : $R_{VV}(X_1, X_2) = E[V(X_1)V(X_2)]$ (3)
- Auto-covariance : $C_{VV}(X_1, X_2) = E\left[\left(V(X_1) - \mu_{V(X_1)}\right)\left(V(X_2) - \mu_{V(X_2)}\right)\right]$ (4)

Les deux fonctions précédentes représentent le degré de dépendance linéaire entre les valeurs des propriétés mesurées en deux points différents dans le domaine considéré.

- Coefficient d'auto-corrélation :

$$\rho_{VV}(X_1, X_2) = \frac{C_{VV}(X_1, X_2)}{\sigma_{V(X_1)}\sigma_{V(X_2)}} \quad (5)$$

Ce coefficient représente l'auto-covariance normalisée, adimensionnelle, dont les valeurs restent comprises entre -1 et +1.

- Fonction de distribution cumulée de probabilité : $F_{V(X_1), V(X_2), \dots, V(X_n)}(v_1, v_2, \dots, v_n; X_1, X_2, \dots, X_n) = P[(V(X_1) \leq v_1, V(X_2) \leq v_2, \dots, V(X_n) \leq v_n)]$ (6)

On suppose le plus souvent en pratique que le champ présente une certaine régularité. On admet qu'il existe une homogénéité statistique, quitte à subdiviser le milieu en plusieurs sous-domaines. L'hypothèse simplificatrice la plus courante est celle de la stationnarité au sens large. L'espérance est considérée comme constante dans le domaine étudié, éventuellement après élimination d'une dérive, et l'on admet que l'auto-

covariance dépend uniquement de la distance vectorielle t entre les points X_1 et X_2 , on a alors :

$$C_{VV}(X_1, X_2) = C_{VV}(X_2 - X_1) = C_{VV}(X_2 - X_1) = C_{VV}(\tau) \text{ (ou plus simplement } C(\tau)) \quad (7)$$

À l'origine, cette fonction n'est autre que la variance de $V(X)$.

$$\text{Var}[V(X)] = C(O) \quad (8)$$

et le coefficient d'auto-corrélation s'écrit comme suit :

$$\rho_{VV}(X_1, X_2) = \rho_{VV}(\tau) \text{ ou plus simplement } \rho(\tau) \quad (9)$$

La représentation graphique de cette fonction reçoit souvent le nom de « corrélogramme ».

Sous le même formalisme mathématique, l'interprétation donnée à ce champ peut être extrêmement variable. Suivant le sens attribué au concept de probabilité, qui doit bien sûr être clairement défini avant toute analyse probabiliste, on est amené en fait à considérer plusieurs types de champs aléatoires.

a) Champs estimatifs

Lorsque les propriétés des sols sont estimées à partir d'opinions basées principalement sur l'expérience et les observations d'un expert, le champ stochastique est de type estimatif. Ses paramètres doivent alors refléter les connaissances mais aussi l'ignorance de l'expert. Cette situation est courante en géotechnique car il n'est pas toujours possible de réaliser un nombre de mesures suffisant pour pouvoir s'affranchir d'opinions largement subjectives. Dans ce cas, les probabilités reflètent avant tout le « degré de confiance » (*degree of belief*) que méritent les estimations avancées par l'expert.

Au moment du projet des ouvrages en terre, on ne dispose bien entendu d'aucunes données sur la variabilité spatiale des propriétés, à l'exception du zonage général prévu par le projet lui-même. Dans cette situation, qui est une de celles où la modélisation stochastique présente le plus grand intérêt, seul l'emploi d'un champ estimatif reflétant les incertitudes *a priori* sur les propriétés des matériaux des différentes zones est alors possible.

b) Champs descriptifs de la variabilité spatiale

Lorsque l'on dispose d'un nombre suffisant de mesures fiables sur la variabilité spatiale des propriétés du sol, le champ aléatoire peut être de type descriptif et être utilisé pour décrire les incertitudes résiduelles. Les paramètres du champ peuvent alors être estimés par analyse statistique des données. C'est le domaine de la *géostatistique* au sens qui lui a été donné durant ces dernières décennies. Dans ce cas, les probabilités doivent être interprétées en termes de fréquence relative. On soulignera que les champs stochastiques décrivant la variabilité spatiale des propriétés des sols ne sont connus que par l'intermédiaire de mesures ou d'estimations qui sont elles-mêmes entachées d'incertitude. On ne peut donc avoir accès qu'à un champ apparent dont les caractéristiques diffèrent de celles du champ réel.

Dans la pratique de la géotechnique, l'utilisation de ces champs descriptifs se heurte souvent à de sérieuses difficultés pour les raisons suivantes :

– l'estimation des paramètres du champ doit suivre les règles applicables à la prise d'échantillons représenta-

tifs aléatoires ou systématiques (Cochran, 1977). Il est courant en géotechnique que le nombre de données disponibles ne soit pas suffisant pour que l'estimation soit précise ; une erreur de type statistique est alors introduite dans l'estimation ;

– à l'inverse, quand les données sont nombreuses et fiables, il devient possible d'estimer les paramètres du champ avec précision mais c'est alors toute l'approche probabiliste de la variabilité spatiale qui perd de son intérêt du point de vue de l'ingénieur ;

– la corrélation entre les valeurs du champ associées à deux points du milieu n'est pas une propriété intrinsèque de ces deux points. Elle dépend bien entendu du domaine (population) dans lequel elle est définie et duquel elle ne peut être dissociée. Ainsi, dans un milieu fortement contrasté du point de vue stratigraphique, les propriétés de deux points appartenant à une même couche peuvent être fortement corrélées, mais ne plus l'être du tout dans une analyse de corrélation ne portant plus que sur la couche où ils se trouvent. Un sondage trop court ou des mesures faites dans un seul plan horizontal ou tout autre zone particulière du sol peuvent donner une image totalement incorrecte de la variabilité et de la structure de corrélation de l'ensemble du milieu ;

– à défaut de mesures, il est parfois possible d'estimer l'ordre de grandeur des distances d'auto-corrélation. On dispose maintenant de nombreux résultats obtenus sur différents types de sols mis en place par diverses techniques et concernant en particulier la portée ou distance d'influence de la corrélation (Rossa, 1988 ; Favre, 1987 ; Cherubini, 1993). Il est toutefois indispensable que les conditions géotechniques du site étudié soient vraiment similaires à celles des sites où les mesures ont été réalisées pour que ces valeurs puissent être adoptées ;

– remarquons par ailleurs que les champs stochastiques tridimensionnels ne peuvent être pris en compte dans les analyses courantes aux éléments finis en déformation plane. En effet ce type d'analyse suppose une homogénéité parfaite dans la direction perpendiculaire au plan de l'analyse, ce qui contredit l'hypothèse de variation aléatoire spatiale.

c) Champs mixtes

En pratique, la situation est souvent intermédiaire entre les deux précédentes. Les propriétés des sols sont définies par une combinaison d'opinions (probabilités *a priori*) et de résultats expérimentaux qui peuvent être introduits dans une analyse bayésienne explicite ou implicite permettant de définir des probabilités *a posteriori* (tenant compte des mesures). L'interprétation donnée au concept de probabilité devient dans ce cas plus ambigu mais ce concept reste malgré tout utile dans la pratique (Ditlevsen, 1996).

2.2.2

Discretisation pour les analyses aux éléments finis

Pour introduire l'incertitude sur les propriétés des sols dans un modèle de calcul, il est nécessaire de passer de la représentation du champ aléatoire continu à un nombre limité de variables aléatoires. C'est ainsi que, pour les calculs aux éléments finis, on devra finalement associer à chaque élément ou groupe d'éléments correspondant à un volume de sol donné, des

paramètres mécaniques aléatoires représentatifs. Le problème de la définition de ces valeurs représentatives est délicat car il est souvent de nature tout autant mécanique que statistique. Les méthodes de discrétisation les plus courantes ignorent en fait l'aspect mécanique du problème qui ressort des techniques d'homogénéisation.

Soit $V(X); X \in \Omega$, un champ aléatoire multidimensionnel, défini dans un domaine Ω et représenté par son espérance $\mu_V(X)$, sa variance $\sigma_V^2(X)$ et son coefficient d'autocorrelation $\rho_{VV}(X, X')$. V est le vecteur des variables qui permettent de représenter le champ stochastique, défini par un vecteur moyen μ et une matrice de covariance T_{VV} . Il existe diverses méthodes de discrétisation reliant $V(X)$ à V (Der Kiureghian *et al.*, 1988). Nous n'examinerons ici que les plus simples.

• Méthode des points moyens

La valeur du champ dans le domaine Ω_e d'un élément ou groupe d'éléments est décrite par une simple variable aléatoire représentant la valeur au point central du sous-domaine considéré :

$$\tilde{V}(X) = V(X_c) \text{ avec } X \in \Omega_e \quad (10)$$

Le vecteur V est formé par toutes les variables aléatoires $V(X_c)$ des sous-domaines considérés. La réalisation du champ ainsi définie est une fonction en escalier présentant des discontinuités aux frontières des éléments. La moyenne et la matrice de covariance du vecteur sont données en fonction de la moyenne, de la variance et des coefficients d'auto-corrélation du champ, évalués au centre de gravité des sous-domaines.

• Méthode des moyennes spatiales

Cette méthode a été proposée par Cornell (1971), puis reprise par Vanmarcke *et al.* (1983). Le champ est décrit à l'intérieur du sous-domaine considéré par la moyenne spatiale du champ :

$$\tilde{V}(X) = \bar{V}_e = \frac{\int_{\Omega_e} V(X) d\Omega_e}{\int_{\Omega_e} d\Omega_e} \text{ avec } X \in \Omega_e \quad (11)$$

L'ensemble des valeurs moyennes \bar{V}_e des variables aléatoires forme le vecteur V ; les valeurs statistiques μ et T_{VV} (espérance et matrice de covariance) du vecteur V sont obtenues par intégration des moments du champ aléatoire.

La variance de la valeur moyenne d'une propriété aléatoire dans un domaine donné tend à diminuer quand les dimensions du domaine considéré augmentent (sauf dans le cas trivial de la corrélation parfaite). Dans la littérature, ce phénomène est connu sous le nom de « réduction de variance ». Dans la mesure où l'on a recours à des moyennes spatiales pour les calculs aux éléments finis, les incertitudes résiduelles les plus significatives sont donc souvent celles correspondant aux biais des tests ou des corrélations, qui eux ne souffrent pas de réduction de variance, ce qui peut justifier dans de nombreux cas, au moins en géotechnique,

l'abandon de la modélisation par champ stochastique et son remplacement par une modélisation plus simple par variables aléatoires.

• Valeurs ponctuelles et moyennes spatiales définies dans des champs conditionnels

Dans les paragraphes précédents, on a supposé que le modèle stochastique de l'incertitude spatiale avait pu être établi à partir de certaines données ponctuelles ou continues (sondages) et on a examiné ses possibilités d'utilisation pour les modélisations stochastiques dans les analyses aux éléments finis. Cette approche ignore le fait que l'existence de ces données réduit considérablement l'incertitude au voisinage des points où elles ont été obtenues. Il est évident que l'on peut faire mieux et estimer les paramètres à attribuer aux éléments finis ou groupes d'éléments finis en prenant en compte directement les résultats des mesures. C'est l'objectif des techniques d'estimation conditionnelle, telles que le *krigeage* (Matheron, 1971) qui peuvent être utilisées pour associer à chaque point du milieu une espérance et une variance conditionnelle de la propriété considérée ou pour attribuer à chaque sous-domaine considéré (élément fini ou groupe d'éléments finis) une espérance et variance conditionnelle de la valeur moyenne de cette propriété sur ce sous-domaine.

2.2.3

Conclusion

Les mesures et estimations et la variabilité spatiale des propriétés des sols introduisent une incertitude importante dans les analyses aux éléments finis. L'incertitude sur les propriétés des sols peut être représentée par des champs stochastiques. Néanmoins, dans la pratique, la description de l'incertitude devra finalement être réalisée au moyen d'un nombre limité de variables aléatoires.

3

La méthode des éléments finis stochastiques

Le couplage entre les méthodes probabilistes et la méthode des éléments finis a donné naissance à la « méthode des éléments finis stochastiques » (MEFS). Les techniques permettant d'appliquer la méthode des éléments finis stochastiques (MEFS), peuvent généralement être divisées en deux catégories (Teigen *et al.* 1991) : les méthodes des perturbations et les méthodes de simulation. Par ailleurs, la méthode peut être couplée avec un calcul de fiabilité (Flores, 1994 ; Lemaire, 1997). La littérature consacrée au développement de ces méthodes et, dans une moindre mesure, à leurs applications dans différents domaines est abondante. Nous citerons entre autres : Vanmarcke *et al.* (1983) et Bittnar (1996).

3.1

Méthode des perturbations

La méthode des perturbations est basée sur une représentation simplifiée des fonctions de distribution

de probabilité de fonctions de variables aléatoires, au moyen des nièmes premiers moments (le nombre dépendant de la technique utilisée). L'estimation de ces premiers moments est obtenue en ajoutant à un terme central moyen les effets « perturbateurs » introduits par chaque variable aléatoire considérée séparément.

3.1.1

Méthode classique des perturbations (Mathews, 1964)

La méthode est basée sur l'approximation de la densité de probabilité d'une fonction de variable aléatoire à partir du développement limité en série de Taylor de cette fonction.

La résolution du système d'équations de base de la méthode des éléments finis en élasticité linéaire permet de déterminer le vecteur du champ des déplacements et par conséquent ceux des contraintes et des déformations

$$\{U\} = [K]^{-1} \{F\} \quad (12)$$

avec :

$\{U\}$: vecteur du champ de déplacements aux nœuds du système ;

$\{F\}$: vecteur des forces extérieures et volumiques appliquées aux nœuds ;

$[K]$: matrice de rigidité fonction du vecteur de variables aléatoires $x = (x_1, x_2, \dots, x_n)^T$ qui représentent les aléas du système.

Si les aléas sont assez petits et présentent une moyenne nulle, la matrice de rigidité peut se mettre sous la forme :

$$[K] = [K^0] + \sum_{i=1}^n [K_i^I] x_i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n [K_{ij}^{II}] x_i x_j + \dots \quad (13)$$

n : nombre de variables aléatoires,

$[K^0]$: matrice de rigidité évaluée en $x = (x_1, x_2, \dots, x_n)^T = 0$.

$[K_i^I]$ et $[K_{ij}^{II}]$: dérivées partielles de la matrice de rigidité $[K]$, définies par :

$$[K_i^I] = \left. \frac{\partial [K]}{\partial x_i} \right|_{x=0} ; [K_{ij}^{II}] = \left. \frac{\partial^2 [K]}{\partial x_i \partial x_j} \right|_{x=0} \quad (14)$$

Le vecteur du champ de déplacements $\{U\}$ prend alors la forme suivante :

$$\{U\} = \{U^0\} + \sum_{i=1}^n \{U_i^I\} x_i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \{U_{ij}^{II}\} x_i x_j + \dots \quad (15)$$

Les vecteurs des coefficients $\{U^0\}$, $\{U_i^I\}$, $\{U_{ij}^{II}\}$ sont évalués après avoir résolu le système d'équations suivant :

$$\{U^0\} = [K^0]^{-1} \{F^0\}$$

$$\{U_i^I\} = [K^0]^{-1} (\{F_i^I\} - [K_i^I] \{U^0\}) \quad (16)$$

$$\{U_{ij}^{II}\} = [K^0]^{-1} (\{F_{ij}^{II}\} - [K_i^I] \{U_j^I\} - [K_j^I] \{U_i^I\} - [K_{ij}^{II}] \{U^0\})$$

Une fois $[K^0]^{-1}$ déterminée, on peut évaluer tous les déplacements $\{U^0\}$, $\{U_i^I\}$, $\{U_{ij}^{II}\}$. De la même façon, on peut déterminer les vecteurs du champ des contraintes $\{\sigma\}$ et du champ des déformations $\{\epsilon\}$ à partir des équations de la théorie d'élasticité.

La détermination des deux premiers moments de la réponse et de leurs intervalles de confiance dépend du degré d'approximation de la fonction considérée :

Si l'approximation souhaitée est du premier ordre, le vecteur du champ des déplacements s'écrit comme suit :

$$\{U\} = \{U^0\} + \sum_{i=1}^n \{U_i^I\} x_i \quad (17)$$

$$\text{avec : } E^I [U] = \{U^0\} \quad (18)$$

$$\begin{aligned} \text{et } Cov^I [U, U] &= E \left[(\{U\} - E^I [U]) (\{U\} - E^I [U])^T \right] \\ &= \sum_{i=1}^n \sum_{j=1}^n \{U_i^I\} \{U_j^I\} E [x_i x_j] \end{aligned} \quad (19)$$

Si l'approximation est du deuxième ordre, le développement en série de Taylor jusqu'au troisième terme nous permet d'écrire :

$$\{U\} = \{U^0\} + \sum_{i=1}^n \{U_i^I\} x_i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \{U_{ij}^{II}\} x_i x_j + \dots \quad (20)$$

$$\text{avec : } E^{II} [U] = E^I [U] + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \{U_{ij}^{II}\} x_i x_j + \dots \quad (21)$$

$$\begin{aligned} Cov^{II} [U, U] &= Cov^I [U, U] + \frac{1}{4} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n \{U_{ij}^{II}\} \{U_{kl}^{II}\} \\ &\quad (E [x_i x_j x_k] + E [x_i x_j x_l]) \end{aligned} \quad (22)$$

L'application de cette méthode demande bien entendu le calcul de toutes les dérivées partielles premières et secondes de la matrice de rigidité.

3.1.2

Méthode du premier ordre – seconds moments (FOSM ; Cornell, 1971)

La méthode FOSM, proposée par Cornell (1971), est également basée sur le développement en série de Taylor. Le caractère aléatoire est analysé en termes d'espérance mathématique, de variance et de covariances des paramètres.

Le principe de cette méthode est d'approximer la densité de probabilité d'une fonction d'une variable aléatoire, $g(x)$, en remplaçant cette dernière par les deux ou trois premiers termes de son développement en série de Taylor au voisinage de la moyenne de x_i .

$$g(x_i) = g(\mu_{x_i}) + (x_i - \mu_{x_i}) \left[\frac{\partial g(x_i)}{\partial x_i} \right]_{\mu_{x_i}} + \frac{(x_i - \mu_{x_i})^2}{2} \left[\frac{\partial^2 g(x_i)}{\partial x_i^2} \right]_{\mu_{x_i}} + \dots \quad (23)$$

μ_{x_i} : la valeur moyenne de la variable aléatoire x_i .

On peut distinguer les cas suivants :

Dans le cas d'une fonction $y = g(x)$ d'une seule variable aléatoire x (et éventuellement d'autres variables non-aléatoires), l'expression (23) prend la forme suivante :

$$y = g(x) = g(\mu_x) + (x - \mu_x) \left[\frac{\partial g(x)}{\partial x} \right]_{\mu_x} + \frac{(x - \mu_x)^2}{2} \left[\frac{\partial^2 g(x)}{\partial x^2} \right]_{\mu_x} + \dots \quad (24)$$

et :

$$E[y] = E[g(x)] \approx g(\mu_x) + \frac{\sigma_x^2}{2} \left[\frac{\partial^2 g(x)}{\partial x^2} \right]_{\mu_x} \quad (25)$$

$$\text{Var}[y] = \text{Var}[g(x)] \approx \sigma_x^2 \left[\frac{\partial g(x)}{\partial x} \right]_{\mu_x}^2 \quad (26)$$

Si la fonction dépend de plusieurs variables aléatoires $y = g(x_1, x_2, \dots, x_n)$, le développement en série de Taylor donne :

$$E[y] = E[g(x_1, x_2, \dots, x_n)] \approx g(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_n}) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \left[\frac{\partial^2 g(x_k)}{\partial x_i \partial x_j} \right]_{\mu_{x_k} \mu_{x_j}} \text{Cov}[x_i, x_j] + \dots$$

$$\text{Var}[y] = \text{Var}[g(x_1, x_2, \dots, x_n)] \approx \sum_{i=1}^n \sum_{j=1}^n \left[\frac{\partial g(x_k)}{\partial x_i} \right]_{\mu_{x_k}} \left[\frac{\partial g(x_k)}{\partial x_j} \right]_{\mu_{x_k}} \text{Cov}[x_i, x_j] \quad (28)$$

avec n : nombre total de variables aléatoires.

La méthode se généralise pour les fonctions multiples de plusieurs variables aléatoires.

Pour chaque fonction :

$$E[y_k] = E[g_k(x_1, x_2, \dots, x_n)] \approx g_k(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_n}) \quad (29)$$

$$\text{Var}[y_k] \approx \sum_{i=1}^n \sum_{j=1}^n \left[\frac{\partial g_k}{\partial x_i} \right]_{\mu_{x_i}} \left[\frac{\partial g_k}{\partial x_j} \right]_{\mu_{x_j}} \text{Cov}[x_i, x_j] \quad (30)$$

et :

$$\text{Cov}[y_k, y_l] \approx \sum_{i=1}^n \sum_{j=1}^n \left[\frac{\partial g_k}{\partial x_i} \right]_{\mu_{x_i}} \left[\frac{\partial g_l}{\partial x_j} \right]_{\mu_{x_j}} \text{Cov}[x_i, x_j] \quad (31)$$

Remarquons que cette méthode ne permet d'obtenir la forme exacte de la densité de probabilité que dans le cas où celle-ci est gaussienne.

• Application de la méthode FOSM en élasticité linéaire (Cambou, 1977)

Pour un milieu continu et isotrope et suivant une loi de comportement élastique linéaire, le principe de base des éléments finis en élasticité est la résolution du système d'équations (12) :

$$\{U\} = [K^{-1}(x_1, x_2, \dots, x_n)] \{F\} = \{g(x_1, x_2, \dots, x_n)\}$$

Il est possible d'estimer les premiers moments, ainsi que les covariances des divers résultats de l'analyse, en se basant sur le développement en série de Taylor d'ordre 2 :

$$E[U_k] = E[g_k(x_1, x_2, \dots, x_n)] \approx g_k(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_n}) \quad (32)$$

$$\text{Var}[U_k] \approx \sum_{i=1}^n \sum_{j=1}^n \left[\frac{\partial g_k}{\partial x_i} \right]_{\mu_{x_i}} \left[\frac{\partial g_k}{\partial x_j} \right]_{\mu_{x_j}} \text{Cov}[x_i, x_j] \quad (33)$$

$$\text{Cov}[U_k, U_l] \approx \sum_{i=1}^n \sum_{j=1}^n \left[\frac{\partial g_k}{\partial x_i} \right]_{\mu_{x_i}} \left[\frac{\partial g_l}{\partial x_j} \right]_{\mu_{x_j}} \text{Cov}[x_i, x_j] \quad (34)$$

En dérivant les équations du système, on obtient :

$$\{F_f\} = [K] \{U_f\} = \{U_f\} \quad (35)$$

$$\text{avec : } \{U_f\} = \frac{\partial \{U\}}{\partial x_i} \text{ et } \{F_f\} = - \frac{\partial \{K\}}{\partial x_i} \{U\}$$

$\{U_f\}$ et $\{F_f\}$ définissent respectivement des déplacements et forces fictives qui peuvent être utilisés pour obtenir les dérivées partielles des déplacements par rapport à chaque paramètre aléatoire en ayant recours au même algorithme que pour la résolution du système (12). Ces dérivées partielles peuvent alors être insérées dans les équations (32) à (34).

On voit que la méthode se ramène en fait à celle des perturbations classiques.

De la même façon, pour l'estimation des incertitudes sur les contraintes, on obtient, en dérivant le système d'équations général de l'élasticité par rapport aux caractéristiques aléatoires du milieu :

$$\frac{\partial \{\sigma\}}{\partial x_i} = \frac{\partial [D][B]}{\partial x_i} \{U\} + [D][B] \frac{\partial \{U\}}{\partial x_i} \quad (36)$$

Avec :

$\{\sigma\}$: tenseur des contraintes ;

$[D]$: matrice d'élasticité ;

$[B]$: matrice géométrique.

Pour l'estimation des incertitudes sur les déformations, on obtient :

$$\frac{\partial \{\epsilon\}}{\partial x_i} = \frac{\partial [B]}{\partial x_i} \{U\} + [B] \frac{\partial \{U\}}{\partial x_i} \quad (37)$$

Les dérivées des contraintes et des déformations obtenues à partir de ces équations peuvent être insérées dans des développements de Taylor analogues à ceux des déplacements. Ces différentes dérivées peuvent être déterminées explicitement en différenciant la matrice de rigidité $[K]$, la matrice $[B]$ et la matrice $[D]$ par rapport aux variables aléatoires x_i (Bouayed, 1997).

• **Application de la méthode FOSM en présence de non-linéarités géométriques ou mécaniques**

Si des non-linéarités géométriques ou mécaniques sont présentes dans les modèles utilisés, la détermination des déplacements en chaque nœud par la méthode des éléments finis se fait de manière itérative. Le calcul explicite des différentes dérivées des fonctions du développement de Taylor devient impossible. La méthode FOSM reste cependant applicable. Il devient alors nécessaire de calculer les dérivées des développements de Taylor numériquement. À cet effet, différents auteurs ont proposé des méthodes numériques qui permettent de calculer les différentes dérivées d'une manière approchée. Les plus utilisées sont celles des approximations d'Evans (1967) et la méthode des rapports polynomiaux (Chowdury et al. 1993). Cette dernière méthode a été utilisée dans de nombreuses études et a donné des résultats satisfaisants, tant pour des analyses de sensibilité (Orlandi, 1996 ; Bouayed, 1997) que pour des analyses de fiabilité (Nechneh, 1994; Benmansour, 1996). On doit souligner cependant que, dans certains cas, la méthode, comme toute technique numérique, peut se heurter localement à certaines difficultés si la fonction implicite reliant les résultats de l'analyse numérique et les paramètres aléatoires n'a pas un comportement régulier. Remarquons aussi que les méthodes numériques pour le calcul des dérivées présentent par rapport aux méthodes de dérivation explicite, l'avantage de pouvoir être appliquées de façon « externe » au code de calcul par éléments finis, sans qu'il soit nécessaire de modifier ce code.

3.13

Méthode des approximations ponctuelles

Rosenblueth (1975) a présenté une méthode qui permet d'estimer les premiers moments d'une fonction de variables aléatoires, en connaissant seulement les trois premiers moments de chaque variable et se basant sur une approximation au moyen de valeurs ponctuelles, c'est-à-dire en utilisant une distribution de probabilité discrète à mêmes moments. Il a montré qu'une formulation générale de la méthode est possible avec un nombre quelconque de points d'estimation pour chaque variable aléatoire. La démarche générale est représentée sur la figure 1.

La méthode devient malheureusement trop lourde dès que le nombre de variables aléatoires dépasse la dizaine, surtout s'il s'agit de variables corrélées entre elles.

Dans la même ligne, Bolle (1988) a proposé une méthode dite des « perturbations indépendantes » basée sur la discrétisation en trois points de la fonction de distribution de probabilité d'une variable aléatoire et faisant appel au développement limité en série de Taylor et à la transformation des variables de base en variables principales non corrélées. Les premiers moments de la distribution de la fonction $f(x)$ sont

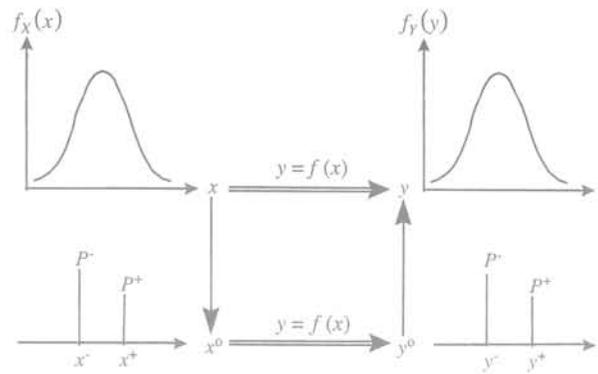


FIG. 1 Approximation bionctuelle de Rosenblueth

Rosenblueth's two-point estimates method.

déterminés en ajoutant à la valeur de la fonction calculée au point central de la distribution discrétisée la somme des différents termes dispersants, ou perturbateurs, liés à chaque variable aléatoire indépendante (additivité des effets de chaque variable aléatoire). La méthode des perturbations indépendantes est très avantageuse si le nombre de variables considérées est important.

3.2

Méthodes de simulation

Pour se libérer des contraintes liées à l'utilisation des méthodes de transformation analytiques et numériques disponibles, dont le champ d'application est parfois trop restreint et qui sont difficiles à mettre en œuvre, il est toujours possible d'avoir recours aux techniques de simulation basées sur la méthode de Monte-Carlo. Un jeu de valeurs des variables aléatoires représentant les champs stochastiques est obtenu au hasard, mais en respectant les densités de probabilité et les corrélations, et l'analyse numérique du problème est effectuée. On répète cette opération jusqu'à ce qu'une tendance statistique se définisse et permette de juger de la dispersion des résultats. Si le vecteur de variables aléatoires considéré est à composantes indépendantes, la simulation se fait pour chacune des composantes. Si le vecteur est à composantes dépendantes, on peut se ramener au cas simple précédent, en utilisant la méthode de Cholesky généralisée qui permet de diagonaliser la matrice de covariance. Cette méthode n'est valable que si le vecteur de variables aléatoires est gaussien. Dans le cas contraire, il existe dans la littérature des méthodes de simulation propres à chacune des lois utilisées. On peut par ailleurs se ramener au cas gaussien en appliquant différentes transformations (Fogli, 1980).

L'inconvénient majeur de la méthode de Monte-Carlo est sa convergence extrêmement lente. Plusieurs techniques existent pour déterminer le nombre de simulations nécessaires pour une approximation fixée par un intervalle de confiance. Parmi les méthodes les plus pratiques, citons la méthode classique de convergence (intervalle de confiance d'une valeur moyenne espérée) et la méthode de Chebishev.

Applications

Pour illustrer les techniques exposées dans les paragraphes précédents nous nous limiterons à présenter deux exemples simples.

4.1

Analyse stochastique d'une structure simple

Considérons une plaque de chant sur deux appuis, l'un fixe et l'autre glissant sur un plan horizontal, et soumise à une force localisée (Fig. 2). La plaque est constituée par deux matériaux à module d'Young aléatoire, mais à même espérance et avec un coefficient de variation (écart type divisé par espérance) $CV = 10\%$ et corrélés négativement ($\rho = -1$).

Les diagrammes de la figure 3 représentent l'espérance des contraintes horizontales et verticales et des cisaillements (colonne de gauche), et leur écart type (colonne de droite) obtenu par la méthode FOSM.

On est frappé par la complexité de la répartition spatiale des incertitudes. Celle-ci obéit en effet à de multiples facteurs dont les effets ne sont pas toujours mis en valeur de façon suffisamment précise par la discrétisation en éléments finis ni par l'algorithme d'interpolation (*Surfer*) utilisé :

– conditions aux limites concernant les contraintes (pas de contraintes verticales sur les faces horizontales de la

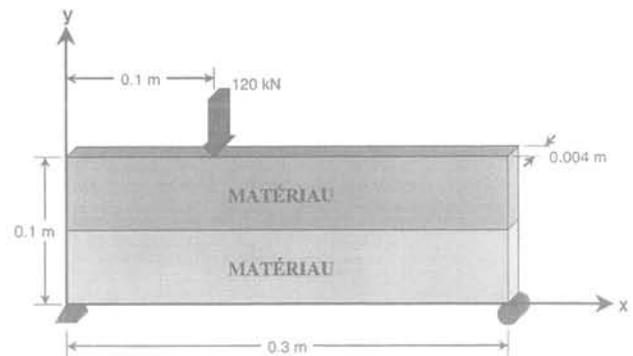


FIG. 2 Plaque sur deux appuis constituée de deux matériaux différents.
Plate on two supports made of two different materials.

plaque, pas de contraintes horizontales sur les côtés verticaux, pas de cisaillements sur tous les côtés de la plaque). Ces conditions aux limites éliminent localement l'incertitude ;

– perte d'homogénéité lors des variations aléatoires des deux modules, créant des sollicitations d'interaction parasites à l'interface, qui amplifient les incertitudes initiales ;

	E(E) MPa	CV(E)	E(v)	CV (v)	$\rho (E_1, E_2)$
MATÉRIAU 1	21×10^4	0.10	0.30	0	-1
MATÉRIAU 2	21×10^4	0.10	0.30	0	

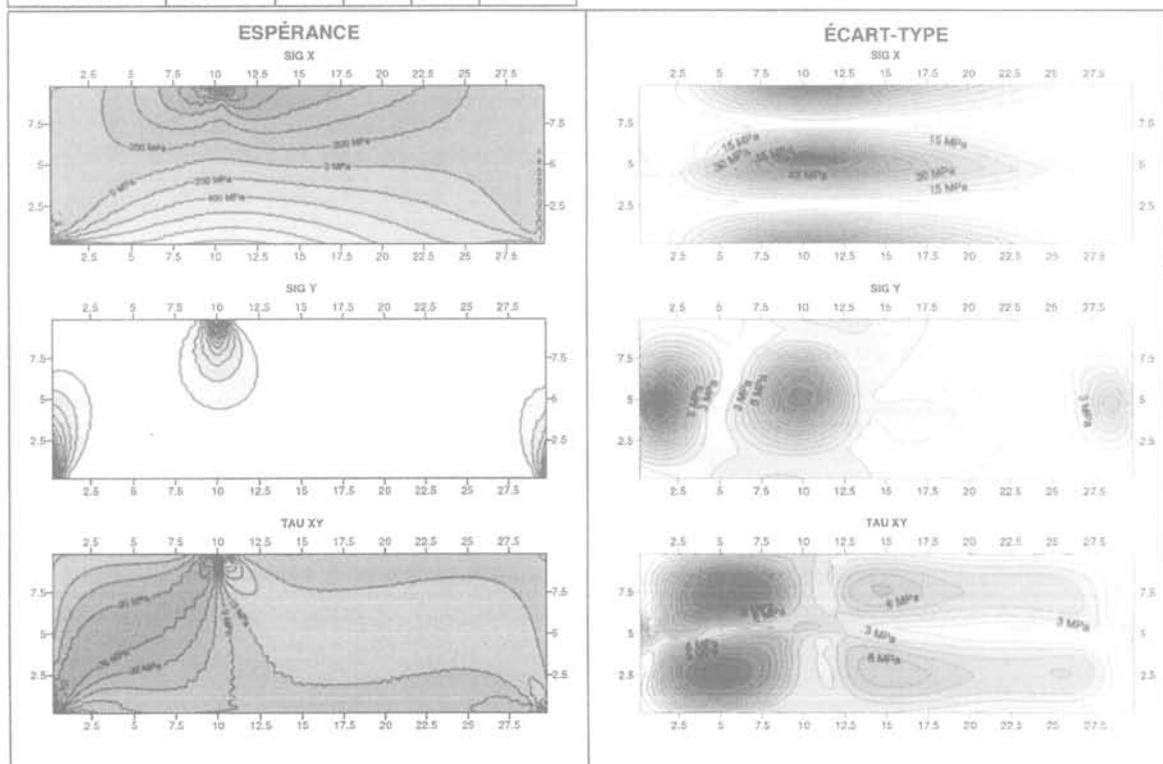


FIG. 3 Espérance et écart type des contraintes verticales (sigy), horizontales (sigx) et de cisaillement (tauxy).
Expected value and standard deviation of vertical (sigy), horizontal (sigx) and shear stresses (tauxy).

– effets de moyenne atténuant l'incertitude au voisinage de l'interface (voir cisaillement).

Seule la technique des éléments finis permet de saisir ce comportement complexe des incertitudes, difficile à prévoir *a priori* et d'interprétation délicate.

4.2

Analyse stochastique d'un remblai

Pour l'analyse des structures en terre, telles que les remblais et barrages en terre, des analyses ont pu être réalisées en utilisant la méthode de Monte-Carlo et la méthode FOSM. La méthode FOSM peut être mise en œuvre en procédant au calcul numérique des dérivées par la méthode des quotients polynomiaux, ce qui permet de ne pas modifier les logiciels classiques aux éléments finis tels que FEADAM84 ou CESAR-LCPC. L'emploi de ces logiciels comme sous-programmes d'une analyse d'incertitude pose toutefois de redoutables problèmes de couplage. Pour illustrer l'application de ces méthodes, nous présenterons certains résultats concernant les déplacements dans un remblai en cours de construction.

On considère ici un remblai constitué par deux matériaux mis en place par couches (Fig. 4). On admet que les deux matériaux suivent une loi non-linéaire simple de type Kondner-Duncan (Duncan, 1984). L'espérance des valeurs des neuf paramètres de cette loi est indiquée dans le tableau. Par simplicité, on a admis ici que ces paramètres présentaient tous un coefficient de variation de 10 %. Pour un matériau donné, les valeurs des coefficients de corrélation entre paramètres ont pu être estimées par une analyse statistique

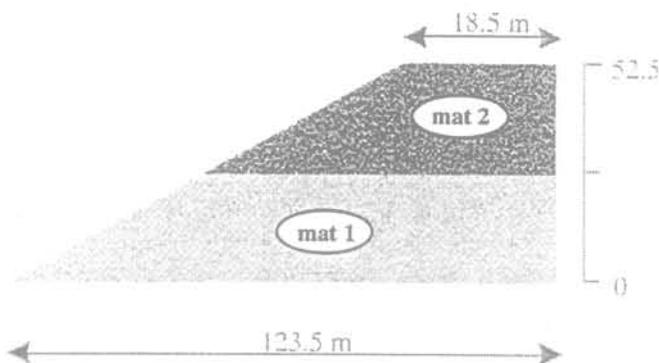
de la base de données de Duncan (1984). Pour des matériaux différents, les corrélations considérées sont subjectives et cherchent à refléter le fait que les mêmes doutes existent sur les mêmes paramètres. La matrice de corrélation finalement utilisée pour les deux matériaux est présentée sous forme de tableau (tableau I).

La modélisation par éléments finis permet de simuler la construction par couches (non-linéarité géométrique). L'incertitude sur les résultats est évaluée par Monte-Carlo et par FOSM. Le programme en éléments finis stochastiques utilisé fournit l'espérance (Fig. 5) et le coefficient de variation des déplacements verticaux dus à l'incertitude sur tous les paramètres du modèle. On observe que les résultats obtenus par FOSM (Fig. 6) et par Monte-Carlo (Fig. 7) sont analogues. Les irrégularités du champ des coefficients de variation montrent par ailleurs que pour une analyse en éléments finis stochastiques, le maillage doit généralement être plus fin que pour une analyse ordinaire.

5

Conclusion

La méthode des éléments finis stochastiques présente un intérêt particulier pour les ingénieurs géotechniciens car elle permet d'évaluer l'influence des incertitudes affectant les propriétés des matériaux sur la validité des résultats de ce type d'analyse numérique. Elle peut être employée pour obtenir les écarts types et les coefficients de variation et pour définir des intervalles de confiance sur les résultats de la modélisation tels que les déplacements, les déformations et les contraintes. Comme nous l'avons montré, malgré ses limites actuelles, la MEFS permet donc aux ingé-



		K	K_{ur}	n	R_f	K_b	m	c (kPa)	ϕ (°)	$\Delta\phi$ (°)	γ (kN/m ³)
Espérance	mat 1	120	180	0.5	0.7	110	0.2	10	25	1.0	17
	mat 2	400	600	0.6	0.7	200	0.5	22	33	1.0	15
Cv		10 %									

FIG. 4 Remblai construit par couches et constitué de deux matériaux différents. Embankment constructed placing successive layers of two different materials.

TABLEAU 1

Matrice de corrélation des propriétés des deux matériaux
Correlation matrix of the two materials.

	Matériau 1									Matériau 2									
	K	K_{ur}	n	R_f	K_b	m	$c(\text{kPa})$	$\varphi(^{\circ})$	$\Delta\varphi(^{\circ})$	K	K_{ur}	n	R_f	K_b	m	$c(\text{kPa})$	$\varphi(^{\circ})$	$\Delta\varphi(^{\circ})$	
Matériau 1	K	1	0,9	-0,7	-0,1	0,5	-0,2	-0,2	0,2	0,4	0,7	0,6	-0,4	0	0,3	-0,1	0	0	0
	K_{ur}	0,9	1	-0,6	-0,1	0,6	-0,3	-0,2	0,2	0,1	0,6	0,7	-0,4	0	0,4	0	0	0	0
	n	-0,7	-0,6	1	-0,2	-0,4	0,3	0,1	0,2	0	-0,4	-0,4	0,7	0	0	0	0	0	0
	R_f	-0,1	-0,1	-0,2	1	-0,2	-0,1	-0,2	0,2	0	0	0	0	0,7	0	-0,1	0	0	0
	K_b	0,5	0,6	-0,4	-0,2	1	-0,6	-0,4	0,4	0,1	0,3	0,4	0	0	0,7	0	0	0	0
	m	-0,2	-0,3	0,3	-0,1	-0,6	1	0,3	-0,2	-0,3	-0,1	0	0	-0,1	0	0,7	0	0	0
	$c(\text{kPa})$	-0,2	-0,2	0,1	-0,2	-0,4	0,3	1	-0,5	-0,3	0	0	0	0	0	0	0,7	-0,3	-0,2
	$\varphi(^{\circ})$	0,2	0,2	0,2	0,2	0,4	-0,2	-0,5	1	0,5	0	0	0	0	0	0	-0,3	0,7	0,3
	$\Delta\varphi(^{\circ})$	0,4	0,1	0	0	0,1	-0,3	-0,3	0,5	1	0	0	0	0	0	0	-0,2	0,3	0,7
Matériau 2	K	0,7	0,6	-0,4	0	0,3	-0,1	0	0	0	1	0,9	-0,7	-0,1	0,5	-0,2	-0,2	0,2	0,4
	K_{ur}	0,6	0,7	-0,4	0	0,4	0	0	0	0	0,9	1	-0,6	-0,1	0,6	-0,3	-0,2	0,2	0,1
	n	-0,4	-0,4	0,7	0	0	0	0	0	0	-0,7	-0,6	1	-0,2	-0,4	0,3	0,1	0,2	0
	R_f	0	0	0	0,7	0	-0,1	0	0	0	-0,1	-0,1	-0,2	1	-0,2	-0,1	-0,2	0,2	0
	K_b	0,3	0,4	0	0	0,7	0	0	0	0	0,5	0,6	-0,4	-0,2	1	-0,6	-0,4	0,4	0,1
	m	-0,1	0	0	-0,1	0	0,7	0	0	0	-0,2	-0,3	0,3	-0,1	-0,6	1	0,3	-0,2	-0,3
	$c(\text{kPa})$	0	0	0	0	0	0	0,7	-0,3	-0,2	-0,2	-0,2	0,1	-0,2	-0,4	0,3	1	-0,5	-0,3
	$\varphi(^{\circ})$	0	0	0	0	0	0	-0,3	0,7	0,3	0,2	0,2	0,2	0,2	0,4	-0,2	-0,5	1	0,5
	$\Delta\varphi(^{\circ})$	0	0	0	0	0	0	-0,2	0,3	0,7	0,4	0,1	0	0	0,1	-0,3	-0,3	0,5	1

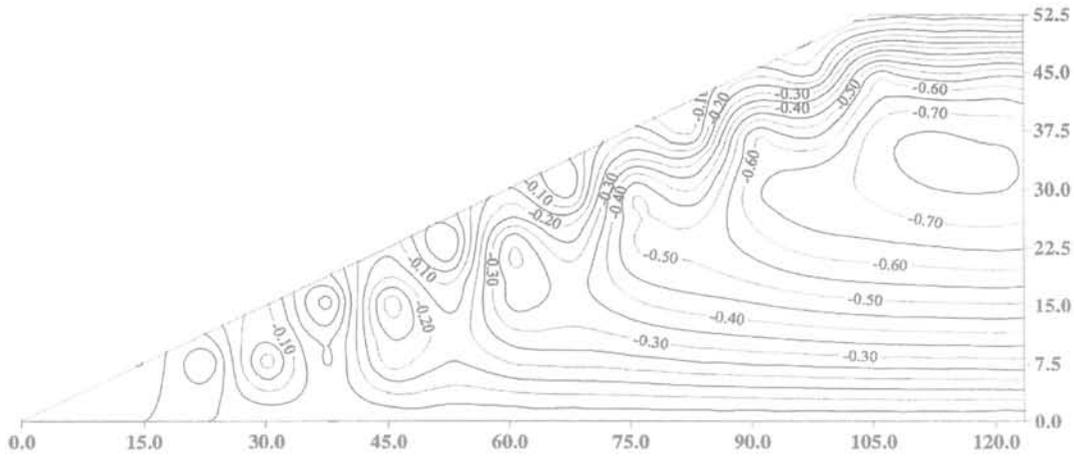


FIG. 5 Déplacements verticaux (m), (espérance mathématique).
Vertical displacements (m), (expected value)

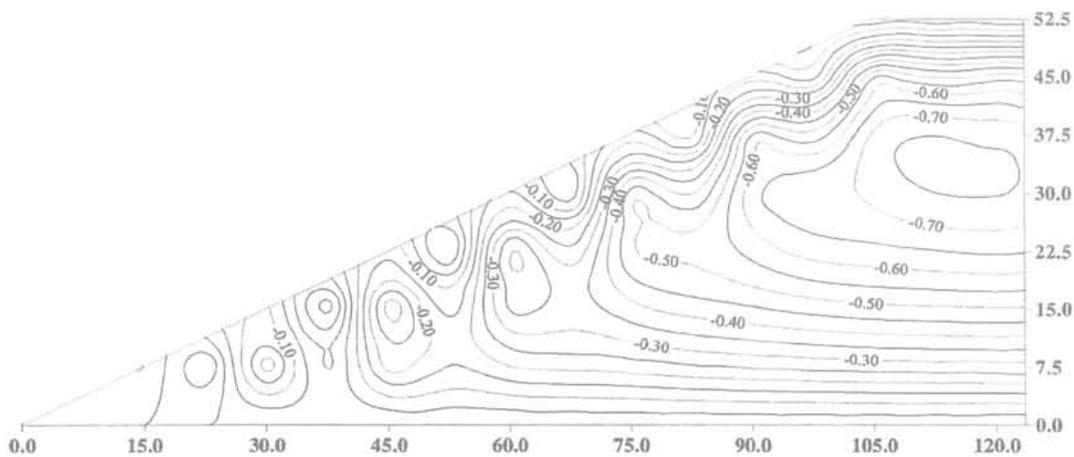


FIG. 6 Coefficient de variation des déplacements verticaux (FOSM).
Coefficient of variation of vertical displacements (FOSM).

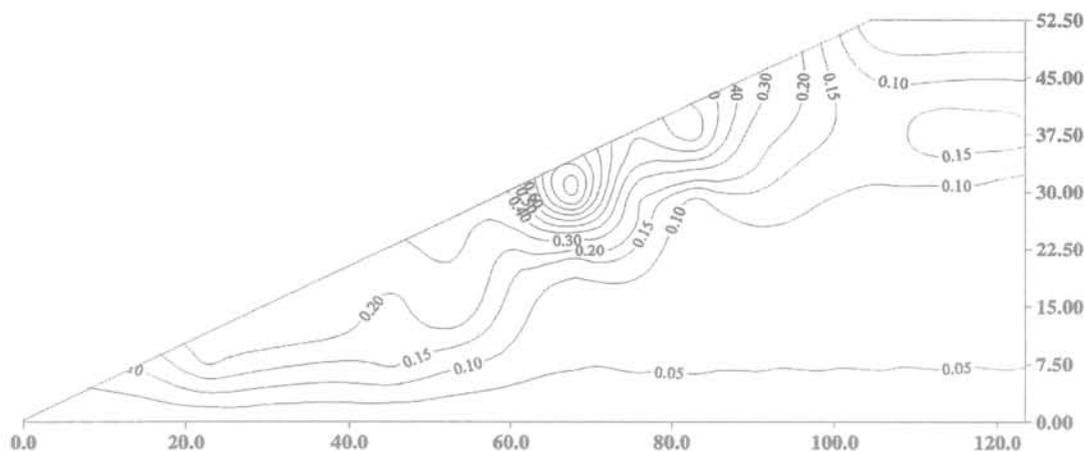


FIG. 7 Coefficient de variation des déplacements verticaux (MC).
Coefficient of variation of vertical displacements (MC).

nieurs géotechniciens d'introduire un degré de réalisme supplémentaire bien nécessaire dans les modélisations aux éléments finis. Elle est également très

utile pour mettre en évidence les paramètres les plus critiques de la loi de comportement retenue pour la modélisation.

Bibliographie

- Alabert F. (1997) – « The practice of fast conditional simulations through the LU decomposition of the covariance matrix ». *Mathematical Geology*, vol. 19, n° 5, p. 369-386.
- Auvinet G., Bouayed A., Ben Cheikh M. (1995) – « Modélisation stochastique par éléments finis en géomécanique ». *Actes du 12^e Congrès français de Mécanique*, Strasbourg, France, vol. 4, p. 201-204.
- Auvinet G., Bouayed A., Orlandi S., Lopez F. (1996) – « Stochastic Finite Element Method in Geomechanics ». *Proceedings, Geotechnical Engineering Congress 1996, « Uncertainty 96 »*, University of Wisconsin, Geotechnical Engineering Division, ASCE *Geotechnical Special Publication*, vol. 2, n°58, p. 1239-1253.
- Auvinet G., Mellah R., Masroufi F. (1999) – « Stochastic finite element analyses in geomechanics ». *Proceedings, VIIIth International Conference on Applications of Statistics and Probability*, ICASP8, Melchers & Stewart, eds., 2000 Balkema, Rotterdam, p. 79-85.
- Baecher G.B., Ingra T.S. (1981) – « Stochastic FEM in settlement predictions ». *Journal of the Geotechnical Engineering Division, ASCE*, vol. 107, n° GT4, p. 449-463.
- Bencheikh M. (1994) – « Éléments finis stochastiques pour les écoulements permanents ». Mémoire de DEA, ENSG, INPL, Nancy, 60 p.
- Benjamin J.R., Cornell C.A. (1970) – *Probability, statistics and decisions for civil engineers*. New York, McGraw-Hill, p. 151-189.
- Benmansour A. (1996) – « Fiabilité des conduites enterrées ». Thèse de doctorat, ENSG-INPL, Nancy, 232 p.
- Bittnar Z., Sejnoha J. (1996) – « Numerical Methods in Structural Mechanics ». ASCE Press, Thomas Telford.
- Bolle A. (1988) – « Approche probabiliste en mécanique des sols avec prise en compte de la variabilité spatiale ». Thèse de doctorat à l'École polytechnique fédérale de Lausanne, n° 743, 265 p.
- Bolle A. (1993) – « How to manage the spatial variability of natural soils ». *Proceedings, PROBAMAT, Probabilités et matériaux, Essais, Modèles et Applications*, NATO ARW930521, ENS Cachan.
- Bouayed A. (1997) – « Modélisation stochastique par éléments finis en géomécanique ». Thèse de doctorat, ENSG, INPL, Nancy, 157 p.
- Cambou B. (1997) – « Método del elemento finito. Análisis de incertidumbre de primer orden ». Publicación n° 339 del Instituto de Ingeniería, UNAM, Mexico, 88 p.
- Cambou B. (1975) – « Application of first-order uncertainty analysis in the finite element method in linear elasticity ». *Proceedings, Second International Conference on Application of Statistics and Probability in Soil and Structure Engineering*, Aachen, p. 67-87.
- Cambou B. (1975) – « L'incertitude sur les résultats d'un problème de mécanique des sols ou des roches traitées par la méthode des éléments finis ». *Revue française de géotechnique*, n° 1, p. 55-64.
- Cherubini C. (1993) – « The variability of geotechnical parameters ». *Probabilities and materials-Tests, Models and Applications*. D. Breyse, ed. Dordrecht, Kluwer Academic Publishers, p. 69-80.
- Chowdhury R.N., Xu D.W. (1993) – « Rational polynomial technique in slope reliability analysis ». *Journal of the Geotechnical Engineering Division, ASCE*, vol. 119, n° 12, p. 1910-1928.
- Cochran W.G. (1977) – *Sampling techniques*. John Wiley and Sons Inc.
- Cornell A.C. (1971) – « First order uncertainty analysis of soils deformation and stability ». *Proceedings, First International Conference on Applications of Statistics and Probability to Soil and Structural Engineering*, Hong Kong, p. 130-144.
- Deplagne F. (1994) – « Simulation de la cohésion pour une digue en argile ». *Actes 12^e Rencontres universitaires de génie civil*, p. 75-82.
- Der Kiureghian A., Ke J.B. (1988) – « The stochastic finite element method in structural reliability ». *Probabilistic Engineering Mechanics*, vol. A3, n° 2, p. 83-91.
- Der Kiureghian A., Li C.C., Zhang Y. (1991) – « Recent developments in stochastic finite elements ». *Proceedings, 4^e IFIP W.G 7.5 Conference*, Munich, Springer-Verlag, p. 19-38.
- Ditlevsen O., Madsen H.O. (1996) – *Structural Reliability Methods*. New York, John Wiley and Sons.
- Duncan J.M., Seed R.B., Wong K.S., Ozawa Y. (1984) – « FEADAM84 : a computer program for Finite element Analysis of Dams ». Research report n° SU/GT/84-03, Department of Civil Engineering, Stanford University, USA, 68 p.
- Evans D.H. (1967) – « An application of numerical integration techniques to statistical tolerancing ». *Technometrics*, vol. 9, n° 3, p. 441-456.
- Favre J.-L., Genevois B. (1987) – « Analyse de la sécurité des fondations superficielles vis-à-vis d'un défaut de portance : effet de la variabilité spatiale des paramètres ». *Revue française de géotechnique* n° 39, p. 33-40.
- Fenton G.A., Vanmarcke E.H. (1990) – « Simulation of random fields via local average subdivision ». *Journal of Engineering Mechanics*, vol. 116, n° 8, p. 1733-1749.
- Flores Macías O. (1994) – « Modèles fiables et mécaniques : éléments finis stochastiques. Méthodes de couplage et applications ». Thèse doctorale, université Blaise Pascal, Clermont-Ferrand.
- Flores Macías O., Lemaire M. (1997) – « Éléments finis stochastiques et fiabilité ». *Revue française de génie civil*, vol. 1, n° 2, p. 247-284.

- Fogli M. (1980) – « L'approche de Monte-Carlo dans les problèmes de sécurité : application à l'estimation probabiliste du risque de ruine des poutres hyperstatiques en béton armé soumises à des actions aléatoires statistiques ». Thèse de docteur ingénieur, INSA de Lyon, 217 p.
- Griffiths D.V., Fenton G.A. (1993) – « Seepage beneath water retaining structures founded on spatially random soil ». *Géotechnique*, vol. 43, n° 4, p. 577-587.
- Kleiber M., Duong Hien T. (1992) – *The stochastic finite element method : basic perturbation technique and computer implementation*. John Wiley & Sons Ltd, 321 p.
- Kondner R.L. (1963) – « Hyperbolic stress-strain response : cohesive soils ». *Journal of the Soil Mechanics and Foundations Division*, ASCE, vol. 89, n° SM1, 115 p.
- Lawrence M. (1987) – « Basic random variables in finite element analysis ». *International Journal of Numerical Methods in Engineering*, 24, , p. 1849-1863.
- Lemaire M. (1997) – « Éléments finis stochastiques et fiabilité ». *Revue française de génie civil*, vol. 1, n° 2, p. 248-284.
- López Acosta P. (1999) – « Flujo de agua en suelos con permeabilidad aleatoria ». Master degree thesis, DEPF, UNAM, Mexico, D.F.
- Magnan J.-P. (1987) – « Programme de calcul par la méthode des éléments finis stochastiques (PROBEF) ». Actes, Greco Géomatériaux, Aussois, France.
- Magnan J.-P., Bouheraoua A. (1995) – « Stochastic analysis of soil consolidation in theory and practice ». *Proceedings*, VIIth International Conference on Applications of Statistics and Probability, Lemaire, Favre & Mébarki ed., Rotterdam, Balkema, p. 77-84.
- Magnan J.-P., Bouheraoua A. (1996) – « Analyse probabiliste de la consolidation unidimensionnelle des sols ». *Revue française de géotechnique*, n° 77, p. 19-30.
- Matheron G. – (1971) « La théorie des variables régionalisées et ses applications ». Fasc. 5, Ecole nationale supérieure des mines, Paris.
- Mathews J., Walker R.L. (1964) – *Mathematical Methods of Physics*. Benjamin Publishers, 2^e ed., p. 286-298.
- Mellah R., Auvinet G., Masroufi F. (1998) – « Estimation des incertitudes sur les résultats de la modélisation par éléments finis d'un remblai ». Actes, 2^e Conférence nationale sur la Fiabilité des matériaux et des structures " JN-Fiab'98 ", p. 209-218.
- Mellah R., Auvinet G., Masroufi F. (1998) – « Stochastic Finite Element Method applied to Non-linear Analysis of Embankments ». Third International Conference on Computational Stochastic Mechanics. CSM'98, *Computational Stochastic Mechanics*, Spanos, ed., Rotterdam, Balkema, p. 531-537.
- Mellah R. (1999) – « Modélisation stochastique par éléments finis en élastoplasticité appliquée à la géomécanique ». Thèse doctorale, ENSG-INPL, Nancy, 160 p.
- Mellah R., Auvinet G., Masroufi F. (1999) – « Analyse probabiliste d'un remblai en terre par la simulation de Monte-Carlo ». *Récents Progrès en génie des procédés*, vol. 13, p. 173-178.
- Mellah R., Auvinet G., Masroufi F. (2000) – « Stochastic Finite Element Method applied to Non-linear Analysis of Embankments ». *Journal of Probabilistic Engineering Mechanics*, vol. 15, n° 3, p. 251-259.
- Mestat P., Prat M. (sous la direction de) (1999) – *Ouvrages en interaction*. Chapitre 12 : « Indicateurs d'erreur a priori » et Chapitre 13 : « Estimateurs d'erreur a posteriori ». Paris, Hermès Science Publications, p. 797-838.
- Mrabet Z. (1999) – « Fiabilité des remblais homogènes compactés ». Thèse de doctorat, ENSG-INPL, Nancy, , 251 p.
- Nechnech A. (1994) – « Étude fiabiliste des murs de soutènement ». Mémoire de DEA, ENSG-INPL, Nancy, 59 p.
- Orlandi S. (1996) – « Método del elemento finito estocástico en geotecnia : análisis no lineal ». Master Degree thesis, DEPF, UNAM, Mexico, 81 p.
- Poirion F., Soize C. (1989) – « Simulation numérique de champs vectoriels stochastiques gaussiens homogènes et non homogènes ». *La Recherche aérospatiale*, n° 1, p. 41-61.
- Righetti G., Harrop-Williams K. (1988) – « Finite element analysis of random soil media ». *Journal of Geotechnical Engineering*, ASCE, vol. 114, n° 1, p. 59-75.
- Rossa, Fry. (1988) – « Exploitation des données recueillies sur la digue Aube ». Électricité de France, rapport interne.
- Rosenblueth E. (1975) – « Point estimates for probability moments ». *Proceedings*, Nat. Acad. Sciences, USA, vol. 72, October, p. 3812-3814.
- Soulié M., Favre M., Konrad J.-M. (1983) – « Analyse géostatistique d'un noyau de barrage tel que construit ». *Canadian Geotechnical Journal*, vol. 20, p. 453-467.
- Spanos P.D., Ghanem R. (1989) – « Stochastic finite element expansion for random media ». *Journal of Engineering Mechanics*, 115, , p. 499-519.
- Su Y.L., Wang Y.J., Stefanko R. (1969) – « Finite element analysis of underground stresses utilizing stochastically simulated material properties. Rock Mechanics: theory and Practice ». *Proceedings*, 11th United States Symposium on Rock Mechanics, Berkeley, Calif., p. 253-266.
- Tang W.H. (1993) – « Recent developments in geotechnical reliability ». *Probabilistic Methods in Geotechnical Engineering*, Li & Lo, 1993, Rotterdam Balkema, p. 3-27.
- Teigen J.G., Frangopol D.M. (1991) – « Probabilistic FEM for non-linear concrete structures. I : theory ». *Journal of Structural Engineering*, ASCE, vol. 117, n° 9, p. 2674-2689.
- Vanmarcke E.H., Grigoriu M. (1983) – « Stochastic finite element analysis of simple beams ». *Journal of Engineering Mechanics*, ASCE, vol. 109, n° 5, p. 1203-1214.
- Vanmarcke E.H. (1994) – « Stochastic finite elements and experimental measurements ». *Probabilistic Engineering Mechanics*, vol. 9, p. 103-114.
- Yongjian R., Guoqing W., Derong W., Jizhong S. (1994) – « Application of stochastic finite element method to strength and stability analysis of earth dams ». *Acta Mechanica Solida Sinica*, vol. 7, n° 1, p. 80-86.

Décision et incomplétude

D. BOISSIER

LERMES/CUST
BP 206

63174 Aubière Cedex

Résumé

La prise de décision en génie civil se fait dans un contexte où imprécis et incertain cohabitent : imprécis des objectifs assignés et de la connaissance du sol, imprécis des modèles utilisés, inadéquation entre la précision des modèles et le niveau de décision, mais aussi incertain des actions qui vont solliciter les ouvrages, incertain du comportement des ouvrages. Dans ce texte, nous montrons tout d'abord l'importance de l'imprécision et de l'incomplétude des informations en géotechnique ; nous proposons l'apport de la théorie des sous-ensembles flous (SEF) et des fonctions de croyance dans l'aide à la résolution des problèmes de géotechnique. Nous rappelons tout d'abord les concepts et bases de ces théories (partie A) puis, nous appuyons sur quatre exemples concernant les fondations pour en proposer différentes utilisations (partie B).

Mots clés : aide à la décision, incertain, imprécis, sous-ensembles flous, géotechnique.

Decision making and incompleteness

Abstract

Decision making in civil engineering takes place in an uncertain and imprecise context. The assigned objectives are imprecise such as soil knowledge and bad adapted models. Loads and load effects deal with uncertainty and the behaviour of the works is less or more predictable. In this paper we show that in the geotechnics field data are very imprecise and uncertain ; we propose to introduce the use of the fuzzy sets and of the belief functions theories. We first present the theoretical bases in Part A, then four examples in part B. in order to illustrate the use of these concepts in the geotechnics field.

Key words : decision making, uncertain, imprecise, fuzzy sets, geotechnics.

NDLE : Les discussions sur cet article sont acceptées jusqu'au 1^{er} août 2001.

Partie A – CONCEPTS ET MÉTHODES

1

Introduction

« Le bon sol », « un sol compressible », « une portance d'environ 0,2 Mpa », « une couche homogène », « une cohésion entre 20 et 30 kPa », « un sol saturé », « la fondation est surdimensionnée », « la nappe est variable »... Toutes ces expressions ont en commun deux caractéristiques :

- 1) elles sont communément utilisées et comprises par des géotechniciens ;
- 2) elles sont toutes à support d'information non précis.

Ce texte montre la nécessité d'approches non traditionnelles pour la prise en compte de ce type d'information par les mécaniciens des sols ; il propose quelques repères pour l'aide au choix de modèles possibles de représentation et de traitement. Il s'appuie sur des travaux de recherche que j'ai dirigés à l'université d'Artois (J. Al-Hajjar, H. Mammeri, I. Semaan, A. Kar-nib), puis à l'université Blaise Pascal (J. Lair) et sur d'autres travaux qui sont en cours sur ce thème au sein de l'équipe Mécanique des sols du Lermes avec C. Bacconnet et R. Gourvès.

2

Constat d'échec et proposition

La théorie des ensembles et la théorie des probabilités ne permettent pas toujours de rendre compte d'une réalité complexe. En effet, l'information disponible est souvent imparfaite (incomplète, imprécise, inadaptée, suspecte...) ; elle ne provient pas toujours non plus d'une mesure (dire d'expert, estimation...) ; enfin, les modèles de l'ingénieur ne permettent pas de modéliser une réalité complexe où les décisions sont prises au regard de considérations diverses.

2.1

Le constat de L.A. Zadeh (1965)

Au fur et à mesure que la complexité des systèmes augmente, notre aptitude à formuler de manière *précise et significative* leur comportement diminue jusqu'à une limite au-delà de laquelle la précision et la signification deviennent des caractéristiques mutuellement exclusives.

Notre connaissance du monde réel est imprécise et celle de son évolution incertaine ; il est vain alors de vouloir le décrire par des modèles précis et déterministes ; cela est vain et surtout dangereux car le sentiment de précision est illusoire, la signification des résultats est très pauvre et les décisions sont prises sur la base d'informations imprécises et sans signification.

Deux arguments sont alors mis en évidence : la fiabilité de l'information et son « informativité » ; il apparaît vite que ces deux arguments sont souvent conflictuels.

• *Exemple 1* : la cohésion du sol dans cette couche est de 20 kPa.

Cette information est très informative et l'on peut faire des calculs ou prendre des décisions sur cette base, par contre elle est très peu fiable ; en effet en pratique l'on ne peut pas assurer que le sol a une cohésion de 20 kPa ; tout au plus peut-on dire que la mesure dans telle condition sur tel appareil de tel échantillon provenant de tel prélèvement a donné une valeur que l'on arrondit à 20 kPa.

• *Exemple 2* : la cohésion du sol dans cette couche est comprise entre 10 kPa et 100 kPa.

Cette information est très fiable ; il y a très peu de chance que dans cette même couche de sol une autre mesure donne une valeur inférieure à 10 kPa mais par contre elle n'est que très peu informative et il est difficile de faire des calculs ou de prendre des décisions sur cette base ; ce sont des valeurs numériques que l'on introduit dans des codes de calcul, parfois des distributions de probabilités et c'est une valeur que l'on compare à une limite ou un index qui sert à prendre une décision ; que faire alors d'une information sous forme d'un intervalle ou d'un ensemble ?

2.2

Décision en ingénierie (Boissier D., Al-Hajjar J., 1990)

2.2.1

Décision et précision

Dans les problèmes d'ingénieur, il est trivial de rappeler que c'est en début de résolution que l'espace des solutions est le plus grand mais que c'est aussi à ce moment-là que l'espace des données disponibles est le plus petit et que les conséquences d'une décision sont les plus importantes car réduisant fortement l'espace des solutions ; à ce moment les données sont souvent de mauvaise qualité et la stabilité de la décision ne peut être assurée (on aurait pris une autre décision si l'on avait eu toutes les informations...).

• *Exemple* : La première reconnaissance visuelle d'un site est un des critères de décision d'achat d'une parcelle en vue d'une opération immobilière ; la connaissance du sol est alors incomplète et de mauvaise qualité ; par contre la décision d'investissement est lourde de conséquences. A l'inverse, en fin de conception, lorsque les études de sol auront été dépouillées, il sera possible « d'optimiser » les fondations et d'économiser quelques m³ de béton ; le gain marginal est faible, le risque aussi car le problème est mieux défini. Il se pourrait aussi qu'avec toutes les informations disponibles à ce moment-là on souhaite revenir sur la décision d'investir alors que cela n'est plus possible ; la décision est donc conditionnée par la qualité de l'information.

2.2.2

Informations disparates

Les informations disponibles pour résoudre un problème donné sont souvent disparates : elles proviennent de sources diverses (observations, mesures

diverses, bibliographie, dire d'experts) et sont exprimées sous formes différentes.

Elles peuvent être incohérentes, incompatibles, redondantes et plus ou moins fiables (la fiabilité est liée à la valeur associée à l'information et non à la variabilité intrinsèque de l'information). Pourtant, pour prendre une décision, il faudra agréger toutes ces informations et souvent les ramener à une information unique.

• *Exemple* : On dispose d'une stratigraphie donnée par le géologue et d'un découpage des couches selon leur résistance mécanique obtenue à partir de l'analyse d'un signal pénétrométrique. Comment choisir la profondeur d'ancrage de fondations ?

2.3

Un constat positif

2.3.1

Perception globale et langage

L'être humain possède des capacités de perception globale des systèmes complexes. Il est souvent capable de porter « intuitivement » ou par « expertise non explicite » un jugement plus sûr, c'est-à-dire plus stable en décision, que celui obtenu par des modèles précis mais trop restrictifs (hypothèses non justifiables ou non crédibles).

L'être humain est aussi capable d'exprimer le vague et l'imprécis par le langage ; or ce langage est éloigné de celui utilisé dans les théories traditionnelles des ensembles et des probabilités.

2.3.2

Quelques exemples en ingénierie

a) la visualisation d'un système constructif permet au concepteur de détecter une erreur de conception (appui sur ou sous abondant, réservation trop petite, croisement de réseaux impossible, superposition de structures incorrecte) alors que le calcul géométrique ou mécanique dans un système de CAO ne le permet pas ;

b) un ingénieur peut pour un problème donné caractériser un sol de « mauvais », alors qu'aucun modèle mécanique ne permettrait de calculer sa résistance (hétérogénéité, présence de cavités...);

c) on peut affirmer que la charge d'une structure sera d'environ x . L'égalité n'a pas de sens car au moment du calcul la structure n'existe pas encore et parce qu'au cours de sa vie la charge peut varier ; ce qualificatif environ est peu précis et on ne sait exactement ce qu'il recouvre mais il montre *explicitement* que l'on est dans une ignorance relative ;

d) on peut affirmer qu'un sol est compressible ; ce qualificatif le sépare de la catégorie des sols incompressibles. La limite « compressible/incompressible » n'est pas nette ; elle est relative à des indicateurs (l'indice de compression est un indicateur parmi d'autres) ; pour un matériau donné, la valeur obtenue pour l'indicateur est liée à la précision de la mesure (technique de mesure, conditions opératoires) et à l'échantillon (variabilité spatiale et temporelle) ;

e) les règles de décision sont souvent basées sur des classifications (sable/gravier, sec/humide) qui en réalité

ont des domaines d'intersection non vides (un même sol peut ne pas appartenir à la même catégorie dans des classifications différentes).

3

Apport de la logique floue

3.1

Définitions

Imprécis : Il concerne des informations mal définies ; on lui associe les qualificatifs de type : « environ », « à peu près », « dans l'intervalle », « compris entre et entre », « approximativement »... ; il représente la qualité de la connaissance que nous avons d'un phénomène.

Incertain : Il résulte essentiellement d'une épreuve dont le résultat n'est pas connu *a priori* et peut changer si on la recommence ; il représente la variabilité d'un phénomène ; il peut être représenté de façon rigoureuse et précise.

Dans notre domaine, certaines données (charges et résistances par exemple) sont, par nature même, incertaines ; leur variabilité spatio-temporelle est irréductible (données climatiques, propriétés des géomatériaux et des matériaux de construction, charges d'exploitation...).

Remarque : Il ne faut pas confondre la qualité d'une donnée et la qualité de l'information que l'on possède sur la donnée.

3.2

La théorie des sous ensembles flous (Fuzzy sets)

3.2.1

Présentation (Bouchon Meunier, 1995)

L'idée de base est que lorsqu'une classe est définie par une notion vague, on ne peut pas utiliser le concept classique d'appartenance ; il faut alors introduire une échelle – la fonction d'appartenance – qui permette de hiérarchiser les éléments d'un univers donné par rapport à cette propriété vague.

Ici c'est l'ensemble qui est flou (ou du moins ses limites) par opposition à un ensemble (« net ») traditionnel dont les limites sont précises ; la fonction d'appartenance exprime si une valeur x précise satisfait pleinement ou non les propriétés de l'ensemble.

L'ensemble classique (ou net) est un cas particulier où l'appartenance est binaire : x appartient ou n'appartient pas à l'ensemble. En théorie des possibilités, c'est l'ensemble qui est net alors que la mesure (donnée) est floue.

3.2.2

Définition

La théorie des sous-ensembles flous (SEF) a été développée par L.A. Zadeh (1965) ; elle permet de

modéliser l'imprécis et l'incertain d'une autre manière que par les modèles stochastiques. Les SEF sont une voie d'introduction de l'incertain et de l'imprécis dans l'aide à la décision.

Soit X , un ensemble de référence ; un sous-ensemble flou A de X peut être défini à partir de la fonction d'appartenance μ_A telle que :

$$\forall x \in X \xrightarrow{\mu_A} \mu_A(x) \in [0, 1] \quad (1)$$

L'application μ_A (éq. 1) est appelée fonction d'appartenance et la valeur $\mu_A(x)$ est appelée niveau d'appartenance.

3.2.3

Propriétés

Le support de A , noté $\text{supp}(A)$ est la partie de X pour laquelle $\mu_A(x) > 0$.

La hauteur, notée $h(A)$ est la plus grande valeur de la fonction d'appartenance. $h(A) = \sup_{x \in X} (\mu_A(x))$

Le noyau, noté $\text{noyau}(A)$, est l'ensemble des valeurs de A pour lesquelles la fonction d'appartenance est maximale.

Lorsque l'ensemble X est à valeurs discrètes, la cardinalité est la somme des valeurs des fonctions d'appartenance.

3.2.4

Cas particulier

Il existe de nombreuses variétés de sous-ensembles flous ; nous nous limitons ici aux SEF de type L_R normés ; leur fonction d'appartenance est paramétrée par deux fonctions L et R et quatre paramètres m, n, α, β et leur hauteur est égale à 1 ; nous pouvons voir un SEF de type L_R trapèze noté $A(m; n; \alpha; \beta; 1)$ (Fig. 1a) et L_R triangle symétrique noté $A(m; \alpha; \beta; 1)$ (Fig. 1b).

Contrairement aux ensembles ordinaires, un SEF n'a pas une appartenance nettement définie mais comporte des transitions graduelles d'appartenance depuis l'appartenance totale (degré 1) jusqu'à l'absence d'appartenance (degré 0).

Ce concept peut être illustré comme suit : quand le niveau d'appartenance est près de 1 alors nous pou-

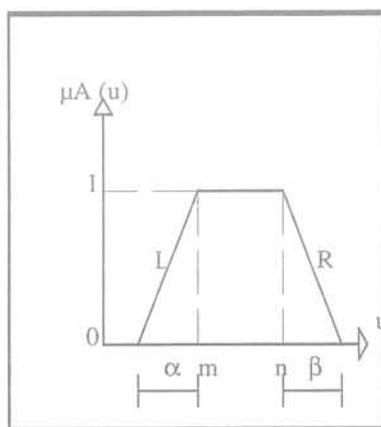


FIG. 1A SEF L_R trapèze.
Trapezium L_R fuzzy Set.

vons assurer que la proposition « x appartient à F » est satisfaite ; lorsque ce niveau est près de 0 la proposition n'est pas satisfaite ; la validité de la proposition peut donc être liée à la valeur $\mu_F(x)$. L'exemple 1 illustre l'utilisation de ce concept en fiabilité.

3.2.5

Compatibilité entre deux sous-ensembles flous

Soit X un ensemble de référence, A et B 2 SEF et $\mu_A(x)$ et $\mu_B(x)$ leurs fonctions d'appartenance respectives ; soit $\mu_A^{-1}(x)$, la fonction inverse de $\mu_A(x)$. Alors, Zadeh L.A. (1978) et Dubois D. et Prade H. (1983) définissent la compatibilité entre A et B sous la forme d'une distribution des valeurs possibles du degré de compatibilité ; celle-ci est alors appelée profil de compatibilité ; elle est considérée comme un SEF C , dont la fonction d'appartenance $\mu_C(x)$ est calculée par l'équation 2 et qui est définie sur le domaine $X_C = [0,1]$.

$$\mu_C(t) = \sup(\mu_A(x) \text{ avec } x \in X \text{ et tel que } \mu_B(x) = t) \quad (2)$$

$$\mu_C(t) = 0 \text{ si } \mu_B^{-1}(t) = 0$$

Pour toute valeur t du degré de compatibilité, la procédure de calcul est la suivante :

- 1) On recherche toutes les valeurs de B qui ont un degré d'appartenance t ;
- 2) On calcule le degré d'appartenance de chacune de ces valeurs au SEF A ;
- 3) Le maximum de ces degrés d'appartenance est le degré d'appartenance du degré de compatibilité t .

Si le degré d'appartenance t ne correspond à aucune des valeurs du SEF B , alors le degré de compatibilité de t est nul.

Cette procédure est illustrée sur la figure 2 et utilisée dans l'exemple 2.

3.2.6

Comparaison de deux sous-ensembles flous

Soient A et B , deux sous-ensembles flous ; l'objectif de cette comparaison est de voir dans quelle mesure une observation A peut être considérée comme une observation de l'ensemble B de référence ; lorsque A

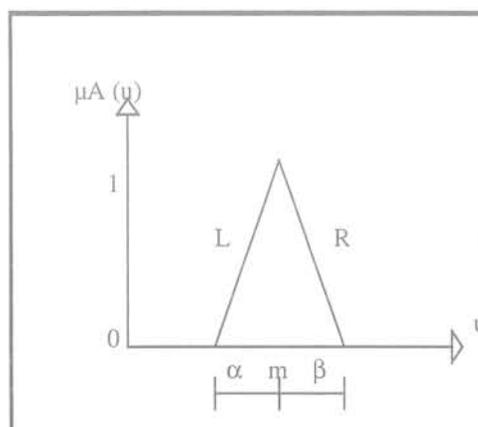


FIG. 1b SEF L_R triangle.
Triangle L_R fuzzy Set.

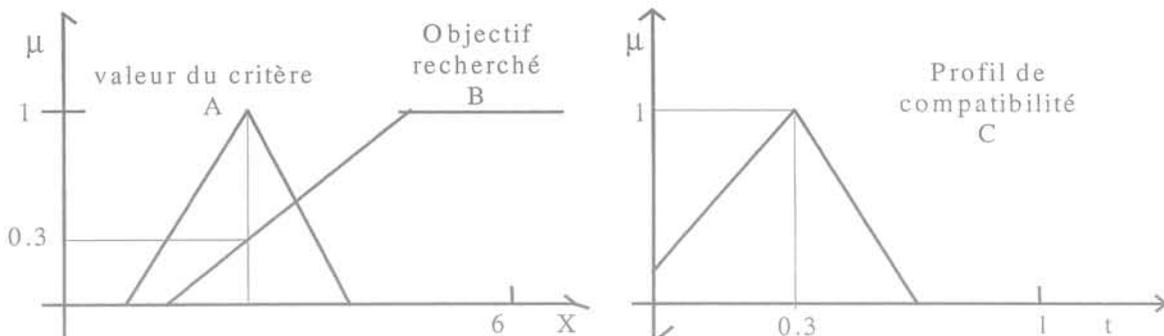


FIG. 2 Profil de compatibilité entre deux sous ensembles flous.
Profile of compatibility between two fuzzy sets.

et B sont deux SEF, il s'agit de comparer des intervalles auxquels sont associées des fonctions d'appartenance. La méthode de comparaison des intervalles flous est basée sur les mesures de cardinalités floues ; $|A|$ la cardinalité floue de A est définie par l'équation 3 :

$$|A| = \sum_{x \in X} \mu_A(x) \quad (3)$$

$C_p(A/B)$, la cardinalité floue relative, est un degré de vraisemblance qui traduit combien A peut être considérée comme une observation de B par le biais de la proportion de paramètres satisfaits dans A, compte tenu de B.

N. Aggarwal, P. Quarrey, C. Berry (1993) calculent C_p par la règle du tableau I :

Pour des SEF continus, le calcul du degré de vraisemblance par la cardinalité floue revient au calcul de l'aire de l'intersection des deux intervalles. Si cette intersection est vide alors la cardinalité relative est nulle.

Cette formulation à deux SEF peut être étendue au cas où l'on veut comparer le SEF A à plusieurs SEF de type B ; un algorithme de discrimination permet alors de trouver quel est l'ensemble B qui est le plus proche du SEF A (ou réciproquement à quel SEF B peut être rattaché avec le plus de vraisemblance le SEF A). L'exemple 3 illustre l'utilisation de cette méthode.

3.3

La théorie des fonctions de croyance (Dempster Shafer)

3.3.1

Présentation

Cette théorie, dont les bases ont été établies en 1976 par G. Shafer (1976), permet de représenter l'ignorance et l'incertain, de prendre en compte toutes les informations existantes sur un sujet donné quels que soient leur provenance, leur format, et de dégager une information consensuelle. Il a proposé des règles de « fusion des données » ou « d'unification de données » qui sont basées sur les hypothèses suivantes :

- de toute information il y a quelque chose à déduire ;
- à chaque information est attachée une masse de croyance, c'est-à-dire simplement un attribut précisant la qualité de cette information.

Le problème alors est de construire (ou de calculer) une information résultante à partir de plusieurs informations ; la méthode est basée sur l'idée qu'il faut prendre en compte en même temps l'information (ou sa modélisation) et sa qualité (ici la masse de croyance) ; c'est ce que le décideur fait, de manière non formelle souvent, lorsque après avoir écouté les experts il se

TABLEAU I Calcul du degré de vraisemblance $C_p(A/B)$.
Calculus of the $C_p(A/B)$ likelihood degree.

Si	Alors	Sinon
Support (A) \subseteq support (B) et noyau (A) \subseteq noyau (B)	$C_p(A/B) = 1$	$C_p(A/B) = \frac{\sum_{i=1}^n \text{Min}(\mu_A(x_i), \mu_B(x_i))}{ A }$

« fait une opinion » et décide. Il prend en compte l'information (donnée, assertion...) mais aussi sa provenance sachant que certaines sources sont sûres alors que d'autres sont moins fiables, que certains modèles ou analogies sont près de la réalité, d'autres plus éloignés.

Une difficulté est que parfois les informations sont contradictoires et donc que des dires d'experts ou des résultats expérimentaux sont infirmés. L'information résultant de deux expériences en partie ou complètement contradictoires a cependant une signification et une qualité qu'il faut déterminer.

3.3.2

Principe

Supposons un ensemble Θ appelé « cadre de discernement » correspondant à l'ensemble des réponses à une question donnée (par exemple si la question est « quelle est la cohésion du sol ? », alors Θ est l'ensemble des cohésions : $\Theta = [0, c]$). Ce cadre de discernement doit être exhaustif, ce qui signifie que la solution (ici la cohésion) appartient obligatoirement au cadre de discernement. Toute réponse à la question est appelée une preuve et pointe sur un sous-ensemble S inclus dans Θ ; les sous-ensembles pointés peuvent être une mesure ponctuelle, un intervalle, une densité de probabilité, un ensemble flou...

Toute preuve est obtenue dans un « contexte d'acquisition » particulier et c'est à partir de l'analyse de ce contexte d'acquisition que l'on peut apprécier la qualité de la preuve S . Dans la théorie de Shaffer, la « qualité d'une preuve » est représentée par une quantité $m(S)$ associée à chacun des sous-ensembles ; c'est un nombre compris entre 0 et 1 qui représente la force de la preuve, sa « fiabilité ». L'estimation de ces masses peut se faire subjectivement en mode déclaratif ou être guidée par une grille d'analyse.

En réponse à une question telle que « quelle est la cohésion du sol », des exemples de preuve sont la réponse donnée par un expert qui annonce une cohésion entre 20 et 40 kPa, le résultat d'un essai in situ grâce auquel on estime une cohésion de 40 kPa, une étude des dossiers sur des bâtiments existants qui montre que les calculs ont été faits avec une cohésion moyenne de 35 kPa, etc. ; toutes les sources dont sont issues ces données ont des « fiabilités différentes ».

Ayant toutes les informations sous forme de couples $[S : \text{information} ; m : \text{croyance}]$, nous pouvons les fusionner au moyen de la règle de Dempster Shaffer (§333), afin d'obtenir un consensus.

3.3.3

Unification de deux informations (J. Lair, 2000)

Soit deux preuves P_1 et P_2 pointant respectivement sur A et B , de masses de croyance m_a et m_b ; distinguons les deux situations suivantes : $A \cap B \neq \emptyset$ et $A \cap B = \emptyset$.

• Preuves concordantes ($A \cap B \neq \emptyset$)

Pour modéliser la connaissance contenue dans la preuve pointant sur A (soit P_1), nous affectons une masse m_a sur le sous-ensemble A , une masse $(1 - m_a)$ sur $\overline{A \cap B}$ et $(1 - m_a)$ sur l'ignorance $\overline{A \cup B}$. L'ignorance

est en fait une indécision : on affecte la masse $(1 - m_a)$ sur Θ puisque l'on ne sait pas.

Par analogie, pour modéliser la connaissance contenue dans la preuve pointant sur B (P_2), nous affectons une masse m_b sur B , une masse $(1 - m_b)$ sur $\overline{A \cap B}$ et $(1 - m_b)$ sur l'ignorance $\overline{A \cup B}$ (Tableau II)

Le consensus est obtenu en affectant des masses aux sous-ensembles comme ci-après :

pour $A \cap B$: $m_a \cdot (1 - m_b)$, pour $\overline{A \cap B}$: $m_a \cdot m_b$, pour

$B \cap \overline{A}$: $m_b \cdot (1 - m_a)$ et pour $\overline{A \cup B}$: $(1 - m_a) \cdot (1 - m_b)$. La

somme des quatre masses consensuelles donne 1. Le produit $(1 - m_a) \cdot (1 - m_b)$ correspond à l'ignorance et reflète le cas où les deux preuves seraient fausses.

• Preuves conflictuelles ($AB = \emptyset$) (Tableau III)

De la même manière :

– pour modéliser la connaissance contenue dans la preuve pointant sur A , nous affectons une masse m_a sur le sous-ensemble A , une masse $(1 - m_a)$ sur B et $(1 - m_a)$ sur l'ignorance ;

– pour modéliser la connaissance contenue dans la preuve pointant sur B , nous affectons une masse m_b sur le sous-ensemble B , une masse $(1 - m_b)$ sur A et $(1 - m_b)$ sur l'ignorance.

De nouveau, le consensus est obtenu en affectant le produit des masses aux sous-ensembles concernés. La somme des quatre masses consensuelles donne $(1 - m_a m_b)$.

Cet écart $m_a m_b$ caractérise le caractère conflictuel des données, c'est-à-dire la masse que l'on affecte à un ensemble inexistant. La règle de Dempster nous dit de répartir la masse conflictuelle proportionnellement sur les trois autres masses en appliquant un facteur de normalisation k (eq. 4).

$$k = \frac{1}{1 - m_c} \quad \text{où } m_c \text{ est la masse affectée à l'ensemble vide} \quad (4)$$

Soit deux preuves P_1 et P_2 pointant respectivement sur les sous-ensembles A_i et B_j , de masses de croyance $m_a(A_i)$ et $m_b(B_j)$, la règle d'assemblage de Dempster, donnant la répartition des masses résultantes est (eq. 5) :

$$m(\theta) = k \sum_{\substack{i,j \\ A_i \cap B_j = \theta}} m_a(A_i) \cdot m_b(B_j) \quad (5)$$

L'application de Dempster dans le cas de deux données conflictuelles aboutit donc à l'équation 6 :

$$m(A) = \frac{1}{1 - m_a \cdot m_b} \cdot m_a \cdot (1 - m_b) \quad (6)$$

Nous obtenons ainsi une nouvelle répartition des masses. La « preuve résultante » pointe maintenant sur A , sur B , sur l'ignorance et sur $A \cap B$ (cas de preuves concordantes). C'est une preuve fictive qui englobe toute la connaissance que nous avons dans les preuves initiales. Le principe de la règle de Dempster est donc que deux éléments concordants se renforcent mutuellement et deux éléments discordants s'érodent.

TABEAU II Affectation des masses. Cas des preuves concordantes.
Masses affectation for heterogeneous evidences.

Preuve 1	m_a	m_a	$(1 - m_a)$	$(1 - m_a)$	$\Sigma = 1$
Preuve 2	$(1 - m_b)$	m_b	m_b	$(1 - m_b)$	
Consensus	$m_a \cdot (1 - m_b)$	$m_a \cdot m_b$	$m_b \cdot (1 - m_a)$	$(1 - m_a) \cdot (1 - m_b)$	

TABEAU III Affectation des masses. Cas des preuves conflictuelles.
Masses affectation for conflicting evidences.

Preuve 1	m_a		$(1 - m_b)$	$(1 - m_b)$	$\Sigma = 1 - m_a \cdot m_b$
Preuve 2	$(1 - m_b)$		m_b	$(1 - m_b)$	
Consensus	$m_a \cdot (1 - m_b)$		$m_b \cdot (1 - m_a)$	$(1 - m_a) \cdot (1 - m_b)$	

En utilisant ces règles d'assemblage, nous fusionnons deux à deux toutes les preuves disponibles. Nous obtenons ainsi une répartition de masses sur les sous-ensembles de Θ pointés par les preuves et sur ceux générés par le processus de fusion (intersections). L'exemple 4 illustre ce concept.

méthodes fiabilistes prennent en compte l'aléa de chaque information initiale à laquelle le décideur peut associer une modélisation stochastique ; il introduit alors ces modélisations dans des codes probabilistes qui évaluent une probabilité de défaillance (P_f) ou un index de fiabilité tel que l'index β_{HL} proposé par Hasofer A.M., Lind N.C. (1974).

Partie B – EXEMPLES EN GEOTECHNIQUE

4

Exemple 1 : optimisation d'une fondation selon sa fiabilité

Sur ce premier exemple nous utilisons le concept de fiabilité possibiliste basé sur l'utilisation de SEF et introduit par Boissier D., Alhajjar J., Mammeri H. (1995).

4.1

Définitions

- **Fiabilité**

La fiabilité d'une structure est son aptitude à remplir ses objectifs de conception pendant une durée de référence spécifiée ; la défaillance d'une structure est son incapacité à remplir ses objectifs.

- **Fiabilité probabiliste**

En contexte probabiliste, la fiabilité est liée à la probabilité de défaillance par référence à un état limite. Les

- **Fiabilité possibiliste**

Dans un contexte incertain et imprécis nous définissons la fiabilité possibiliste comme la compatibilité entre une proposition de solutions et des états limites (objectifs).

L'analyse possibiliste consiste alors à associer à cette compatibilité un degré de vraisemblance. Elle prend en compte l'imprécis des informations à laquelle le décideur associe un modèle possibiliste, en général un sous-ensemble flou (SEF) ; un code de calcul possibiliste permet alors d'associer une vraisemblance aux résultats à comparer aux objectifs.

4.2

Cas du dimensionnement d'une fondation

4.2.1

Cas d'étude et fiabilité déterministe

L'exemple est celui du calcul de la portance ultime q_u d'un élément de fondation superficielle à partir des résultats d'essais de laboratoire. Nous considérons une semelle filante soumise à une charge verticale et uniformément répartie sur sa longueur. Les hypothèses de l'application sont : théorie de plasticité parfaite, fondations parfaitement rigides, actions verticales et centrées et cas de sol naturel et homogène.

La portance ultime q_u (éq. 7) du facteur de sécurité F (éq. 8) et de la probabilité de défaillance P_f s'exprime par :

$$q_u = \frac{1}{2} \gamma * B * N_\gamma + c * N_c + \gamma * D * N_q \quad (7)$$

$$F = q_u / q \quad (8)$$

$$P_f = \text{Probabilité } (q_u / q < 1) \quad (9)$$

avec B : largeur (ici $B = 1,50$ m) ; D : profondeur (ici $D = 0,50$ m) ; C : cohésion ; α : angle de frottement interne ; φ : poids volumique des terres ; q : charge ; N_γ , N_c , N_q : facteurs de surface, de cohésion et de profondeur.

Les causes d'incertain sont la variation spatio-temporelle des propriétés du sol et de la charge appliquée ; les causes d'imprécision sont l'empirisme et la simplification des modèles mécaniques et des méthodes de calcul, la mauvaise qualité de la reconnaissance (incomplétude, imprécision des mesures, irrégularité des sondages...).

4.2.2

Fiabilité probabiliste

Les espérances mathématiques des données sont résumées dans le tableau IV (CV est le coefficient de variation) ; les variables aléatoires sont supposées gaussiennes et indépendantes.

TABLEAU IV Données probabilistes.
Data : laws of probability.

q (kN/m)	CV	C (kPa)	CV	φ d°	CV	γ (kN/m ³)	CV
200	0,10	30	0,20	20	0,10	20	0,05

Le calcul probabiliste du critère $F = q_u / q$ (éq. 8) donne une probabilité de défaillance $P_f = \text{Proba}[F < 1] = 3,90.10^{-3}$ (éq. 9) et un indice de Lind Hasofer $\beta_{HL} = -2,66$.

Deux solutions ou deux variantes de dimensionnement peuvent alors être comparées à partir de ces critères ; l'interprétation est simple mais les approximations, estimations et hypothèses inhérentes à cette méthode ne sont pas explicites.

4.2.3

Fiabilité possibiliste

Nous décrivons la connaissance que nous avons de chaque caractéristique par un intervalle de valeurs numériques auquel nous associons un sous-ensemble flou ; ce modèle quantitatif peut être alors introduit dans un code de calcul. Le degré de vraisemblance associé à chacune des valeurs, bien qu'il soit subjectif, donne une information sur la qualité de la caractéristique ; il provient du jugement que nous pouvons porter sur la représentativité de la caractéristique analysée dans le massif (par référence à des caractéristiques bibliographiques, en fonction de la qualité de la mesure...) ou de l'action.

Ici, nous avons associé un SEF de type trapézoïdal à chaque donnée (tableau V).

La portance ultime q_u et le facteur de sécurité F sont calculés à partir des équations (7) et (8) par les opérations de convolution sur les SEF associés à C , φ et γ ; l'imprécision de l'angle de frottement φ induit des imprécisions sur les facteurs de portance N_c , N_q et N_γ .

TABLEAU V Données possibilistes.
Data : fuzzy sets.

Paramètre	Valeur inférieure	Valeur supérieure	Étalement à gauche	Étalement à droite
C (kPa)	20	35	5	5
φ (d°)	15	25	2	1
γ (kN/m ³)	18	22	1	2
q (kN/m)	160	260	10	20

Le tableau VI récapitule les résultats possibilistes.

TABLEAU VI Résultats possibles.
Fuzzy results.

Critère	Valeur inférieure	Valeur supérieure	Étalement à gauche	Étalement à droite
Q_u (kN/m)	303,39	1 049,48	92,74	266,88
$F = q_u / q$	1,2	6,5	0,5	2,2

L'évaluation possibiliste fournit une distribution de possibilité du coefficient q_u / q . Cette distribution est caractérisée par un noyau et un support de valeurs possibles. L'objectif est représenté par une valeur (3 dans cet exemple, tableau VI et figure 3).

La figure 3 montre qu'il n'y a pas de conflit apparent entre la solution et l'objectif parce que la valeur de l'objectif - 3 ici - appartient au noyau de q_u / q . Si l'objectif avait été de 7 par exemple, la solution n'aurait eu

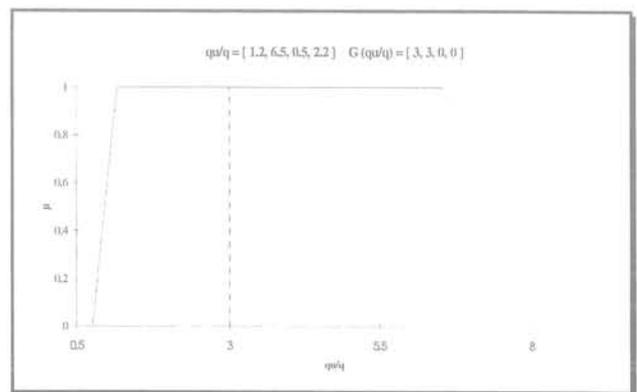


FIG. 3 Comparaison de l'évaluation imprécise et de l'objectif déterministe.

Unprecise assessment and deterministic objective comparison.

qu'un degré de 0,78 de compatibilité avec l'objectif. Pour un objectif supérieur à 8,8 (sur dimensionnement) ou inférieur à 0,5 (sous dimensionnement) la solution aurait été totalement incompatible.

4.3

Conclusion sur l'exemple 1

L'application au cas de la fondation des deux approches possibilistes et probabilistes pour l'analyse de la fiabilité montre qu'elles sont complémentaires.

L'approche possibiliste nécessite peu d'informations ; la modélisation est plus proche des pratiques des concepteurs ; elle est fiable mais peu informative ici.

L'approche probabiliste nécessite des informations beaucoup plus précises ; le langage de modélisation est éloigné des pratiques des concepteurs et cette approche est complexe à mettre en œuvre ; par contre elle se ramène à un critère facilement utilisable dont la précision explicite cache le fait que les lois de probabilités retenues ne sont en rien justifiées.

5

Exemple 2 : application au tassement d'une fondation superficielle

Ce deuxième exemple présenté par D. Boissier, J. Alhajjar et H. Mammeri (1995) porte sur l'aptitude d'une proposition technique (ici une solution par semelle filante) à respecter un cahier des charges (ici minimiser le tassement). L'évaluation du tassement, obtenue à partir de mesures, est imprécise, et le tassement limite que l'on ne souhaite pas dépasser est lui aussi imprécis.

Il faut alors voir quelle est la compatibilité entre une solution proposée et un objectif à atteindre tous deux imprécis. Une réponse est alors possible sous la forme de la compatibilité entre deux SEF ; ici il s'agit de la compatibilité entre le tassement prévisionnel (SEF A) et le tassement admissible (SEF B).

5.1

Tassements prévisibles et souhaités

5.1.1

Tassements prévisibles

Il s'agit de l'estimation du tassement d'une semelle filante par la méthode pressiométrique de Ménard. Le tassement pressiométrique total (s_t) s'exprime dans l'équation 11 par la somme de deux termes : le tassement de consolidation (s_c) et déviatorique (s_d) (éq. 10) :

$$\begin{aligned}
 s_c &= \frac{\alpha}{9E_{Mc}} (q - q_0) \lambda_c B s_d \\
 &= \frac{2}{9E_{Md}} (q - q_0) \lambda_d B_0 \left(\lambda_d \frac{B}{B_0} \right)^\alpha
 \end{aligned}
 \tag{10}$$

et

$$s_t = s_c + s_d ; \tag{11}$$

avec B : largeur (B = 1,50 m), D : profondeur d'ancrage (D = 1,0 m), ql : charge (ql = 1 800 kN/ml) qui donne une contrainte q (ici q = 1,2 MPa).

À partir d'une analyse géostatistique du profil du module pressiométrique, nous associons à E_{Mc} et E_{Md} deux intervalles flous de représentation trapézoïdale (Fig. 1a et éq. 12) :

$$E_{Mc} : (18 ; 20 ; 2 ; 2 ; 1) \text{ et } E_{Md} : (20 ; 25 ; 6,5 ; 28 ; 1) \tag{12}$$

Les autres paramètres sont déterministes. Les deux termes du tassement s_c et s_d sont obtenus grâce à l'équation 10 ; le tassement total (s_t) calculé à partir de l'équation 11 est évalué après linéarisation des fonctions L et R (éq. 13), (Fig. 4) ; nous obtenons :

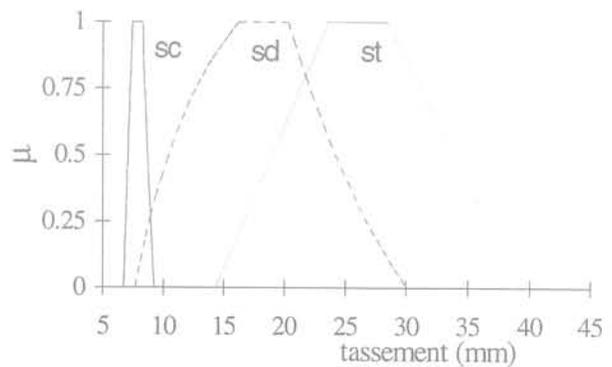


FIG. 4 Intervalles flous associés aux tassements. Fuzzy sets for settlement.

$$\begin{aligned}
 s_c &: (7,39 ; 8,21 ; 0,67 ; 1,06 ; 1) \\
 s_d &: (16,23 ; 20,28 ; 8,57 ; 9,77 ; 1) \\
 s_{ct} &: (23,61 ; 28,49 ; 9,25 ; 10,83 ; 1)
 \end{aligned}
 \tag{13}$$

5.1.2

Tassements préconisés

L'objectif associé au tassement (s_t) est défini par le sous-ensemble flou B (Fig. 5). La déclaration de cet objectif signifie que pour un tassement < 20 mm, la solution est parfaitement compatible ($\mu = 1$), que pour un tassement > 50 mm la solution est incompatible ($\mu = 0$) et que pour un tassement $\in [20 \text{ mm}, 50 \text{ mm}]$ l'adéquation de la solution est donnée par un degré de compatibilité $\mu \in [0, 1]$.

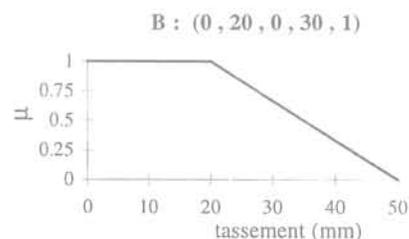


FIG. 5 Représentation de l'objectif lié au tassement. Fuzzy set for objective of settlement.

Adéquation de la semelle à atteindre l'objectif

L'adéquation d'une solution à un objectif (fiabilité) est basée sur l'intersection logique des deux intervalles ; la position des noyaux respectifs (Fig. 6) montre l'existence d'un conflit ; la nature du conflit est déterminée par la comparaison des positions des supports respectifs (Al-Hajjar J., Boissier D., Boulemia J., 1991).

Les différents degrés de compatibilité sont calculés à partir de l'expression (1) : A = tassement (s_s) ; B = objectif associé. Le profil de compatibilité entre l'objectif et s_s est représenté par la figure 7. La solution n'est que partiellement compatible avec l'objectif ; le degré de compatibilité maximal est de 0,85.

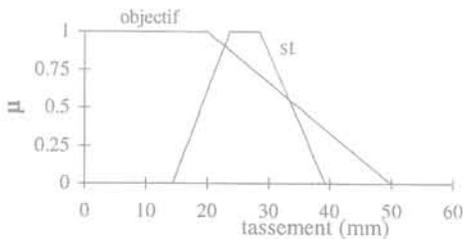


FIG. 6 Intersection de l'évaluation (s_s) et de l'objectif (B).
s, Assessment and B objective intersection.

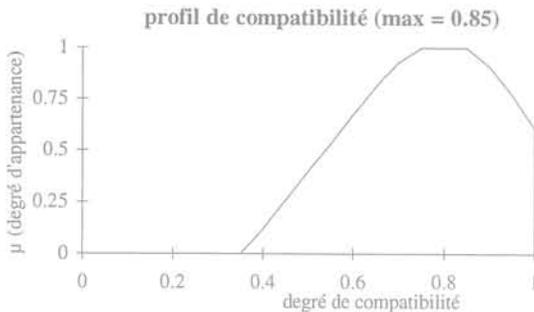


FIG. 7 Profil de comptabilité entre l'évaluation et l'objectif.
Profile of compatibility between objective and assessment.

6

Exemple 3 : Qualités des caractéristiques de sol dans une base de données

6.1

Problème

Nous disposons d'une base de données qui contient, pour chaque type de matériau, un intervalle de variation pour une caractéristique donnée ; à cha-

cun de ces intervalles est associé un indicateur de la qualité de l'estimation de cet intervalle normé entre 0 et 1 (noté conf. : confiance accordée à l'estimation). Un utilisateur fournit une caractéristique d'un sol ainsi que la confiance qu'il accorde à la mesure : à quelle classe peut-on associer ce matériau ?

6.2

Exemple (Boissier D., Boulemia C., Semman I., 1998)

Les données fournies par l'utilisateur concernant le module pressiométrique pour une couche d'argile sont : $E = 4,5$ MPa et $\text{conf}_E = 0,3$. Les cas de référence trouvés dans la bibliographie ou dans une base de données concernant le module pressiométrique pour une couche d'argile sont présentés dans le tableau suivant ; E_{\min} et E_{\max} sont les bornes des intervalles estimés (Tableau VII).

TABLEAU VII Domaine de variation de E pour une argile (Nuyens, 73), (Fascicule, 93).
Range of E for a clay (Nuyens, 73), (Fascicule, 93).

Caractéristique	Nature-couche	Consistance_couche	E_{\min} E_{\max}
		Molle	[0, 3] Mpa
E	Argile	Plastique	[3, 8] Mpa
		Raide	[8, 40] Mpa

À chaque intervalle, on associe une modélisation par SEF donnée dans le tableau VIII avec les notations de la figure 1a où U est l'ensemble des cohésions possibles pour les sols.

TABLEAU VIII Résultats de la modélisation des domaines de variation.
Fuzzy parameters.

Consistance-couche	$m - \alpha$	m	n	$n + \beta$
Molle	0,00	0,00	1,50	4,50
Plastique	1,33	4,67	4,95	9,67
Raide	12,67	18,67	29,33	50,67

Nous constatons sur la figure 8 que le cas décrit par l'utilisateur chevauche les trois cas de référence.

Le calcul de la confiance accordée donne les résultats suivants (tableau IX).

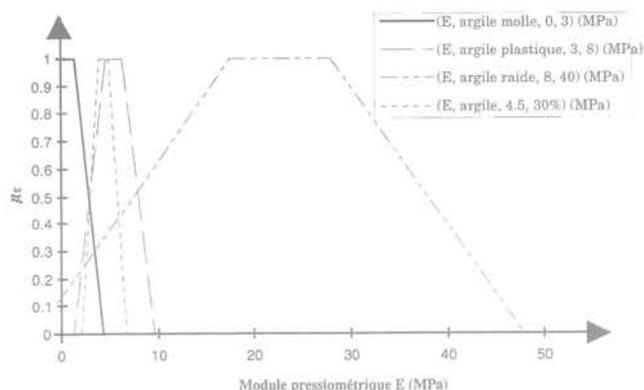


FIG. 8 Comparaison entre le cas décrit et les cas de référence.
Objective and assessment comparison.

TABLEAU IX Résultats de la discrimination et décision.
Results of the discrimination and decision making.

Caractéristique	Nature _couche	Consistance _couche	$[E_{\min}, E_{\max}]$ Mpa	Confiance
		Moille	[0, 3] Mpa	0,20
E	Argile	Plastique	[3, 8] Mpa	0,95
		Raide	[8, 40] Mpa	0,48

La proposition « la couche considérée est une argile plastique » est retenue, car elle a le degré de vraisemblance le plus élevé et proche de 1 (confiance = 0,95) est retenue ; la proposition « la couche considérée est une argile raide » est possible (confiance = 0,48) ; la proposition « la couche considérée est une argile molle » est rejetée (confiance = 0,2).

6.3

Conclusion de l'exemple

La méthode que nous avons proposée permet de quantifier la possibilité qu'une caractéristique appartienne à une classe donnée ; l'exemple très simple présenté ici peut être étendu à des combinaisons d'informations successives permettant chaque fois d'infirmer ou de confirmer cette appartenance.

7

Exemple 4 : fusion de données

Nous illustrons la règle de fusion de Dempster pour la question « quelle est la cohésion du sol » ; nous proposons deux exemples simples sous condition de concordance des données (Fig. 9) et sous condition de conflit (Fig. 10)

Considérons pour un sol donné $c = 80$ kPa comme borne supérieure de l'ensemble « ignorance » (la valeur maximale des cohésions). Supposons que nous ayons obtenu deux preuves (réponses) à cette question.

7.1

Cas de preuves concordantes

Preuve 1 : l'expert annonce une cohésion entre 20 et 50 kPa ; la masse de croyance associée à cette information est de 0,6.

Preuve 2 : un deuxième expert annonce une cohésion entre 40 et 80 kPa ; la masse de croyance associée à cette information est de 0,7.

La preuve consensus a donc une masse maximale pour l'intervalle [40, 50] kPa (Fig. 9).

7.2

Cas de preuves discordantes

Preuve 1 : l'expert annonce une cohésion entre 20 et 40 kPa ; la masse de croyance associée à cette information est de 0,6.

Preuve 2 : un deuxième expert annonce une cohésion entre 50 et 80 kPa ; la masse de croyance associée à cette information est de 0,7.

Le principe de normalisation apparaît clairement dans cet exemple (Fig. 10). La preuve consensus a donc une masse maximale pour l'intervalle [50, 80] kPa ; la masse d'ignorance (0,21) est plus importante que dans le cas 1 (0,12).

7.3

Commentaires

7.3.1

Sur les données

Les valeurs peuvent provenir de sources diverses que nous avons regroupées sous le vocable « dire d'expert » ; en particulier elles peuvent être le résultat d'un essai ou d'un calcul.

Les masses de croyance sont de type déclaratif et estimées soit subjectivement soit à l'aide d'une grille d'analyse construite à partir de critères tels que la fiabilité de la méthode d'évaluation, les références de l'organisme donnant l'information, la ressemblance entre les conditions d'obtention de la donnée et les conditions du cas réel (labo/in situ).

7.3.2

Sur la méthode

Nous avons appliqué la méthode au cas de deux preuves ; elle est bien sûr généralisable au cas de plusieurs preuves ; son application impose à l'utilisateur de s'interroger non seulement sur l'information mais aussi sur sa fiabilité, car il est obligé de déclarer la confiance (ou la méfiance) qu'il a en une information.

7.3.3

Sur les résultats

Les masses de croyance obtenues peuvent être utilisées directement en aide à la décision (affectation d'une

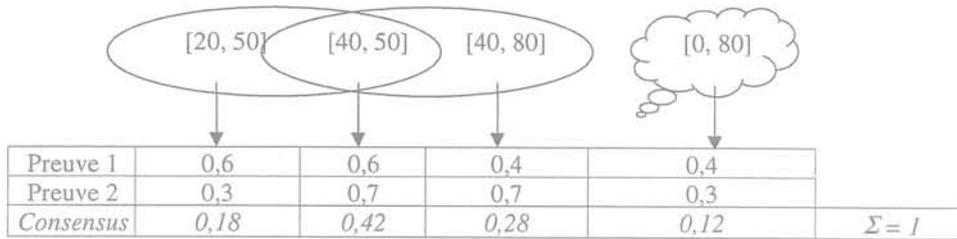


FIG. 9 Affection des masses. Cas des preuves concordantes.
Masses allocation. Heterogeneous evidences.

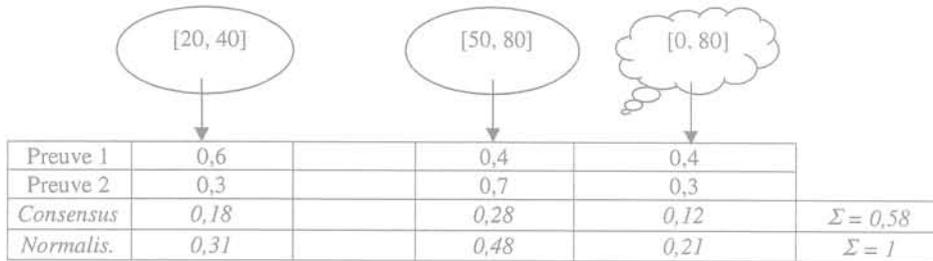


FIG. 10 Affection des masses. Cas des preuves conflictuelles.
Masses allocation. Conflicting evidences.

valeur à une classe) ou introduites dans un calcul. Elles n'ont qu'une valeur comparative.

8

Conclusion

• Intérêts et limites

L'intérêt principal de ces méthodes possibilistes est d'obliger l'utilisateur à expliciter la confiance qu'il accorde à une information ; on ne peut plus faire comme si les données étaient exactes et on accepte de traiter des informations approximatives ; cet intérêt est aussi un des facteurs de rejet de ce type de méthode jugée parfois dérangement.

Techniquement, ces méthodes peuvent utiliser aussi bien des informations qualitatives que quanti-

tatives, des statistiques pauvres ou des champs d'expérimentation complets et les utilitaires d'aide à la modélisation (transformation en SEF par exemple) ne demandent pas de connaissances très sophistiquées.

• SEF ou modèles probabilistes ?

Les modélisations probabilistes et possibilistes ne sont pas conflictuelles mais alternatives ou complémentaires (La définition fréquentielle de la possibilité permet éventuellement de faire le lien entre les deux distributions). La modélisation possibiliste paraît mieux adaptée à la gestion du qualitatif et de l'information incomplète ; elle correspond aux phases primaires de conception et à l'aide à la prise des décisions principales (choix de parti). La modélisation probabiliste demande de l'information quantitative et complète ; elle est mieux adaptée à la phase de calcul détaillé.

Bibliographie

Aggarwal N., Quarrey P., Berry C. (1993) - « Application des ensembles flous et de l'aide à la décision au domaine de l'orientation professionnelle ». Colloque « Les Applications des ensembles flous », Nîmes, 09/1993, Éditions EC2, p. 41-50.

Al-Hajjar J., Boissier D., Boulemia C. (1991) - « Building substructural design as fuzzy decision model », Mexico, CR de ICASP6, vol. 2, p. 1076-1083.

Boissier D., Al-Hajjar J. (1990) - « Logique floue et aide à la décision ; application à l'infrastructure des bâtiments ». Première table ronde sur la modélisation des processus de conception créative, Cabourg.

Boissier D., Al-Hajjar J., Mammeri H. (1995) - « Fiabilité et possibilité en mécanique des sols : application à l'infrastructure des bâtiments ». 12^e Congrès de mécanique, Strasbourg, 4-8 sept.

Boissier D., Boulemia C., Semaan I. (1998) - « Traitement de l'incertain dans une base de connaissances sur le sol ». Première Conférence internationale sur les nouvelles technologies de l'information pour l'aide à la décision dans le domaine du génie civil, Montreal, Canada, 11-13 octobre 1998. Miresco E.T., éd., p. 745-756.

Bouchon-Meunier B. (1995) - *La logique floue et ses applications*. Amsterdam, Addison-Wesley, p. 84-92.

Dubois D., Prade H. (1983) - « Ranking fuzzy numbers in the sitting of possibility theory ». *Information Sciences*, n° 30, New York, Elsevier Science Publishing Co., Inc., p. 183-224.

Haofer A.M., Lind N.C. (1974) - « Exact and second invariant second moment code format ». *Engineering Mechanics Division ASCE*, 100.

Lair J. - « Evaluation de la durabilité de systèmes constructifs du bâtiment ». Thèse de doctorat de l'université Blaise Pascal, septembre 2000.

Shafer G. (1976) - *A Mathematical Theory of Evidence*. Chichester, Princeton University Press, 297 p.

Zadeh L.A. (1965) - « Fuzzy sets ». *Information and Control*, n° 8, p. 338-353.