

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITE MOULOUD MAMMERI

- TIZI OUZOU -

Faculté du Génie de la construction
Département de Génie Civil

MÉMOIRE DE MAGISTER

Spécialité : *génie civil*
Option : *structures et matériaux*

Présenté par :

Fatma ZIBANI

THÈME

***Effet de la variabilité des paramètres de sol sur
le comportement des ouvrages de
géotechnique***

Devant la commission d'examen :

Mr Youcef BOUAFIA	Professeur - UMMTO	<i>Président</i>
Mr Ali BOUHERAOUA	Maître de conférences (A) - UMMTO	<i>Rapporteur</i>
Mr Ali HAMZA	Maître de conférences (A) - UMMTO	<i>Examineur</i>
Mr Hacène AIT AIDER	Maître de Conférences (A) - UMMTO	<i>Examineur</i>

Soutenu le : 01/07/ 2012

AVANT PROPOS

Que toutes les personnes qui m'ont aidé à réaliser ce travail de recherche trouvent ici l'expression de ma gratitude.

Les dettes de reconnaissances contractées tout au long de la confection de ce travail vont d'abord à Monsieur **BOUHERAOUA Ali**, qui m'a proposé le thème de ce travail. Je lui suis reconnaissante pour son aide et son encadrement.

Je tiens à remercier Monsieur **BOUAFIA Youcef**, Professeur à l'UMMTO d'avoir accepté de présider le jury.

Mes remerciements vont aussi à monsieur **HAMZA Ali**, Maître de conférence à l'UMMTO, et à monsieur **AIT AIDER** Hacène, Maître de conférence à l'UMMTO, de m'avoir fait l'honneur d'examiner ce travail.

J'adresse mes chaleureux remerciements à mes parents pour leur soutien tout le long de mes études, mes frères et sœurs ainsi qu'à mon mari, **HANAFI**, pour son aide, sa patience, ses conseils, ses encouragements et son soutien moral surtout pendant les moments difficiles.

SOMMAIRE

INTRODUCTION	01
Chapitre I <i>ELEMENTS DE STATISTIQUES ET PROBABILITES</i>	
I.1 Généralités	05
I.2 Distributions statistiques et représentations associées	06
I.3 Fréquences cumulées et fonction de répartition	12
I.4 Caractéristiques d'une distribution. Tendance centrale et dispersion	13
I.5. Modèle théorique de distribution. Variables aléatoires d'échantillonnage	20
Chapitre II <i>SIMULATION ET GENERATION DE VARIABLES ALEATOIRES</i>	
II.1 Simulation	26
II.2 Génération de variables et vecteurs aléatoires	32
II.3 Simulation de Monte Carlo	40
Chapitre III <i>LES TECHNIQUES DE GEOSTATIQUE</i>	
III.1 Introduction	44
III.2 Bases de la géostatistique	45
III.3 Propriétés du variogramme	47
III.4 Variation d'estimation et variance de dispersion	50
III.5 Le Krigeage	51
Chapitre IV <i>VARIABILITE SPATIALE DES PARAMETRES GEOTECHNIQUES</i>	
IV.1 Introduction	59
IV.2 sources d'incertitudes sur les paramètres de sol	60
IV.3 Approche probabiliste	64
IV.4 Modélisation stochastique de la variabilité spatiale	67
IV.5 Conclusion	72
Chapitre V <i>CORRELATIONS ENTRE LES PROPRIETES DES SOLS</i>	
V.1 Introduction	73
V.2 Relations et corrélations dans les sols : généralités	73
V.3 Principales techniques d'étude des corrélations	76
V.4 Exemples de corrélations	83
V.5 Domaines de validité des corrélations	88
Chapitre VI <i>METHODE PRATIQUE DE CALCUL PROBABILISTE</i>	
VI.1 Introduction	91
VI.2 Approximation par les séries de Taylor	92
VI.3 Approximation par intégration numérique	93
VI.4 Approximation par une loi normale ou log normale (Lamb, 1974)	95
VI.5 Simulation : méthode de MONTE-CARLO	96
VI.6 Estimation ponctuelle : méthode de Rosenblueth	100

Chapitre VII ANALYSE PROBABILISTE DE QUELQUES OUVRAGES

VII.1 Introduction	105
VII.2 Coefficient de sécurité et probabilité de rupture	105
VII.3 Méthode de calcul probabiliste du comportement des ouvrages	110
VII.4 Analyse probabiliste de la stabilité des pentes	111
VII.5 Processus de mobilisation de la résistance au cisaillement	124
VII.6 Principes généraux des méthodes d'analyse probabiliste de la stabilité des pentes	131
VII.7 perspectives	133
VII.8 Analyse de la consolidation bidimensionnelle des sols	134
VII.9 Capacité portante	151
VII.10 Autres calculs	152

Chapitre VIII ANALYSE PROBABILISTE DE QUELQUES CAS D'ETUDE

VIII.1 Introduction	156
VIII.2 Méthodologie et outils de modélisation	156
VIII.3 La consolidation	156
VIII.4 Analyse probabiliste du comportement des murs de soutènement	204
VIII.5 Analyse probabiliste d'une semelle	217

CONCLUSION	236
-------------------	-----

LISTE DES FIGURES

Figure I.1 Représentations graphiques des variables nominales

Figure I.2 Représentations graphiques des variables continues

Figure I.4 Fonction de répartition

Figure I.5 Classe médiane

Figure I.6 Coefficients d'asymétrie

Figure I.7 Distribution t de Student (type VII de Pearson)

Figure I.8 Exemples de distribution gamma.

Figure I.9 Classification des distributions selon la valeur de K

Figure I.10 La courbe en cloche

Figure II.1 Procédure de génération de valeurs aléatoires d'une variable X dont la fonction de répartition G (x) est connue.

Figure III.1 Exemple de variogramme

Figure III.2 Ajustement du variogramme- Modèle théorique.

Figure III.3 Modèles de variogramme

Figure III.4 Variance d'estimation (a) et variance de dispersion (b)

Figure III.4 Variance d'estimation (a) et variance de dispersion (b)

Figure IV Schéma général de calcul par un modèle de type probabiliste

Figure IV Les différentes formes de variation des paramètres de sol en fonction de la profondeur (Jaksa, 1995).

Figure IV Modification des caractéristiques d'autocorrélation par séparation de la dérive (Magnan, 1982)

Figure V.1 Estimation de la valeur la plus probable de y pour une valeur donnée de x, sur la base de l'analyse de régression de X et Y

Figure V.2 Relations entre la teneur en eau w et le coefficient de compressibilité $C_c/(1+e_0)$ pour les tourbes de Normandie

Figure V.3 Relations entre le taux de variation de la perméabilité C_k et l'indice des vides initial e_0 des argiles du Canada

Figure V.4 Relations entre la limite de liquidité w_L et l'indice de compression C_c

Figure V.5 Relations entre la pression limite nette pressiométrique et la résistance de cône statique q_c pour les sables de Dunkerque

Figure V.6 Relations entre la pression limite nette pressiométrique p_l-p_0 et la résistance de cône statique q_c pour des argiles de provenances diverses

Figure V.7 Relation entre les résistances de cône statique q_c et dynamique q_d pour des sables argileux

Figure V.8 Limitations des corrélations : variabilité à l'intérieur d'un site [données du site expérimental des Laboratoires des Ponts et Chaussées à Cubzac-les-Ponts]

Figure VI.1 Procédure de génération de valeurs aléatoires d'une variable X dont la fonction de répartition $G(x)$ est connue.

Figure VII.1 Exemple de situation complexe pour le calcul du coefficient de sécurité

Figure VII.2 Distribution des probabilités de C et R

Figure VII.3 Hypothèses géométriques et mécaniques de l'analyse de stabilité

Figure VII.4 Hypothèses géométriques

Figure VII.5 Etapes du calcul de stabilité à l'aide d'abaques

Figure VII.6 Valeurs de la cohésion non drainée mesurées dans chaque couche

Figure VII.7 Histogramme des résultats des calculs

Figure VII.8 Géométrie de la surface de glissement

Figure VII.09 Variation de la cohésion du sol

Figure VII.10 Procédure de détermination du cercle de rupture le plus défavorable

Figure VII.11 Géométrie et paramètre du cas étudié par Alonso (1976)

Figure VII.12 Position des piézomètres et conclusions des mesures

Figure VII.13 Procédure de génération de valeurs aléatoires d'une variable X dont la fonction de répartition $G(x)$ est connue

Figure VII.14 Courbes d'égale probabilité de rupture et d'égale coefficient de sécurité

Figure VII.15 Contributions de différents facteurs à la variance des moments (cercle critique)

Figure VII.16 Influence de la pression interstitielle sur la probabilité de rupture

Figure VII.17 Influence de la cohésion et du frottement sur la probabilité de rupture

Figure VII.18 Influence du degré d'autocorrélation des propriétés du sol sur la probabilité de rupture.

Figure VII.19 Relation entre le coefficient de sécurité et la probabilité de rupture.

Figure VIII.20 Comparaison des mesures et du résultat de l'analyse probabiliste (Remblai B de Cubzac-les-Ponts, Magnan et Baghery, 1982).

Figure VII.21 Maillages utilisés pour les calculs en éléments finis : maillage (a) pour le calcul de la variabilité uniquement verticale ; maillage (b) pour le cas de la variabilité verticale et horizontale.

Figure VII.22 Schéma montrant les différentes étapes nécessaires à la détermination de la solution probabiliste

Figure VII.23 Evolution au cours du temps du tassement de plusieurs points du massif (a), des écarts-types (b) et des coefficients de variation (c) (variabilité unidimensionnelle).

Figure VII.24 Evolution au cours du temps des surpressions interstitielles moyennes en plusieurs points du massif (a), des écarts-types (b) et des coefficients de variation (c) (variabilité unidimensionnelle).

Figure VII.25 Evolution au cours du temps du tassement de plusieurs points du massif (a), des écarts-types (b) et des coefficients de variation correspondants (c) (variabilité bidimensionnelle).

Figure VII.26 Evolution au cours du temps des surpressions interstitielles moyennes en plusieurs points du massif (a), des écarts-types (b) et des coefficients de variation correspondants (c) (variabilité bidimensionnelle).

Figure VII.27 Comparaison des calculs probabilistes et déterministes :

- (a) Comparaison de l'évolution des tassements de la surface du sol selon les calculs (échelle linéaire du temps logarithmique) ;
- (b) Comparaison de l'évolution des tassements de la surface du sol selon les calculs (échelle du temps logarithmique) ;
- (c) Comparaison de l'évolution des écarts-types selon les calculs (valeurs à 0 m et 3m de profondeur)

Figure VII.28 Comparaison des calculs probabilistes et déterministes :

- a) Comparaison de l'évolution des surpressions interstitielles en fonction de la profondeur aux temps 1000 et 20000 jours ;
- (b) Comparaison de l'évolution des écarts types des surpressions interstitielles ;
- (c) et des coefficients de variation correspondants.

Figure VII.29 Surfaces stochastiques de rupture.

Figure VII.30 Coefficients stochastiques de sécurité.

Figure VIII.1 Maillage utilisé dans l'étude de la consolidation unidimensionnelle dans le cas de la variabilité unidimensionnelle

Figure VIII.2 Diagramme de chargement

Figure VIII.3 Evolution, au cours du temps de la moyenne du tassement en plusieurs points du sol

Figure VIII.4 Evolution, au cours du temps de la moyenne du tassement en fonction de la profondeur

Figure VIII.5 Evolution, au cours du temps de la moyenne des surpressions interstitielles en plusieurs points du sol

Figure VIII.6 Evolution, au cours du temps, des incertitudes du tassement, en fonction de la profondeur

Figure VIII.7 Evolution des incertitudes du tassement, en fonction de la profondeur, à des temps différents.

Figure VIII.8 Evolution, au cours du temps, des incertitudes des surpressions interstitielles, à différentes profondeurs.

Figure VIII.9 Maillage utilisé dans l'étude de la consolidation unidimensionnelle dans le cas de la variabilité bidimensionnelle

Figure VIII.10 Evolution, au cours du temps de la moyenne du tassement en plusieurs points du sol.

Figure VIII.11 Evolution, en fonction de la profondeur, de la moyenne du tassement, à différents temps.

Figure VIII.12 Evolution, au cours du temps, de la moyenne des surpressions interstitielles, à différentes profondeurs.

Figure VIII.13 Evolution, au cours du temps, des incertitudes des surpressions interstitielles, à différentes profondeurs.

Figure VIII.14 Evolution, en fonction de la profondeur, du tassement, à des temps différents.

Figure VIII.15 Evolution, au cours du temps, des incertitudes des surpressions interstitielles en plusieurs points du sol.

Figure VIII.16 Evolution, au cours du temps, du tassement, à différentes profondeurs.

Figure VIII.17 Evolution, en fonction de la profondeur, à des temps différents.

Figure VIII.18 Evolution, au cours du temps, des surpressions interstitielles, à différentes profondeurs.

Figure VIII.19 Evolution, au cours du temps, des écarts types du tassement, à différentes profondeurs

Figure VIII.20 Evolution du tassement, en fonction de la profondeur, à des temps différents.

Figure VIII.21 Evolution, au cours du temps, des incertitudes des surpressions interstitielles en plusieurs points du sol.

Figure VIII.22 Evolution, au cours du temps, du tassement, à différentes profondeurs

Figure VIII.23 Evolution, en fonction de la profondeur, à différents temps.

Figure VIII.24 Evolution, au cours du temps, des surpressions interstitielles à différentes profondeurs.

Figure VIII.25 Evolution, au cours du temps, des écarts types du tassement, à différentes profondeurs.

Figure VIII.26 Evolution des écarts types du tassement, à différents temps, en fonction de la profondeur.

Figure VIII.27 Evolution, au cours du temps, des écarts types des surpressions interstitielles, à différentes profondeurs.

Figure VIII.28 Evolution, au cours du temps, du tassement, à différentes profondeurs

Figure VIII.29 Evolution du tassement, au cours du temps, à différents temps.

Figure VIII.30 Evolution, au cours du temps, des surpressions interstitielles, à différentes profondeurs

Figure VIII.31 Evolution, au cours du temps, des écarts types du tassement, à différentes profondeurs.

Figure VIII.32 Evolution du tassement, en fonction de la profondeur, à différents temps.

Figure VIII.33 Evolution, au cours du temps, des écarts types des surpressions interstitielles,

Figure VIII.34 Evolution, en fonction de la profondeur, du tassement, à des temps différents

Figure VIII.35 Evolution, au cours du temps, du tassement, à différentes profondeurs à différentes profondeurs.

Figure VIII.36 Evolution, au cours du temps, des surpressions interstitielles, à différentes profondeurs.

Figure VIII.37 Evolution, au cours du temps, des écarts types du tassement, à différentes profondeurs

Figure VIII.38 Evolution, en fonction de la profondeur, des écarts types du tassement, à des temps différents.

Figure VIII.39 Evolution, au cours du temps, des écarts types des surpressions interstitielles, à différentes profondeurs

Figure VIII.40 Evolution, au cours du temps, du tassement, à différentes profondeurs

Figure VIII.41 Evolution du tassement, au cours du temps, à des temps différents.

Figure VIII.42 Evolution des surpressions interstitielles, au cours du temps, à des temps différents.

Figure VIII.43 Evolution des écarts types du tassement, au cours du temps, à des temps différents.

Figure VIII.44 Evolution des écarts types du tassement, en fonction de la profondeur, à des temps différents.

Figure VIII.45 Evolution des écarts types des surpressions interstitielles, au cours du temps, à des temps différents.

Figure VIII.46 Comparaison des calculs probabilistes uni, bi et tridimensionnelle dans le cas de la variabilité unidimensionnelle (V1) :

- (a) Comparaison de l'évolution des tassements de la surface du sol en fonction du temps
- (b) Comparaison de l'évolution des surpressions interstitielles à 3 m de profondeur

Figure VIII.47 Comparaison des calculs probabilistes uni, bi et tridimensionnelle dans le cas des variabilités bidimensionnelle (V2) et tridimensionnelle (V3) :

- (a) Comparaison de l'évolution des tassements de la surface du sol en fonction du temps
- (b) Comparaison de l'évolution des surpressions interstitielles à 3 m de profondeur

Figure VIII.48 Présentation du mur de soutènement poids étudié.

Figure VIII.49 Maillage et conditions aux limites.

Figure VIII.50 Maillage et conditions aux limites du cas analysé.

Figure VIII.51 Maillage et conditions aux limites du cas analysé (E et ϕ variables aléatoirement d'une couche à une autre).

Figure VIII.52 Maillage et conditions aux limites du cas analysé (E et ϕ variables aléatoirement d'un bloc à un autre).

Figure VIII.53 Contraintes horizontales moyennes derrière le mur pour les différents calculs

Figure VIII.54 Contraintes horizontales max et min derrière le mur pour les différents calculs

Figure VIII.55 Evolution des écart-types des contraintes horizontales derrière le mur pour les différents calculs

Figure VIII.56 Déplacements horizontaux moyens derrière le mur pour les différents calculs

Figure VIII.57 Déplacements horizontaux max et min derrière le mur pour les différents calculs

Figure VIII.58 Evolution des écart-types des déplacements horizontaux derrière le mur pour les différents calculs

Figure VIII.59 Contraintes verticales moyennes à la base du mur pour les différents calculs

Figure VIII.60 Contraintes verticales max et min à la base du mur pour les différents calculs

Figure VIII.61 Evolution des écart-types des contraintes verticales à la base du mur pour les différents calculs

Figure VIII.62 Déplacements verticaux moyens à la base du mur pour les différents calculs

Figure VIII.63 Déplacements verticaux max et min derrière le mur pour les différents calculs

Figure VIII.64 Evolution des écart-types des contraintes verticales à la base du mur pour les différents calculs

Figure VIII.65 Schéma montrant les différentes étapes nécessaires à la détermination de la solution probabiliste

Figure VIII.66 Présentation de la fondation superficielle étudiée

Figure VIII.67 Maillage réalisé pour l'étude du mur de soutènement poids

Figure VIII.68 Maillage, conditions aux limites et chargement du cas étudié

Figure VIII.69 La position de ligne de coupe

Figure VIII.70 déplacements verticaux «V»

Figure VIII.71 Contraintes de cisaillement max « σ max»

Figure VIII.72 les Isovaleurs

Figure VIII.73 Maillage, conditions aux limites et chargement cas statistiquement homogène

Figure VIII.74 Valeurs moyenne maximales et minimales des déplacements verticaux du cas statistiquement homogène

Figure VIII.75 Valeurs moyenne des Contraintes de cisaillement maximales du cas statistiquement homogène

Figure VIII.76 Maillage, conditions aux limites et chargement du cas Stratifié

Figure VIII.77 Valeurs moyenne des déplacements verticaux du cas statistiquement Stratifié

Figure VIII.78 Valeurs moyenne des Contraintes de cisaillement maximales du cas statistiquement Stratifié

Figure VIII.79 Maillage, conditions aux limites et chargement du cas hétérogène

Figure VIII.80 Valeurs moyenne des déplacements verticaux du cas statistiquement Stratifié

Figure VIII.81 Valeurs moyenne des Contraintes de cisaillement maximales du cas statistiquement Stratifié

Figure VIII.82 Valeurs moyennes des déplacements horizontaux des différents calculs comparé au cas déterministe

Figure VIII.83 Ecart type des déplacements horizontaux pour les différents calculs

Figure VIII.84 Valeurs moyennes des contraintes de cisaillement max des différents calculs comparé au cas déterministe

Figure VIII.85 Ecart type des contraintes de cisaillement max pour les différents calculs

LISTE DES TABLEAUX

Tableau I.1 Différents types de variables statistiques

Tableau I.2 Présentation des variables statistiques

Tableau I.3 Variables continues : amplitudes et fréquences

Tableau I.4 Distributions cumulées

Tableau I.5 Caractéristiques d'une distribution

Tableau VI.1 Valeurs aléatoires pour distribution uniforme $R_U(0,1)$

(moyenne 0,4992, écart type 0,2989)

Tableau VI.2 Valeurs aléatoires pour une distribution réduite uniforme $R_U(0,1)$

(moyenne 0,1403, écart type 0,9047).

Tableau VII.1 Valeurs aléatoires pour distribution uniforme $R_U(0,1)$ (moyenne 0,4992, écart type 0,2989)

Tableau VII.2 Valeurs aléatoires pour une distribution réduite uniforme $R_U(0,1)$ (moyenne 0,1403, écart type 0,9047).

Tableau VIII.1 Caractéristiques des matériaux pour la modélisation

Tableau VIII.2 Hypothèse de calcul

Tableau VIII.3 Valeurs générées des paramètres variables c et φ

NOTATIONS PRINCIPALES

σ_h	: Contrainte latérale totale.
σ_v	: Contrainte verticale totale.
K	: Coefficient de la pression des terres.
σ'_h	: Contrainte latérale totale effective.
σ'_v	: Contrainte verticale totale effective.
K_0	: Coefficient des terres au repos.
φ	: Angle de frottement interne du sol.
c	: Cohésion totale du sol.
γ	: Poids volumique total du sol.
c'	: Cohésion effective du sol.
φ'	: Angle de frottement interne effectif du sol.
E	: Module d'Young.
ν	: Coefficient de poisson.
ψ	: Angle de dilatance.
K_a	: Coefficient de la pression active.
K_p	: Coefficient de la pression passive.
p_a	: Pression active.
p_p	: Pression passive.
F_a	: Résultante de la pression active par mètre linéaire de largeur.
F_p	: Résultante de la pression passive par mètre linéaire de largeur.
δ	: L'angle de frottement écran-sol.
c_w	: Adhérence sol-mur.
W	: Poids du mur.
S	: Sous-Pression.
P_p	: butée des terres.
P_a	: Poussée des terres.
R	: Réaction du sol de fondation.
F_r	: Coefficient de stabilité au renversement.
F_g	: Coefficient de sécurité au glissement.
e	: Excentrement de la composante normale de R .
$U(x)$	: Fonction approchée
$U_{ex}(x)$	: Fonction exacte.
$N_i(x)$	: Fonctions d'interpolation nodales ou fonctions de forme.

Ω	: Domaine.
Ω^e	: Sous domaines.
$[N]$	: Matrices des fonctions d'interpolation nodale.
$[B]$	: Matrice de déformation.
$[D]$	: Matrice qui contient les propriétés élastiques du matériau.
$[K]$	: Matrice de rigidité globale.
$[k^e]$	: Matrice de rigidité de l'élément.
$\{U\}$	: Vecteur déplacement global.
$\{F\}$	: Vecteur des forces nodales.
f_i	: Fréquence.
$\bar{x}$	: Moyenne arithmétique.
$V(X)$	: Variance.
$\sigma(X)$	: Ecart-type.
$C_v(X)$	: Coefficient de variation.
$E(X)$	: Espérance mathématique.

RESUME

La mécanique des sols repose sur des hypothèses simplificatrices qui ignorent le caractère hétérogène des sols. Les méthodes de calcul actuellement utilisées en géotechnique (dites méthodes déterministes) se basent sur des lois de comportement du sol ou des valeurs fixes sont attribuées aux paramètres figurant dans les équations du modèle mathématique adopté.

Les variations d'un point à un autre des propriétés physiques et mécaniques du sol crée pourtant des incertitudes sur les valeurs représentatives des paramètres de calcul, ce qui se traduit par des incertitudes sur la prévision du comportement des ouvrages.

Et malgré le développement d'outils de calculs (méthodes numériques) de plus en plus performants et une description de sol de plus en plus proche de la réalité, il existe toujours un écart entre les résultats de calcul et les valeurs réelles.

C'est dans l'espoir de réduire ce biais existant, ou, tout au moins de progresser dans la compréhension de ces divergences, que les géotechniciens poussent leurs investigations dans l'application de techniques diverses, dont les statistiques et probabilités, avec comme objectifs de prendre en compte l'effet de la variabilité des propriétés physiques et mécaniques de sol dans les divers calculs d'ouvrages.

En d'autres termes, on doit résoudre le problème suivant : comment, à partir des résultats d'essai de laboratoire et in situ, peut-on évaluer au mieux la variabilité naturelle des paramètres de sol, dont l'échantillonnage est souvent fragmentaire et limité, et comment prendre en compte cette variabilité dans les calculs pour déterminer la solution d'un problème de mécanique des sols ?

La description et l'analyse de l'incertain en géotechnique, avec pour champs d'application la sécurité des ouvrages et les risques, sont des questions à la mode en ce début du 21^{ème} siècle. Le présent rapport, composé de sept chapitres, dresse un petit état de l'art des techniques d'analyse des sols hétérogènes et apporte un éclairage sur une façon différente d'estimer les solutions de problèmes géotechniques.

Le premier chapitre a trait à la description de la variabilité des paramètres géotechniques. Il justifie l'opportunité d'une approche nouvelle pour prendre en charge le problème que pose l'hétérogénéité des sols.

Les méthodes pratiques de calcul probabiliste qui sont exposées au chapitre VI, exigent la maîtrise d'un certain nombre de notions telles que les statistiques et probabilités (chapitre I), les techniques de géostatistique (chapitre III) et les notions de simulation que nous avons tenté de vulgariser au chapitre II. Pour une meilleure description des propriétés des sols, il nous a semblé utile d'exposer les corrélations entre propriétés des sols (chapitre V).

Le chapitre VII donne une illustration de quelques cas d'application concernant les calculs d'ouvrages de géotechniques, comme le problème de stabilité qui représente une des plus fréquentes applications de l'approche probabiliste dans le domaine géotechnique ou l'analyse probabiliste de la consolidation. Cela met en œuvre des techniques diverses pour venir à bout des difficultés inhérentes à la variabilité des paramètres de sols.

On termine ce travail par une quelques cas de calculs probabilistes.

ABSTRACT

Soil mechanics is based on simplifying assumptions which ignore the heterogeneity of soils. The calculation methods currently used in geotechnical engineering (known deterministic methods) are based on laws of behaviour from the ground or fixed values are assigned to parameters in the equations of the mathematical model adopted.

The variations from one point to another of the physical and mechanical properties of soil, however, creates uncertainty about the values of representative parameters of calculation, which translates into uncertainty about predicting the behaviour of structures. And despite the development of computational tools (numerical methods) more efficient and a description of soil increasingly close to reality, there is still a gap between results and the actual values.

In the hope of reducing this existing bias, or at least progress in understanding these differences, that geotechnicians pushing their investigations in the application of various techniques, including statistics and probability, with the aim to take into account the effect of variability of physical and mechanical properties of soil in the various calculations of books.

In other words, we must solve the following problem : how, from the results of laboratory testing and in situ, can we best assess the natural variability of soil parameters, including sampling is often fragmentary and limited, and how to take this variability into account in the calculations to determine the solution to a problem of soil mechanics?

The description and analysis of uncertainty in geotechnical engineering, with the scopes of works and safety hazards, are matters to be fashionable at the beginning of 21st century. This report, consisting of seven chapters, provides small state of the art techniques for analyzing heterogeneous soils and sheds light on a different way to estimate the solutions of geotechnical problems.

The first chapter deals with the description of the variability of geotechnical parameters. It justifies the desirability of a new approach to handle the problem of the heterogeneity of soils. These analysis techniques outlined in Chapter VI requires the mastery of a number of concepts such as statistics and probability (Chapter II), geostatistical techniques (Chapter IV) and the notions of simulation that we have tried to popularize in the Chapter III. For a better description of soil properties, it seemed useful to explain the correlations between soil properties (Chapter V).

Chapter VII is an illustration of some cases of application for the calculation of geotechnical structures, such as the stability problem which is one of the most common applications of the probabilistic approach in the geotechnical analysis or probabilistic of consolidation. It implements various techniques to overcome the difficulties inherent in the variability of soil parameters.

We concludes this work by exempls

INTRODUCTION GENERALE

INTRODUCTION GENERALE

Traditionnellement, l'analyse et le dimensionnement des ouvrages en géotechnique sont basés sur des approches déterministes. Dans ces approches, les aléas et incertitudes des différents paramètres (caractéristiques du sol, chargement, etc...) sont pris en compte de manière simplifiée sous la forme d'un facteur de sécurité global. Ce facteur est souvent appelé facteur d'ignorance car il représente notre ignorance des incertitudes des différents paramètres. Pour tenir compte des aléas et incertitudes inhérents aux différents paramètres, la théorie de la fiabilité est actuellement de plus en plus utilisée en géotechnique. Ceci est devenu possible grâce aux avancées importantes au niveau de la quantification des incertitudes des paramètres du sol.

Le travail de cette thèse concerne l'utilisation des méthodes fiabilistes dans le calcul des ouvrages géotechniques.

Les méthodes de calcul actuellement utilisées en géotechnique (dites méthodes déterministes) se basent sur des lois de comportement du sol ou des valeurs fixes sont attribuées aux paramètres figurant dans les équations du modèle mathématique adopté. Les facteurs de sécurité ainsi calculés sont comparés à des différentes sources d'incertitudes comme par exemple l'échantillonnage limité, les erreurs de mesure inévitables, l'imperfection des modèles mathématiques et la variabilité dans le temps et dans l'espace des principaux paramètres géotechniques. Il est évident que ces derniers ne peuvent jamais être évalués de manière entièrement déterministe.

C'est pourquoi les résultats finaux d'une analyse géotechnique deviennent à leur tour affectés par une certaine quantité d'incertitude.

Il a semblé utile, dans cette étude, de tenter d'énumérer l'ensemble des techniques d'analyse des massifs de sols hétérogènes. Ces techniques prennent en compte le caractère hétérogène des sols dans la description des propriétés physiques

et mécanique des sols et dans les calculs d'ouvrages géotechniques en estimant les risques et les incertitudes qui en découlent.

La variabilité élevée des paramètres géotechniques, leur caractère évolutif dans le temps et l'espace et l'incertitude importante qui les affectent, rendent nécessaire leur prise en compte sous forme de champs aléatoires. Ces derniers sont définis comme des variables aléatoires auxquelles on associe une localisation dans l'espace.

L'approche probabiliste, et c'est la méthode principale dans il s'agit, tient compte du caractère aléatoire des variables. Elle fournit un résultat exprimé en probabilité de rupture. Celle-ci constitue une mesure homogène de la fiabilité qui peut en outre s'intégrer dans une étude économique globale.

Les méthodes de calcul probabiliste existantes sont ainsi passées en revue et une analyse critique en est faite dans l'optique de leur application aux problèmes de mécanique des sols. Sur la même base de la théorie de l'information, d'autres méthodes sont examinées telle la méthode d'estimation ponctuelle de Rosenblueth ou la méthode d'approximation par les séries de Taylor. Auparavant, il était nécessaire d'insister sur quelques outils élémentaires mais essentiels sans quoi toutes ces techniques n'auraient pas été possibles.

Le chapitre I donne quelques éléments et autres notions de ce qu'il est obligatoire de maîtriser si l'on veut aller au bout des projets : les notions de statistiques et probabilistes. Celles-ci sont une façon pénétrer l'inconnu qui découle de l'hétérogénéité des sols en termes de l'incertitude qui affecte l'estimation des paramètres de calculs.

La technique statistique utilisée dans ce type d'approche et qui reste la plus populaire et la plus performante est évoquée et détaillée au chapitre II. Il s'agit de la simulation de Monte Carlo. Par ailleurs, la notion de simulation y est largement vulgarisée.

Le chapitre III se distingue du chapitre II par son aspect original, que l'on doit à Matheron de l'Ecole des Mines et que l'on peut considérer comme une extension des

notions développées dans ce chapitre II. Ces outils, dits de géostatistique, permettent une description des variations des variables dites régionalisées.

Le chapitre IV traite de la variabilité spatiale des paramètres géotechniques. Dans ce chapitre il est démontré la nécessité de la prise en compte de cette variabilité et de ses effets sur les solutions des problèmes de génie civil.

Pour pouvoir prendre en compte le caractère hétérogène du sol au travers de ses propriétés physiques et mécaniques, et donc pouvoir décomposer le massif en éléments plus petits avec des valeurs des paramètres de calcul, nous utilisons presque toujours la méthode des éléments finis, bien souvent au travers de logiciels comme le CESAR-LCPC. C'est la méthode la plus utilisée tant elle est performante.

Dans la pratique du génie civil et particulièrement de la mécanique des sols, les corrélations entre paramètres sont utilisées comme moyen de contrôle des résultats des essais en place et en laboratoire, et comme moyen de fabrication de valeurs complémentaires de certains paramètres en fonction des autres. Ces notions, très utiles dans la recherche de la description de la variabilité des sols, sont passées en revue dans le chapitre V.

Le chapitre VI traite d'une question essentielle dans cette étude : les méthodes pratiques de calcul probabiliste. Diverses méthodes existent plus ou moins performantes et permettent la résolution de problèmes complexes.

L'avant dernier chapitre, soit donc le chapitre VII, quelques cas d'application sont alors évoqués ou passés en revues de manière détaillée comme l'analyse probabiliste de la stabilité des pentes, l'étude stochastique de la consolidation bi et tridimensionnelle et l'effet de la variabilité spatiale du sol sur l'analyse et le dimensionnement d'une fondation superficielle filante en se basant sur des mécanismes de ruine en analyse limite. La méthode de discrétisation des champs utilisée dans ce chapitre est basée sur le concept du moyennage local.

Dans toutes les études rapportées les auteurs testent diverses méthodes pour venir à bout de la variabilité des propriétés physiques et mécaniques des sols.

Ces exemples montrent que ces techniques sont appliquées aussi bien pour la modélisation statistique des variables que pour les méthodes de calcul qui y sont ensuite appliquées.

Il faut, en outre, noter que toutes ces approches proposées ne sont pas limitées à ce domaine particulier de la mécanique des sols où et elles fournissent un cadre pour l'analyse des problèmes résultant de l'hétérogénéité des sols, mais applicables aussi à pratiquement tous les domaines de l'art de l'ingénieur.

Malgré le développement d'outils de calculs (méthodes numériques) de plus en plus performants et une description de sol de plus en plus proche de la réalité, il existe toujours un écart entre les résultats de calcul et les valeurs des solutions réels à la fois en laboratoire et in situ.

C'est dans l'espoir de réduire ce biais existant, ou, tout au moins de progresser dans la compréhension de ces divergences, que les géotechniciens poussent leurs investigations dans l'application de diverses techniques dont les méthodes statistiques et probabilistes, avec comme objectif de prendre en compte l'effet de la variabilité des propriétés physiques et mécaniques de sol sur le comportement des ouvrages de génie civil.

Enfin ce travail réalisé dans le cadre de la préparation d'un mémoire de magister se termine par une conclusion. L'accent est ainsi mis sur l'importance de ces techniques et de leur développement éventuel. L'espoir est qu'un jour les ingénieurs accorderont plus d'importance à toutes ces méthodes dans leurs diverses applications.

Chapitre I

ELEMENTS DE STATISTIQUES ET PROBABILITES

Chapitre I

ELEMENTS DE STATISTIQUES ET PROBABILITES

I.1 GENERALITES

I.1.1 Besoin de statistiques

Du fait de la variabilité, on est dans le domaine de l'incertain. Cette science de l'incertain, c'est le défi qu'a relevé la statistique en s'appuyant sur le concept de probabilité. Plutôt qu'une seule valeur, la prise en compte de l'incertain permet de déterminer un intervalle à l'intérieur duquel on a une certaine probabilité de se situer et donc un risque de ne pas y être.

Dans ce chapitre nous présentons les notions essentielles de la statistique descriptive, indispensables dans l'analyse de la variabilité des sols et le calcul probabiliste des ouvrages. On apprend comment décrire de façon claire et concise l'information apportée par des observations nombreuses et variées sur un phénomène donné.

Il s'agit de trier ces données, les décrire, les résumer sous forme de tableaux, de graphiques, et sous forme d'un petit nombre de paramètres-clés (moyenne, médiane par exemple).

I.1.2 Statistique descriptive et statistique inférentielle

De manière approximative, il est possible de classer les méthodes statistiques en deux groupes : celui des méthodes descriptives et celui des méthodes inférentielles.

- La statistique descriptive. On regroupe sous ce terme les méthodes dont l'objectif est la description des données étudiées ; cette description des données se fait à travers leur présentation (la plus synthétique possible), leur représentation graphique, et le calcul de résumés numériques. Dans cette optique, il n'est pas fait appel à des modèles probabilistes. On notera que les termes de statistique descriptive, statistique exploratoire et analyse des données sont quasiment synonymes.

- La statistique inférentielle. Ce terme regroupe les méthodes dont l'objectif principal est de préciser un phénomène sur une population globale, à partir de son observation sur une partie restreinte de cette population ; d'une certaine manière, il s'agit donc d'induire (ou

encore d'inférer) du particulier au général. Le plus souvent, ce passage ne pourra se faire que moyennant des hypothèses de type probabiliste. Les termes de statistique inférentielle, statistique mathématique, et statistique inductive sont eux aussi synonymes.

D'un point de vue méthodologique, la statistique descriptive précède en général la statistique inférentielle dans une démarche de traitement de données : les deux aspects se complètent bien plus qu'ils ne s'opposent.

I.2 DISTRIBUTIONS STATISTIQUES ET REPRESENTATIONS ASSOCIEES

I.2.1 Définitions de base

On appellera :

- **Individu**, l'unité d'observation (exemples : entreprise, chaîne de production) ;
- **Population**, l'ensemble des individus concernés par l'étude (exemples : ensemble des entreprises algériennes, ensemble des pièces sortant de la chaîne) ;
- **Échantillon**, un sous-ensemble de la population dont les individus feront l'objet de l'étude. Le choix de l'échantillon se fait en respectant certaines règles ;
- **Variable**, ou caractère statistique, l'aspect de l'unité statistique que l'on va étudier (exemples : situation géographique de l'entreprise, diamètre de la pièce...). On dira que cette variable prend des valeurs (ou modalités).

I.2.2 Types de variables statistiques

On peut définir quatre classes (ou types) dans lesquelles se répartissent les variables statistiques selon la nature de leurs valeurs. Les différents types de variables sont présentés dans le tableau I.1.

Tableau I.1 Différents types de variables statistiques

Différents types de variables statistiques			
Ensemble des valeurs prises par la variable	Type de variable	Exemples	Genre de variable
Amorphe (sans structure)	Catégorielle (ou nominale)	▪ Sexe ▪ Nationalité ▪ Catégorie socioprofessionnelle ▪ Contrôle Qualitatif d'une pièce ▪ Situation de famille	Qualitatives
Ordonné	Ordinale	▪ Tout jugement qualitatif ▪ Mention à un examen	
Une partie de l'ensemble des entiers	Discrète	▪ Nombre d'enfants ▪ Nombre de diplômes	Quantitatives
Une partie de l'ensemble des réels	Continue	▪ Poids ▪ Température ▪ Fréquence d'un signal ▪ Amplitude d'un bruit thermique ▪ Valeur boursière	

I.2.3 Distributions statistiques. Effectifs, fréquences

Lorsque le recueil des données a été effectué, on dispose, pour chacun des individus de l'échantillon (ou de la population), de la valeur de la variable étudiée. Le premier traitement consiste alors à relever cette valeur pour chaque individu et ensuite à compter le nombre d'individus pour lesquels la variable prend une valeur donnée.

On associe, à chaque valeur prise par le caractère statistique étudié, son **effectif**.

Notation : les variables seront notées par des lettres majuscules $X, Y, Z, \dots$; on note leurs modalités (valeurs) par des lettres minuscules x_i, y_j, z_l et les effectifs associés par n_i, n_j, n_l .

Exemple : X = sexe, x_1 = féminin, x_2 = masculin, n_1 = nombre de femmes, n_2 = nombre d'hommes

Ce traitement n'est bien sûr directement possible que pour les variables qualitatives ou discrètes, qui n'ont qu'un nombre limité de valeurs possibles, discernables entre elles. Pour les variables continues, on commence par ranger les observations en classes, celles-ci étant des intervalles de la forme $[a_{i-1}, a_i[$. Ensuite, pour chaque classe, on compte le nombre d'individus dont le caractère appartient à la classe : ce nombre est l'effectif de la classe. On note k le nombre de modalités.

Définition : on appellera **distribution statistique des effectifs** de la variable X :

- L'ensemble des données (x_i, n_i) , $i = 1, \dots, k$, si X est une variable qualitative ou discrète,
- L'ensemble des données $([a_{i-1}, a_i[, n_i)$, $i = 1, \dots, k$, si X est une variable continue.

Les résultats sont généralement présentés dans un tableau du type du tableau II.2.

Tableau I.2 Présentation des variables statistiques

Présentation des variables statistiques			
X est catégorielle, ordinale ou discrète		X est continue	
Classes	Effectifs	Classes	Effectifs
x_1	n_1	$[a_0, a_1[$	n_1
x_2	n_2	$[a_1, a_2[$	n_2
$\cdot$	$\cdot$	$\cdot$	$\cdot$
$\cdot$	$\cdot$	$\cdot$	$\cdot$
x_k	n_k	$[a_{k-1}, a_k[$	n_k
Total (1)	$N = n_1 + n_2 + \dots + n_k$	Total	$N = n_1 + n_2 + \dots + n_k$
(1) N est l'effectif total de l'échantillon			

Remarque : dans le cas des données individuelles (c'est-à-dire lorsque l'on ne regroupe pas les réponses), on a : $n_1 = n_2 = \dots = n_k$.

À la distribution d'effectifs définie ci-dessus, on préfère souvent la **distribution de fréquences** associée.

Définition. La **fréquence** (ou proportion) associée à la valeur du caractère (resp. à la classe $[a_{i-1}, a_i[$) est la valeur f_i définie par :

$$f_i = \frac{n_i}{N}$$

La fréquence f_i représente donc la part de l'échantillon pour laquelle la valeur de la variable est x_i (ou appartient à $[a_{i-1}, a_i[$). On peut par exemple l'exprimer sous forme de pourcentage (le pourcentage sera alors $100 \times f_i$).

Remarque. Cette quantité est indépendante de la taille de l'échantillon, ce qui permet de comparer les résultats obtenus sur plusieurs échantillons.

Les fréquences vérifient les propriétés suivantes :

$$0 \leq f_i \leq 1 \quad i = 1, \dots, k$$

$$f_1 + f_2 + \dots + f_k = 1$$

I.2.4 Représentations graphiques des distributions statistiques

Très souvent, on préfère des représentations graphiques à des tableaux. Les graphes apparaissent comme plus « parlants ». Ces représentations sont adaptées au type de variable étudiée : nominale, ordinale, discrète ou continue.

a. Variables nominales

On dispose pour ces variables de **diagrammes en bâtons**, ainsi que de **diagrammes circulaires** (ou **en secteurs**, ou en « **camembert** »).

- **Diagramme en bâtons** (figure II.1a)

À chaque modalité x_i , on associe un « bâton » de longueur h_i proportionnelle à la fréquence f_i (ou, si l'on veut, à l'effectif n_i). On a donc $h_i = \text{Cte} \times f_i$.

Pour une variable nominale, seules les hauteurs sont significatives ; l'ordre et l'écart des x_i ne sont pas significatifs.

- **Diagramme circulaire** (figure II.1b)

L'angle de chaque secteur α_i est proportionnel à la fréquence f_i .

En degrés, on a $\alpha_i = 360 \times f_i$.

C'est la représentation la plus utilisée pour les variables nominales. C'est la représentation la plus utilisée pour les variables nominales. De surcroît, elle est plus fidèle que la précédente.

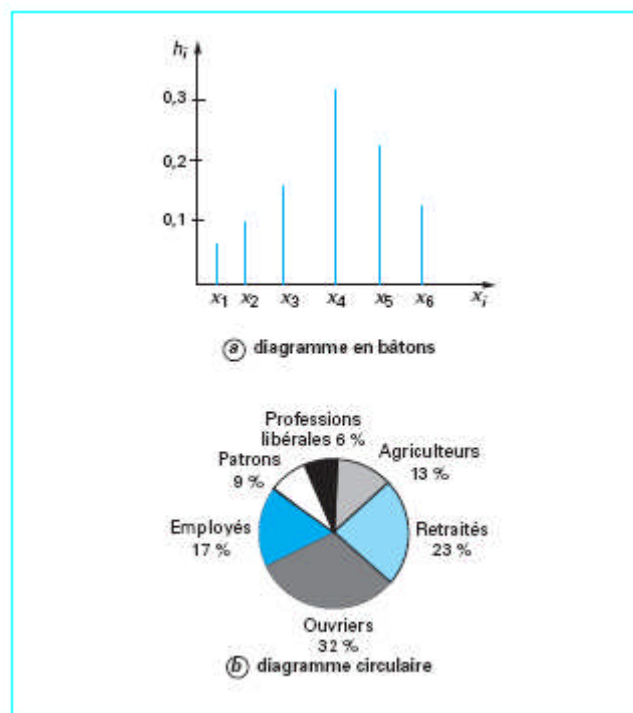


Figure I.1 Représentations graphiques des variables nominales

b. Variables ordinales et variables discrètes

▪ Variables ordinales

On utilise les mêmes représentations que pour les variables nominales. Toutefois, il convient de noter que, pour le diagramme en bâtons, l'ordre des modalités à un sens concret, car il doit correspondre à l'ordre existant entre les valeurs.

▪ Variables discrètes

Pour ce type de variables, on préfère le diagramme en bâtons car, dans ce cas, l'ordre et l'écart entre les bâtons sont significatifs.

c. Variables continues : histogramme, polygone des fréquences, diagramme « branche et feuille »

On considère une variable statistique continue dont les valeurs ont été rangées en classes $[a_{i-1}, a_i[$. L'**amplitude** de la classe $[a_{i-1}, a_i[$ est $A_i = a_i - a_{i-1}$.

▪ Pour représenter graphiquement la distribution statistique d'une telle variable, on a recours à un **histogramme**. Le principe est le suivant : à chaque classe, on fait correspondre un rectangle de base l'intervalle $[a_{i-1}, a_i[$ (pour la classe i) et de hauteur h_i , de sorte que la surface du rectangle soit proportionnelle à l'effectif. Ainsi, on calcule la hauteur h_i du rectangle au moyen de la formule suivante :

$$h_i = \frac{n_i}{a_i - a_{i-1}}$$

D'un point de vue pratique, on constituera un tableau du type du tableau II.3.

Tableau I.3 Variables continues : amplitudes et fréquences

Variables continues : amplitudes et fréquences					
i	Classes	Effectifs n_i	Fréquences f_i	Amplitudes A_i	Hauteurs h_i
1	$[a_0, a_1[$	n_1	f_1	$a_1 - a_0$	$n_1 / (a_1 - a_0)$
2	$[a_1, a_2[$	n_2	f_2	$a_2 - a_1$	$n_2 / (a_2 - a_1)$
·	·	·	·	·	·
k	$[a_{k-1}, a_k[$	n_k	f_k	$a_k - a_{k-1}$	$n_k / (a_k - a_{k-1})$

On obtient ainsi le graphique de la figure II.2 :

- en abscisse, on porte l'**ensemble des valeurs** prises par la variable, découpé en classes ;
- en ordonnée, on porte les **hauteurs** :

$$h_i = \frac{n_i}{a_i - a_{i-1}}$$

- on trace enfin des rectangles.

Remarque. Si les amplitudes sont toutes égales, on porte les **effectifs** en ordonnée.

La construction de l'histogramme s'opère de la façon suivante :

- On calcule la différence de la distribution, différence entre la valeur la plus élevée et la valeur la plus faible.
- On partage l'étendue de la distribution, en k classes d'amplitudes égales.
- On compte le nombre de valeurs comprises dans chacune des classes.
- Puis on reporte ces nombres N_i sur un graphique où l'on porte en abscisse les valeurs du paramètre étudié et en ordonnée les effectifs de chaque classe.

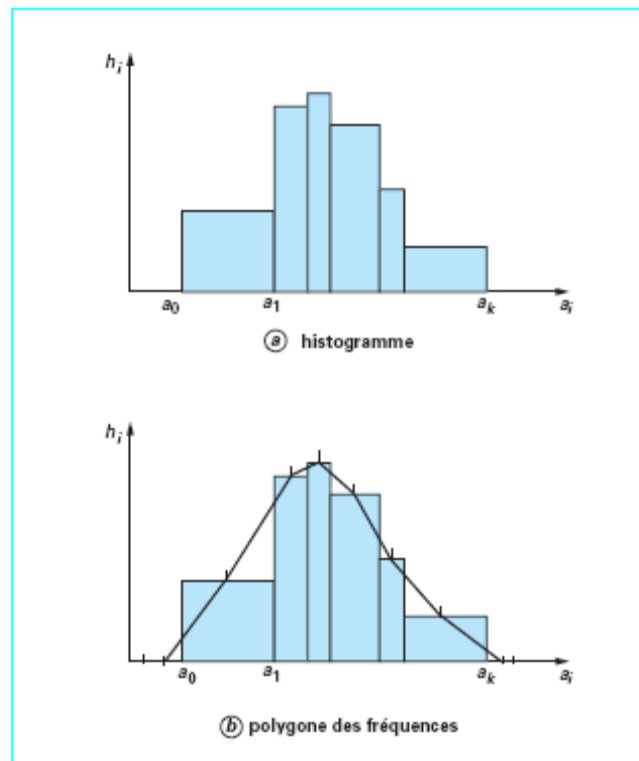


Figure I.2 Représentations graphiques des variables continues

- À partir de l'histogramme d'une variable statistique continue, on peut tracer le **polygone des fréquences** associé (figure II.2b) en procédant de la manière suivante :
 - on joint par des morceaux de droites les milieux des segments horizontaux supérieurs des rectangles de l'histogramme ;

- on ajoute à droite et à gauche de l'histogramme des classes fictives, toutes deux de même amplitude et d'effectif nul, ce qui donne alors lieu à deux nouveaux segments.

Remarque. On ne doit pas « lisser » la courbe.

▪ En dernier lieu, lorsque l'on étudie un échantillon de faible taille ($N < 100$) et que l'on dispose des données individuelles, on peut dresser, pour la variable étudiée, un diagramme dit **branche et feuille** (en anglais **stem and leaf**), qui a l'avantage de conserver l'information de la répartition à l'intérieur des classes.

Principe. Dans tout nombre, on peut distinguer deux parties : un chiffre de « plus haut poids » (branche) et un chiffre de « plus bas poids » (feuille).

I.3. FREQUENCES CUMULEES ET FONCTION DE REPARTITION

I.3.1. Fréquences cumulées

Pour les variables qualitatives ordinales et pour les variables quantitatives, on peut exploiter la relation d'ordre existant entre les valeurs possibles de la variable. On définit ainsi les **distributions cumulées** (Figure II.3 et tableau II.4).

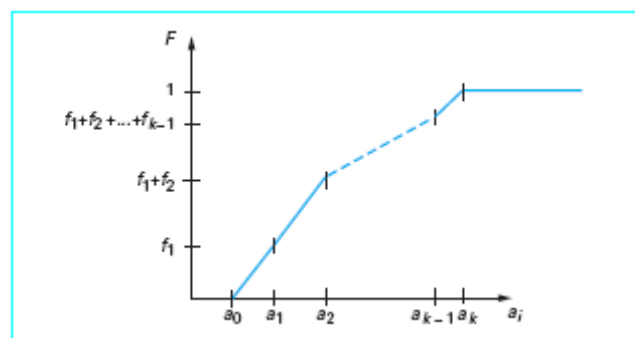


Figure I.4 Fonction de répartition

Tableau I.4 Distributions cumulées					
i	Valeurs	Effectifs	Fréquences	Effectifs cumulés	Fréquences cumulées
1	x_1	n_1	f_1	n_1	f_1
2	x_2	n_2	f_2	n_1+n_2	f_1+f_2
·	·	·	·	·	·
·	·	·	·	·	·
k-1	x_{k-1}	n_{k-1}	f_{k-1}	$n_1+n_2+\dots+n_{k-1}$	$f_1+f_2+\dots+f_{k-1}$
k	x_k	n_k	f_k	$n_1+n_2+\dots+n_k=N$	$f_1+f_2+\dots+f_k=1$

I.3.2 Fonction de répartition

Cette notion ne concerne que les variables quantitatives.

Définition. La **fonction de répartition du caractère X** est la fonction F , allant de l'ensemble des réels vers $[0,1]$, définie par : $F(x) =$ proportion d'individus de l'échantillon dont la valeur de X est $< x$

Soit X une variable continue, dont les valeurs sont rangées en classes $[a_0, a_1[, \dots, [a_{k-1}, a_k[$ avec des fréquences $f_1, \dots, f_k$.

- On commence par calculer les valeurs de F aux points du découpage :

$$F(a_0) = 0, F(a_1) = f_1, \dots, F(a_{k-1}) = f_1 + f_2 + \dots + f_{k-1}, \\ f(a_k) = f_1 + f_2 + \dots + f_k$$

- Ensuite, dans chaque classe $[a_{i-1}, a_i[$, on fait une interpolation linéaire (on relie les points extrêmes par un segment de droite).
- Puis on prolonge la courbe par 0 à gauche de a_0 et par 1 à droite de a_k (figure 3).

I.4 CARACTERISTIQUES D'UNE DISTRIBUTION. TENDANCE CENTRALE ET DISPERSION

I.4.1 Généralités

Jusqu'à présent, nous nous sommes intéressés uniquement à la représentation des données statistiques. Cependant, s'il est vrai que les divers tableaux et graphes définis plus haut «résument» la distribution, ils ne permettent aucune quantification. Le but de ce paragraphe est donc de définir, pour chaque type de distribution statistique, un certain nombre de **caractéristiques** (ou **indicateurs**), c'est-à-dire quelques nombres permettant de résumer de manière **quantitative** (et non plus qualitative) chaque distribution. Bien entendu, n'importe quelle quantité ne peut pas être un indicateur.

En 1950, le statisticien Yule a donné un certain nombre de propriétés de « bon sens » que doivent, a priori, vérifier les indicateurs statistiques.

Selon lui, ceux-ci doivent :

- être défini de manière objective (et donc être indépendant de l'observateur) ;
- utiliser toutes les observations ;
- avoir une signification concrète, afin d'être compris par les non-spécialistes ;
- être simple à calculer ;
- être peu sensibles aux fluctuations d'échantillonnage (notion introduite dans l'article suivant de ce traité) ;

- se prêter aisément aux opérations mathématiques simples.

Nous nous limiterons ici à 2 types de caractéristiques statistiques :

- celles dites de **tendance centrale**, qui donnent un « ordre de grandeur » de la variable étudiée en dégagant la modalité de la variable la plus représentative ;
- celles dites de **dispersion** qui, elles, fournissent des informations sur la façon dont les individus se répartissent (se « dispersent ») autour de la tendance centrale.

Le tableau II.5 donne les caractéristiques étudiées pour chaque type de variable.

Tableau I.5 Caractéristiques d'une distribution

CARACTERISTIQUES D'UNE DISTRIBUTION		
Type de variable	Tendance centrale	Dispersion
Nominale	Mode	
ordinaire	Mode, médiane, quantiles	Ecart interquartile
Quantitative	Mode, médiane, quantiles, moyenne	Ecart-type, écart interquartile

I.4.2 Caractéristiques de tendance centrale

a. Mode

Il est défini pour tous les types de variables. On le définit comme suit.

- Si X est une variable statistique **nominale**, **ordinaire** ou **discrète**, le **mode** de la distribution associée est la modalité de X la plus représentée, c'est-à-dire celle pour laquelle l'**effectif** est le **plus grand** ;
- Si X est une variable **continue**, le **mode** (ou **classe modale**) de la distribution associée est la classe dont la **hauteur dans l'histogramme** est la **plus élevée**.

b. Médiane et quantiles

Ces indicateurs sont définis pour toutes les variables sauf les variables nominales.

La médiane est la valeur de la variable telle que le nombre d'observations supérieures ou égales à cette valeur est égal au nombre d'observations strictement inférieures à cette valeur.

On voit que, par exemple, pour les variables continues, cela revient à chercher un x tel que $F(x) = 0,5$. En règle générale, cette valeur de x n'existe pas dans le tableau de données dont on dispose.

C'est pourquoi on adopte la définition suivante : la médiane de la distribution de X est donnée par :

- pour les variables ordinales ou discrètes :

- o Si la fréquence cumulée en x_{i-1} est $< 0,5$ et celle en x_i est $> 0,5$, alors la médiane vaut x_i ,
- o Si la fréquence cumulée en x_{i-1} est égale à $0,5$, alors la médiane vaut x_i ;
- Pour les variables continues, réparties en classes $[a_{i-1}, a_i[$:
 - o Si $F(a_{i-1}) < 0.5$ et $F(a_i) > 0.5$, la **classe médiane** est $[a_{i-1}, a_i[$ et on calcule la médiane par interpolation linéaire sur l'intervalle $[a_{i-1}, a_i[$:

$$M_{éd} = a_{i-1} + (a_i - a_{i-1}) \frac{0.5 - F(a_{i-1})}{F(a_i) - F(a_{i-1})}$$

avec F fonction de répartition de X (figure II.5),

- o Si $F(a_{i-1}) = 0.5$, la médiane vaut a_{i-1} .

Remarque. La médiane est peu sensible aux valeurs extrêmes de la variable, donc aux erreurs de mesure qui, bien souvent, produisent des valeurs aberrantes. On dit que la médiane est robuste ou résistante.

Celle notion de médiane peut se généraliser à celle de quantile.

Soit α dans l'intervalle $]0, 1[$. Si $F(a_{i-1}) < \alpha$ et $F(a_i) > \alpha$, on définit le nombre Q_α , quantile d'ordre α , par

$$Q_\alpha = a_{i-1} + (a_i - a_{i-1}) \frac{\alpha - F(a_{i-1})}{F(a_i) - F(a_{i-1})}$$

Les cas particuliers les plus cités sont :

- o Les quartiles
- o Les déciles
- o Les centiles.

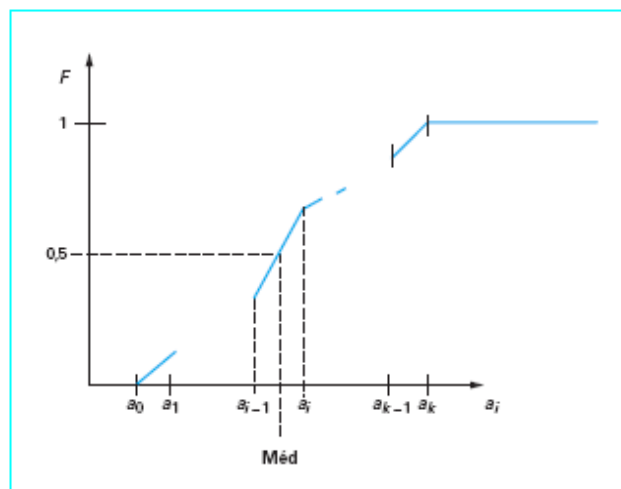


Figure I.5 Classe médiane

c. Moyenne arithmétique

Elle n'est définie que pour les variables quantitatives et, pour celles-ci, c'est la caractéristique de tendance centrale la plus « naturelle » et la plus utilisée.

La moyenne (arithmétique) d'une variable X sera notée $\bar{x}$ et $N = n_1 + n_2 + \dots + n_k$

On définit la moyenne arithmétique de la manière suivante.

- Si X est une variable quantitative **discrète**, donnée par sa distribution d'effectifs $(x_i, n_i), i = 1, \dots, k$, alors la **moyenne** de X est donnée par

$$\bar{x} = \frac{1}{N} (n_1 x_1 + n_2 x_2 + \dots + n_k x_k)$$

- Si X est une variable **continue rangée en classes** $[a_{i-1}, a_i[$, la **moyenne** de X est

$$\bar{x} = \frac{1}{N} (n_1 c_1 + n_2 c_2 + \dots + n_k c_k)$$

Où, pour tout i, c_i est le centre de la classe $[a_{i-1}, a_i[$, soit

$$c_i = \frac{a_{i-1} + a_i}{2}$$

On dira qu'une variable est **centrée** si sa moyenne est **nulle**.

Il faut noter les remarques suivantes :

- la moyenne peut être définie à l'aide des fréquences $\bar{x} = f_1 x_1 + f_2 x_2 + \dots + f_k x_k$: pour les variables discrètes et $\bar{x} = f_1 c_1 + f_2 c_2 + \dots + f_k c_k$ pour les variables continues ;
- il existe d'autres sortes de moyennes (géométrique, harmonique...) dont nous ne parlerons pas ici ;
- la moyenne, prenant en compte toutes les valeurs observées, est très sensible aux observations aberrantes ;
- chaque fois que la répartition est assez symétrique (ce qui se traduit par un histogramme proche d'une courbe « en cloche »), la moyenne, la médiane et le mode sont proches. La moyenne est plus élevée que le mode ou la médiane si la répartition est dissymétrique, avec un accent vers les valeurs élevées ; si l'accent est, par contre, sur les valeurs faibles, la moyenne est plus petite que le mode ou la médiane.

I.4.3 Caractéristiques de dispersion

Les caractéristiques de tendance centrale donnent un ordre de grandeur du caractère statistique observé. Il est intéressant d'obtenir des informations sur la **variabilité** des observations et de leur **dispersion** autour de la tendance centrale. Intuitivement, une « bonne »

caractéristique de dispersion doit être telle que, plus la variabilité est grande autour de la tendance centrale correspondante, plus cette caractéristique doit être grande, et inversement lorsqu'il y a peu de dispersion, la caractéristique doit être voisine de 0. De plus, une caractéristique de dispersion doit toujours être positive.

1. Écart interquartile

Il est défini pour toutes les variables, excepté les variables nominales.

Définition. L'**écart interquartile** est la distance entre le 1er et le

3e quartile. Il vaut donc $Q_{0.75} - Q_{0.25}$. Il représente les valeurs extrêmes d'une dispersion de 50 % des effectifs autour de la médiane.

2. Écart-type. Variance

Ils ne sont définis que pour les variables quantitatives.

Définition. La **variance** est la moyenne des carrés des écarts à la moyenne, c'est-à-dire :

- pour une variable discrète :

$$V(X) = \frac{1}{N} \left(\sum_{i=1}^{i=k} n_i (x_i - \bar{x})^2 \right) = \left(\frac{1}{N} \sum_{i=1}^{i=k} n_i x_i^2 \right) - \bar{x}^2$$

- pour une variable continue rangée en classes $[a_{i-1}, a_i[$, de centres c_i

$$V(X) = \frac{1}{N} \left(\sum_{i=1}^{i=k} n_i (c_i - \bar{x})^2 \right) = \left(\frac{1}{N} \sum_{i=1}^{i=k} n_i c_i^2 \right) - \bar{x}^2$$

Dans chaque cas, c'est la seconde expression qui sera le plus souvent utilisée pour effectuer les calculs.

L'**écart-type** est alors la racine carrée de la variance :

$$\sigma(X) = \sqrt{\text{Var}(X)}$$

3. Coefficient de variation $C_v(X)$

Si l'écart type mesure l'erreur absolue dans l'estimation de la moyenne $\bar{x}$, alors le coefficient de variation, noté $C_v(X)$ est :

$$C_v(X) = \frac{\sigma(X)}{\bar{x}}$$

$C_v(X)$ est un facteur adimensionnel utile. Il caractérise la dispersion intrinsèque de la variable.

I.4.4 Moments et caractéristique de forme

1. Moments empiriques

Si l'on dispose d'un échantillon de n valeurs $x_1, x_2, \dots, x_n$, on peut d'abord calculer les moments empiriques m'_r de cet échantillon.

- Moments empiriques d'ordre 1 (moyenne arithmétique empirique) :

$$m'_1 = m' = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$$

- Moment empirique d'ordre r :

$$m'_r = \frac{1}{n} \sum_{i=1}^n x_i^r$$

Toutefois les moments empiriques ne constituent pas des estimateurs sans biais des moments correspondants. En effet un estimateur n'est en général pas égal à la grandeur à estimer et l'on appelle biais d'un estimateur la différence entre sa valeur et celle de la grandeur considérée.

On utilise en pratique les estimateurs sans biais suivants pour les moments centrés d'ordre r .

- Estimateur de moments centré d'ordre 1.

$$\mu'_1 = 0$$

- Estimateur du moment centré d'ordre 2.

$$\mu'_2 = s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - m')^2$$

- Estimateur du moment centré d'ordre 3.

$$\mu'_3 = \frac{n}{(n-1)(n-2)} \sum_{i=1}^n (x_i - m')^3$$

- Estimateur du moment centré d'ordre 4.

$$\mu'_4 = \frac{n}{(n-1)(n-2)(n-3)} \left\{ (n+1) \sum_{i=1}^n (x_i - m')^4 - \frac{3(n-1)}{n} \left[\sum_{i=1}^n (x_i - m')^2 \right]^2 \right\} + 3s^4$$

2. Caractéristique de forme

Les moments centrés d'ordre 3 et 4 donnent les informations sur la forme de la densité de probabilité. Ils sont souvent présentés sous forme adimensionnelle :

$$\sqrt{\beta_1} = \frac{\mu'_3}{\sigma^3} = \frac{1}{\sigma^3} \sum_{i=1}^{i=n} n_i (x_i - m')^3 = E[T_3]$$

$$\beta_1 = E[T_4] = \frac{\mu'_4}{\sigma^4}$$

Si la distribution est symétrique par rapport à la moyenne arithmétique m' , on a $\sqrt{\beta_1} = 0$. Pour une distribution asymétrique, $\sqrt{\beta_1}$ peut être positif ou négatif suivant le signe de μ'_3 . Il mesure alors « l'asymétrie » de la distribution (figure ci-dessous) $\sqrt{\beta_1}$ est pour cette raison appelée « coefficient d'asymétrie ».

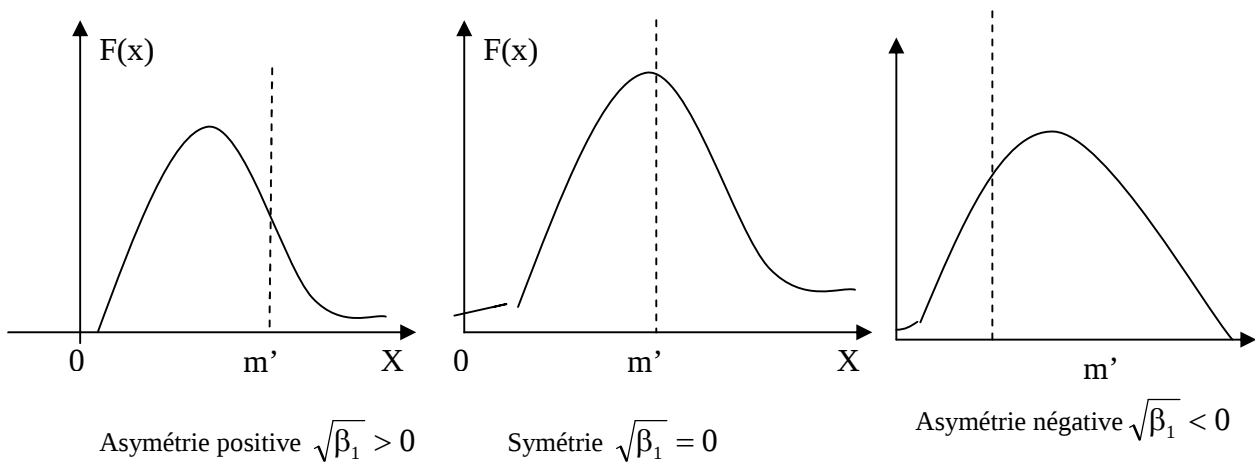


Figure I.6 Coefficients d'asymétrie

$\sqrt{\beta_1} = 0$: la distribution des fréquences est symétrique par rapport à m' .

$\sqrt{\beta_1} > 0$: la distribution est plus étalée à droite de m' qu'à sa gauche.

$\sqrt{\beta_1} < 0$: la distribution est plus étalée à gauche.

Le coefficient β_2 donne une indication sur la façon dont la distribution est concentrée autour de la moyenne ; de faibles valeurs de β_2 impliquent une courbe plus plate pour la densité de probabilité ; β_2 est appelé « coefficient d'aplatissement ».

En vue de comparer toute distribution à la distribution normale, on forme

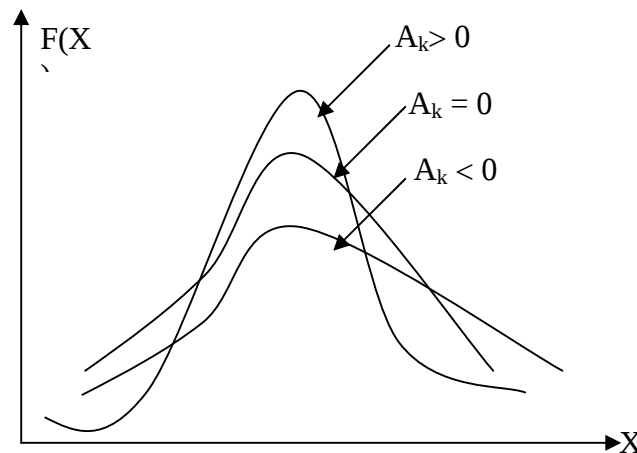
$$A_k = \frac{\mu_4}{\sigma^4} - 3 = \frac{1}{\sigma^4} \sum_{i=1}^k n_i (x_i - m')^4 - 3$$

Le rapport $\frac{\mu_4}{\sigma^4}$ pour une distribution normale (gaussienne) est égal à 3.

Ainsi :

- pour une distribution gaussienne $A_k = 0$,

- pour une distribution pointue que la distribution gaussienne $A_k > 0$,
- pour une distribution moins pointue que la distribution gaussienne: $A_k < 0$.



Remarque

$T = \frac{X - m}{\sigma}$ est dite variable réduite. Elle est écrite sous forme adimensionnelle. Elle

est de moyenne nulle et a pour variance 1.

I.5. MODELE THEORIQUE DE DISTRIBUTION. VARIABLES ALEATOIRES D'ECHANTILLONNAGE

La loi de distribution des valeurs d'un paramètre physique ou mécanique dans un massif de sol ne peut naturellement pas être parfaitement connue, puisqu'il est impossible de mesurer ce paramètre en tous points. On ne dispose donc en pratique que de l'histogramme des valeurs mesurées et des moments empiriques de l'échantillon statistique fourni par les essais.

On conçoit pourtant aisément que l'utilisation des lois de distribution expérimentales n'est pas très pratique : on essaie toujours pour cette raison de modéliser la loi observée par l'une des lois de distribution théoriques développées par les spécialistes des statistiques.

Il est donc intéressant de rappeler quelques notions générales concernant les lois de distribution des variables aléatoires.

I.5.1 Variables aléatoires

On appelle variable aléatoire toute grandeur non déterminée à priori dont on sait avec quelle probabilité elle peut prendre telle ou telle valeur parmi une population (ensemble de valeurs).

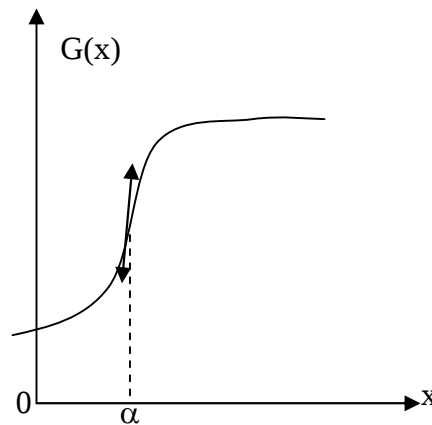
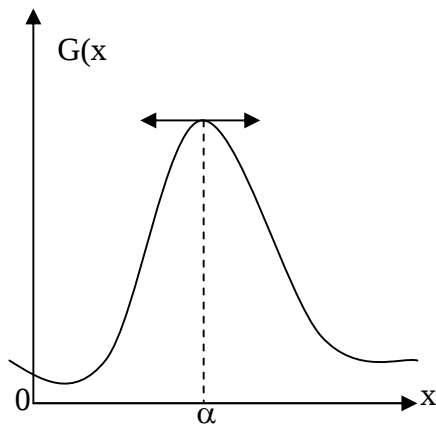
Par exemple, la teneur en eau dans une couche de sol, qui prend des valeurs $x_1, x_2, \dots, x_n$ avec des probabilités $p_1, p_2, \dots, p_n$ respectivement. La variable aléatoire peut être caractérisée par une densité de probabilité $f(x)$ (modèle mathématique d'un diagramme de distribution) défini de la façon suivante :

$$f(x) = \text{prob}\{x < X < x + dx\}$$

Elle peut être aussi définie par sa fonction de répartition $F(x)$ (modèle mathématique du diagramme de répartition) qui est la fonction définie en tout point x comme la probabilité que la variable aléatoire X soit inférieure ou égale à x .

$$F(x) = \text{prob}\{X \leq x\}$$

$$G(x) = \text{prob}\{X \leq x\} = \int_{-\infty}^x g(x) d(x) = \int_{-\infty}^x \alpha p(x)$$



La variable aléatoire X peut être indifféremment définie par sa densité de probabilité $g(x)$ ou sa fonction de répartition $G(x)$.

I.5.2 Espérance et moments

Nous avons vu qu'à la variable statistique X , nous avons par analogie fait correspondre la variable aléatoire X de même nous avons caractérisé une distribution observée grâce à deux sorte de paramètres.

- Paramètre de position, situant la valeur centrale des observations, pour cela nous avons retenu pour valeur centrale, la moyenne arithmétique $\bar{x}$ ou m ou Moy (x),
- Paramètres de dispersion, précisant la répartition des écarts entre les diverses valeurs observées et la moyenne, nous avons principalement retenu la variable $\text{Var}(X)$ et l'écart type s ou $\sigma(x)$, racine carrée de la précédente.

En nous donnant une loi de probabilité d'une variable aléatoire X , nous décrirons simplement une distribution théorique des différentes valeurs de cette variable aléatoire.

Il est donc tout à fait normal que nous définissons des paramètres de position et de dispersion : ils portent la dénomination générale de moments.

1. Espérance mathématique $E(X)$ qui correspond à la moyenne $Moy(X)$.

C'est un paramètre de position. L'espérance mathématique d'une variable aléatoire X de densité $g(x)$ est sa valeur moyenne définie par

$$E(X) = \int_{-\infty}^{+\infty} x \cdot g(x) dx = \int_0^1 x \cdot dp(x)$$

2. Le moment d'ordre r de la variable aléatoire X est égale à :

$$E(X^r) = m_r = \int_{-\infty}^{+\infty} x^r g(x) dx.$$

On note que l'espérance mathématique est égal au moment d'ordre 1 de X , noté m_1 ou plus généralement m .

Les moments centrés calculés par rapport à l'espérance mathématique et d'usage plus fréquent que les moments calculés par rapport à l'origine ($[X^r]$)

$$E[(X - E[X])^r] = \mu_r = \int_{-\infty}^{+\infty} (x - m)^r g(x) dx$$

Le premier moment centré est nul et le second est égal à la variance qui est un paramètre de dispersion

$$V(x) = \sigma^2 = \mu^2 = E[(x - E(x))^2]$$

où σ est l'écart type.

I.5.3 Lois de distribution théorique

1. Introduction

En pratique, on ne dispose que de l'histogramme des valeurs mesurées et des moments empiriques (moyenne arithmétique, variance,...) de l'échantillon statistique fourni par les essais.

On essaye alors de modéliser la loi de distribution observée par l'une des lois de distributions développées par les spécialistes des statistiques.

La loi la plus connue est la « loi de Gauss » qui a été la plus employée pour représenter les variations des propriétés des sols à l'intérieur d'une couche homogène. Cette loi n'est pourtant pas toujours la mieux adaptée aux problèmes de mécanique des sols.

2. Fonction de distribution

Toute fonction $g(x)$ vérifiant la condition $\int_{-\infty}^{+\infty} g(x) dx = 1$ peut servir de densité de probabilité pour une variable aléatoire. Il existe toutefois un certain nombre de fonctions employées de façon classique et qui ont de ce fait une importance pratique beaucoup plus grandes que les autres.

La plus part des densités de probabilité $g(x)$ classiques peuvent être générées à partir de l'équation différentielle

$$\frac{dg(x)}{dx} = \frac{(a_0 + b)x}{b_0 + b_1x + b_2x^2} \quad (*)$$

Avec : a_0, b_0, b_1 et b_2 constantes.

Les distributions correspondantes sont regroupées sous le nom de « systèmes de Pearson ».

Pearson a défini trois principaux types de courbes d'après les valeurs de K , appelé « Critère » qui s'exprime en fonction des quatre premiers moments de $g(x)$ (Harr, 1977)

$$K = \frac{\beta_1(\beta_2 + 3)}{4(2\beta_2 + 3\beta_1 - 6)(4\beta_2 - 3\beta_1)}$$

$\sqrt{\beta_1}$ et β_2 étant respectivement les coefficients d'asymétrie et d'aplatissement.

Les trois types de courbes sont :

Type I : $K < 0 \rightarrow$ Distribution bêta

Type IV : $0 < K < 1 \rightarrow$ Courbes non bornées et asymétriques

Types VI : $1 < K \rightarrow$ Borné d'un côté.

Ces trois types de courbes de distribution couvrent l'ensemble des cas possibles.

Il existe toutefois des distributions de transition entre les différents types.

Par exemple, si $\beta_1 = 0$ et $\beta_2 = 3 \Rightarrow K = 0$ (dans l'équation * on a $b_1 = b_2 = 0$) : distribution normale. C'est une distribution de transition entre les types I et IV.

Toujours pour $K=0$, mais $\beta_2 \neq 3$, Pearson a défini deux types de courbes symétriques : le type II ($\beta_2 < 3$) et le type VII ($\beta_2 > 3$). La distribution de Student est un exemple de distribution de type VII.

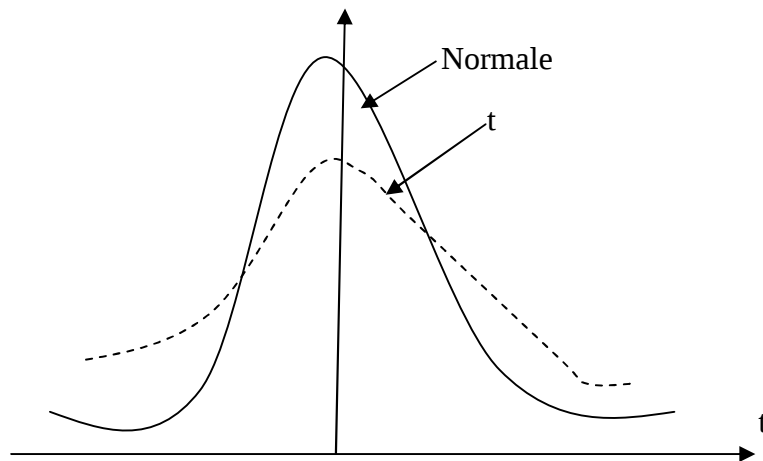


Figure I.7 Distribution t de student (type VII de Pearson)

La distribution de type III de Pearson correspond à une valeur infinie de K ($K = \pm\infty$). C'est la généralisation d'une distribution appelée « Gamma ».

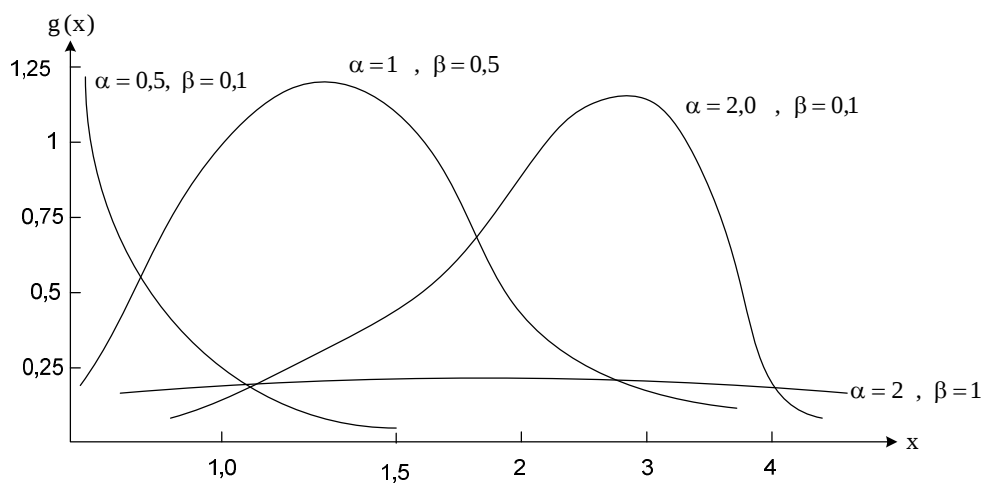


Figure I.8 Exemples de distribution gamma.

Les distributions X^2 et exponentielle sont deux exemples de distributions de type III. Pour $K = 1 \rightarrow$ type V (par exemple, la loi lognormale). La figure I.6 ci-dessous présente une classification schématique des types de distribution de Pearson en fonction des variations du critère K .

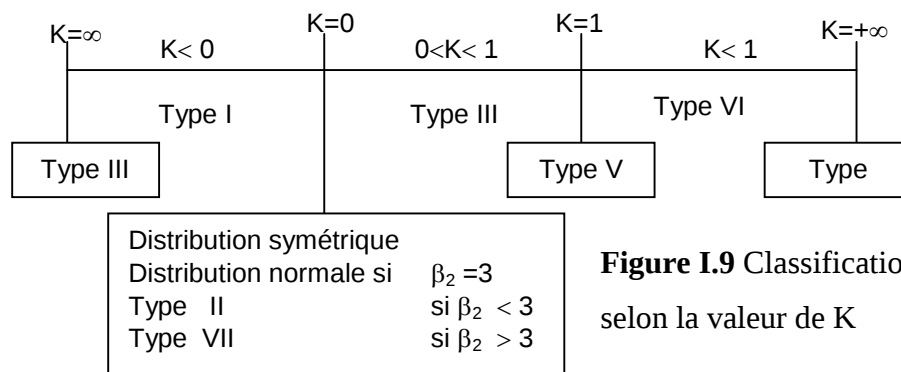


Figure I.9 Classification des distributions selon la valeur de K

I.5.4 Définition générale d'une variable Gaussienne

On dit très généralement qu'une variable aléatoire (continue ou assimilable à une variable probabiliste de l'événement $\{X(x)\}$) s'exprime par la fonction de répartition

$$F(x) = \frac{1}{\sigma(x)\sqrt{2\pi}} \int_{-\infty}^x \exp\left[-\frac{1}{2}\left(\frac{x - E(x)}{\sigma(x)}\right)^2\right] dx$$

Expression dans laquelle $E(x)$ est l'espérance mathématique et $\sigma(x)$ l'écart type de la variable aléatoire x considérée.

Si l'on substitue à la variable aléatoire X la variable centrée réduite $T = \frac{X - E(X)}{\sigma(X)}$,

On obtient une expression plus simplifiée :

$$\text{prob} \left\{ \frac{x - E(X)}{\sigma(X)} < t \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t \exp\left(-\frac{t^2}{2}\right) dt$$

Et la fonction de densité de probabilité est :

$$f(x) = \frac{1}{\sigma(n)\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{n - e(n)}{\sigma(n)}\right)^2\right]$$

Ou :

$$f(t) = \frac{1}{\sqrt{2C}} \exp\left(-\frac{t^2}{2}\right)$$

La courbe représentative des variations de $f(t)$ est la très fameuse courbe en cloche.

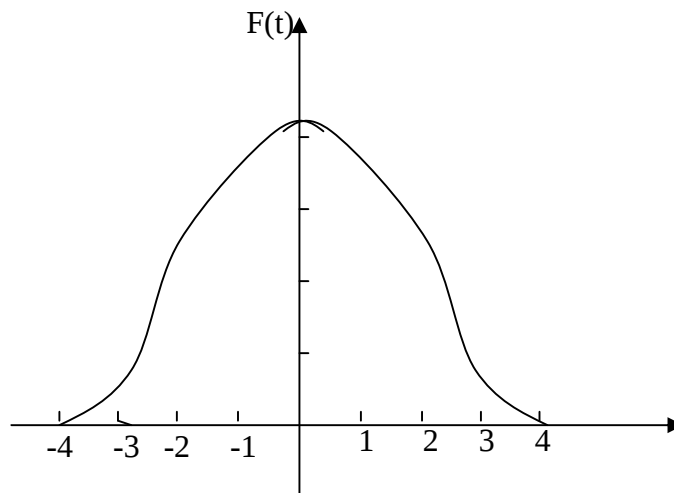


Figure I.10 La courbe en cloche

Chapitre II

SIMULATION ET GÉNÉRATION DE VARIABLES ALÉATOIRES

Chapitre II

SIMULATION ET GÉNÉRATION DE VARIABLES ALÉATOIRES

II.1 SIMULATION

II.1.1 GÉNÉRALITES

Simuler, dit le dictionnaire, *"c'est donner pour réel ce qui ne l'est pas, en imitant l'apparence de la chose réelle à laquelle on veut faire croire"*. La simulation, dans le langage courant, a souvent un sens négatif : elle a à voir avec la feinte, le déguisement, la ruse, la comédie. On la distingue du mime en ce qu'elle n'est ni une pure copie, ni une simple reproduction. En général, elle consiste en une imitation d'un original par rapport auquel elle introduit toujours des variations. De là vient que la simulation est toujours double, à la fois jeu et exercice, tromperie et prouesse, faux-semblant et art, dissimulation et apprentissage. Sur une face, elle apparaît comme faiseuse d'illusions. Sur l'autre, elle est gage de perfectionnement et de conquête.

Certains de ces aspects se retrouvent dans l'usage que les sciences font de la simulation. Celle-ci y renvoie, là encore, à l'idée d'imitation. Elle consiste même en une imitation "au carré" de la nature au sens où ce n'est pas la nature elle-même qu'elle reproduit, mais le modèle par lequel nous avons préalablement choisi de la représenter. Elle exhibe ainsi un "surcroît de réalité" en rendant explicite ce qui, dans une théorie ou un modèle, n'est qu'implicite ou camouflé. Une bonne simulation procède de la révélation.

Mais il arrive aussi que, mal conduite ou mal interprétée, elle ne montre au contraire qu'une infra-réalité illusoire ou qu'elle caricature à l'excès l'image de ce qu'elle est censée illustrer. Elle peut par exemple dépasser les bornes d'un modèle, sortir de son domaine de validité, oublier un phénomène essentiel, et cela sans qu'aucun signal d'alerte ne soit aussitôt émis. Une mauvaise simulation procède du simulacre.

Puisque le modèle qui lui sert de support est toujours incomplet, une modélisation ne peut jamais elle-même prétendre à la complétude.

Au mépris du sens des mots, la simulation a donc toujours partie liée avec la dissimulation. Est-ce grave ?

Non, car rien n'est perdu dès lors que ce qu'elle cache ou néglige n'est pas pertinent dans le

cadre du problème qu'on se pose.

D'ailleurs, qu'est-ce qu'un bon modèle, sinon celui qui contient le minimum d'ingrédients nécessaires pour exhiber de l'universalité à propos de ce que l'on veut décrire ou comprendre ?

D'une façon générale, pour un scientifique, simuler, c'est faire des "expérimentations sur un modèle" : il réalise une reproduction artificielle du phénomène qu'il désire étudier, puis observe le comportement de cette reproduction lorsqu'il fait varier les actions que l'on peut exercer sur celle-ci, et en induit ce qui se passerait dans la réalité sous l'influence d'actions analogues. Dans la pratique, cette méthodologie se décline en plusieurs variantes qui offrent diverses vertus.

Bien simuler, c'est au bout du compte se donner les moyens de mieux comprendre, de mieux concevoir, et surtout de mieux agir.

La simulation est également déterminante dans les cas où l'on ne dispose pas de bases théoriques solides, et où l'on cherche précisément à élaborer une théorie rendant compte des données d'observation. On peut alors définir avec précision les conséquences concrètes des différents modèles théoriques possibles, déterminer lequel fournit l'approximation la plus fiable, et ensuite comprendre, voire optimiser, certains processus.

Une bonne simulation peut apporter plus d'informations qu'une expérience globale insuffisamment préparée ou difficilement reproductible, ne serait-ce qu'en raison de son coût.

Entre théorie et expérience

Comme on voit, la simulation présente de multiples facettes. Mais quel est son statut épistémologique ? Ni celui de la théorie, ni celui de l'expérience. La simulation a en effet introduit, sans conteste possible, un nouveau rapport vis-à-vis de notre représentation du monde. Jusque dans les années 1950, théorie et expérience étaient présentées comme des activités dialectiquement corrélées (Gaston Bachelard, Karl Popper).

L'idée était que la science naissait d'un incessant va-et-vient entre l'une et l'autre. Cette donne a changé avec l'arrivée massive des ordinateurs et du calcul numérique, au point que de nouvelles questions sont aujourd'hui posées : les simulations pourraient-elles avoir le même statut que les expériences ? Deviendront-elles capables, à l'occasion, de falsifier une théorie ? Permettent-elles toujours de mieux comprendre les modèles qu'elles illustrent ? Quant à l'effet positif de la simulation, qui tient à l'économie de temps et d'argent, voire à l'accès à des réalités hostiles ou dangereuses par l'apport des réalités virtuelles, ne doit-il pas être contrebalancé par le risque de fuite devant le réel et devant l'expérimentation concrète ?

II.1.2 MÉTHODOLOGIE DE LA SIMULATION

Une des façons de faire progresser la connaissance scientifique est de proposer des explications rationnelles à des phénomènes naturels par l'élaboration de modèles. Ces schémas mentaux simplifiés par rapport à la réalité sont suggérés par des observations expérimentales et suscitent eux-mêmes des expériences. Un modèle est donc une abstraction de la réalité qui, au fur et à mesure de la maturation des idées, peut être progressivement affiné (de la théorie de la chute des corps de Newton à la relativité générale d'Einstein, par exemple). Il s'exprime au travers d'équations mathématiques. Dans les cas les plus simples, les équations peuvent se résoudre à l'aide de solutions analytiques. En général, ce n'est pas possible. Pour démontrer le degré de pertinence d'un modèle par rapport aux données expérimentales, il faut souvent en résoudre les équations sur ordinateur.

Il y a ainsi irruption de la simulation numérique dans la dialectique théorie/expérience qui est au cœur de la démarche scientifique

II.1.3 UN TRIPTYQUE NON ORDONNE

Le cycle de la recherche scientifique devient alors le suivant : théorie et modélisation, simulation numérique et vérification expérimentale. C'est ce que nous appelons ici la méthodologie de la simulation. Concrètement, la connaissance du processus physique ou de l'ensemble de processus étudiés est rassemblée dans un logiciel permettant de résoudre les équations des modèles préalablement établis. Comment valide-t-on ce logiciel ? En simulant, justement, le processus dont on pense qu'il a donné lieu au phénomène observé. Si les résultats numériques et expérimentaux sont proches, le modèle traduit en logiciel est déclaré valide, avec le degré d'approximation choisi. Cette validation est souvent complexe, car de nombreux processus sont impliqués. Il faut alors procéder pas à pas, en vérifiant que chaque modèle élémentaire de la description globale est satisfaisant. C'est une validation *par parties*. Il faut ensuite valider l'ensemble, ce qui implique de prendre en compte les interactions entre les modèles élémentaires. C'est la validation *globale*.

II.1.4 QU'EST-CE QU'UNE SIMULATION NUMÉRIQUE

La simulation numérique consiste à reproduire par le calcul le fonctionnement d'un système, préalablement décrit par un ensemble de modèles. Elle s'appuie sur des méthodes mathématiques et informatiques spécifiques. Les principales étapes de la réalisation d'une étude par simulation numérique sont communes à de nombreux

secteurs de la recherche et de l'industrie, en particulier le nucléaire, l'aérospatial ou l'automobile

En chaque point de l'"objet" considéré, plusieurs grandeurs physiques (vitesse, température ...) décrivent l'état et l'évolution du système étudié. Celles-ci ne sont pas indépendantes, mais reliées et régies par des équations, généralement aux dérivées partielles. Ces équations constituent la traduction mathématique des lois de la physique qui modélisent le comportement de l'objet. Simuler l'état de ce dernier, c'est déterminer - idéalement en tout point - les valeurs numériques de ses paramètres. Comme il y a un nombre infini de points, donc une infinité de valeurs à calculer, cet objectif est inaccessible (sauf dans des cas bien particuliers où l'on peut résoudre les équations de départ à l'aide de formules analytiques). Une approximation naturelle consiste donc à ne considérer qu'un nombre fini de points. Les valeurs des paramètres à calculer sont ainsi en nombre fini et les opérations nécessaires deviennent abordables grâce à l'ordinateur. Le nombre effectif de points traités dépendra bien sûr de la puissance de celui-ci : plus il sera élevé, meilleure sera finalement la description de l'objet. A la base du calcul des paramètres comme à la base de la simulation numérique, il y a donc la réduction de l'infini au fini, la discrétisation.

Comment opère-t-on précisément à partir des équations mathématiques du modèle? Deux méthodes sont très souvent utilisées, respectivement représentatives des méthodes de calcul déterministe, qui résolvent les équations régissant les phénomènes étudiés après avoir **discretisé** les variables, et des méthodes de calcul statistique ou probabiliste.

Le principe de la première, connue sous le nom de méthode des volumes finis, est antérieur à l'usage des ordinateurs. Chacun des points de l'objet est assimilé simplement à un petit volume élémentaire (un cube par exemple), d'où le nom de *volume fini*. Un plasma, par exemple, est ainsi vu comme un ensemble ou un réseau de volumes contigus qui, par analogie avec la trame d'un tissu, sera dénommé maillage. Les paramètres de l'état de l'objet sont maintenant définis dans chaque maille du maillage. Pour chacune d'elles, en reformulant les équations mathématiques du modèle par des moyennes volumiques, il sera alors possible de construire des *relations algébriques* entre les paramètres de la maille et ceux de ses voisines. Au total, il y aura autant de relations que de paramètres inconnus et ce sera à l'ordinateur de résoudre le système de relations obtenu. Il faudra pour cela recourir aux techniques de l'analyse numérique et programmer des algorithmes spécifiques. L'accroissement de la puissance des

ordinateurs a permis d'augmenter la finesse de discrétisation, permettant de passer de quelques dizaines de mailles dans les années soixante à plusieurs dizaines de milliers dans les années quatre-vingt, à des millions dans les années quatre-vingt-dix et jusqu'à la dizaine de milliards de mailles aujourd'hui, chiffre qui devrait décupler à la fin de la décennie.

Un raffinement du maillage, le remaillage adaptatif, consiste à ajuster la taille des mailles en fonction des circonstances, par exemple en les rendant plus petites et plus serrées aux interfaces entre deux milieux, là où les phénomènes physiques sont les plus complexes, où les variations sont les plus importantes.

La méthode des volumes finis s'applique dans des contextes physiques et mathématiques très variés. Elle autorise toute forme de maille (cube, hexaèdre, tétraèdre ...) et le maillage peut être modifié durant le calcul, en fonction de critères géométriques ou physiques. Enfin, elle est aisée à mettre en œuvre dans le contexte des ordinateurs parallèles (en cadre B, *Les moyens informatiques de la simulation numérique hautes performances*), le maillage pouvant en effet faire l'objet d'un découpage pour des calculs sur ce type de machines.

Appartiennent à la même famille la méthode des différences finies, cas particulier de la méthode des volumes finis où les cotes des mailles sont orthogonales, et la méthode aux éléments finis, qui peut juxtaposer divers types de mailles. La deuxième grande méthode, dite de Monte-Carlo.

L'efficacité de cette méthode, mise en œuvre à Los Alamos dès les années 1940, dépend bien sûr de la qualité statistique des tirages au hasard.

Il existe pour cela des méthodes de nombres aléatoires, bien adaptées au traitement par un ordinateur.

Les méthodes des volumes finis et de Monte-Carlo ont suscité et suscitent de nombreuses études mathématiques. Ces études s'attachent notamment à préciser la convergence de ces méthodes, c'est-à-dire comment la précision de l'approximation varie avec le nombre de mailles ou de particules. Cette question est naturelle lors de la confrontation des résultats de la simulation numérique à ceux de l'expérience.

II.1.5 COMMENT SE DÉROULE UNE SIMULATION NUMÉRIQUE

Il est souvent question d'*expérience numérique* pour souligner l'analogie entre la pratique d'une simulation numérique et la conduite d'une expérience de physique.

Brièvement, cette dernière utilise un dispositif expérimental, configure selon des conditions initiales (de température, de pression ...) et des paramètres de contrôle (durée de l'expérience, des mesures ...). Durant l'expérience, le dispositif produit des points de mesures qui sont enregistrés. Ces enregistrements sont ensuite analysés et interprétés.

Dans une simulation numérique, le dispositif expérimental consiste en un ensemble de programmes informatiques exécutés sur des ordinateurs. Les codes ou logiciels de calcul sont la traduction, à travers des algorithmes numériques, des formulations mathématiques des modèles physiques étudiés. En amont et en aval du calcul, les *logiciels d'environnement* effectuent la gestion de plusieurs opérations complexes de préparation des calculs et de leur dépouillement,

Les données initiales de la simulation comporteront d'abord la délimitation du domaine de calcul à partir d'une représentation approchée des formes géométriques (produite par le dessin et la CAO, conception assistée par ordinateur), suivie de la discrétisation de ce domaine de calcul sur un maillage, ainsi que les valeurs des paramètres physiques sur ce maillage et les paramètres de contrôle du bon déroulement des programmes... Toutes ces données (produites et gérées par les logiciels d'environnement) seront saisies et vérifiées par les codes. Les résultats des calculs proprement dits, c'est-à-dire les valeurs numériques des paramètres physiques, seront sauvegardés au fur et à mesure. En fait, un protocole spécifique structurera les informations produites par l'ordinateur afin de constituer une base de données numériques,

À l'issue des calculs (résolution numérique des équations décrivant les phénomènes physiques qui se déroulent dans chaque maille), l'analyse des résultats par des spécialistes reposera sur l'exploitation de la base de données numériques. Elle comportera plusieurs étapes : extraction sélective des données (selon le paramètre physique recherché) et visualisation, extraction et transfert des données pour calculer et visualiser des diagnostics.

Le parallèle entre la conduite d'un cas de calcul, d'une expérience numérique et la conduite d'une expérience physique ne s'arrête pas là : les résultats numériques seront comparés aux résultats expérimentaux. Cette analyse comparative, effectuée sur la base de critères quantitatifs standardisés, fera appel et à l'expérience et à l'art de l'ingénieur, du physicien, du mathématicien. Elle débouchera sur de nouvelles améliorations des modèles physiques et des programmes informatiques de simulation.

II.2 GÉNÉRATION DE VARIABLES ET VECTEURS ALÉATOIRES

II.2.1 INTRODUCTION

Un des aspects importants de la méthode de Monte-Carlo est qu'elle nécessite la génération de séries de valeurs aléatoires de chacun des paramètres des modèles de calcul, conformément à la loi de distribution observée.

II.2.2 SIMULATION D'UNE VARIABLE ALÉATOIRE

Soit X une variable aléatoire ayant pour fonction de répartition F_x . comment peut-on réaliser dans la pratique un ensemble d'épreuves tel que la fréquence avec laquelle « soit » l'intégrale $X \leq x$ soit une bonne approximation de $F_x(x)$?

Du point de vue pratique, simuler une variable aléatoire X de fonction de répartition F_x donnée, revient à construire une suite $\{x_k\}_{k=1, \dots, m}$ de réalisations de la variable X . ceci nécessite les deux opérations suivantes :

- a. Construction d'une suite de nombres pseudo- aléatoires $\{r_k\}_{k=1, \dots, m}$ équirépartie sur l'intervalle $[0,1]$. Notons ici qu'en vertu des propriétés des suites equiréparties, cette suite est tel que :
 - Chaque nombre r_k est une réalisation d'une variable aléatoire uniformément distribué sur $[0,1]$
 - Les m valeurs de r_k sont indépendantes dans leurs ensembles.
- b. A partir de cette suite, construction de la suite $\{x_k\}_{k=1, \dots, m}$ des réalisations de la variable aléatoire X .

Si $F_x^{-1}(r)$ peut être exprimée algébriquement, chaque terme x_k est calculé par

$$x_k = F_x^{-1}(r).$$

Sinon une technique numérique appropriée doit être recherchée.

II.2.3 SUITES ÉQUIRÉPARTIES

Comme il a été montré précédemment, toute variable aléatoire peut se ramener à une variable uniformément distribuée sur $[0,1]$. Le problème se pose alors de construire une suite de nombres compris entre 0 et 1 qui puissent représenter des réalisations d'une variable aléatoire R uniforme sur $[0,1]$.

Les méthodes proposées pour créer des nombres aléatoires reposent sur des calculs de nature arithmétique dans lesquelles le hasard ne joue aucun rôle. Les nombres aléatoires sont entièrement reproductibles et se calculent par des procédés compatibles avec les techniques habituelles de programmation des machines à calculer.

Dans ce cas, la suite $\{\Gamma_k\}$ de nombres pseudo- aléatoires générés est définie par une suite de récurrence :

$$r_{k+1} = f(r_k, r_{k1}, \dots, r_{k-m}) \quad ; m \geq 0$$

Où $(r_0, r_1, \dots, r_m)$ est le vecteur de nombres aléatoires.

A titre d'exemple nous citerons la méthode du milieu du carré et la méthode de Lehmer .

1. Méthode du milieu du carré (VON NEWMAN, 1951)

La relation de récurrence est définie comme suit :

$$Y_{k+2} = (Y_k / \alpha^m) - \left(\frac{Y_k^2}{\alpha^{3m}} \right) \cdot 10^{2m}$$

Y_0 est un nombre de 2m chiffres dans une base quel conque. En base 10, il devient

$$Y_{k+2} = \left(\frac{Y_k}{10^m} \right) - \left(\frac{Y_k^2}{10^{3m}} \right) \cdot 10^{2m}$$

Cependant, la suite de nombres engendrés dégénère en un cycle souvent très petit et la méthode ne permet donc pas d'obtenir une longue suite de nombre avant de trouver une périodicité qui dépend de la base de m et du nombre Y_0 de départ.

2. Méthode de LEHMER (1951) généralisée

La méthode de LEHMER est une méthode de congruence multiplicative définie par :

$$\lambda_{k+1} = \lambda_{y_k} \pmod{P}$$

Où λ est un entier arbitrairement fixé et P une constante (on pourra se ramener à une suite dans $[0,1]$ en divisant par P).

Une condition qui maximise la période du cycle obtenu est que Y_0 et P soient premiers entre eux, par exemple si $P = 2^m$ et Y_0 impair la période de la suite est 2^{m-2}

La méthode de Lehmer peut être généralisée par : $\lambda_{k+1} = \lambda_{y_k} + \mu \pmod{P}$.

En choisissant λ et μ correctement, la longueur maximale de la période peut être portée à P.

Une fois la suite construite, il est nécessaire de vérifier que les propriétés espérées sont réalisées (indépendance equirepartition). Pour cela, un ensemble de tests statistiques a été mis au point. Parmi ceux-ci le plus connu est le test de fréquence.

3. Méthodes relatives à la loi normale

A partir des suites numériques représentant des réalisations d'une variable aléatoires uniformément distribuée sur $[0,1]$ (si les nombres générés appartiennent à l'intervalle $[0,P]$, il suffit de les diviser par P) déterminées par des procédés décrits ci-dessous, il est possible d'obtenir des représentations de variables aléatoires ayant une loi donnée.

En effet, comme on l'a déjà rappelé ; si R est une variable aléatoire uniforme sur $[0,1]$, alors $F_x^{-1}(R)$ a pour fonction de répartition F_x . cette inversion se fera soit directement (si cela est impossible), soit par construction intermédiaire.

Il existe pour cela plusieurs méthodes, mais pour la loi normale, des techniques très performantes ont été mises au point. Nous citerons entre autre la méthode directe et la méthode logarithmique trigonométrique.

▪ Méthode directe

La densité de probabilité d'une variable aléatoire normale a pour expression

$$f_x(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\}$$

Ou sous la forme réduite (moyenne $m=0$ et écart type $\sigma = 1$)

$$f_x(x) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left\{-\frac{x^2}{2}\right\}$$

On calcule $x_k = F_x^{-1}(r_k)$ en approchant la fonction de répartition par des approximations rationnelles ou des polynômes différents dans des intervalles successifs $[0,1]$.

Ainsi, à $\pm 22 \cdot 10^{-5}$ après sur $]-\infty, +\infty[$

$$\frac{1}{\sqrt{2\pi}} \cdot \exp\left\{-\frac{r^2}{2}\right\} = \frac{1}{b_0 + b_2x^2 + b_4x^4 + b_6x^6 + b_8x^8 + b_{10}x^{10}}$$

Avec :

$$b_0 = 2,5052367$$

$$b_2 = 1,283104$$

$$b_4 = 0,2264718$$

$$b_6 = 0,1306469$$

$$b_8 = 0,0202490$$

$$b_{10} = 0,0039132$$

▪ Méthodes logarithmique- trigonométrique

Cette méthode de simulation très performante repose sur le théorème suivant :

« Soient R et θ deux variables aléatoires respectivement uniforme sur $[0,1]$ et $[0,2\pi]$. La variable aléatoire G définie par $G = \sqrt{-2 \lg R} \cdot \cos \theta$ est gaussienne centrée normée. »

Toute variable aléatoire X normale, de moyenne m et d'écart type σ peut alors s'écrire :

$$X = m + \sigma \cdot G$$

▪ Description de la méthode

Considérons une suite de nombres pseudo aléatoires $\{r_k\}$ équirépartie sur $[0,1]$, soit $(r_1^{(j)}, r_2^{(j)})$ le $j^{\text{ème}}$ couple tiré à partir de $\{r_k\}$ et représentant la $j^{\text{ème}}$ réalisation des vecteurs $\vec{R} = (R_1, R_2)$ à composants indépendants et uniformes sur $[0;1]$

Dans ces conditions la $j^{\text{ème}}$ réalisation de la variable aléatoire G s'écrit :

$$g^{(j)} = \sqrt{-2 \lg r_1^{(j)} \cdot \cos 2\pi r_2^{(j)}}$$

On déduit alors la $j^{\text{ème}}$ réalisation de la variable aléatoire X $x^{(j)} = m + \sigma g^{(j)}$

II.2.4 SIMULATION D'UN VECTEUR ALÉATOIRE

Soit $\vec{X} = (x_1, x_2, \dots, x_n)$ un vecteur aléatoire R^n ; que nous supposons dans un premier temps à composantes indépendantes et notons F_{xi} la fonction de répartition de la composante X_i . Associons d'autre part à chaque argument aléatoire X_i une variable aléatoire R_i uniformément distribué sur $[0,1]$. Dans l'hypothèse où les variables R_i sont indépendantes nous pouvons poser $\vec{X} = (X_1 = F_{x1}^{-1}(R_1), X_2 = F_{x2}^{-1}(R_2), \dots, X_n = F_{xn}^{-1}(R_n))$,

La variable R_i étant uniforme sur $[0,1]$, le vecteur $\vec{X}$ est distribué suivant la loi donnée qui a pour fonction de répartition $F_x(x) = \prod_{i=1}^n F_{xi}(X_i)$.

Pratiquement, simuler un vecteur aléatoire $\vec{X} = (x_1, \dots, x_n)$ de loi normale revient à construire une suite $\{x^{(j)}\}_{j=1, \dots, m}$ de réalisation de ce vecteur. Comme une variable aléatoire, cette opération se fait en 2 étapes :

- Construire d'une suite vectorielle $\{\vec{R}^{(j)}\}_{j=1, \dots, m}$ équirépartie dans l'hyper cube unitaire $[0,1]^n$. en vertu des propriétés des suites vectorielles équiréparties, cette suite est telle que tout vecteur $\vec{R}^{(j)} = (r_1^{(j)}, r_2^{(j)}, \dots, r_n^{(j)})$ est une réalisation du vecteur aléatoire $\vec{R} = (R_1, R_2, \dots, R_n)$ uniformément distribué sur $[0,1]^n$ et à composantes indépendantes.

- Construction à partir de cette suite de la suite vectorielle $\left\{ \overrightarrow{X}^{(j)} \right\}_{j=1, \dots, m}$ de du vecteur réalisation du vecteur aléatoire $\overrightarrow{X}$.

Si le vecteur $\overrightarrow{X} = (\overline{X}_1, X_2, \dots, X_n)$ est à composantes indépendantes et si chaque fonction $F_{xi}^{-1}(r_i)$ peut être exprimée algébriquement, la suite des réalisations de e vecteur est engendrée comme suit :

$$\overrightarrow{X}^{(j)} = (x^{(j)}_1 = F_{x1}^{-1}(r_1^{(j)}); x^{(j)}_2 = F_{x2}^{-1}(r_2^{(j)}); \dots, x^{(j)}_n = F_{xn}^{-1}(r_n^{(j)}))$$

Avec $j = 1, \dots, m$.

Si une fonction $F_{x2}^{-1}(r_i)$, au moins, n'est pas algébriquement formulable ou ni le vecteur est à composantes intercorrélées, le recours à des procédés numériques appropriés est nécessaire.

II.2.5 GÉNÉRATION DE VARIABLES CORRÉLÉES (NGUYEN ET CHOWDHURY 1985)

Dans les procédures de génération de système de variables aléatoires généralement disponibles, la corrélation entre les variables stochastiques n'est pas prise en compte. Pour créer une méthode qui permet de générer des variables aléatoires corrélées, une hypothèse relative à la distribution des systèmes à deux variables, doit être faite. Si chacune des variables est supposée de distribution normale, il est alors logique de supposer une distribution normale jointe ou une distribution normale bi variable pour les deux variables corrélées.

En effet, pour une distribution normale impliquant deux variables aléatoires corrélées X et Y, la fonction de densité jointe gaussienne est donnée par :

$$f(x, y) = \frac{1}{2\pi \sigma_x \sigma_y (1 - \rho^2)^{1/2}} \cdot \exp \left\{ -\frac{1}{2(1 - \rho^2)} (x^2 - 2\rho xy + y^2) \right\}$$

Où σ_x^2 et σ_y^2 sont les variances des distributions X et Y respectivement et ρ le coefficient de corrélation entre X et Y défini par

$$\rho = \rho_{x,y} = \frac{\sigma_{x,y}}{\sigma_x \sigma_y} = \frac{E\{(X - \mu_x)(Y - \mu_y)\}}{\sigma_x \sigma_y}$$

avec $\sigma_{x,y}$: espérance mathématique du produit $(X - \mu_x)(Y - \mu_y)$ (ou covariance de X et Y)

Il faut noter que :

$$x = (X - \mu_x) / \sigma_x \quad \text{et} \quad y = (Y - \mu_y) / \sigma_y$$

x et y sont appelés variables réduites ou variables standardisées et μ_x et μ_y sont les moyennes des distributions X et Y respectivement.

Dans le cas d'un système à n variables $X_1, X_2, \dots, X_n$ de distribution gaussienne multi variable, la matrice de covariance 1 est de la forme :

$$\Lambda = [\sigma_{ij}] = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_{nn} \end{bmatrix}$$

L'élément $\sigma_{X,Y}$ est la covariance des variables aléatoires X_j et Y_k .

$$\sigma_{jk} = \text{cov}(X_j, X_k) = E\{(X_j - \mu_j)(X_k - \mu_k)\}$$

Les termes de la diagonale principale sont les variances et si toutes les variables sont indépendantes et non corrélées, tous les termes de la matrice excepté ceux de la diagonale principale sont nuls.

En notation matricielle :

$$\Lambda = E\{(X - M)(X - M)^T\} \quad (\text{II.1})$$

Or il est connu en statistique et probabilités, que certaines transformations permettent de passer d'un ensemble de variables aléatoires normales non corrélées de matrice de covariance diagonale à un ensemble de variables aléatoires de matrice de covariance définie positive. Soit donc la transformation linéaire

$$X = AY + B$$

Où A est une matrice $n \times n$ dont les éléments a_{ij} sont constants et B un vecteur colonne d'éléments b_i constants ainsi dans le cas d'une variable unique :

$$X = \sigma \cdot Y + m$$

X est une variable aléatoire normale (m, σ^2) et Y une variable aléatoire normale unité $[0,1]$

Dans le cas d'un système à 2 variables, il vient :

$$X_1 = a_{11}y_1 + a_{12}y_2 + b_1$$

$$X_2 = a_{21}y_1 + a_{22}y_2 + b_2$$

Cela implique que : $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

Si Y est le vecteur, colonne de variables aléatoires de distribution normale unité $(0,1)$, alors

$$E(X) = E(AY) + E(B) = AE(Y) + B = B \quad (\text{II.2})$$

C'est-à-dire que B est le vecteur colonne des moyennes des éléments du vecteur X

$$B=M$$

de même, $E(Y Y^T) = I \quad (\text{II.4})$

Où I est la matrice unité, car les éléments du vecteur Y colonne sont non corrélées et ont des distributions unités

En combinant les équations (1) et (2), il vient :

$$\begin{aligned} \Lambda &= E[(AY)(AY)^T] \\ \Lambda &= E[AYY^T A] = AE[YY^T]A^T \end{aligned} \quad (\text{II.5})$$

Des équations (III.4) et (III.5), on déduit une nouvelle expression de la matrice de covariance du vecteur colonne variable transformée :

$$\Lambda = AA^T \quad (\text{II.6})$$

En conséquence de la transformation linéaire d'un ensemble de variables normales unités Y en un autre ensemble de variables normales corrélées X, de matrice de covariance spécifiques Λ , est possible si une matrice constante A satisfaisant l'équation (III.6) peut être trouvée.

Dans le cas bivariable, on a :

$$\Lambda = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

σ_1^2 et σ_2^2 Sont les variables de X_1 et X_2 (respectivement et ρ le coefficient de corrélation entre X_1 et X_2 en élevant au carré la matrice A, on obtient :

$$AA^T = \begin{bmatrix} a_{11}^2 + a_{12}^2 & a_{11} \cdot a_{21} + a_{12} \cdot a_{22} \\ a_{21} \cdot a_{11} + a_{22} \cdot a_{12} & a_{21}^2 + a_{22}^2 \end{bmatrix}$$

En égalant les éléments de AA^T et Λ on a alors l'ensemble des équations simultanées suivantes :

$$a_{21}^2 + a_{12}^2 = \sigma_1^2 \quad (\text{II.7})$$

$$a_{21} \cdot a_{11} + a_{22} \cdot a_{12} = \rho \sigma_1 \sigma_2 \quad (\text{II.8})$$

$$a_{21}^2 + a_{22}^2 = \sigma_2^2 \quad (\text{II.9})$$

$$a_{11} \cdot a_{21} + a_{12} \cdot a_{22} = \rho \sigma_1 \sigma_2 \quad (\text{II.10})$$

Il y a trois équations indépendantes pour quatre inconnus. On égale alors à 0 un des termes a_{ij} par exemple $a_{12}=0$

Il vient alors finalement :

$$X_1 = \sigma_1 Y_1 + m_1 \quad (\text{II.11})$$

$$X_2 = \rho \sigma_1 \sigma_2 + \sigma_2 (1 - \rho^2)^{1/2} \cdot Y_2 + m_2 \quad (\text{II.12})$$

Seule l'équation (III.11) est nécessaire si toutes les variables sont indépendantes. Des procédures de transformation similaires peuvent être trouvées pour des cas généraux impliquant n variables.

Ainsi, pour le cas d'un système à trois variables z_1, z_2 , et z_3 corrélées, Freeze (1975) cite des travaux de Mood et Gray Bill (1963) qui indiquent que la fonction de distribution multi variable est dans ce cas de forme :

$$f(Z) = \frac{1}{|V|^{1/2} (2\pi)^{p/2}} \cdot \exp \left\{ -\frac{1}{2} (Z - \mu)^T V^{-1} (Z - \mu) \right\}$$

Avec :

$$\rho = 3; \quad Z = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \quad \mu = \begin{bmatrix} \mu_{z1} \\ \mu_{z2} \\ \mu_{z3} \end{bmatrix} \quad v = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12}^2 & \sigma_{13}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 & \sigma_{23}^2 \\ \sigma_{31}^2 & \sigma_{32}^2 & \sigma_{33}^2 \end{bmatrix}$$

Z_1, Z_2, Z_3 : variables aléatoires de moyennes $\mu_{z1}, \mu_{z2}, \mu_{z3}$

σ_{ii}^2 : Variance de Z_{ii}

σ_{ij}^2 : Covariance de Z_i et Z_j

Les valeurs de Z_3, Z_2, Z_1 sont générées de façon séquentielle en utilisant l'expression de générateur normal pour une variable indépendante (équation (11)) et les distributions suivantes :

$$\begin{aligned} Z_3 &= N[\mu_3, \sigma_{33}] \\ Z_2 &= N[\mu_2 + a, b \cdot \sigma_{22}] \\ Z_1 &= N[\mu_1 + c, b \cdot \sigma_{11}] \end{aligned}$$

Avec :

$$\begin{aligned} a &= \rho_{23} \frac{\sigma_{22}}{\sigma_{33}} \cdot (Z_3 - \mu_3) \\ b &= \rho_{23} \frac{\sigma_{22}}{\sigma_{33}} \cdot (Z_3 - \mu_3) \end{aligned}$$

$$c = \frac{(\rho_{12} - \rho_{13} \cdot \rho_{23})}{1 - \rho_{23}^2} \cdot \frac{\sigma_{11}}{\sigma_{22}} \cdot (Z_2 - \mu_2) + \frac{(\rho_{13} - \rho_{12} \rho_{23})}{(1 - \rho_{23}^2)} \cdot \frac{\sigma_{11}}{\sigma_{33}} \cdot (Z_3 - \mu_3)$$

$$d = \sqrt{1 - \rho_{13}^2 - \frac{(\rho_{12} - \rho_{13} \rho_{23})^2}{(1 - \rho_{23}^2)}}$$

II.2.6 GÉNÉRATION DE VARIABLES CORRÉLÉES ET PRÉSENTANT UNE DÉRIVE

Toujours d'après Freeze (1975) qui cite Mood et Gray Bill (1963) les valeurs générées des variables Z_i s'obtiennent en utilisant les équations suivantes :

$$Z_{1j} = Z'_{1j} + B_1(X_j - 1/2)$$

$$Z_{2j} = Z'_{2j} + B_2(X_j - 1/2)$$

$$Z_{3j} = Z'_{3j} + B_3(X_j - 1/2)$$

Les valeurs des Z'_i sont générées de façon séquentielle à l'aide d'un générateur normal à une variable et d'un algorithme de récurrence comme décrit ci-dessus

$\beta_1, \beta_2, \beta_3$ sont les pentes des droites de régression linéaire de Z_i/X_i .

II.3 SIMULATION DE MONTE CARLO

Il est toujours possible de déterminer empiriquement la densité de probabilité d'une variable $Y = f(X_d)$ en calculant les valeurs de y correspondant à des ensembles de valeurs des x_i générés de façon aléatoire conformément aux densités de probabilité de chacune des variables aléatoires X_i et en déterminant la densité de probabilité de Y d'après la distribution des fréquences des y obtenus. La précision de cette simulation augmente avec la racine carrée de la taille de l'échantillon et de ce fait il faut disposer d'un échantillon très important pour obtenir des résultats utiles. Ceci nécessite en pratique l'utilisation d'un ordinateur. Lamb (1974), Benjamin et Corneille (1970) et Harry (1977), par exemple, donnent des indications sur l'emploi de cette méthode.

Un des aspects importants de la méthode de Monte-Carlo est qu'elle nécessite la génération de séries de valeurs aléatoires de chacun des paramètres des modèles de calcul. Il existe dans les bibliothèques de sous-programmes statistiques des programmes capables de générer des suites de nombres aléatoires en suivant des lois de distribution simples (densité uniforme sur un intervalle donné, loi normale, loi logomachique, etc.).

Pour les lois de distribution uniformes, les programmes de calcul utilisent des algorithmes tels

que :

$$x_{n+1} = (Ax_n) \bmod M \quad (\text{II.13})$$

ou

$$x_{n+1} = (Ax_n + B) \bmod M \quad (\text{II.14})$$

qui permettent de calculer une série de nombres Xu à partir d'une valeur initiale x_0 , quelconque. La notation $X = (y) \bmod M$ indique que x est le reste de la division du nombre entier y par le nombre entier M . Le nombre M est la période de la séquence aléatoire generis. On le choisit très grand. Le nombre A est compris entre 0 et $M - 1$, ainsi que le nombre B . Les nombres générés sont uniformément repartis entre 0 et $M - 1$. On ajuste la série $\{x_n\}$ à l'intervalle $[a, b]$ désiré en utilisant la formule.

$$x'_n = \frac{b - a}{M} x_n + a$$

Le tableau III.1 contient une série de 100 nombres aléatoires uniformément repartis sur l'intervalle $[0,1]$, obtenus à l'aide de la relation

$$x'_n = \frac{x_n}{199017} = \frac{1}{199017} (24298x_{n-1} + 99991) \bmod 199017$$

Avec $x_0 = 3157$

Tableau 1 Valeurs aléatoires pour une distribution uniforme $r_u(0,1)$
(moyenne 0,4992 ; écart type 0,2989)

0,94078	0,21584	0,84210	0,61357	0,02350	0,91285	0,71604	0,65892	0,57355	0,3798
0,67091	0,07410	0,97724	0,21177	0,70248	0,09452	0,89312	0,99525	0,72393	0,0158
0,43931	0,10304	0,64419	0,21744	0,39051	0,17125	0,70640	0,20549	0,57834	0,3864
0,01838	0,20799	0,24438	0,075%	0,23613	0,59846	0,63295	0,70570	0,21014	0,3467
0,24788	0,45595	0,46758	0,34120	0,16860	0,97289	0,93245	0,75541	0,64645	0,8490
0,54450	0,29001	0,97433	0,00728	0,27009	0,94595	0,735	0,54936	0,06715	0,8061
0,91757	0,32528	0,85926	0,84568	0,19368	0,31652	0,51090	0,01134	0,26668	0,1181
0,63820	0,28335	0,92873	0,86820	0,67134	0,50875	0,45717	0,09495	0,48185	0,1468
0,6%44	0,48675	0,02402	0,19270	0,78788	0,32044	0,84654	0,76848	0,73332	0,14788
0,72449	0,68472	0,34764	0,93623	0,49643	0,60870	0,75928	0,09536	0,88138	0,84515

Pour obtenir des valeurs aléatoires suivant une loi normale de moyenne μ et d'écart type σ , on utilise les valeurs générées sur l'intervalle $[0,1]$ pour une distribution uniforme, soit $R_N(0,1)$: on calcule d'abord la valeur d'un nombre $R_N(0,1)$ suivant une loi normale réduite (moyenne 0 et écart type 1) à l'aide de l'expression

$$R_N(0,1) = \sqrt{-2 \ln u_1} \cos(2\pi u_2) \quad (\text{II.15})$$

qui fait intervenir deux nombres u_1 et u_2 générés dans l'intervalle $[0,1]$ comme indique ci-dessus. Puis on ajuste

la loi normale à sa moyenne et à son écart type à l'aide de l'expression :

$$R_N(\mu, \sigma) = \sigma R_N(0,1) + \mu \quad (\text{II.16})$$

Le tableau III.2 contient une série de 100 nombres aléatoires suivant une loi normale réduite $R_N(0,1)$.

TABLEAU II.2 Valeurs aléatoires pour une distribution normale réduite $R_N(0,1)$ (moyenne 0,1403 ; écart type 0,9047)

-0,16658	0,58029	-0,80567	0,63986	-0,17195	-0,10918	1,52516	-0,53041	1,04284	-0,32753
1,27407	0,03311	0,11936	-0,55938	0,25937	0,37464	-0,09567	-0,58433	-1,70448	0,41919
-1,60535	1,21703	-0,23755	-0,26834	0,85217	-0,98083	-0,25826	0,86542	-0,06157	0,12251
-0,26800	0,49648	-0,85951	-1,11724	-1,61530	-0,11159	0,53155	-0,14963	0,65953	-0,24840
-0,13576	-1,57224	-0,69035	0,03364	0,57880	1,78357	-0,64373	0,66009	-0,26376	1,68138
1,56473	0,23514	0,35392	-0,91300	1,38452	0,05203	-0,64234	1,62489	0,62041	-0,52594
0,55623	-1,55169	-1,53046	-0,48880	-0,78719	1,74257	-1,01330	-0,05342	0,09909	0,94039
-0,31176	1,31568	-0,22106	-0,71326	0,19764	-0,25877	0,43608	0,70446	-0,20704	0,50900
-0,31175	0,39155	-1,51453	2,47609	1,24760	0,42002	-0,15974	2,29259	-0,52884	-1,35866
-0,47852	1,67100	-1,17028	0,59931	1,10049	-1,19535	-0,51306	-0,29220	0,22532	0,64171

Pour les valeurs entières des paramètres des lois bêta, on peut utiliser l'expression suivante, due à

$$\text{Hahn et Shapiro : } R_B(0,1) = \frac{\sum_{i=1}^{2\alpha+2} R_{Ni}^2(0,1)}{\sum_{i=1}^{2(\alpha+\beta+2)} R_{Ni}^2(0,1)} \quad (\text{II.17})$$

qui utilise $2(\alpha + \beta + 2)$ valeurs aléatoires normales réduites $R_N(0,1)$. Pour les autres valeurs de α et β il faut procéder par interpolation.

Il faut pour terminer indiquer que l'on peut «facilement générer des valeurs aléatoires de toute variable dont on connaît la fonction. de répartition $G(x)$ en utilisant l'algorithme suivant, qui repose sur le fait que les valeurs de $G(x)$ sont uniformément réparties sur l'intervalle $[0,1]$: on génère des valeurs aléatoires de $R_N(0,1)$ et l'on retient les valeurs X_i de la variable qui sont telles que $G(x_i) = R_{ui}(0,1)$. La figure 36 illustre la procédure adoptée.

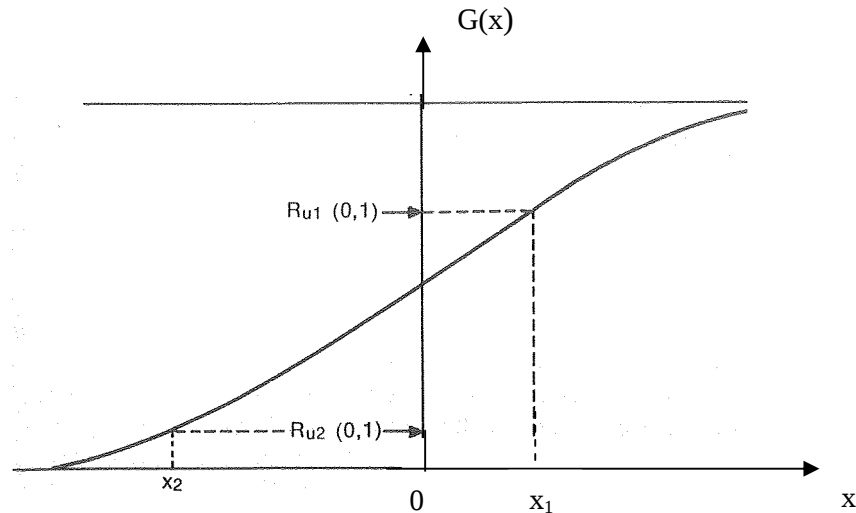


Figure II.1 Procédure de génération de valeurs aléatoires d'une variable X dont la fonction de répartition $G(x)$ est connue.

Les techniques de génération de valeurs aléatoires qui viennent d'être décrites s'appliquent à des variables isolées indépendantes. Dans le cas de plusieurs variables dépendantes, la génération d'ensembles de données est plus difficile. Il existe néanmoins différentes techniques utilisables. Nous citerons seulement celle qui consiste, dans le cas de deux variables X_1 et X_2 dont on connaît la densité de probabilité jointe $h(X_1, X_2)$, à tirer d'abord une valeur aléatoire de X_1 pour pouvoir définir la densité de probabilité $g(X_2 | x_1)$ et générer ensuite la valeur de X_2 .

Par exemple, dans le cas d'une distribution bi-normale, on obtient

$$x_1 = \mu_1 + R_{N_1}(0,1) \cdot \sigma_1$$

$$x_2 = \mu_2 + \left\{ \rho R_{N_1}(0,1) + \sqrt{1 - \rho^2} R_{N_2} \right\} \sigma_2$$

Chapitre III

LES TECHNIQUES DE GEOSTATIQUE

Chapitre III

LES TECHNIQUES DE GEOSTATIQUE**III.1 INTRODUCTION**

Les paramètres de variations spatiales montrent deux caractéristiques : une fluctuation locale aléatoire et un comportement global systématique. Ce comportement global possède une certaine structure, qui est naturellement différente d'un paramètre à l'autre et d'un site à l'autre. Les études statistiques classiques s'intéressent seulement à l'aspect aléatoire des valeurs du paramètre et ne peuvent prendre en compte la corrélation éventuelle entre les valeurs du paramètre.

En même temps, les traits structuraux présentent une irrégularité et une variabilité locale telles qu'ils échappent à toute représentation fonctionnelle simple. Cet aspect chaotique suggère une interprétation probabiliste.

Ce choix méthodologique donne alors l'idée de considérer la distribution spatiale d'une variable de sol comme une réalisation unique d'une fonction aléatoire.

Une telle variable en géostatistique est dite « variable régionalisée » (Matheron, 1971). C'est une fonction du point.

A ce propos MATHERON écrit : « un phénomène est régionalisé s'il se déploie dans l'espace et y montre une certaine structure ».

L'idée principale est donc d'associer à l'ensemble des points x_i de l'espace, un ensemble de variables aléatoires ou autrement dit une fonction aléatoire. Mais en pratique, la valeur $z(x)$ du paramètre du sol n'est connue qu'en certains points x_i ($i=1, \dots, N$), à partir de sondages ou prélèvements.

A chaque point x_k de l'espace R^n est associée une valeur $z(x_k)$ du paramètre, qui est considérée comme une réalisation particulière de la variable aléatoire $z(x_k)$

Plus généralement, on associe à l'ensemble fini de points x_i de l'espace un ensemble de variables aléatoires ou autrement dit une fonction aléatoire $Z(x_i)$.

Mais en chaque point x_k , on ne dispose qu'une d'une seule réalisation $z(x_k)$ et l'ensemble fini des points de mesure x_i est donc considéré comme une réalisation particulière de la fonction aléatoire $Z(x_i)$. Si par exemple, on avait effectué le même nombre d'essais en des points localisés différemment, on aurait trouvé des valeurs différentes.

Le modèle utilisé pour représenter les ensembles de valeurs va consister à trouver ce qu'il peut y avoir de commun entre ces ensembles. Le problème de l'étude de la fonction $Z(x)$, vu sous l'angle probabiliste, se ramènera à ajuster une loi de probabilité sur les données de manière qu'elle rende compte de l'échantillonnage.

Dans notre cas, il faut définir les caractéristiques de la fonction $Z(x)$ dont la distribution spatiale est :

$$F_{x_1 \ x_2 \dots x_k}(z_1, z_2, \dots, z_k) = P[Z(x_1) < z_1, \dots, Z(x_k) < z_k]$$

ou de l'ensemble de toutes les lois de distribution, pour tout entier positif k et pour tous les choix possibles de k points d'appui dans l'espace.

Caractériser ce modèle à partir d'une réalisation unique est impossible. On se heurte donc, au problème de l'inférence statistique qui n'est possible qu'à la condition d'introduire des hypothèses limitatives ayant pour but de réduire le nombre de paramètres de la fonction aléatoire pour reconstituer en partie la loi de la fonction aléatoire dont l'information fragmentaire du site est une réalisation.

On est donc conduit à formuler des hypothèses limitatives d'homogénéité spatiale du point de vue statistique.

L'hypothèse la plus courante est l'hypothèse de stationnarité du second ordre de la fonction aléatoire que l'on veut analyser.

Soit $Z(x)$ la variabilité régionalisée étudiée. Si les valeurs de $Z(x)$ en deux points quelconques du champ de réalisation sont indépendantes, le phénomène n'est pas structuré et l'on est en présence d'un modèle aléatoire.

Toutefois ceci est rare dans la nature et la valeur $Z(x)$ en un point est liée en général à celle des points situés au voisinage. Le champ de régionalisation est plus au moins organisé et l'on distingue plusieurs degrés dans stationnarité.

III.2 BASES DE LA GEOSTATISTIQUE

III.2.1 Hypothèse de stationnarité

a) Stationnarité au sens strict

Une fonction aléatoire est stationnaire, au sens strict si la loi spatiale est invariante par translation. Dans ce cas les deux variables aléatoires vectorielles à k composants $\{Z(x_1), \dots, Z(x_k)\}$ et $\{Z(x_1 + h), \dots, Z(x_k + h)\}$ présentent la même loi de distribution à k variables, quel que soit le vecteur translation h .

Cette hypothèse permet de résoudre le problème posé par l'inférence statistique, car à partir d'une réalisation on peut obtenir plusieurs.

En effet, une réalisation de la variable aléatoire vectorielle $[T(x_1), \dots, T(x_n)]$ est fournie par les valeurs du paramètre données par les sondages réalisés en $x_1, x_2, \dots, x_n$. Mais les valeurs données par les sondages implantés aux points :

$$x_1 + \vec{h}, \quad x_2 + \vec{h}, \dots, x_n + \vec{h}$$

constituent également une réalisation de la variable vectorielle et ceci pour toute valeur de $\vec{h}$

Mais ceci est rare en mécanique des sols. En pratique les lois de variation des paramètres changent fréquemment d'un point à un autre.

C'est une hypothèse forte.

b) Stationnarité au sens large (d'ordre 2)

C'est l'hypothèse la plus courante. Elle implique que la moyenne est constante

$$E\{Z(x)\} = m = \text{Constante}$$

et que pour toute coupe $\{Z(x), Z(x+h)\}$, la covariance existe et ne dépend que de la distance h .

$$C(h) = E\{Z(x+h) \cdot Z(x)\} - m^2, \quad \forall x$$

L'existence et la stationnarité de la covariance impliquent l'existence et la stationnarité de la variance.

En effet :

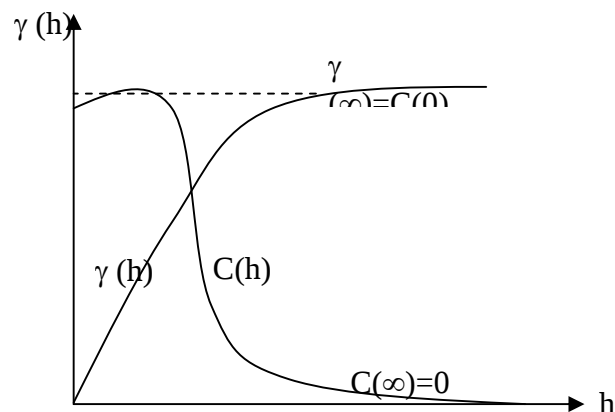
$$\text{Var} \{Z(x)\} = E\{Z(x) - m\}^2 = C(0), \quad \forall x$$

Lorsque la variance à priori est finie $C(0)$ existe et l'on a :

$$\gamma(h) = \frac{1}{2} E\{[Z(x+h) - Z(x)]^2\} = C(0) - C(h); \quad \forall x$$

$\gamma(h)$ est appelé **variogramme**

La covariance et le variogramme sont deux outils équivalents pour caractériser la variabilité de la variable $Z(x)$.



Mais certains phénomènes physiques présentent une capacité de dispersion illimitée, c'est-à-dire qu'ils ne présentent ni covariance ni variance à priori finie. Pour les traiter, il convient de considérer leur accroissement ce qui conduit à l'hypothèse intrinsèque qui ne suppose que l'existence du variogramme.

c. l'hypothèse intrinsèque

Une fonction aléatoire $Z(x)$ est dite intrinsèque si

- L'espérance mathématique existe et ne dépend que du point d'implantation x :
 $E\{Z(x)\} = m, \quad \forall x.$

- Pour tout vecteur h , l'accroissement $[Z(x+h) - Z(x)]$ aura une variance finie qui ne dépend pas de x :

$$\text{Var} \{Z(x+h) - Z(x)\} = E\{[Z(x+h) - Z(x)]^2\} = 2\lambda(h), \quad \forall x$$

En pratique la fonction structurale (covariance ou variogramme) n'est utilisée que pour des distances limitées $h < b$. la limite b représente la distance du voisinage d'estimation ; pour estimer la valeur inconnue $Z(x_0)$, seules les données situées dans un rayon b sont pris se compte.

III.3 PROPRIETES DU VARIOGRAMME

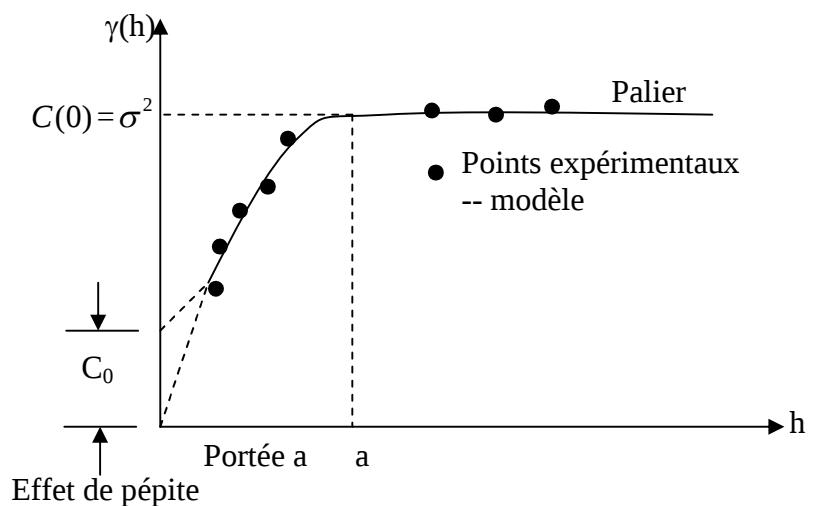
Dans le cas de l'hypothèse intrinsèque, la fonction semi- variogramme $\gamma(h)$ est définie par la relation :

$$\gamma(h) = E\{[Z(x+h) - Z(x)]^2\}$$

$Z(x)$ étant la fonction aléatoire représentant la variable étudiée.

L'étude de la structure par le variogramme consiste à suivre l'évolution de « variation quadratique moyenne » de l'accroissement de la fonction $Z(x)$ en fonction de h d'amplitudes, croissante. On obtient ainsi le variogramme dans une direction donnée.

La figure ci-dessous représente une courbe de la variation typique du variogramme en fonction de la distance h .



Le variogramme croît avec h . ceci provient du fait que plus les points sont éloignés, plus les valeurs des paramètres en ces points ont des chances d'être différentes.

En absence de dérive et lorsque la capacité de dispersion du milieu est finie, le variogramme se stabilise autour d'une valeur limite $\gamma(\infty)$ pour des distances h supérieures à une certaine limite à appeler portée.

$\gamma(\infty)$ n'est d'autre que la variance à priori de la fonction aléatoire, soit :

$$\gamma(\infty) = \text{Var}\{Z(x)\} = C(\theta).$$

Il peut arriver que le variogramme présente une discontinuité à l'origine, c'est-à-dire que $\gamma(0^+) = C_0 > 0$

C'est l'effet de pépite qui traduit la variabilité due aux erreurs de mesure et /ou la variabilité à petite échelle.

L'étude des variations de $[\gamma(x, h)]$ pour différentes directions met en évidence l'éventuelle anisotropie de la variable $Z(x)$. À titre d'exemple, la portée du variogramme dans la direction verticale est en général différente de celle obtenue dans la direction horizontale.

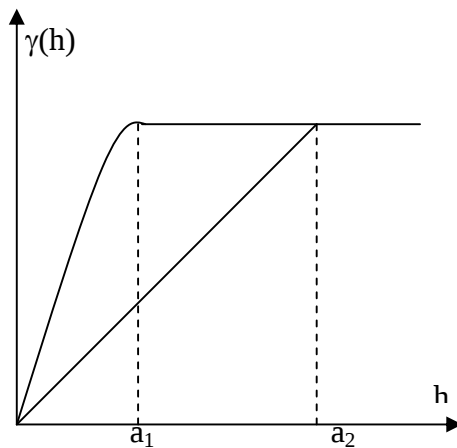


Figure III.2 Ajustement du variogramme- Modèle théorique.

En principe, pour estimer le variogramme expérimental $\gamma(h)$ à partir des données disponibles, on utilise la formule suivante :

$$\hat{\gamma}(h) = \frac{1}{2N} \sum_{i=1}^N [Z(x_i + h) - Z(x_i)]^2$$

Dans laquelle N représente le nombre de couples de valeurs de $Z(x)$ mesurées en des points distincts de h .

Le variogramme expérimental étant obtenu, on détermine le variogramme théorique qui s'ajuste le mieux aux points du variogramme expérimental.

Différents modèles théoriques ont été élaborés pour tenir compte des traits caractéristiques du comportement du variogramme. On distingue les modèles sans paliers et modèles avec palier. Pour ces derniers citons :

- Le modèle sphérique, d'équation :

$$\gamma(h) = C(0) \cdot \left[\frac{3h}{2a} - \frac{h^3}{2a^3} \right] \quad h < a$$

$$\gamma(h) = C(0), \quad h \geq a$$

a est la portée et $C(0)$ le palier.

- Le modèle exponentiel d'équation

$$\gamma(h) = C(0) \cdot (1 - e^{-h/a})$$

Ce palier $C(0)$ est atteint théoriquement quand $h \rightarrow \infty$ mais en pratique, on prendra la portée égale à $3a$.

La figure ci-dessous donne l'allure des modèles de variogramme théoriques sphérique et exponentiel.

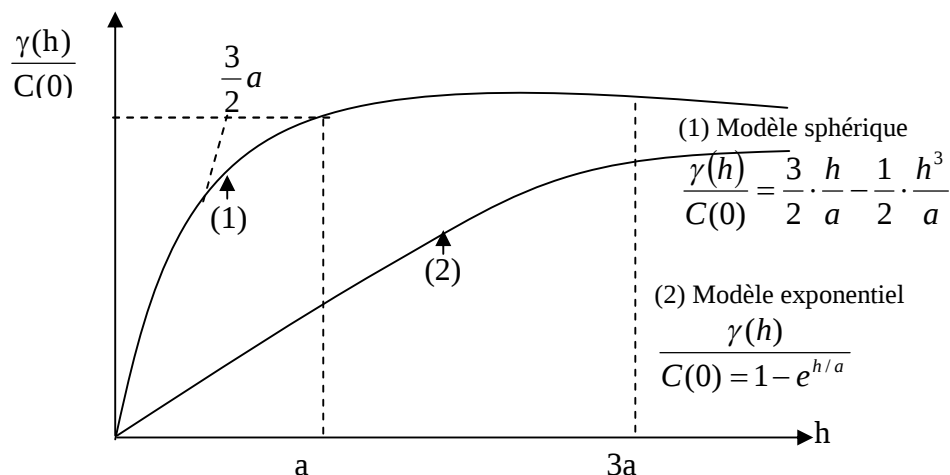


Figure III.3 Modèles de variogramme

Le variogramme théorique sert d'une part à l'analyse structurale du phénomène étudié (effet de pépite, portée, existence de palier,...) et d'autre part à aborder certains problèmes de variabilité spatiale et d'estimation.

III.4 VARIATION D'ESTIMATION ET VARIANCE DE DISPERSION

En fait la vraie valeur de la variable Z n'est connue qu'en certains points où l'on dispose de mesures ponctuelles par des sondages. Pour connaître la vraie valeur de Z en tout autre point, on doit procéder à une estimation à partir des données disponibles.

Il convient alors de connaître l'erreur commise lorsqu'on utilise la valeur estimée Z^* en un point au lieu de la vraie valeur Z inconnue. Pour caractériser cette erreur, on fait appel aux notions de variance de dispersion et de variance d'estimation.

Considérons un sous domaine v de l'espace V .

Par exemple, V est un dépôt d'argile et v l'ensemble fini des points de mesure.

$$v = \{x_i / i = 1, 2, \dots, N\}$$

où N est le nombre de points de mesure. Par suite, la moyenne des valeurs mesurées est :

$$Z^* = \frac{1}{N} \sum_{i=1}^N z(x_i)$$

Lorsqu'on estime la vraie valeur de la moyenne z par z^* on commit une erreur d'estimation

$$\bar{V} = z_v - z_v^*$$

Quand on adapte comme valeur moyenne de la variable dans un domaine V celle du domaine v , on fait une erreur d'estimation dont la variance (d'estimation) est :

$$\sigma_e^2(v, V) = \text{Var}(Z_v - Z^*) = 2\bar{\gamma}(v, V) - \bar{\gamma}(V, V) - \bar{\gamma}(v, v)$$

$\bar{\gamma}(v, V)$, $\bar{\gamma}(V, V)$ et $\bar{\gamma}(v, v)$ sont des valeurs moyennes de variogramme.

$\gamma(x_i - x)$, $\gamma(x - y)$ et $\gamma(x_i - x_j)$ respectivement, quand x_i , x_j décrivent indépendamment le domaine v et x et y décrivent le domaine V .

Exemple :

$$\begin{aligned}\bar{\gamma}(v, V) &= \frac{1}{v \cdot V} \int_v dx_i \int_V \gamma(x_i - x) dx \\ \bar{\gamma}(V, V) &= \frac{1}{V^2} \int_V dx \int_V \gamma(x - y) dy \\ \bar{\gamma}(v, v) &= \frac{1}{V^2} \int_v dx_i \int_v \gamma(x_i - x_j) dx_j\end{aligned}$$

La variance d'estimation dépend de la forme et de la taille de v et V , ainsi que de la position relative de v par rapport à V .

De même, la variance de la valeur moyenne Z_v^* des $z(x_i)$ lorsque v occupe toutes les positions possibles dans un domaine peu vaste V a une valeur moyenne appelée variance de dispersion de v dans V et est désignée par : $\sigma_D^2(v/V)$. On démontre que :

$$\sigma_D^2(v/V) = \gamma(V, V) - \gamma(v, v)$$

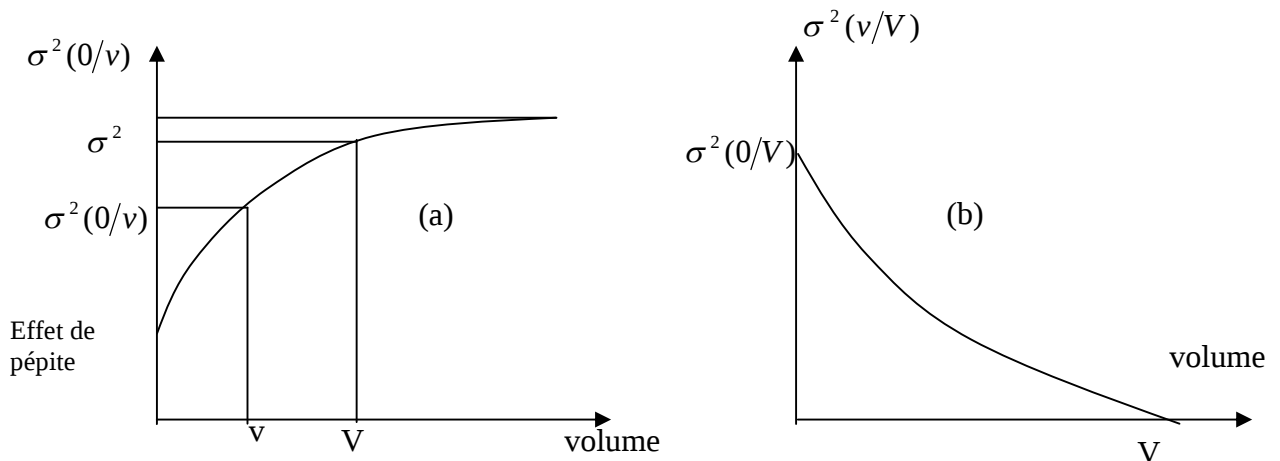


Figure III.4 Variance d'estimation (a) et variance de dispersion (b)

III.5 LE KRIGEAGE

Comme on l'a laissé entendre précédemment le variogramme permet d'aborder deux classes de problèmes : les problèmes d'estimation (théorie du krigeage) et les problèmes de fluctuation (thème des simulations).

Les observations sont généralement faites de manière discontinue et irrégulière, à une ou plusieurs dimensions.

Il est donc souvent nécessaire de procéder à des interpolations pour connaître la valeur la plus probable de la variable en dehors des points de mesure souvent disparates et en nombre limité.

Deux types d'interpolation sont généralement employées : d'une part les méthodes d'interpolation par ajustement global, qui consistent à « ajuster » une fonction (polynôme simple ou trigonométrique) à l'ensemble des valeurs expérimentales, et d'autre part les méthodes d'interpolation par ajustement local comprenant les procédés classiques basés sur la méthode des moindres carrés et la technique de krigeage. Mais ces procédés, excepté la technique du krigeage, ont un inconvénient majeur : l'ajustement conduit en quelque sorte à « forcer » les données à entrer dans le cadre rigide d'une fonction analytique et la continuité de la variable est parfois surestimée. En plus, pour ce qui est des méthodes classiques, l'isotropie est souvent admise et la condition d'optimalité n'est pas respectée.

Seule la méthode d'interpolation par krigeage repose sur une méthode statistique satisfaisante et permet d'obtenir la vraie variance d'estimation.

Brièvement, car nous ne pouvons le faire dans ce cadre de travail, nous dirons que le krigeage est une méthode d'estimation de la dérive spatiale d'une variable aléatoire qui opère par combinaisons linéaires des différents échantillons disponibles en minimisant l'erreur moyenne.

Cette méthode est particulièrement bien adaptée pour tenir compte des différents facteurs qui influent la précision de l'estimation :

- nombre de points de mesure de la propriété et qualités des mesures,
- position de ces points dans le domaine étudié et distance entre eux et la zone intéressante,
- continuité spatiale des variables interpolées.

Le problème se pose ainsi : on dispose des valeurs expérimentales $\{z(x_\alpha), \alpha = 1, 2, \dots, N\}$, mesurées aux points x_α du sous ensemble v du domaine V .

On connaît le variogramme $\gamma(h)$ de $z(x)$, on veut déterminer au point $x_0 \in V$ le meilleur estimateur linéaire non biaisé $z^*(x_0)$.

L'estimateur de la fonction aléatoire $z(x)$ de moyenne m inconnue et dont les $z(x_\alpha)$ sont une réalisation particulière, au point x_0 , est de la forme :

$$z^*(x_0) = \sum_{\alpha=1}^N \lambda_\alpha Z(x_\alpha)$$

L'estimateur $z^*(x_0)$ est une combinaison linéaire des n valeurs connues ; les coefficients λ_α , dits pondérateurs, sont à déterminer de telle façon que :

- L'estimateur soit non biaisé (condition d'universalité)

$$E\{Z(x_0) - Z^*(x_0)\} = 0$$
- La variance d'estimation soit minimale (condition d'optimalité)

$$E\{[Z(x_0) - Z^*(x_0)]^2\} \text{ minimale}$$

III.5.1 Absence de dérive (krigeage ponctuel simple)

Dans ce cas, la moyenne m est constante, soit :

$$E\{Z(x_\alpha)\} = m \text{ et } E\{Z(x_0)\} = m$$

- La condition de non biais entraîne que :

$$E\left\{\sum_{\alpha=1}^N \lambda_{\alpha} (x_{\alpha}) - Z(x_0)\right\} = \sum_{\alpha=1}^N \lambda_{\alpha} m - m = m \left(\sum_{\alpha=1}^N \lambda_{\alpha} - 1\right)$$

Cette condition est vérifiée pourvu que :

$$\sum_{\alpha=1}^N \lambda_{\alpha} = 1$$

▪ La condition de variance d'erreur minimale s'exprime en fonction de la fonction de covariance :

$$E\left\{[Z(x_0) - Z^*(x_0)]^2\right\} = \sum_{\alpha=1}^N \sum_{\beta=1}^N \lambda_{\alpha} \lambda_{\beta} \cdot C(x_{\alpha}, x_{\beta}) + C(0) - 2 \sum_{\alpha=1}^N \lambda_{\alpha} \cdot C(x_{\alpha}, x_0)$$

ou en termes de variogramme :

$$E\left\{[Z(x_0) - Z^*(x_0)]^2\right\} = - \sum_{\alpha=1}^N \sum_{\beta=1}^N \lambda_{\alpha} \lambda_{\beta} \cdot \gamma(x_{\alpha}, x_{\beta}) + 2 \sum_{\alpha=1}^N \lambda_{\alpha} \cdot \gamma(x_{\alpha}, x_0)$$

La variance d'estimation apparaît ainsi comme une forme quadratique en λ_{α} , λ_{β} qu'il faut minimiser sous la contrainte de non biais :

$$\sum_{\alpha=1}^N \lambda_{\alpha} = 1$$

Pour cela, on introduit le multiplicateur de Lagrange μ et on cherche le minimum de l'expression :

$$\phi = \text{Var}\{Z(x_0) - Z^*(x_0)\} - 2\mu \left(\sum_{\alpha=1}^N \lambda_{\alpha} - 1\right)$$

En annulant les dérivées partielles :

$$\frac{\partial \phi}{\partial \lambda_{\alpha}}, \alpha = 1, 2, \dots, N \text{ et } \frac{\partial \phi}{\partial \mu}$$

On obtient le système de (N+1) équations à (N+1) inconnues (les pondérateurs λ_{α} et le paramètre de Lagrange μ), dit « système de krigeage » :

$$\begin{cases} \sum_{\beta=1}^N \lambda_{\beta} \cdot \gamma(x_{\alpha}, x_{\beta}) + \mu = \gamma(x_{\alpha}, x_0), \alpha = 1, 2, \dots, N \\ \sum_{\alpha=1}^N \lambda_{\alpha} = 1 \end{cases}$$

La variance de krigeage est donnée par l'expression :

$$\sigma_K^2 = \text{Var}\{Z(x_0) - Z^*(x_0)\} = \sum_{\alpha=1}^N \lambda_{\alpha} \gamma(x_0 - x_{\alpha}) + \mu$$

Pour chaque point à kriger, il existe donc un ensemble d'équations du type :

$$\sigma_{CK}^2 = \left(1 - \sum_{\alpha=1}^N \lambda_{\alpha}\right) \cdot C(0) + \sum_{\alpha=1}^N \lambda_{\alpha} \cdot \gamma(x_0 - x_{\alpha})$$

III.5.2 Prise en compte d'une dérive (krigeage universel)

Quand il existe une dérive, le variogramme expérimental ne tend généralement pas vers une limite : les valeurs de $z(x)$ sont interdépendantes dans tout le champ d'étude.

Dans le cas général, la variable étudiée peut être considérée comme la réalisation $Z(x)$ d'une fonction aléatoire, non stationnaire en général. Elle peut être décomposée en la somme d'un terme de dérive et d'un terme résiduel $y(x) = Z(x) - m(x)$, d'espérance nulle :

$$Z(x) = m(x) + y(x), \text{ avec } E\{Z(x)\} = 0$$

On suppose connue la forme de la dérive $m(x)$. Par exemple $m(x)$ peut être une combinaison linéaire de k fonctions quelconques mais connues ($f_l(x)$, $l = 1$ à k), les coefficients a_l restant bien sûr inconnus, de sorte que la dérive $m(x)$ reste inconnue :

$$m(x) = \sum_{l=1}^k a_l \cdot f_l(x)$$

On peut ainsi adopter les premiers termes du développements de Taylor, soit :

$$m(x) = a_1 + a_2 x \text{ pour une dérive dite « linéaire »,}$$

$$m(x) = a_1 + a_2 x + a_3 x^2 \text{ pour une dérive dite « quadratique ».}$$

Dans un espace à deux dimensions, la dérive quadratique prendrait la forme :

$$m(u, v) = a_1 + a_2 u + a_3 v + a_4 u^2 + a_5 v^2 + a_6 uv$$

L'estimateur $Z^*(x_0)$ de la variable à interpoler au point x_0 est donnée par la même expression que précédemment :

$$z^*(x_0) = \sum_{\alpha=1}^N \lambda_{\alpha} z_{\alpha}$$

Les conditions à respecter sont aussi les mêmes :

- Condition de non biais

$$\sum_{\alpha=1}^N \lambda_{\alpha} \sum_{l=0}^k a_l f_l(x_{\alpha}) - \sum_{l=0}^k a_l f_l(x_0) = 0$$

Donc :

$$\sum_{l=0}^k a_l \left(\sum_{\alpha=1}^N \lambda_{\alpha} \cdot f_l(x_{\alpha}) - f_l(x_0) \right) = 0$$

Les coefficients a_l étant inconnus, on doit aussi vérifier que :

$$\sum_{\alpha=1}^N \lambda_{\alpha} \cdot f_l(x_{\alpha}) = f_l(x_0), \text{ quel que soit } l.$$

l variant de 0 à k, ce qui impose (k+1) conditions supplémentaires.

- Variance d'estimation minimale

Compte tenu des conditions de non biais précédentes, les termes faisant intervenir la dérive inconnue $m(x)$ s'éliminent du développement de la variance d'estimation :

$$E\left\{[Z(x_0) - Z^*(x_0)]^2\right\},$$

Et il reste :

$$E\left\{[Z(x_0) - Z^*(x_0)]^2\right\} = 2 \sum_{\alpha=1}^N \lambda_{\alpha} \cdot \gamma(x_{\alpha}, x_0) - \sum_{\alpha=1}^N \sum_{\beta=1}^N \lambda_{\alpha} \lambda_{\beta} \cdot \gamma(x_{\alpha}, x_{\beta})$$

La variance d'estimation sera donc minimale si les λ_{α} vérifient :

$$\begin{cases} \sum_{\alpha=1}^N \lambda_{\alpha} \cdot \gamma(x_{\alpha} - x_{\beta}) = \gamma(x_{\alpha} - x_0) - \sum_{l=0}^k \mu_l \cdot f_l(x_{\alpha}), \forall \alpha = 1 \text{ à } N \\ \sum_{\alpha=1}^N \lambda_{\alpha} \cdot f_l(x) = f_l(x_0), l = 0 \text{ à } k \end{cases}$$

Il y a donc k multiplicateurs de Lagrange associés à la condition de non biais. Il en existe autant que de termes dans l'expression de la dérive.

Notons que :

$\gamma(x_{\alpha} - x_{\beta})$ est la valeur de la demi variance correspondant à la distance entre deux points connus x_{α} et x_{β} dans l'espace de régionalisation et $\gamma(x_{\alpha} - x_0)$ est la valeur de la demi-variance correspondant à la distance entre le point x_{α} et le point à interpoler x_0 .

Les valeurs $\gamma(x_{\alpha} - x_{\beta})$ et $\gamma(x_{\alpha} - x_0)$ sont calculées à l'aide de l'expression correspondant au modèle de variogramme sous-jacent choisi. La variance d'estimation est :

$$\sigma^2 = \sum_{\alpha=1}^N \lambda_{\alpha} \cdot \gamma(x_{\alpha} - x_0) + \sum_{l=0}^k \mu_l \cdot f_l(x_0)$$

III.5.3 Krigage étendu (estimation des valeurs moyennes)

C'est la généralisation du krigage ponctuel. Le support est ici un segment, une surface ou un volume. Le krigage est évidemment effectué là aussi qu'il existe ou non une dérive (krigage étendu simple ou krigage étendu universel). La valeur moyenne peut être calculée de deux manières : ou bien les observations faites dans l'élément étudié servent seules au calcul (krigage étendu local), ou bien celles faites en des points du voisinage sont elles aussi prises en compte (krigage étendu global).

Considérons l'estimation de la valeur moyenne :

$$Z_v(x_0) = \frac{1}{V} \int_V z(x) dx$$

définie sur le segment V centré au point x_0 , à partir de N informations $z(x_\alpha)$ définies sur les supports v_α .

La technique de calcul est identique à celle du krigeage ponctuel.

Ainsi, l'estimateur est de la forme :

$$Z_v(x_0) = \sum_{\alpha=1}^N \lambda_\alpha \cdot z(x_\alpha)$$

N désigne le nombre de points de mesure, et les coefficients λ_α sont choisis tels que :

- $E\{Z_v(x_0) - Z_v^*(x_0)\} = 0$ (estimateur sans biais)
- $E\{Z_v(x_0) - Z_v^*(x_0)\}^2$ minimale.

Dans le cas du krigeage local universel, par exemple, la dérive inconnue est une fonction polynomiale de forme générale :

$$m(x) = \sum_{l=1}^k a_l f_l(x)$$

Le formalisme classique de Lagrange, à l'aide duquel la variance de l'erreur $(Z_v - Z_v^*)$ est minimisée sous la contrainte de non biais,

$$\sum_{\alpha=1}^N \lambda_\alpha = 1$$

Conduit à un système de $(N+k)$ équations linéaires à $(N+k)$ inconnues (les N pondérateurs λ_α et les k paramètres de Lagrange μ_l) appelé « système de krigeage sans biais d'ordre k » ou encore système de krigeage universel :

$$\sum_{\beta=1}^N \lambda_\beta \cdot \gamma(x_\alpha - x_\beta) + \sum_{l=1}^k \mu_l \cdot b_{v_\alpha}^l = \bar{\gamma}(v_\alpha, V) \quad \forall \alpha = 1 \text{ à } N$$

$$\sum_{\beta=1}^N \lambda_\beta \cdot b_{v_\beta}^l = b_v^l, \quad \forall l = 1 \text{ à } k$$

Avec :

$$b_v^l = \frac{1}{V} \int_V f_l(x) dx$$

$$b_{V_\alpha}^1 = \frac{1}{V} \int_{V_\alpha} f_1(x_\alpha) dx$$

Les $f_1(x)$ sont les $(k+1)$ fonctions d'estimation connues de la dérive $m(x)$.

La variance de krigeage est :

$$\sigma_{ke}^2 = \sum \lambda_\beta \cdot \bar{\gamma}(v_\alpha, V) + \sum_{l=0}^k \mu_l \cdot b_v^1 - \bar{\gamma}(v, v)$$

Toutefois les supports sélectionnés, à l'intérieur desquels sont recueillis les échantillons, sont souvent des supports riches en ce sens que la caractéristique étudiée y a souvent des valeurs élevées ; le krigeage local conduit pour cette raison souvent à surestimer les estimations et cela d'autant plus que les échantillons extérieurs sont considérés comme d'influence nulle. Il faut donc dans ce cas faire des estimations en tenant compte aussi des échantillons recueillis au voisinage. C'est le krigeage étendu global.

Dans le krigeage étendu global, le principe de calcul de l'estimation de la valeur moyenne d'un paramètre sur un support V est le même que celui utilisé pour le krigeage local, à une différence près : les N points pris en compte dans le système de krigeage sont ceux contenus dans le support V , pour lequel la moyenne est calculée, et ceux se trouvant au voisinage de V .

III.6 COKRIGEAGE

Dans la plupart des situations pratiques, une des variables régionalisées peut ne pas avoir été suffisamment échantillonnée (difficultés expérimentales, coût élevé...).

La technique de cokrigeage, tout à fait analogue à celle du krigeage, permet d'estimer cette variable, non seulement à partir de données disponibles sur cette variable mais aussi sur d'autres variables qui lui sont corrélées.

Chapitre IV

VARIABILITÉ SPATIALE DES PARAMÈTRES GEOTECHNIQUES

Chapitre IV

VARIABILITE SPATIALE DES PARAMETRES GEOTECHNIQUES

IV.1 INTRODUCTION

Une reconnaissance géotechnique est la première phase d'étude d'un projet de génie civil. Cette reconnaissance permet de décrire les différentes couches de sols, leurs natures, leurs caractéristiques mécaniques et physiques, etc. Ces dernières sont déterminées par des mesures in situ, directes ou indirectes, ou par des mesures en laboratoire sur des échantillons remaniés ou non remaniés.

Le rapport de cette campagne comprend, outre les résultats bruts des observations, essais et mesures, une interprétation géologique et géotechnique et éventuellement des avis et des conseils. On notera que, de manière consciente ou non consciente, un certain nombre d'hypothèses sont introduites à ce stade. C'est ainsi que l'on va procéder à l'interpolation entre les différents sondages ou forages en supposant ainsi la continuité des couches géologiques, ce qui se fait souvent alors que ce n'est pas toujours justifié. De même, on introduira implicitement des conditions d'homogénéité de ces couches, ou du moins de faible variabilité.

L'ingénieur doit ensuite analyser ce rapport en collaboration, selon la complexité et l'importance du projet, avec des spécialistes d'autres disciplines : géologues, hydrogéologues, urbaniste, etc. Il pourra ainsi se forger sa propre idée des conditions de site, pour finalement utiliser ces résultats pour le dimensionnement de l'ouvrage dont il a la responsabilité.

L'examen des résultats d'une campagne d'essais géotechniques met en évidence que, même à l'intérieur d'une couche géologique considérée comme homogène en grand, les divers paramètres mesurés manifestent toujours une certaine variabilité, parfois très élevée.

Cette variabilité est quelques fois tellement importante que le choix de valeurs caractéristiques semble illusoire.

En effet, il se peut que la variation imprévisible soit d'une importance telle qu'une description déterministe s'avère insuffisante. Il est alors plus indiqué de faire usage de la théorie

des probabilités. Cette dernière permet de développer des modèles plus proches de la réalité. Ces modèles sont d'autant plus exploitables qu'il est possible de rassembler des caractéristiques du phénomène stochastique complexe en un minimum de paramètres physiques, expérimentalement accessibles.

Avant d'aborder la manière de prendre en compte cette variabilité des paramètres, on peut d'abord s'interroger sur ses causes profondes. L'examen critique et approfondi des variations entre les valeurs mesurées d'un même paramètre dans une couche de sol "homogène", montre qu'elles sont attribuables à plusieurs sources d'incertitude.

L'objectif de ce chapitre réside dans la description des sources de variations des paramètres de sol lesquelles présentent des caractères fortement aléatoires. A cet effet, une description probabiliste de ces paramètres sera établie ce qui va nous permettre de mettre en place une modélisation stochastique de la variabilité des propriétés du sol par simulation.

IV.2 SOURCES D'INCERTITUDES SUR LES PARAMETRES DE SOL

Les incertitudes sont présentées dans la plupart des problèmes de génie civil et plus spécifiquement dans le domaine de la géotechnique. Le mot "incertitude" y est utilisé et interprété de plusieurs façons selon les personnes et le domaine d'application. Dans la plupart des cas, il englobe les concepts d'ambiguïté et de variabilité ; il est aussi utilisé pour décrire l'état de ce qui ne peut être établi avec exactitude, qui laisse place au doute.

Certain auteurs (Denis, 1998 ; cité par Smith et al 2004) distinguent deux niveaux d'incertitude. Le premier niveau est celui où un individu perçoit qu'il ne peut résoudre un problème à cause d'une trop grande complexité ou d'un manque d'information. Le deuxième niveau est celui où plusieurs individus perçoivent différemment une même situation et proposent des solutions contradictoires.

Les incertitudes présentes dans les problèmes de géotechnique proviennent généralement de trois sources : la variabilité inhérente à certains processus naturels, la caractérisation des sols et l'imperfection des modèles rhéologiques.

Partant des études du l'NRC, 2000 (cité par Smith, 2004), de Favre, (2000) et de Magnan, (2000), une synthèse a été élaborée représentant les différentes sources d'incertitude avec leurs composantes principales.

IV.2.1 Incertitude liée à la variabilité naturelle des paramètres des sols

Le désordre qui affecte les propriétés des sols 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, c'est-à-dire 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.

Cet incertain naturel peut avoir divers degrés et donc entraîner différents moyens d'investigation, de prise en compte et de modélisation (Favre et al 2000 et Favre et al 1995b) :

- 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.
- On peut se trouver dans des terrains peu désordonnés présentant une stratification relativement régulière. Aussi, la reconnaissance géotechnique, guidée par le modèle géologique du terrain, prend ici toute son importance pour proposer un modèle mécanique de terrain.
- Les terrains peuvent avoir été fabriqués par l'homme, qu'il s'agisse de barrages en terres, de remblais, de terre armée, etc. Dans ces cas, il faut des contrôles de fabrication et des tests de réception.

IV.2.2 Erreurs et incertitudes liées à la caractérisation des sols

1. Erreurs d'observation

a) Erreurs de mesure, e_{ao}

Elles sont liées à l'appareil de mesure et à l'opérateur.

Avec le progrès de la métrologie et l'acquisition automatique, les erreurs d'imprécision et d'opérateur ont été considérablement réduites. 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, 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 à un autre.

Les problèmes organisationnels tel une mauvaise communication entre l'opérateur des essais (in situ et/ou laboratoire) et ses supérieurs, sont aussi une source potentielle d'incertitude humaine qui est souvent rencontrée en pratique. Ces incertitudes influence directement ou à un second degré tous les aspects d'un problème, de sa définition jusqu'au choix de la solution car le risque est toujours porté et géré par l'humain.

b) Erreurs de représentativité, e_{Rep}

Elles proviennent de la transformation de la mesure physique.

On ne mesure guère directement que des longueurs, des masses et des températures. La plupart des appareils d'essais géotechniques 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 ? Etc.

c) Erreurs de l'instant, e_{inst}

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

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. Erreurs d'enquête

a) Erreurs d'enquêtes proprement dites

Ce sont les erreurs liées à une mauvaise conduite des reconnaissances. Les mesures effectuées ne sont pas représentatives du problème : l'exemple d'un remblais récent mal identifié au droit de deux fondations superficielles qui les sollicite, un remblai ancien en biseau pouvant exister sous une des deux fondations ; un autre exemple illustre le problème d'une fondation superficielle reconnue au moins de 3 à 5 fois la largeur, ignorant ainsi une couche faible qui peut piloter les tassements.

b) Erreurs d'échantillonnage

Ce ne sont pas, à proprement parler, des erreurs. Imaginons que chaque mesure est parfaite mais que la propriété varie d'un point à un autre de l'espace. Cette variabilité naturelle sera modélisée comme une variable aléatoire dont il est à estimer les caractéristiques avec les mesures. Il y a ainsi un échantillon statistique de taille n pour estimer les caractéristiques d'une population infinie. Ces caractéristiques vont être estimées par un intervalle de confiance : la moyenne (ou l'écart type) de la propriété, qui est sûre mais inconnue à une probabilité de 95% par exemple d'être encadrée par deux valeurs dépendant des observations.

IV.2.3 Incertitudes liées à l'imperfection des modèles rhéologiques

Il est à noter que les résultats des calculs sont également affectés par de nombreuses incertitudes, parmi lesquelles, l'imperfection des lois de comportement (les modèles) et l'imprécision des calculs réalisés sur la base de ces modèles. Les propriétés des sols rentrent dans ces lois rhéologiques qui représentent plus ou moins bien la réalité de la relation contraintes-déformations. Ainsi le modèle élasto-plastique linéaire isotrope-plastique parfait ou fragile, alors que le sol est mieux représenté par un modèle élasto-plastique orthotrope (Favre, 2000). Quant au comportement réel du sol, personne ne le connaîtra ; 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.

En particulier, on ne peut pas nier que certaines sollicitations naturelles appliquées aux constructions présentent un caractère essentiellement aléatoire : les actions sismiques, le vent, les sollicitations des vagues sur les structures marines, les efforts de la température ou des gradients thermiques dans les chaussées ou les barrages en béton en sont des exemples typiques.

Les exemples sont innombrables, mêmes en se limitant au domaine des infrastructures, et une réflexion critique conduit rapidement à admettre que les actions connues avec une incertitude pouvant être considérée comme négligeable constituent plutôt l'exception : poids propre des ouvrages, certaines surcharges permanentes, etc.

IV.3 APPROCHE PROBABILISTE

IV.3.1 Principe général

L'approche probabiliste aborde les problèmes d'une manière toute différente, en postulant a priori le caractère aléatoire des grandeurs qui interviennent dans les phénomènes étudiés et dans les modèles de comportement utilisés pour décrire ces phénomènes.

Dans une telle approche, on quantifie toutes les incertitudes attachées au modèle, aux paramètres, aux facteurs humains, etc. On en déduit ensuite des résultats qui sont eux-mêmes affectés d'incertitudes. On obtient ainsi, quelle que soit la nature du problème étudié, des résultats exprimés sous une forme homogène, à savoir la probabilité de voir se produire certains événements.

Dans le modèle de calcul (figure 3.1), les données et les résultats sont affectés d'incertitudes. Ce ne sont plus des grandeurs uniques bien définies, mais des variables aléatoires.

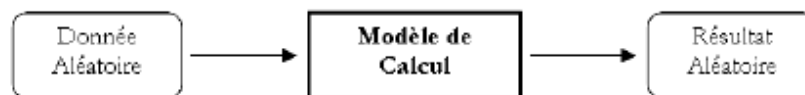


Figure IV.1 Schéma général de calcul par un modèle de type probabiliste

La nature exacte du modèle n'est pas précisée à ce stade, pas plus que les méthodes de calcul particulières qui devront être mises en oeuvre pour utiliser des quantités aléatoires.

IV3.2 Principales étapes de l'approche probabiliste

Le schéma général de calcul par un modèle de type probabiliste a déjà été présenté plus haut (figure I.1). On va maintenant examiner de plus près quelques-unes des nombreuses questions formulées et à résoudre :

- Comment obtenir les données du problème sous la forme de variables aléatoires de telle manière qu'elles soient une représentation assez fidèle de la réalité ?
- Comment choisir rationnellement un modèle de calcul utilisable dans le contexte probabiliste ?

- Comment transmettre, au travers du modèle de calcul, le caractère aléatoire (les variations) de ces données ?
- Comment enfin utiliser les résultats, eux aussi exprimés sous la forme de variable aléatoires, pour obtenir les réponses recherchées ?

1. Modélisation des données aléatoires

Dans le domaine de la mécanique des sols, la quantification des paramètres d'entrée (données) s'appuie souvent sur un traitement statistique de données expérimentales. En l'absence d'un nombre suffisant de mesures relatives au site étudié, on s'appuie sur des valeurs considérées comme vraisemblables sur la base de considérations plus générales, géologiques et autres. En réalité donc, les paramètres de type géotechnique seraient considérés à priori comme des variables aléatoires, c'est-à-dire affectés d'une incertitude, mais dont on pourra pour certains négliger la dispersion.

Notons dès à présent que le traitement statistique dont il est question est presque toujours un simple processus, qui traite les données comme des valeurs extraites au hasard d'une seule population d'éléments tous de nature identique.

Ce traitement peut aussi être un processus plus complexe, prenant en compte la position dans l'espace des valeurs mesurées. De telles grandeurs, sont ainsi définies à la fois par :

- La valeur du paramètre mesuré,
- La position dans l'espace où ce paramètre a été mesuré (ou la position dans l'espace où l'échantillon étudié en laboratoire a été prélevé),

Ces grandeurs forment un champ aléatoire.

Dans les traitements statistiques particuliers pour traiter les caractéristiques du sol, on insistera particulièrement sur la nécessité de toujours prendre en compte la variabilité spatiale de ces grandeurs, même si le nombre de données disponibles est très faible et peut sembler à première vue insuffisant.

2. Choix du modèle et incertitudes associées

La quantification des incertitudes liées aux paramètres est difficile à estimer à ce stade du travail, c'est le cas aussi des incertitudes liées aux modèles. Même si cet aspect n'est pas très développé dans cette étude, il ne sera pas ignoré. Il en sera tenu compte à travers la variation des

paramètres qui sont considérés comme des variables aléatoires. Il en sera de même pour les résultats des modèles.

Dans l'approche qui nous concerne, la prédominance est donnée à la variabilité spatiale des paramètres plutôt qu'aux incertitudes. Ceci est justifié par les problèmes qui nous préoccupent. On pourra ainsi considérer que leurs conséquences sont négligeables devant les causes liées à la variabilité spatiale des paramètres géotechniques. Pour cela, on adoptera, comme dans une approche classique, des modèles de calcul déterministes ayant fait leurs preuves. La principale différence consistera dans le type de données à introduire et le type de résultats obtenus, qui seront tous des variables aléatoires. Les calculs ainsi menés sont des calculs probabilistes.

3. Calculs probabilistes

Une étape cruciale de la mise en oeuvre des méthodes de calcul probabiliste est celle qui consiste à transférer, à travers le modèle de calcul (déterministe) choisi, les incertitudes des différents paramètres. Ces derniers sont en effet des variables aléatoires, et les méthodes rigoureuses permettant de réaliser des calculs exacts sur de telles variables ne s'appliquent qu'à quelques cas extrêmement simples sans aucun intérêt pratique.

4. Exploitation des résultats

Les résultats obtenus par les calculs probabilistes sont donc des variables aléatoires. Mais les problèmes posés aux ingénieurs ont besoin de réponses qui ne peuvent pas être facilement exprimés sous cette forme probabiliste. Ainsi, il est nécessaire de pouvoir formuler des réponses à des interrogations précises, qui pourraient être entre autre :

- La probabilité de voir se produire un événement bien déterminé, comme la rupture ou la mise hors service d'un ouvrage ou d'une partie d'un ouvrage ;
- Les formes, dimensions et caractéristiques des matériaux de l'ouvrage ou de la partie de l'ouvrage conduisant à un niveau de probabilité de rupture ou de mise hors service choisi à priori (calcul de dimensionnement) ;
- Connaissant les coûts respectifs des opérations de constructions, d'entretiens et ceux entraînés par un éventuel accident, on peut optimiser le dimensionnement en terme économique en recherchant la solution conduisant au coût global minimum, etc

IV.4 MODELISATION STOCHASTIQUE DE LA VARIABILITE SPATIALE

La formation d'un sol dépend de nombreux facteurs variables dans le temps et dans l'espace. On peut utiliser des techniques très variées pour analyser la variabilité naturelle des sols. Les méthodes statistiques et probabilistes ont permis aux géotechniciens de s'intéresser à l'analyse de l'hétérogénéité des couches de sol. Cette analyse a deux objectifs :

- Caractériser la variabilité naturelle des propriétés physiques et mécaniques des sols pour en tenir compte dans les modèles de calcul ;
- Optimiser la reconnaissance géotechnique en jouant sur le nombre et position des sondages.

IV.4.1 Différentes formes de variation des paramètres de sol

Le praticien de la mécanique des sols est appelé quotidiennement à choisir, à partir d'un ensemble de mesures expérimentales et d'informations diverses, les valeurs de calcul à utiliser dans les méthodes de calcul de stabilité, de tassement, etc. Ce choix est souvent simplement guidé par l'expérience et le bon sens. Il est rarement le résultat d'un processus totalement objectif.

Il est reconnu que les caractéristiques des sols sont souvent dispersées et qu'elles peuvent, même dans une couche considérée comme homogène en grand, varier fortement avec la profondeur et d'un endroit à un autre dans le plan. Ces variations ne sont prises éventuellement en compte que par une description déterministe, liant par exemple la valeur de la propriété étudiée à la profondeur par une fonction linéaire.

La distribution dans l'espace des valeurs des propriétés physiques et mécaniques des sols est une donnée extrêmement importante pour le dimensionnement des ouvrages. La figure.3.2 ci-dessous montre les trois cas rencontrés en pratique pour la variation des paramètres de sol en fonction de la profondeur :

- Le premier cas correspond à une distribution des paramètres de sol autour d'une moyenne constante.
- Le deuxième cas correspond à des paramètres de sol dont la moyenne dépend de la profondeur Z .
- Le troisième cas correspond à des paramètres de sol dont la variance dépend aussi de la profondeur.

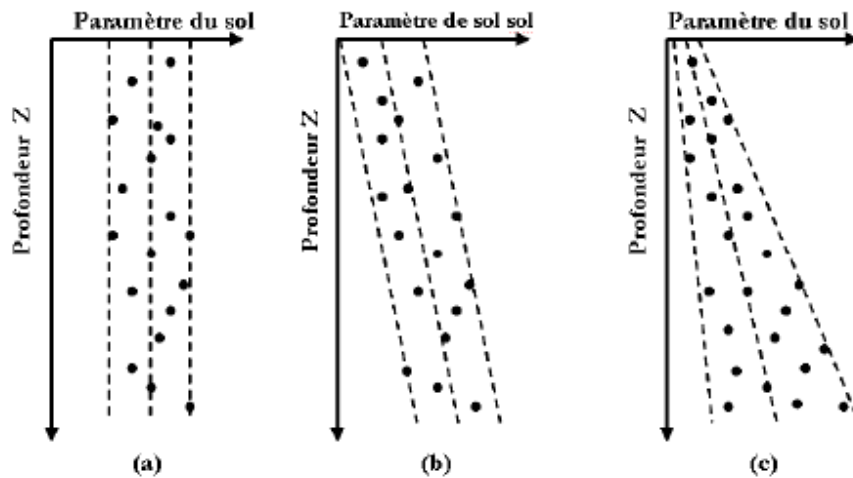


Figure IV.2 Les différentes formes de variation des paramètres de sol en fonction de la profondeur (Jaksa, 1995).

La technique d'analyse la plus courante pour les variations spatiales des propriétés des sols consiste à décomposer les variations observées de chaque propriété en une fonction déterministe des coordonnées du point de mesure dans l'espace appelée dérive (ou tendance), et une composante aléatoire de moyenne nulle et de variance égale à une constante. Dans une telle décomposition, la dérive représente l'effet des phénomènes qui ont influencé la formation du sol pendant de longues périodes. La composante aléatoire traduit plutôt les effets des fluctuations des conditions de formation du sol de courte durée. Les variations de cette composante aléatoire peuvent être plus ou moins importantes suivant les sols et les sites. En général, les valeurs d'une propriété donnée ont plutôt tendance à être voisines en des points voisins. La vitesse de variation de ces valeurs quand on passe d'un point à un autre est aussi un paramètre très important pour la description de la variabilité naturelle des propriétés physiques et mécaniques des sols. On la caractérise classiquement par la fonction d'autocorrélation.

IV.4.2 Dérive et autocorrélation des propriétés des sols

1. Décomposition des variations spatiales des variables aléatoires

La décomposition d'une variable " C " en la somme d'une fonction déterministe " $D(x,y,z)$ " qui décrit les variations de la valeur moyenne de " X ", dans l'espace et une variable aléatoire $A(x,y,z)$ de moyenne nulle en chaque point de l'espace comme le montre la formule suivante :

$$C(x, y, z) = D(x, y, z) + A(x, y, z) \quad (I.1)$$

Cette opération est souvent pratiquée en mécanique des sols pour des observations effectuées en différents points du massif étudié. Une schématisation de cette décomposition en fonction seulement de la profondeur Z est illustrée sur la figure 3.3 ci-dessous.

Si la fonction $D(x, y, z)$ est connue en tous points avec précision, indépendamment des valeurs observées dans le massif (ce qui habituellement pas le cas), la variabilité des paramètres du sol peut être étudiée sur la seule variable $A(x, y, z)$ dont on connaît la valeur aux points de mesure.

Dans la pratique, on ne connaît pas de façon exacte la dérive $D(x, y, z)$. Cette dernière est estimée à partir de l'échantillon, ce qui se traduit par des répercussions sur l'analyse de la composante aléatoire $A(x, y, z)$.

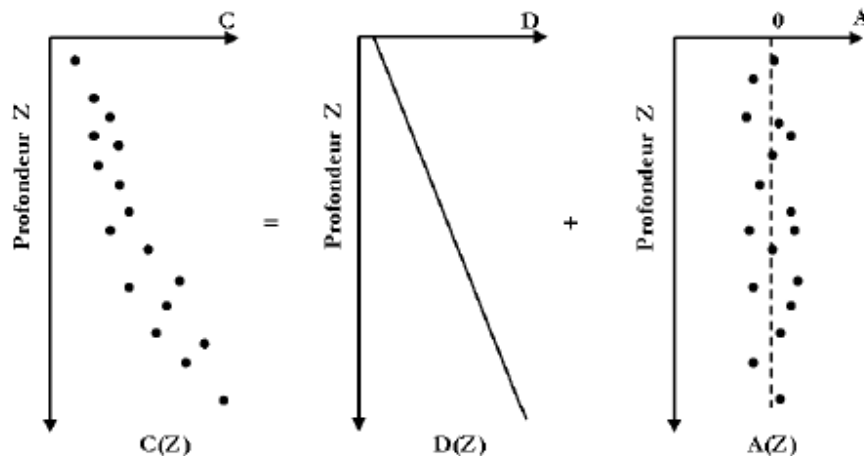


Figure IV.3 Modification des caractéristiques d'autocorrélation par séparation de la dérive
(Magnan, 1982)

2. Etude de la dérive $D(x, y, z)$

A priori, la forme de la fonction $D(x, y, z)$ peut être quelconque. Deux facteurs viennent toutefois limiter l'éventail des formes mathématiques des fonctions $D(x, y, z)$ utilisées en pratique :

- Les techniques courantes d'estimation des relations entre la variable dépendante $C(x,y,z)$ et les variables indépendante x , y et z ne sont adaptées qu'aux fonctions de forme polynomiale,
- Le nombre des points de mesures de la variable " C " est souvent trop faible pour permettre la détermination de nombreux coefficients dans la fonction $D(x, y, z)$.

Suite à cela, la technique utilisée pour la détermination des paramètres de la fonction $D(x,y,z)$ est identique à celle des régressions linéaires ou polynomiales entre paramètres (Baguelin, 2000). On se donne a priori la forme de la fonction $D(x, y, z)$, par exemple :

$$D(x, y, z) = \beta_0 + \beta_1 x + \beta_2 y + \beta_3 z \quad (IV.2)$$

Ou :

$$D(x, y, z) = \beta_0 + \beta_1 x + \beta_2 y + \beta_3 xz + \beta_4 x^2 + \beta_5 y^2 \quad (IV.3)$$

on détermine la matrice des coefficients β_i et le coefficient de corrélation multiple R qui donne une idée du niveau de représentativité de la relation obtenue à l'aide de tests statistiques.

Dans cette étude, c'est le logiciel StatGraphPlus qui a été utilisé pour déterminer les paramètres de la fonction $D(x,y,z)$ et le coefficient de corrélation multiple R .

3. Etude de la composante aléatoire $A(x,y,z)$

La composante aléatoire s'obtient par soustraction de la dérive aux valeurs mesurées :

$$A(x, y, z) = C(x, y, z) - D(x, y, z) \quad (IV.4)$$

Une fois l'influence de la dérive supprimée, le résidu peut être considéré comme la réalisation d'une fonction aléatoire stationnaire au sens large. Nous pouvons alors calculer la fonction d'autocorrélation qui permettra d'étudier la variabilité spatiale du paramètre considéré.

4. Génération du vecteur des variables aléatoires corrélées

Après la discrétisation du champ aléatoire du paramètre du sol étudié $U(X)$ en éléments (zones), il est remplacé par un vecteur de variables aléatoires dont les valeurs sont générées pour effectuer la simulation de Monté Carlo. Cette dernière consiste à effectuer un grand nombre de

réalisations du paramètre étudié en considérant les paramètres de la formule utilisée dans le calcul comme des variables aléatoires.

La génération des valeurs aléatoires doit être effectuée au moyen d'un algorithme approprié pour que leurs propriétés statistiques reflètent le plus fidèlement possible les propriétés statistiques des mesures, la moyenne, la variance, la matrice de covariance, la matrice de corrélation, etc.

L'algorithme utilisé a été proposé par Fenton (Fenton, 1994). Cet algorithme permet de générer un vecteur de valeurs aléatoires de distribution normale, de moyenne nulle et de variance égale à l'unité.

L'algorithme en question génère des valeurs positives et négatives, ces dernières sont inappropriées pour représenter les paramètres du sol. Pour contourner cette difficulté, l'algorithme a été adapté pour que les valeurs générées soient de distribution lognormale.

La méthodologie suivie dans cet algorithme est décrite ci-après.

Pour générer S réalisations (simulations) d'un vecteur $\{V\}$ de N variables aléatoires v_i ($i = 1, 2, \dots, N$) qui doivent vérifier les propriétés suivantes :

- Chaque variable v_i est de distribution normale.
- Les N variables v_i corrélées entre elles ont une matrice de covariance $[COV]$ (de dimension $N \times N$) et dont chaque élément COV_{ij} représente le coefficient de covariance entre les variables v_i et v_j .

Les étapes suivantes sont réalisées :

- Détermination de la matrice de covariance $[COV]$ en utilisant le modèle de la fonction de variance qui correspond à la fonction de corrélation triangulaire. Pour déterminer cette matrice, une échelle de fluctuation est utilisée pour le cas unidimensionnel, alors que pour le cas bidimensionnel, deux échelles de fluctuation (verticale et horizontale) et une surface de corrélation (Vanmarcke, 1983) ont été utilisées.
- Décomposition de la matrice $[COV]$ par la méthode de Cholesky en une matrice triangulaire inférieure $[L]$ et sa matrice transposée $[L]^T$:

$$[COV] = [L][L]^T \quad (IV.6)$$

- Génération d'un vecteur $\{G\}$ de N variables aléatoires G_i non corrélées entre elles et de distribution normale réduite.

- Calcul du vecteur $\{V\}$ par l'expression suivante :

$$\{V\} = [L] \cdot \{G\} \quad (IV.7)$$

Pour obtenir à partir du vecteur $\{V\}$ un vecteur $\{U\}$ de N variables aléatoires U_i dont la distribution est lognormale de moyenne m et de variance σ^2 , la transformation suivante est appliquée à chaque variable U_i :

$$U_i = \exp \left(m_{\ln U} + \sigma_{\ln U}^2 \cdot V_i(x_i) \right) \quad (IV.8)$$

où $m_{\ln U}$ et $\sigma_{\ln U}^2$ sont respectivement la moyenne et la variance du logarithme de U . Ils s'expriment en fonction de m et σ^2 par : avec x_i sont les coordonnées de la variable $U_i(x_i)$.

Par la suite, la technique de simulation de Monte Carlo sera utilisée pour générer les propriétés du sol et pour calculer le paramètre de dimensionnement concerné en plusieurs réalisations.

IV.5 CONCLUSION

Lors d'un examen de données géotechniques, il faut d'abord différencier les différentes zones au sens statistique, c'est-à-dire séparer les couches nettement distinctes les unes des autres. Ces couches doivent être de dimension significative pour le problème étudié. Il est peu utile de décrire finement une succession de couches quand le problème posé concerne des couches de dimensions importantes.

Les résultats d'essais géotechniques utilisés comportent une dispersion aléatoire propre aux méthodes et aux appareils de mesure eux-mêmes, voire aux facteurs humains ajoutés à l'imperfection des lois de comportement. Par conséquent, cette dispersion provoque une imprécision des calculs réalisés sur la base de ces modèles. Pour palier à cette lacune, il est recommandé d'utiliser les approches probabilistes qui prennent en compte la variabilité des paramètres géotechniques.

Le modèle stochastique par simulation présenté dans ce chapitre est l'un des modèles les plus utilisés pour représenter des phénomènes appliqués à la mécanique des sols (tassement, gonflement, etc.) et des paramètres de dimensionnement (capacité portante, etc.).

Par le biais de ce modèle stochastique, une analyse uni et bidimensionnelle sera menée dans les prochains chapitres sur deux sites différents : site de Cubzac-les-Ponts (France) et le site de l'ex-ITE Tlemcen (Algérie).

Chapitre V

CORRELATIONS ENTRE LES PROPRIÉTÉS DES SOLS

Chapitre V

CORRELATIONS ENTRE LES PROPRIETES DES SOLS

V.1 INTRODUCTION

Les paramètres utilisés pour décrire les propriétés physiques et mécaniques sont de nature très variée :

- paramètres d'**identification** et d'**état** (porosité, indice des vides, densité, densité relative, limites d'Atterberg, etc.) ;
- paramètres de **déformabilité** (indices de compression et de gonflement, module œdométrique, module pressiométrique, etc.) ;
- paramètres de **résistance** (cohésion et angle de frottement interne, pression limite pressiométrique, résistance de cône statique ou dynamique, etc.) ;
- paramètres de **perméabilité**.

Il est très rare que, sur un même site, tous ces paramètres soient mesurés en un nombre de points suffisant pour que l'on puisse juger bien connu l'ensemble du massif de sol. Habituellement, la reconnaissance géotechnique est limitée au strict minimum, et l'on dispose des valeurs de certains paramètres en certains points et d'autres paramètres en d'autres points. L'ingénieur géotechnicien doit tirer le meilleur parti possible de ces informations éparses et établir une coupe géotechnique représentative du site étudié.

C'est dans ce cadre général que l'utilisation de corrélations entre les propriétés physiques et mécaniques des sols peut contribuer efficacement au travail de synthèse du géotechnicien.

V.2 RELATIONS ET CORRÉLATIONS DANS LES SOLS : GÉNÉRALITÉS

V.2.1 Origine des relations et corrélations dans les sols

S'il est difficile, voire impossible, de donner une justification théorique quantitative de l'existence de relations entre les propriétés d'un massif de sol naturel, il est facile d'admettre que les différents paramètres d'un sol donné doivent avoir des

relations : la déformabilité comme la résistance au cisaillement ou la perméabilité dépendent à l'évidence de la forme et de la nature des particules, de la densité de leur empilement, de la quantité d'eau présente dans les pores...

De plus, à l'intérieur d'une même catégorie de paramètres, par exemple les paramètres de résistance, il existe à l'évidence des relations entre les paramètres mesurés dans les différents types d'essais en place ou en laboratoire, même si l'on ne peut pas les exprimer de façon explicite. Et si les paramètres de résistance dépendent des mêmes propriétés physiques que les paramètres de déformabilité, il doit également exister des relations entre ces deux catégories de paramètres.

Cette réflexion purement qualitative est confirmée par l'expérience : il existe effectivement, dans chaque dépôt de sol, des relations entre les paramètres géotechniques, ainsi que des relations plus générales, valables pour un type de sol, ou même pour plusieurs types de sols.

Si l'on poursuit l'analyse des relations qui peuvent exister entre les propriétés géotechniques d'un sol, on est conduit à distinguer trois types de relations :

- les **relations mathématiques exactes**, qui existent par exemple entre les paramètres décrivant l'état du sol. On peut illustrer ce type de relations par toutes les formules mathématiques reliant

- L'indice des vides e et la porosité n :

$$e = n / (1 - n)$$

- La teneur en eau w , le poids volumique du sol γ et le poids volumique du sol sec γ_d :

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

- La teneur en eau w , l'indice des vides e , le degré de saturation S_r et les poids volumiques de l'eau γ_w et des grains γ_s :

$$w = e \gamma_w S_r / \gamma_s$$

- Les poids volumiques γ , γ_d , γ_s et γ_w d'un sol saturé :

$$\gamma = (\gamma_s \gamma_w + \gamma_s \gamma_d - \gamma_d \gamma_w) / \gamma_s$$

- etc. ;

- Les **lois d'évolution en fonction de la profondeur**, dues à l'effet de la pesanteur et dont l'origine est liée à l'augmentation des contraintes quand on s'enfonce dans le sol. Par exemple, dans les dépôts homogènes de sols fins dont l'état s'est stabilisé, les

contraintes effectives, pressions de préconsolidation, modules et résistances augmentent avec la profondeur ;

- Les **relations empiriques (ou corrélations)** entre propriétés d'un même volume élémentaire de sol, par exemple la porosité et le coefficient de perméabilité, l'indice de densité d'un sable et son angle de frottement interne, la pression de préconsolidation et la cohésion non drainée d'une argile, etc. Ces relations, qu'il n'est possible de caractériser que de façon statistique, s'expliquent par la raison déjà citée que toutes les propriétés d'un même empilement de particules évoluent de façon coordonnée et traduisent l'existence d'une loi de comportement générale pour chaque grande classe de sol.

V.2.2 Domaines d'utilisation des corrélations

Dans la pratique de la mécanique des sols, les corrélations entre paramètres sont utilisées comme moyen de contrôle des résultats des essais en place et en laboratoire, et comme moyen de fabrication de valeurs complémentaires de certains paramètres en fonction des autres.

Par exemple, sur un site donné, on peut analyser la relation entre deux paramètres mesurés sur une même carotte de sol (indice des vides e et indice de compression C_c , etc.) ou mesurés en place dans le même essai (module pressiométrique E_M et pression limite pressiométrique, etc.) et détecter les variations de la nature ou de l'histoire des sols d'après les modifications de leurs relations. Dans un tel cas, les corrélations servent d'outil de contrôle de l'homogénéité des sols (ou de la qualité des essais, si l'on sait de façon certaine que le sol est le même que celui qui a servi à établir la corrélation).

On utilise aussi les corrélations pour estimer certaines propriétés des sols (souvent, des propriétés mécaniques) en fonction des caractéristiques qui ont été mesurées (souvent, des propriétés physiques, comme la densité ou la teneur en eau). On peut ainsi, lors des études préliminaires et dans certaines situations de projets, disposer de valeurs des paramètres nécessaires au dimensionnement des ouvrages sans les avoir déterminées par des essais.

Les conditions d'utilisation de corrélations dans les études géotechniques dépendent de la fiabilité des corrélations utilisées.

Certains paramètres sont liés, à l'intérieur d'une couche de sol d'un site déterminé, par des relations proches d'une relation mathématique exacte. Par contre, si l'on analyse simultanément des données provenant de deux sites, pour des sols de même nature, on trouve en général que les valeurs des paramètres sont plus dispersées, et cette dispersion augmente quand le nombre de sites s'accroît et quand on regroupe des données relatives à différents types de sols. Les erreurs expérimentales, lors de la détermination des paramètres qui servent à établir les corrélations, exercent également une influence défavorable sur la qualité des corrélations obtenues. Il est, pour cette raison, indispensable de connaître l'origine des corrélations que l'on envisage d'utiliser dans le cadre d'une étude géotechnique, et d'être conscient de la variabilité possible des paramètres autour de leur relation moyenne affichée, notamment quand les corrélations ont été établies entre des fonctions logarithmiques des paramètres.

V.3 PRINCIPALES TECHNIQUES D'ETUDE DES CORRELATIONS

L'étude des relations existant entre les propriétés des sols s'effectue au moyen des outils classiques de la statistique pour l'analyse des données. Les méthodes classiques de l'analyse statistique ont été exposées dans de nombreux ouvrages auxquels le lecteur pourra se reporter pour une description détaillée de ces méthodes. Dans le présent paragraphe, seront rappelés seulement les définitions essentielles et les principes des méthodes couramment utilisées pour les études de corrélations en mécanique des sols.

V.3.1 Définitions et caractéristiques des variables aléatoires

Pour l'application des techniques de l'analyse statistique, chaque paramètre géotechnique du sol doit être considéré comme une variable aléatoire, c'est-à-dire comme une grandeur non déterminée *a priori*, dont on sait qu'elle peut prendre telle ou telle valeur dans un ensemble de valeurs possibles, avec une certaine probabilité.

Cette assimilation des propriétés du sol à des variables aléatoires n'implique pas qu'en un point donné les propriétés du sol ne soient pas parfaitement déterminées. Elle représente seulement l'ignorance de l'ingénieur vis-à-vis des valeurs exactes de chaque propriété en chaque point.

Toute variable aléatoire X peut être caractérisée par une densité de probabilité $g(x)$, qui représente la probabilité de chaque valeur possible x de la variable, ou, de façon parfaitement équivalente, par une fonction de répartition $G(x)$, variant de 0 à 1 et égale à la probabilité que X soit inférieur à x .

Connaissant la fonction $g(x)$ ou $G(x)$, on peut calculer les moments de la variable aléatoire, qui sont, d'une part, les moments d'ordre r :

$$m_r = \int_{-\infty}^{+\infty} x^r g(x) dx$$

et, d'autre part, les moments centrés d'ordre r :

$$\mu_r = \int_{-\infty}^{+\infty} (x - m)^r g(x) dx$$

Le moment d'ordre 1 est appelé **espérance mathématique**, ou **moyenne**, et noté $E[X]$ ou m . Le moment centré d'ordre 1 est nul.

Le moment centré d'ordre 2 est appelé **variance** et noté $\text{Var}[X]$ ou σ^2 . Sa racine carrée positive est appelée **écart type** et notée σ . Le rapport de l'écart type à la moyenne est appelé **coefficient de variation** et noté $CV[X]$ ou Cx .

Les notions précédentes sont définies pour des fonctions mathématiques appelées variables aléatoires. Dans la pratique, quand on analyse un ensemble de données, on ne connaît généralement pas les lois de probabilité des propriétés étudiées. On raisonne alors sur des valeurs estimées des paramètres statistiques (estimées d'après l'ensemble des données dont on dispose). Différents ensembles de données (différents «échantillons», dans le vocabulaire des statistiques) conduisent à des estimations différentes de ces paramètres, si bien que ces paramètres estimés peuvent eux-mêmes être traités comme des variables aléatoires...

V.3.2 Relations entre variables aléatoires. Régression linéaire

Pour analyser simultanément les valeurs de plusieurs propriétés d'un même sol, on fait en général l'hypothèse que les relations cherchées sont linéaires. Cette hypothèse n'exclut pas l'existence de relations de type non linéaire entre les propriétés du sol : les variables aléatoires liées par des relations linéaires peuvent être des fonctions non linéaires des propriétés du sol (logarithmes, fonctions puissances, exponentielles, etc.), ce qui donne une grande souplesse à ce type d'analyse linéaire.

Pour décrire la simultanéité des variations de deux variables aléatoires X et Y , on utilise une fonction voisine de la variance, appelée **covariance**, notée $\text{Cov}[X, Y]$ ou σ_{XY} et définie comme suit :

$$\begin{aligned}\text{Cov}[X, Y] &= \sigma_{XY} = E\{(X - E[X])(Y - E[Y])\} \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - m_X)(y - m_Y) g(x, y) dx dy\end{aligned}$$

où $g(x, y)$ est la densité de probabilité de (X, Y) . La variance de X est égale à $\text{Cov}[X, X]$.

Le **coefficient de corrélation linéaire** ρ_{XY} représente sous forme adimensionnelle cette même variabilité :

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

en notant σ_X et σ_Y , respectivement, les écarts types de X et Y .

Dans le cas de deux variables aléatoires X et Y , la procédure de recherche de la meilleure relation linéaire entre ces variables commence par le choix de la **variable explicative**, qui sera notée X , et de la **variable expliquée**, qui sera notée Y :

$$Y = aX + b$$

Ce choix préliminaire inévitable introduit une dissymétrie entre X et Y et l'on n'obtient pas le même résultat en écrivant $Y = aX + b$ et $X = cY + d$, bien que le coefficient de corrélation soit le même dans les deux cas. Cette différence vient de la procédure utilisée pour estimer les valeurs des coefficients a et b (respectivement c et d).

Si l'on dispose d'un ensemble (échantillon) de n couples de valeurs $(x_i, y_i)_{i=1, n}$ de X et Y pour déterminer la relation entre ces deux variables, on recherche ensuite les **valeurs estimées** de a et b , notées ici $\underline{a}$ et $\underline{b}$, qui minimisent l'écart quadratique moyen entre les y_i et les expressions calculées $(\underline{a}x_i + \underline{b})$ [méthode des moindres carrés], c'est-à-dire :

$$\delta^2 = \sum_{i=1}^n [y_i - (\underline{a}x_i + \underline{b})]^2$$

Tous calculs faits, on obtient :

$$\underline{a} = m_Y - \underline{b} m_X$$

Avec :

$$[V_{XX}] = [\text{Cov}(X_i, X_j)] = \frac{1}{n} \sum_{i,j} [X_i - E(X_i)][X_j - E(X_j)]$$

$$[V_{XY}] = [\text{Cov}(X_j, Y)] = \frac{1}{n} \sum_{j} [X_j - E(X_j)][Y - E(Y)]$$

$$\sigma_Y = \text{écart type de } Y$$

traduit l'écart relatif moyen entre les y_i et les $(\underline{a}_0 + \sum \underline{a}_j x_j)$ de l'ensemble des données $(y_i, x_{1i}, \dots, x_{ki})$ disponibles.

Le coefficient de corrélation peut varier entre -1 et $+1$. Les valeurs proches de zéro indiquent une forte dispersion des valeurs de Y par rapport à la relation linéaire estimée, donc une mauvaise représentativité de l'équation :

$$Y = \underline{a}_0 + \sum_{j=1}^k \underline{a}_j X_j$$

Néanmoins, cela n'exclut pas l'existence d'une meilleure relation non linéaire entre les variables [On cite souvent l'exemple de points (x_i, y_i) répartis sur un cercle et dont le coefficient de corrélation (par rapport à une relation linéaire) est nul]. Il est pour cette raison toujours recommandé de représenter, dans la mesure du possible, les données à analyser.

Dans le cas de deux variables X et Y , le coefficient de corrélation ρ , comme indiqué précédemment, pour expression :

$$\rho = \frac{\text{Cov}[X, Y]}{\sigma_X \sigma_Y} = \frac{\underline{b}}{\sigma_Y}$$

Une fois établie la relation entre la variable expliquée Y et les variables explicatives $\{X_j\}$, on peut :

- déterminer l'erreur moyenne sur l'équation obtenue, caractérisée par la **variance de la régression** :

$$\sigma_d^2 = \text{Var}[Y - (\underline{a}X + \underline{b})] = \frac{1}{n-2} \sum_{j=1}^n [y_j - (\underline{a}x_j + \underline{b})]^2$$

- tester la représentativité des valeurs estimées des coefficients $\underline{a}$ et $\underline{b}$ pour décider s'ils sont significativement différents de zéro. Pour ce type d'analyse, le lecteur pourra se reporter à l'un des ouvrages cités en références bibliographiques ;

- estimer la valeur de Y correspondant à une valeur déterminée x_0 de X , ainsi que la variance correspondante. Ces deux valeurs sont données par les formules suivantes :

$$E[Y(x_0)] = \underline{a} + \underline{b}x_0$$

$$\sigma_{E[Y(x_0)]}^2 = \frac{\sigma_{\varepsilon}^2}{n} \left[1 + \frac{(x_0 - m_x)^2}{\sigma_x^2} \right]$$

On observe sur la figure V. l'effet du second terme de l'expression de la variance : l'incertitude (l'écart type) sur la valeur estimée de $Y(x_0)$ est minimale lorsque x_0 est égal à la valeur moyenne m_x des x_i de l'ensemble des données ayant servi à estimer les paramètres $\underline{a}$ et $\underline{b}$ de la régression linéaire. Elle augmente progressivement, selon une loi parabolique, lorsque x_0 s'éloigne de cette valeur moyenne m_x . Ce résultat est important pour les applications pratiques, où l'on ne peut donc estimer avec une égale fiabilité les valeurs de Y correspondant aux différentes valeurs possibles de X .

V.3.3 Analyse factorielle

En pratique, on s'intéresse souvent aux relations qui peuvent exister à l'intérieur d'un groupe de m variables, et l'utilisation des techniques de régression linéaire conduit à répéter l'analyse décrite au paragraphe précédent en donnant tour à tour à chacune des variables le rôle de variable expliquée et en étudiant l'ensemble des relations qui la lient aux autres, prises isolément, puis par deux, par trois, etc. Pour limiter le nombre des opérations nécessaires, différentes procédures ont été développées. Par exemple, la méthode de régression « pas à pas » ne teste qu'une partie des combinaisons possibles des variables en recherchant la variable X_j la mieux corrélée avec Y , soit X_a , puis la variable qui maximise le coefficient de corrélation multiple de Y avec X_a et une seconde variable X_j , etc.

Mais cette méthode ne garantit pas que l'on n'oublie pas une combinaison éventuellement plus favorable, mais dont aucune variable n'est la plus corrélée avec Y .

L'analyse factorielle, qui recherche les « facteurs » (combinaisons linéaires des variables) représentant le mieux les variations des données analysées, constitue une alternative efficace aux méthodes précédentes. Son principe est de construire un ensemble de nouvelles variables indépendantes en procédant pas à pas et en retenant à chaque étape, parmi les facteurs possibles, celui qui fait diminuer le plus la variance résiduelle. Les applications de l'analyse factorielle en géotechnique sont encore assez rares et ce thème ne sera pas développé dans le présent article, mais l'analyse factorielle

offre des possibilités intéressantes pour guider les études sur le comportement des sols et des roches.

V.3.4 Variabilité spatiale

Il est bien établi que les couches de sols naturels sont rarement homogènes et que leurs propriétés physiques et mécaniques fluctuent avec des amplitudes variables selon les propriétés considérées, la nature et l'origine des sols. Le coefficient de variation donne une mesure utile de cette variabilité. Ainsi, la teneur en eau a souvent un coefficient de variation de l'ordre de 20 %, le poids volumique de 5 %, les paramètres de résistance au cisaillement de 30 %, avec des valeurs plus fortes pour la cohésion non drainée (souvent 50 %). Dans certains sols, les variations sont très rapides et l'on peut considérer, par exemple, qu'à 50 cm de distance les propriétés du sol n'ont pas de lien. Dans d'autres cas, les valeurs d'une même propriété restent voisines sur quelques mètres, voire quelques dizaines de mètres.

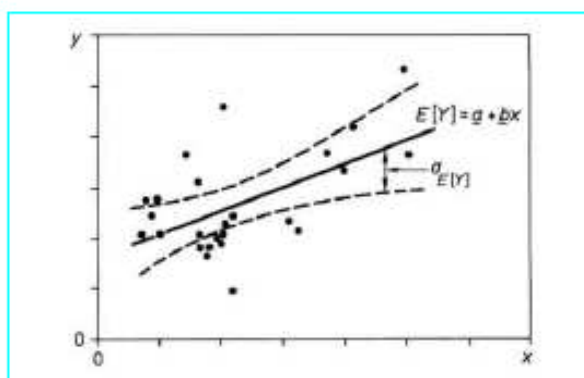


Figure V.1 Estimation de la valeur la plus probable de y pour une valeur donnée de x , sur la base de l'analyse de régression de X et Y

Ces variations spatiales des propriétés des sols exercent une influence sur les résultats des études de corrélations. Cette influence se traduit par :

- la plus faible corrélation des propriétés mesurées en des points éloignés qu'en des points voisins (beaucoup de corrélations sont malheureusement établies avec des données provenant de sondages ou essais assez distants les uns des autres, de telle sorte qu'elles incluent non seulement la corrélation réelle des paramètres en un même point, mais aussi une certaine partie de leur variabilité spatiale. La seule solution pour éviter ce

phénomène est de faire des campagnes d'essais spéciales comportant des essais ou sondages très voisins) ;

- la diminution de la variabilité des paramètres du sol lorsque le volume du sol concerné par l'essai augmente. Ce phénomène peut influencer les corrélations établies, par exemple, entre des propriétés mesurées sur de très petits volumes de sol (teneur en eau, coefficient de perméabilité d'éprouvettes de laboratoire, compressibilité ou résistance au cisaillement mesurée en laboratoire, etc.) et des propriétés mesurées sur de plus grands volumes de sols (pression limite ou module pressiométrique, perméabilités mesurées en place, essais de plaque, etc.).

V.4 EXEMPLES DE CORRELATIONS

De très nombreuses corrélations ont été publiées pour les propriétés des sols. Beaucoup d'entre elles n'existent que sous la forme d'une relation entre paramètres, sans accès possible aux données étudiées ni même d'indication du coefficient de corrélation correspondant, et il convient d'être prudent quand on les utilise.

Nous nous limiterons ici à quelques exemples de corrélations entre les paramètres des sols déterminés en place et en laboratoire pour lesquels les données expérimentales seront présentées en même temps que les fonctions de régression entre les paramètres.

V.4.1 Relation entre la compressibilité et la teneur en eau des tourbes

La durée importante des essais oedométriques conduit à utiliser, chaque fois que c'est possible, des corrélations avec des paramètres de détermination plus rapide, comme la teneur en eau, pour compléter la caractérisation des sols compressibles sur les sites de projets de grande ampleur. Il existe, pour cette raison, de nombreuses corrélations entre ces paramètres. L'exemple présenté sur la figure V.2 concerne différentes tourbes de Normandie [9]. Les figures V.2a et V.2b montrent les relations observées sur deux sites, relations assez marquées et pratiquement linéaires, mais nettement différenciées d'un site à l'autre. Les figures V.2c et V.2d rassemblent tous les points disponibles sur les sites de tourbes de la région : on observe que les relations linéaires prévalant sur chaque site disparaissent au profit d'un nuage de points dont la meilleure approximation n'est pas linéaire (figure V.2c), mais exponentielle (figure V.2d), avec une corrélation nettement

moins forte. En pratique, l'utilisation de corrélations est donc très recommandable au niveau d'un site, mais l'est moins si l'on passe d'un site à un autre.

V.4.2 Relation entre la perméabilité et l'indice des vides des argiles

Il est généralement admis que le coefficient de perméabilité k des argiles est lié à l'indice des vides e par une relation de la forme

$$\Delta e = C_k \Delta(\lg k)$$

Le coefficient C_k de cette relation est lui-même lié à l'indice des vides initial du sol e_0 , comme on le voit sur les figures V.3a, V.b et V.c.

On observe, dans ce cas, que les différences entre les droites de régression linéaire sont peu importantes, avec des coefficients de corrélation élevés dans chaque cas. Une telle corrélation peut donc, à la différence des précédentes, être utilisée sur des sites autres que ceux où elle a été établie.

V.4.3 Relation entre la limite de liquidité et l'indice de compression des vases

La figure V.4 présente les droites de régression obtenues par Vidalie entre la limite de liquidité w_L et l'indice de compression C_c de sols fins organiques (vases) d'origines variées. La droite d'équation $C_c = 0,009 (w_L - 10)$ est celle donnée par Terzaghi pour présenter le comportement moyen des argiles. Cette fois aussi, le coefficient de corrélation est élevé et la relation obtenue peut être considérée comme assez fiable.

V.4.4 Relation entre la pression limite pressiométrique et la résistance de cône au pénétromètre statique

Les figures V.5 et V.6 montrent les relations existant entre la pression limite nette pressiométrique et la résistance de cône déterminée au pénétromètre statique q_c , pour deux ensembles de données publiés par Cassan. Pour les sables de Dunkerque (figure V.5), la relation entre les deux paramètres, bien que différente suivant le sens dans lequel on la recherche, est associée à un coefficient de corrélation assez élevé. Pour les argiles (figure V.6), de provenances diverses, les points sont beaucoup plus dispersés dans le graphique et la qualité de la corrélation est plus faible.

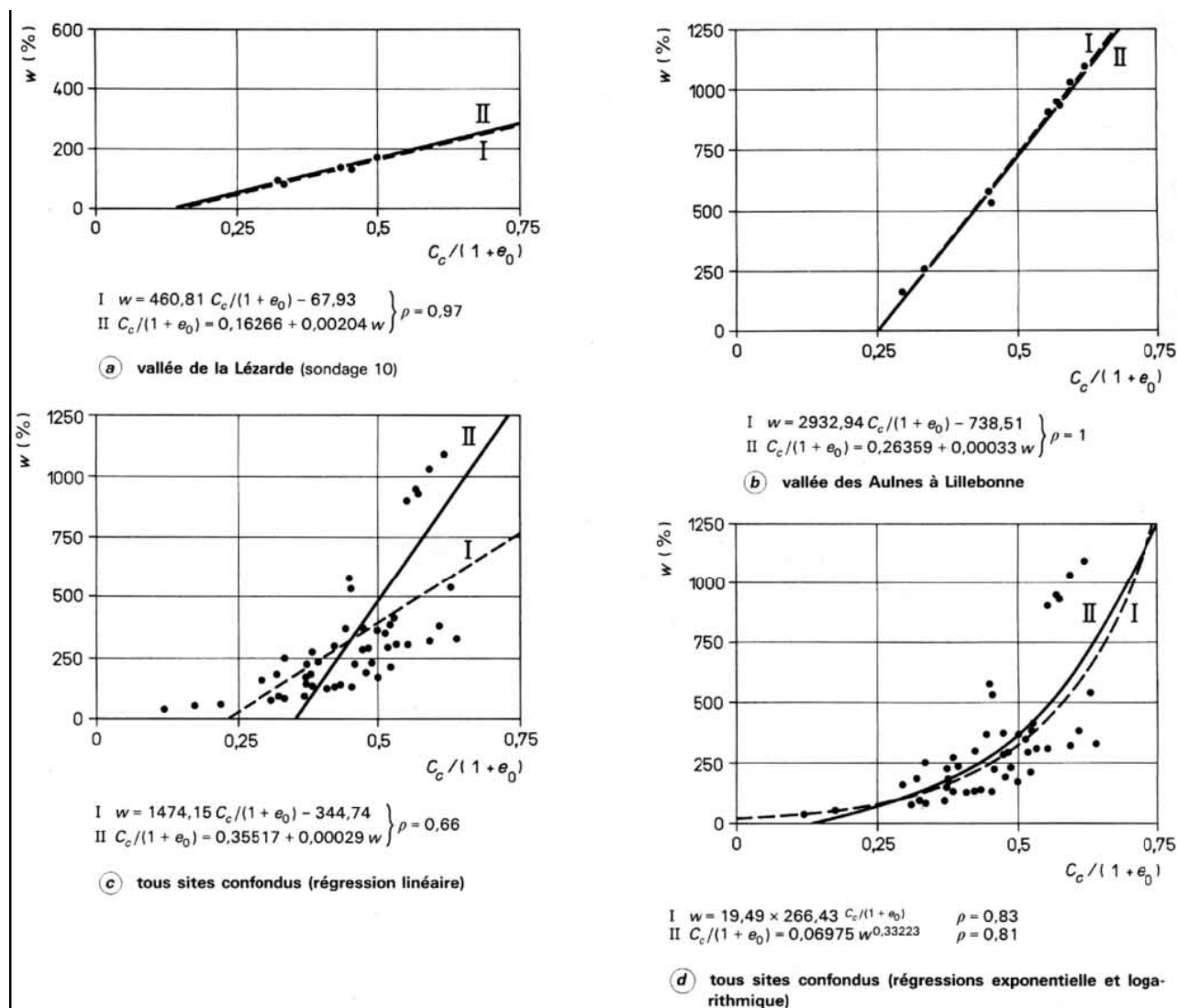


Figure V.2 Relations entre la teneur en eau w et le coefficient de compressibilité $C_c/(1+e_0)$ pour les tourbes de Normandie

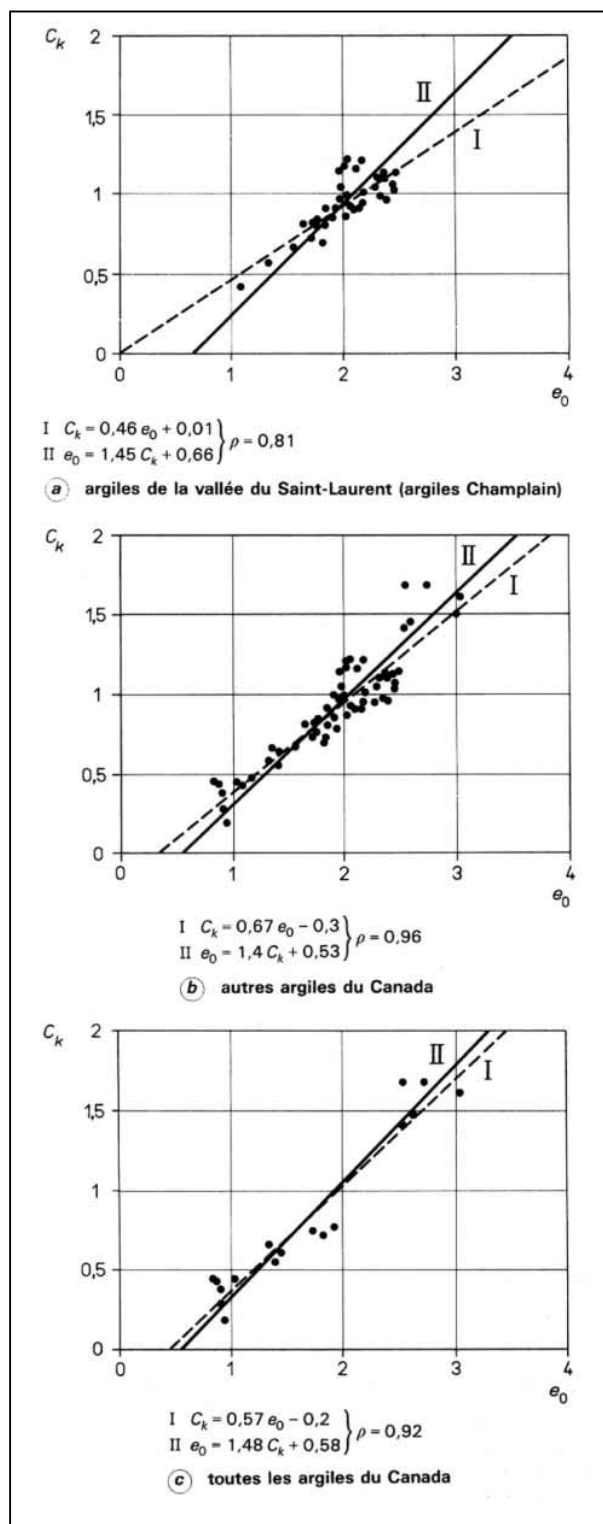


Figure V.3 Relations entre le taux de variation de la perméabilité C_k et l'indice des vides initial e_0 des argiles du Canada

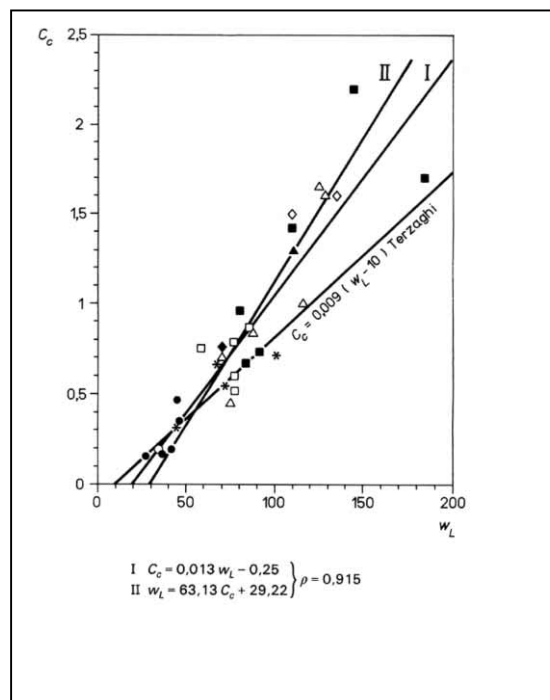


Figure V.4 Relations entre la limite de liquidité w_L et l'indice de compression C_c

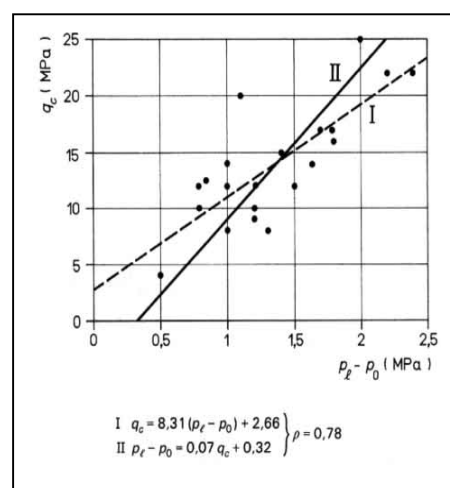


Figure V.5 Relations entre la pression limite nette pressiométrique et la résistance de cône statique q_c pour les sables de Dunkerque

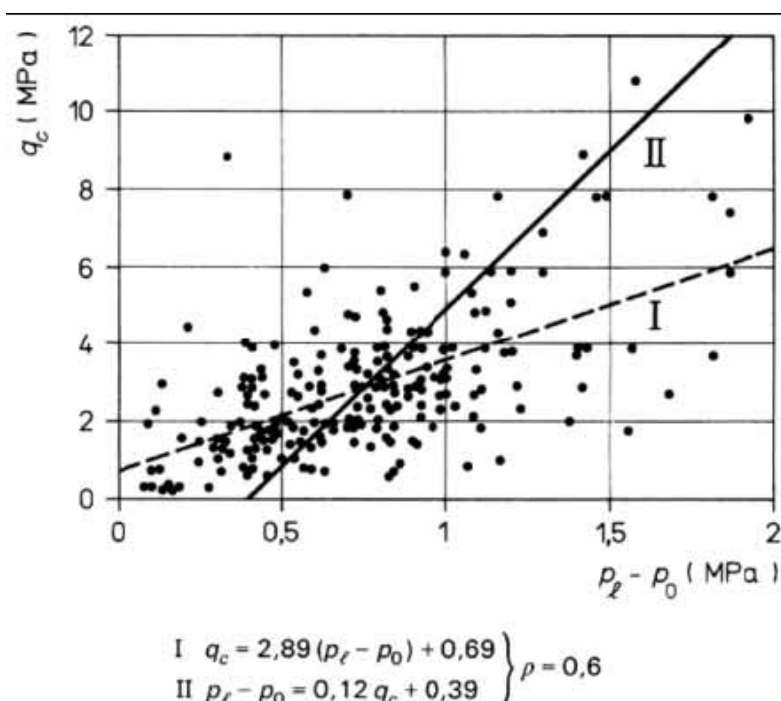


Figure V.6 Relations entre la pression limite nette pressiométrique $p_\ell - p_0$ et la résistance de cône statique q_c pour des argiles de provenances diverses

V.4.5 Relation entre les résistances de cône statique et dynamique

Cassan a publié des résultats d'essais comparatifs sur les résistances de pointe déterminées au pénétromètre statique (q_c) et au pénétromètre dynamique (q_d) dans des sables argileux à Châlon-sur-Saône, au-dessus du niveau de la nappe. Les droites de régression obtenues sur ces données (figure V.7a) correspondent à une forte valeur du coefficient de corrélation sur ce site.

Toutefois, les valeurs mesurées au-dessous du toit de la nappe, à des profondeurs où le sol est saturé (figure V.7b), correspondent à une relation différente entre les deux paramètres, ce qui illustre les limites du domaine de validité des corrélations dans ce cas.

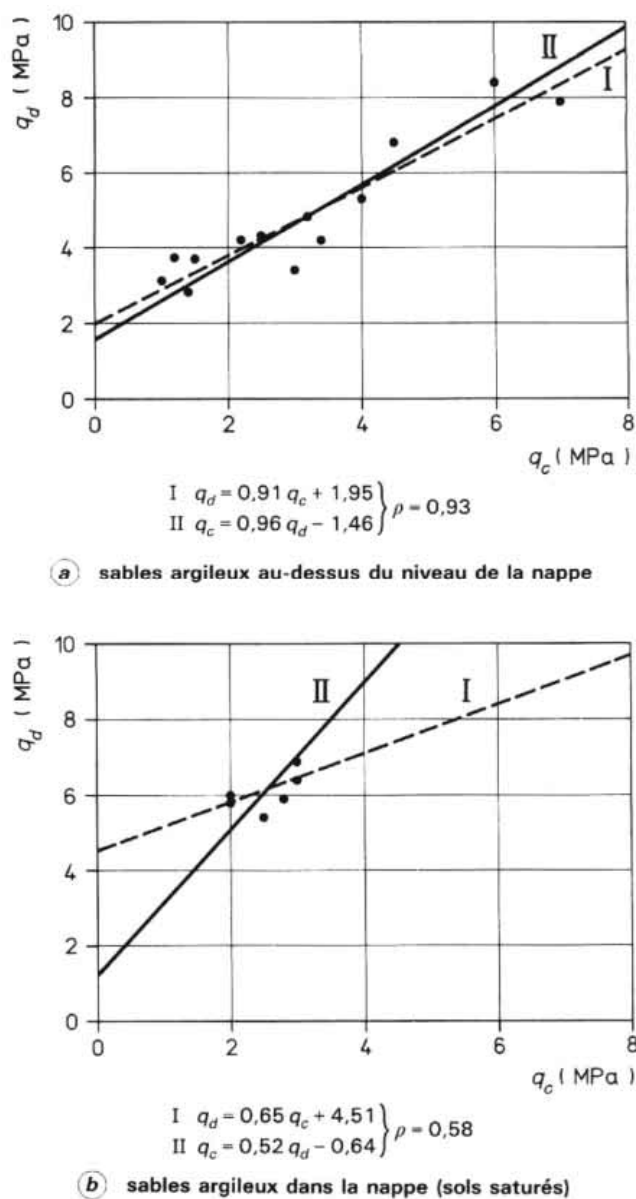


Figure V.7 Relation entre les résistances de cône statique q_c et dynamique q_d pour des sables argileux

V.5 DOMAINES DE VALIDITE DES CORRELATIONS

Les corrélations que l'on peut établir entre les paramètres physiques et mécaniques des sols sont plus ou moins générales suivant les paramètres concernés. Habituellement, la validité d'une corrélation est limitée à la nature du sol étudié : les propriétés des sables, des tourbes ou des argiles n'obéissent pas aux mêmes lois, elles sont d'ailleurs

souvent décrites par des paramètres spécifiques et il n'est pas étonnant que les corrélations établies pour un type de sol ne soient pas valables pour les mêmes propriétés d'un autre type de sol. Les figures V.5 et V.6 illustrent ce fait dans le cas des paramètres mesurés au pressiomètre et au pénétromètre statique.

Certaines corrélations établies sur un site et parfaites pour ce site (par exemple les corrélations des figures V.2a et V.b), peuvent aussi être totalement inadaptées sur un autre site, même constitué d'un sol de même nature. Cette divergence traduit habituellement l'influence d'autres paramètres que ceux qui sont analysés, par exemple l'influence de l'état du sol, en plus de sa nature. Si la relation obtenue par régression linéaire entre deux paramètres dépend d'autres facteurs, elle peut varier non seulement d'un site à l'autre, mais aussi à l'intérieur d'un même site. Les figures V.8a, V.8b et V.8c illustrent une telle variation dans le cas du site expérimental de Cubzac-les-Ponts, où les corrélations ont été étudiées dans plusieurs sous-ensembles du site, notés « remblais A à C » et « hors remblai HR » sur les figures.

Si un tel phénomène est observé sur le site d'un grand projet, il est indispensable de poursuivre l'analyse géotechnique du site afin d'éviter des erreurs d'analyse statistique des données.

Sous réserve d'une certaine prudence quant à la généralisation des corrélations établies sur un site au reste du site ou à d'autres sites, l'utilisation de corrélations constitue une technique très utile pour le progrès des études géotechniques de terrain et son usage peut être recommandé, tant pour compléter des données que pour contrôler la vraisemblance des résultats des essais réalisés en place comme en laboratoire : les données conformes aux corrélations, tant générales qu'établies sur le site, sont en effet plus plausibles que celles qui en sont trop éloignées et pour lesquelles des vérifications complémentaires sont toujours souhaitables.

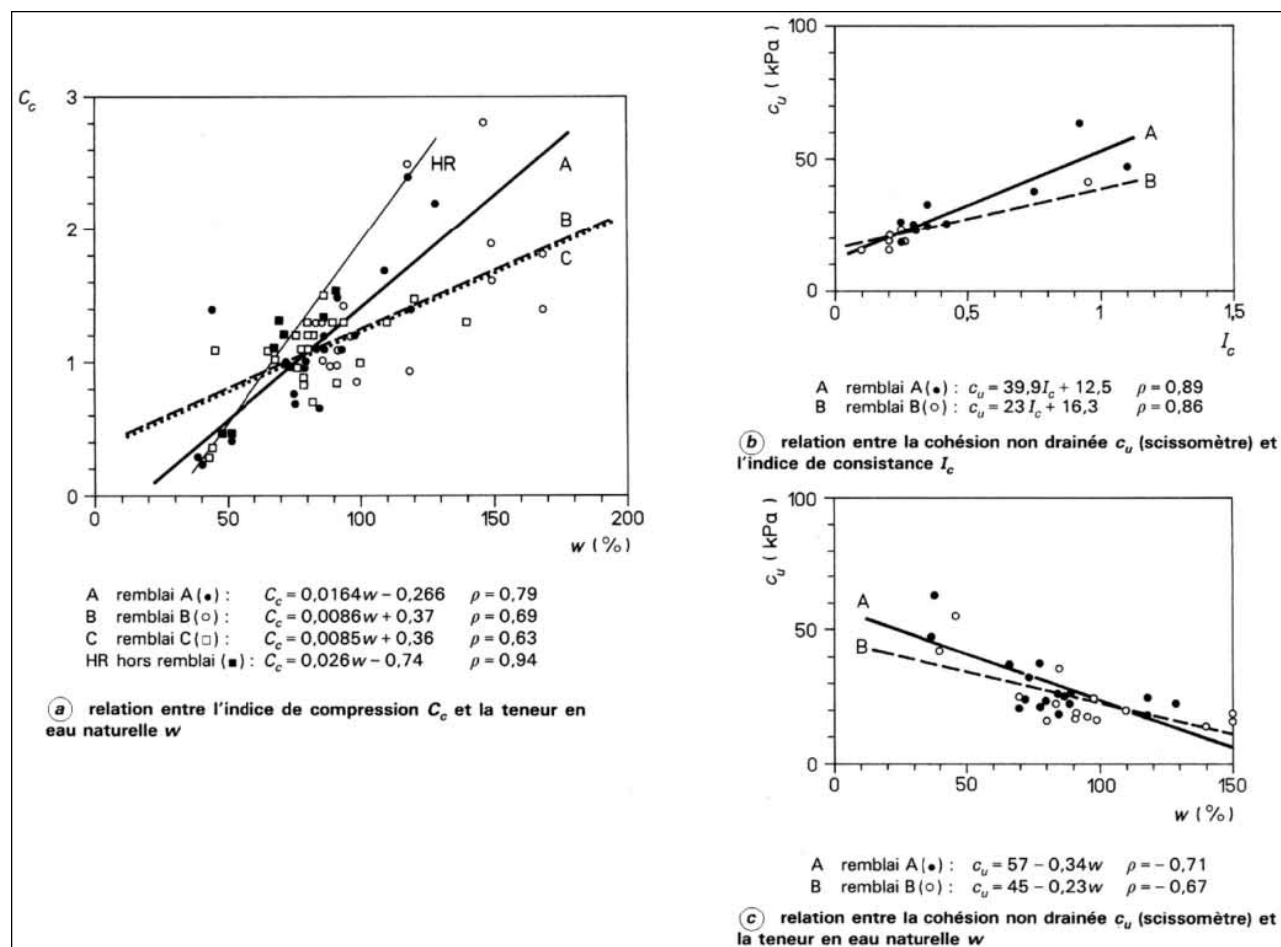


Figure V.8 Limitations des corrélations : variabilité à l'intérieur d'un site [données du site expérimental des Laboratoires des Ponts et Chaussées à Cubzac-les-Ponts]

Chapitre VI

*METHODES PRATIQUES DE CALCUL
PROBABILISTE*

Chapitre VI

METHODES PRATIQUES DE CALCUL PROBABILISTE

VI.1 INTRODUCTION

L'élaboration des méthodes de calcul probabilistes est encore loin de son terme, malgré le nombre élevé des publications consacrées à ce sujet.

Les calculs mécaniques des sols nécessaires au dimensionnement des ouvrages (ouvrages de soutènement, ouvrages en terre, fondations, etc.) ont tous pour objectif de déterminer la valeur de paramètres (coefficient de sécurité, temps de consolidation, amplitude des tassement, etc.) qui dépendent des propriétés physiques et mécaniques des sols, de la géométrie des problèmes et des conditions initiales et aux limites imposées.

Si l'on traite les variations dans l'espace des propriétés des sols, les variations dans le temps des conditions aux limites et les fluctuations de la géométrie des sols et des ouvrages comme des phénomènes aléatoires dont chacun peut être représenté par une variable aléatoire X_i , les résultats des calculs nécessaires au dimensionnement des ouvrages sont eux-mêmes des variables aléatoires Y , fonctions des variables aléatoires X_i .

Chacune des variables aléatoires Y a sa propre fonction de répartition $G(y)$ et sa densité de probabilité $g(y)$, qu'il s'agit d'évaluer.

Le calcul probabiliste des ouvrages de mécanique des sols est donc « simplement » un problème de calcul de fonctions de variables aléatoires : $Y = f(X_i)$.

Pour deux variables indépendantes X_1 et X_2 de densités de probabilité g_1 et $g_2(x)$, la forme exacte de la densité de probabilité de la somme $Y = X_1 + X_2$, produit $Z = X_1 \cdot X_2$ ou du quotient $W = X_1/X_2$ est la suivante (Lumb, 1974) :

$$\begin{aligned}
 g(y) &= \int_{-\infty}^{+\infty} g_1(y - \alpha) \cdot g_2(\alpha) d\alpha \\
 g(y) &= \int_{-\infty}^{+\infty} g_1(z/\alpha) \cdot |\alpha|^{-1} g_2(\alpha) d\alpha \\
 g(y) &= \int_{-\infty}^{+\infty} g_1(w \cdot \alpha) \cdot |\alpha|^{-1} g_2(\alpha) d\alpha
 \end{aligned}$$

Et la combinaison de ces équations permet en principe de calculer la densité de probabilité de toute fonction comportant uniquement des sommes, des produits et des

quotients.. Mais cette méthode peut être très laborieuse en pratique, notamment parce qu'elle nécessite l'évaluation d'intégrales parfois complexes.

Par ailleurs, il n'existe pas toujours de relation explicite entre les paramètres physiques, mécaniques et géométriques des ouvrages et les résultats qu'il s'agit d'évaluer.

Enfin, la plupart des méthodes de calcul utilisées pour le dimensionnement des ouvrages ne sont qu'approchées, ce qui introduit à la fois un biais et une incertitude supplémentaire qu'il faut incorporer dans l'analyse de la fonction $Y = f(X_i)$.

Différentes méthodes peuvent être utilisées pour déterminer de façon approchée la densité de probabilité $g(Y)$ à partir des lois de distribution des paramètres X_i du modèle de calcul. Nous décrirons ci-après quatre d'entre elles :

- Approximation par les séries de Taylor
- Approximation par intégration numérique
- Approximation par une loi normale ou lognormale
- Simulation par la méthode de Monte-Carlo.

VI.2 APPROXIMATION PAR LES SERIES DE TAYLOR

Si la fonction $y = f(x_i)$ est différentiable par rapport aux x_i , analytiquement ou numériquement, elle peut être développée en séries de Taylor autour de la valeur moyenne $f(m_1, m_2, \dots, m_n)$, m_i désignant l'espérance mathématique de la variable aléatoire X_i . En limitant le développement au premier ou second ordre, ou plus, on obtient une expression approchée de y , qui permet de déterminer plus simplement les valeurs des paramètres de la densité de probabilité de Y .

Ainsi (Benjamin et Cornell, 1970) l'espérance mathématique a pour expression approchée au second ordre :

$$E[Y] = f(m_1, m_2, \dots, m_n) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f(x_1, \dots, x_n)}{\partial x_i \partial x_j} \bigg|_{(m_1, m_2, \dots, m_n)} \sigma_{X_i X_j} \quad (VI.4)$$

($\sigma_{X_i X_j}$ désigne la covariance des deux variables aléatoires X_i et X_j) ; la variance a pour expression approchée du premier ordre :

$$V[Y] \approx \sum_{i=1}^n \sum_{j=1}^n \frac{\partial f(x_1, \dots, x_n)}{\partial x_i} \bigg|_{(m_1, m_2, \dots, m_n)} \cdot \frac{\partial f(x_1, \dots, x_n)}{\partial x_j} \bigg|_{(m_1, m_2, \dots, m_n)} \cdot \sigma_{X_i X_j} \quad (VI.5)$$

$(m_1, m_2, \dots, m_n) \times \sigma_{X_i X_j}$

Dans le cas où les variables aléatoires X_i sont indépendantes, on a $\sigma_{X_i X_j} = 0$ pour $i \neq j$ et

$$\sigma_{X_i X_j} = \sigma_{X_i}^2 = V(X_i)$$

D'où

$$V(y) \approx \sum_{i=1}^n \left[\left. \frac{\partial f(x_1, \dots, x_n)}{\partial x_i} \right|_{(m_1, m_2, \dots, m_n)} \right]^2 \cdot V[X_i] \quad (\text{VI.6})$$

Les moments d'ordre 3 et 4 sont donnés par les expressions suivantes (Lumb, 1974) :

$$\begin{aligned} \mu_3[Y] &= \sum_{i=1}^n \left[\left. \frac{\partial f(x_1, \dots, x_n)}{\partial x_i} \right|_{(m_1, m_2, \dots, m_n)} \right]^3 \mu_3[X_i] \\ \mu_4[Y] &= \sum_{i=1}^n \left[\left. \frac{\partial f(x_1, \dots, x_n)}{\partial x_i} \right|_{(m_1, m_2, \dots, m_n)} \right]^4 \mu_4[X_i] + \\ &+ 6 \sum_{i=1}^n \sum_{j < i} \left[\left. \frac{\partial f(x_1, \dots, x_n)}{\partial x_i} \right|_{(m_1, m_2, \dots, m_n)} \right]^2 \times \left[\left. \frac{\partial f(x_1, \dots, x_n)}{\partial x_j} \right|_{(m_1, m_2, \dots, m_n)} \right]^2 (\sigma_{X_i}^2 \cdot \sigma_{X_j}^2 + 2\sigma_{X_i X_j}^2) \end{aligned} \quad (\text{VI.7})$$

La connaissance des moments d'ordre 1, 2, 3 et 4 ($E[Y]$, $V[Y]$, $\mu_3[Y]$, $\mu_4[Y]$) permet d'estimer de probabilité de Y dans le système de Pearson.

VI.3 APPROXIMATION PAR INTEGRATION NUMERIQUE

Lorsque l'expression de $Y = f(X_i)$ n'est pas connue sous forme analytique, mais peut être déterminée sous forme numérique ou expérimentale, le calcul des dérivées partielles précédentes est pratiquement impossible mais les moments de Y peuvent cependant être déterminés de façon précise par intégration numérique (Lumb, 1974).

La méthode repose sur l'utilisation de la formule suivante, qui donne l'espérance mathématique $E[H]$ de la fonction $H = h(X_i)$:

$$\begin{aligned} E[H] &= C \cdot h(\bar{x}) + \sum_i N_i \left[\frac{h(a_i^+ \cdot \sigma_i)}{a_i^+} - \frac{h(a_i^- \cdot \sigma_i)}{a_i^-} \right] + \\ &+ \sum_{ij} P_{ij} \left[\frac{h(a_i^+ \cdot \sigma_i; a_j^+ \cdot \sigma_j)}{a_i^+ \cdot a_j^+} - \frac{h(a_i^- \cdot \sigma_i; a_j^- \cdot \sigma_j)}{a_i^- \cdot a_j^-} \right] \\ &- \frac{h(a_i^- \cdot \sigma_i; a_j^+ \cdot \sigma_j)}{a_i^- \cdot a_j^+} + \frac{h(a_i^+ \cdot \sigma_i; a_j^- \cdot \sigma_j)}{a_i^+ \cdot a_j^-}, \end{aligned}$$

Avec les notations suivantes :

- $C, N_i, P_{ij}, a_i^+, a_i^-$ sont de constantes ;
- $h(X)$ est la valeur de $h(x_i)$ au point $(m_1, m_2, \dots, m_n)$, les m_i désignant l'espérance mathématique des X_i ;
- $h(a_i^+ \cdot \sigma_i)$ est la valeur de $h(x_i)$ au point $[m_1, \dots, (m_i + \sigma_i \cdot a_i^+), \dots, m_n]$
- $h(a_i^+ \cdot \sigma_i; a_j^+ \cdot \sigma_j)$ est la valeur de $h(x_i)$ au point $[m_1, \dots, (m_i + a_i^+ \cdot \sigma_i), \dots, (m_j + a_j^+ \cdot \sigma_j), \dots, m_n]$

Les sommes sont calculées pour toutes les combinaisons possibles des indices. Il y a par conséquent n termes dans la première somme et $\frac{1}{2}n(n-1)$ termes dans la seconde somme.

Cette expression (VI.9) s'applique à toute fonction $h(x_i)$ et notamment à la fonction « espérance mathématique de $Y = f(x_i)$ » et aux moments d'ordre r de $Y = f(x_i)$. Dans le premier cas, on remplace simplement $h(x_i)$ par $f(x_i)$ dans l'équation (9) ; dans le second, on remplace

$$h(x_i) \text{ par } \{f(x_i) - E[f(x_i)]\}$$

Si l'on introduit, pour simplifier, les deux notations suivantes :

$$\delta_i = \sqrt{\beta_1[X_i]} \text{ (coefficient d'asymétrie de la densité de probabilité de la variable aléatoire } X_i),$$

$$\Delta_i = \beta_2[X_i] \text{ (coefficient d'aplatissement de la probabilité de la variable } X_i),$$

Les constantes de l'expression (VI.9) ont les valeurs suivantes :

$$\left. \begin{aligned} a_i^+ &= \frac{1}{2} \delta_i + R_i, \\ S + \frac{1}{2}(S)^2 - \frac{1}{2}T, \\ N_i &= (1 - S_i)/(2 R_i), \\ P_{ij} &= 1/(4 R_i R_j), \end{aligned} \right\}$$

avec :

(VI.10)

$$\left. \begin{aligned} R_i &= \sqrt{\Delta_i - \frac{3}{4}\delta_i^2} \\ S &= \sum_{i=1}^n (\Delta_i - \delta_i^2)^{-1} \\ S_i &= S - (\Delta_i - \delta_i^2)^{-1} \\ T &= \sum_{i=1}^n (\Delta_i - \delta_i^2)^{-2} \end{aligned} \right\}$$

Ces expressions ont été obtenues dans le cas où les variables X_i sont indépendantes, c'est-à-dire que la covariance de X_i et X_j est égale à $\sigma_{X_i X_j} = 0$.

Dans ce cas où toutes les variables X_i sont normales, on trouve

$$\begin{aligned} a_i^+ &= -a_i^- = \sqrt{3} \\ C &= 1 + n(n-7)/18 \\ N_i &= -(n-4)/(6\sqrt{3}) \\ P_{ij} &= 1/12 \end{aligned} \tag{VI.11}$$

VI.4 APPROXIMATION PAR UNE LOI NORMALE OU LOGNORMALE (LAMB, 1974)

Si Y est la somme de n variables aléatoires indépendante, $Y = \sum_{i=1}^n X_i$, d'après le théorème de la limite centrale, quand le nombre n devient grand, la densité de probabilité de Y tend asymptotiquement vers une loi normale c'est-à-dire que :

$$\lim g(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{(y - \bar{y})^2}{2\sigma^2} \right], \tag{VI.12}$$

à la condition que tous les moments des X_i aient des valeurs finies.

De même si Y est le produit de n variables aléatoires indépendantes, $Y = \prod_{i=1}^n X_i$, la densité de probabilité de Y tend asymptotiquement vers une distribution lognormale.

Par conséquent, pour les sommes et les produits de variables aléatoires indépendantes, on n'a

besoin de connaître que l'espérance mathématique et la variance de $Y = \sum_{i=1}^n X_i$

et de $Z = \sum_{i=1}^n \lg X_i$

Pour les combinaisons linéaires de variables indépendantes X_i , on a :

$$Y = \sum_{i=1}^n \beta_i X_i \quad (\text{VI.13})$$

$$E[Y] = \sum_{i=1}^n \beta_i E[X_i], \quad (\text{VI.14})$$

$$V[Y] = \sum_{i=1}^n \beta_i^2 V[X_i] \quad (\text{VI.15})$$

Pour les produits de la forme $Y = \prod_{i=1}^n (X_i)^{\alpha_i}$, on peut écrire :

$$Z = \lg Y = \sum_{i=1}^n \alpha_i \lg X_i = \sum_{i=1}^n \alpha_i W_i \quad (\text{VI.16})$$

$$E[Z] = \sum_{i=1}^n \alpha_i E[W_i], \quad (\text{VI.17})$$

$$V[Z] = \sum_{i=1}^n (\alpha_i)^2 V[W_i] \quad (\text{VI.18})$$

Pour obtenir les expressions des $E[W_i]$ en fonction des $E[X_i]$, on peut utiliser par exemple la méthode du développement en séries de Taylor.

Il vient après mise de la variable X_i sous forme réduite

$$[X_i = \bar{x}_i (1 + C_i U_i)]:$$

$$[X_i] = \lg \bar{x}_i + 0,434 \ln(1 + C_i U_i) = \lg \bar{x}_i + 0,434 \left[C_i U_i - \frac{C_i^2 U_i^2}{2} + \frac{C_i^3 U_i^3}{2} \dots \right] \quad (\text{VI.19})$$

$$\text{d'où : } E[W_i] = \lg \bar{X}_i - \frac{0,434}{2} C_i^2 \quad (\text{VI.20})$$

$$V[W_i] = (0,434 C_i^2)^2 \quad (21)$$

Remarque : Lumb (1974, p.63) donne des formules légèrement différents :

$$E[W_i] = \lg \bar{X}_i - \frac{1}{2} (C_i)^2 - \frac{3}{4} (C_i)^4 \quad (\text{VI.21})$$

$$E[W_i] = \lg \bar{X}_i + \frac{5}{2} (C_i)^4 \quad (\text{VI.22})$$

VI.5 SIMULATION : METHODE DE MONTE-CARLO

Il est toujours possible de déterminer empiriquement la densité de probabilité d'une variable $Y = f(X_i)$ en calculant les valeurs y en correspondant à des ensembles de valeurs

des x_i générés de façon aléatoire conformément aux densités de probabilité de chacune des variables aléatoires x_i en déterminant la densité de probabilité de Y d'après la distribution des fréquences des y obtenus la précision de cette simulation augmente avec la racine carrée de la taille de l'échantillon et de ce fait il faut disposer d'un échantillon très important pour obtenir des résultats utiles. Ceci nécessite en pratique l'utilisation d'un ordinateur. Lumb (1974), Benjamin et Corell (1970) et Harr (1977), par exemple, donnent des indications sur l'emploi de cette méthode.

Un des aspects importants de la méthode de Monte-Carlo est qu'elle nécessite la génération de séries de valeurs aléatoires de chacun des paramètres des modèles de calcul. Il existe les bibliothèques de sous programmes statistiques des programmes capables de générer des suites de nombres aléatoires suivant des lois de distribution simples (densité uniforme sur un intervalle donné loi normale, loi lognormale, etc.).

Pour les loi de distribution uniformes, les programmes de calcul utilisent des algorithmes tels que :

$$X_{n+1} = (A x_n) \bmod M \quad (\text{VI.24})$$

$$X_{n+1} = (A x_n + B) \bmod M \quad (\text{VI.25})$$

qui permettent de calculer une série de nombre x_n à partir d'une valeur initiale x_0 quelconque. La notation $x = (y) \bmod M$ indique que x est le reste de la division du nombre entier y par le nombre entier M . le nombre M est la période de la séquence aléatoire générée. On le choisit très grand. Le nombre A est compris entre 0 et $M-1$, ainsi que le nombre B . les nombres générés sont uniformément répartis entre 0 et $M-1$.

On ajuste la série $\{x_n\}$ à l'intervalle $[a, b]$ désiré en utilisant la formule

$$x'_n = \frac{b-a}{M} x_n + a \quad (\text{VI.26})$$

Le tableau VI.1 contient une série de 100 nombres aléatoires uniformément répartis sur l'intervalle $[0,1]$, obtenus à l'aide de la relation

$$x'_n = \frac{x_n}{199017} = \frac{1}{199017} (24298 x_{n-1} + 99991) \bmod 19997, \quad \text{avec } x_0 = 3157.$$

Tableau VI.1 Valeurs aléatoires pour distribution uniforme $R_U(0,1)$
(moyenne 0,4992, écart type 0,2989)

0,90478	0,21584	0,84210	0,61357	0,02350	0,91285	0,71604	0,65892	0,57355	0,37984
0,67091	0,07410	0,97724	0,21177	0,70248	0,09452	0,89313	0,99525	0,72393	0,01584
0,43931	0,10304	0,64419	0,21744	0,39051	0,17125	0,70640	0,20549	0,57834	0,38648
0,01388	0,20799	0,24438	0,07596	0,23613	0,59846	0,63295	0,70570	0,21014	0,34670
0,24778	0,45595	0,46758	0,34120	0,16860	0,97289	0,93245	0,75541	0,64645	0,84905
0,54450	0,29001	0,97433	0,7281	0,27009	0,94595	0,37735	0,45936	0,06715	0,80614
0,91757	0,32528	0,85926	0,84568	0,19368	0,31652	0,51090	0,01134	0,26668	0,11814
0,63820	0,28335	0,92873	0,86820	0,67134	0,50875	0,45717	0,9495	0,48185	0,14686
0,69644	0,48675	0,02402	0,19270	0,78788	0,32044	0,84654	0,76848	0,73332	0,14788
0,72449	0,68472	0,34764	0,93623	0,46943	0,60870	0,75928	0,09536	0,88138	0,84515

Pour obtenir des valeurs aléatoires suivant une loi normale de moyenne μ et d'écart type σ , on utilise les valeurs générées sur l'intervalle $[0,1]$ pour une distribution uniforme, soit $R_U(0,1)$: on calcul d'abord la valeur d'un nombre $R_N(0,1)$ suivant une loi normale réduite (moyenne 0 et écart type 1) à l'aide de l'expression

$$R_N(0,1) = \sqrt{-2 \ln u_1} \cos(2 \pi u_2) \quad (\text{VI.27})$$

qui fait intervenir deux nombres u_1 et u_2 générés dans l'intervalle $[0,1]$ comme indiqué ci-dessous. Puis on ajuste la loi normale à sa moyenne et à son écart type à l'aide de l'expression :

$$R_N(\mu, \sigma) = \sigma R_N(0,1) + \mu \quad (\text{VI.28})$$

Le tableau II contient une série de 100 nombres aléatoires suivant une loi normale réduite $R_N(0,1)$.

Tableau VI.2 Valeurs aléatoires pour une distribution réduite uniforme $R_U(0,1)$
(moyenne 0,1403, écart type 0,9047).

-0,16658	0,58029	-0,80567	0,63986	-0,17195	-0,10918	1,52516	-0,53041	0,04284	-0,32753
1,27407	0,03311	0,11936	-0,55938	0,25937	0,37464	-0,09567	-0,58433	-1,70448	0,41919
-1,60535	1,21703	-0,23755	-0,26834	0,85217	-0,98083	-0,25826	0,86542	-0,06157	0,12251
-0,26800	0,49648	-0,85951	-1,11724	-1,61530	-0,11159	0,53155	-0,14963	0,65953	-0,24840
-0,13576	-1,57224	-0,69035	0,03364	0,57880	1,78357	-0,64737	0,66009	-0,26376	1,68138
1,56473	0,23514	0,35392	-0,91300	1,38452	0,02503	-0,64234	1,62489	0,62041	-0,52594
0,55623	-1,55169	-1,53046	-0,48880	-0,78719	1,74257	-1,01330	-0,05342	0,09909	0,94039
-0,31176	1,31568	0,22106	-0,71326	0,19764	-0,25877	0,43608	0,70446	-0,20704	0,50900
-0,31175	0,39155	-1,51453	2,47609	1,24760	0,42002	-0,15974	2,29259	-0,52884	-1,35866
-0,47852	1,67100	-1,17028	0,59931	1,10049	-1,19535	-0,51306	-0,29220	0,22532	0,64171

Pour les valeurs entières des paramètres α et β de la loi bêta, on peut utiliser l'expression suivante, due à Hahn et Shapiro (1967) :

$$R_B(0,1) = \frac{\sum_{i=1}^{2\alpha+2} R_{Ni}^2(0,1)}{\sum_{i=1}^{2(\alpha+\beta+2)} R_{Ni}^2(0,1)} \quad (\text{VI.29})$$

qui utilise $2(\alpha + \beta + 2)$ valeurs aléatoires normales réduites $R_{Ni}(0,1)$.

Pour les autres valeurs de α et β , il faut procéder par interpolation.

Il faut pour terminer indiquer que l'on peut « facilement » générer des valeurs aléatoires de toute variable dont on connaît la fonction de répartition $G(x)$ en utilisant l'algorithme suivant, qui repose sur le fait que les valeurs de $G(x)$ sont uniformément réparties sur l'intervalle $[0,1]$: on génère des valeurs aléatoires de $R_U(0,1)$ et l'on retient les valeurs X_i de la variable qui sont telles que $G(x_i) = R_{Ui}(0,1)$. La figure 3 illustre la procédure adoptée.

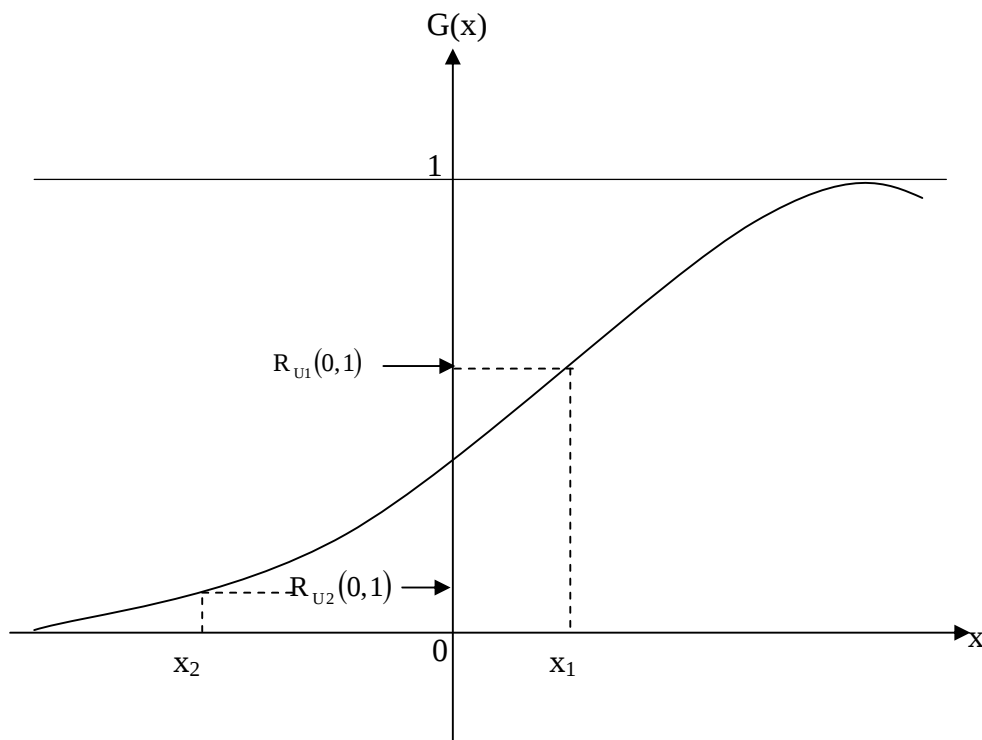


Figure VI.1 Procédure de génération de valeurs aléatoires d'une variable X dont la fonction de répartition $G(x)$ est connue.

Les techniques de génération de valeurs aléatoires qui viennent d'être décrites s'appliquent à des variables isolées indépendantes. Dans le cas de plusieurs variables

dépendantes, la génération d'ensembles de données est plus difficile. Il existe néanmoins différentes techniques utilisables X_1 et X_2 dont on connaît la densité de probabilité jointe $h(x_1, x_2)$, à tirer d'abord une valeur aléatoire de x_1 pour pouvoir définir la densité de probabilité $g(x_2|x_1)$ et générer ensuite la valeur de x_2 . Par exemple dans le cas d'une distribution bi normale de paramètres $(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$, on obtient :

$$\left. \begin{aligned} x_1 &= \mu_1 + R_{N_1}(0,1) \\ x_2 &= \mu_2 + \left\{ \rho R_{N_1}(0,1) + \sqrt{1-\rho^2} R_{N_2}(0,1) \right\} \sigma_2 \end{aligned} \right\} \quad (VI.30)$$

VI.6 ESTIMATION PONCTUELLE : méthode de Rosenblueth

De toutes les méthodes d'approximation d'une fonction de variables aléatoires, la plus simple est celle de Rosenblueth (1975). Elle ne nécessite ni la connaissance de l'expression analytique (développements en séries de Taylor), ni des moyens de calculs très puissants (méthode de simulation de Monte Carlo). Rosenblueth a en effet présenté une méthode basée sur une approximation par des valeurs ponctuelles.

Elle permet l'estimation des moments statistiques d'une fonction d'une ou plusieurs variables aléatoires. C'est une méthode d'intégration numérique approximative d'une fonction aléatoire :

$$E(y^n) = \int_{-\infty}^{+\infty} y^n df(y)$$

Soit la variable aléatoire x , sa fonction de densité de probabilité $f(x)$ et deux valeurs de x , x_+ et x_- , situées de part et d'autre de son espérance mathématique (moyenne $\bar{x}$) (figure ci-dessous).

Le postulat de Rosenblueth est que x_0 , variable aléatoire discrète dont x_+ et x_- sont le domaine de définition, est une approximation ponctuelle de la variable continue x , si les trois premiers moments de x_0 sont égaux à ceux de la distribution de x .

Les conditions d'approximations donnent :

$$- \text{ l'espérance de } x_0 \text{ est équivalente à celle de } (\bar{x}_0 = \bar{x}),$$

$$\bar{x} = P_+ x_+ + P_- x_- \quad (VI.31)$$

- La variance de x_0 est équivalente à celle de x ($\sigma_{x_0}^2 = \sigma_x^2$)

$$\sigma_x^2 = P_+ (x_+ - \bar{x})^2 + P_- (x_- - \bar{x})^2 \quad (VI.32)$$

- Le troisième moment central de x_0 est équivalent à celui de x :

$$\beta_{1x} \sigma_x^2 = P_+ (x_+ - \bar{x})^3 + P_- (x_- - \bar{x})^3 \quad (\text{VI.33})$$

β_1 est le coefficient de dissymétrie de Pearson et P_+ et P_- sont les probabilités correspondantes de x_+ et x_- :

$$P_+ = \text{Pr ob}(x_0 = x_+)$$

$$P_- = \text{Pr ob}(x_0 = x_-)$$

Les valeurs x_+ et x_- étant mutuellement exhaustives et exclusives, on a :

$$P_+ + P_- = 1 \quad (\text{VI.34})$$

Le système des quatre équations (VI.31) à (VI.4) a pour solution :

$$P_+ = \frac{1}{2} \left[1 - \frac{\beta_{1x}}{|\beta_{1x}|} \cdot \sqrt{1 - \frac{1}{1 + (\beta_{1x}/2)^2}} \right]$$

$$P_- = 1 - P_+$$

$$x_+ = \bar{x} + \sigma_x \cdot \sqrt{P_-/P_+}$$

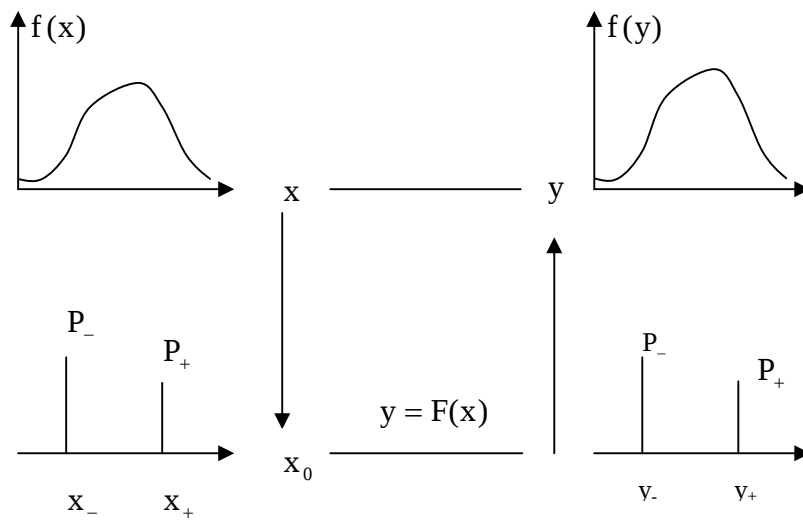
$$x_- = \bar{x} + \sigma_x \cdot \sqrt{P_+/P_-}$$

Considérons maintenant une fonction $y = F(x)$, univoque, dans laquelle x est une variable aléatoire.

Rosenblueth conçoit que la quantité aléatoire y puisse également être représentée par une approximation ponctuelle y_0 , définie par les valeurs y_+ et y_- , de probabilités P_+ et P_- , égales à celles attribuées à x_+ et x_- :

$$\begin{aligned} y_+ &= F(x_+) \\ y_- &= F(x_-) \end{aligned} \quad (\text{VI.35})$$

La démarche correspondante est illustrée par la figure ci-dessous :



Connaissant x_+ et x_- , P_+ et P_- , les paramètres de la distribution $f_y(y)$ s'obtiennent par application des relations (VI.1) à (VI.3) :

$$\bar{y} = P_+ y_+ + P_- y_- \quad (\text{VI.36})$$

$$\sigma_y^2 = P_+ (y_+ - \bar{y})^2 + P_- (y_- - \bar{y})^2 \quad (\text{VI.37})$$

$$\beta_{1y} \sigma_y^3 = P_+ (y_+ - \bar{y})^3 + P_- (y_- - \bar{y})^3 \quad (\text{VI.38})$$

Et d'une manière plus générale :

$$E(y^N) = P_+ y_+^N + P_- y_-^N$$

(Moments non centrés)

$$E[(y - \bar{y})^N] = P_+ (y_+ - \bar{y})^N + P_- (y_- - \bar{y})^N$$

(Moments centrés)

Cas d'une fonction de plusieurs variables :

Dans le cas d'une fonction de plusieurs variables $y = F(x_1, x_2, \dots, x_n)$ dans laquelle $x_1, x_2, \dots, x_n$ sont des variables indépendantes ou dépendantes (corrélées), les premiers moments s'obtiennent comme pour les fonctions d'une seule variable :

- Dans le cas de deux variables ($y = F(x_1, x_2)$) :

$$E(y) = P_{++}y_{++} + P_{+-}y_{+-} + P_{-+}y_{-+} + P_{--}y_{--}$$

avec :

$$y_{\pm\pm} = F\left[\bar{x}_1 \pm \sigma_{x_1} \cdot (P_-/P_+)^{\pm 1/2}, \bar{x}_2 \pm \sigma_{x_2} \cdot (P_-/P_+)^{\pm 1/2}, \right]$$

$$E(y^N) = P_{++}y_{++}^N + P_{+-}y_{+-}^N + P_{-+}y_{-+}^N + P_{--}y_{--}^N$$

$$V(y) = \sigma_y^2 = E(y^2) - [E(y)]^2$$

- Dans le cas de trois variables ($y = F(x_1, x_2, x_3)$) :

$$E(y^N) = P_{+++}y_{+++}^N + P_{++-}y_{++-}^N + P_{+-+}y_{+-+}^N + P_{+--}y_{+--}^N + \\ P_{-++}y_{-++}^N + P_{-+-}y_{-+-}^N + P_{-+-}y_{-+-}^N + P_{---}y_{---}^N$$

$$V(y) = \sigma_y^2 = E(y^2) - [E(y)]^2$$

- dans le cas de n variables $y = F(x_1, x_2, \dots, x_n)$

$$E(y) = \sum_1^{2n} P_{i,j,k,l,n} \cdot y_{i,j,k,l,n}$$

où les suites i, j, k, l, n sont les permutations de signes $\pm$; il y a 2^n permutations.

Les termes :

$$y_{i,j,k,l,n} = F\left(\bar{x}_1 \pm \sigma_{x_1} \cdot (P_-/P_+)^{\pm 1/2}, \dots, \bar{x}_n \pm \sigma_{x_n} \cdot (P_-/P_+)^{\pm 1/2}\right)$$

sont les fonctions ponctuelles de la fonction.

En générale :

$$E(y^N) = \sum_1^{2n} P_{i,j,k,l,n} \cdot y_{i,j,k,l,n}^N$$

$$V(y) = \sigma_y^2 = E(y^2) - [E(y)]^2$$

Les concentrations ont les valeurs suivantes :

- dans le cas où les variables sont symétriques et indépendantes :

$$P_{++} = P_{+-} = P_{-+} = P_{--} = 1/4$$

$$P_{+++} = \dots = P_{---} = 1/8$$

$$P_{i,j,k,l,n} = (1/2)^n$$

- dans le cas où les variables sont symétriques mais corrélées :

$$P_{++} = P_{--} = (1 + \rho)/4$$

$$P_{+-} = P_{-+} = (1 - \rho)/4$$

où ρ est le coefficient de corrélation entre x_1 et x_2 .

$$P_{+++} = P_{---} = (1 + \rho_{12} + \rho_{23} + \rho_{31})/8$$

$$P_{++-} = P_{--+} = (1 + \rho_{12} - \rho_{23} - \rho_{31})/8$$

$$P_{+-+} = P_{-+-} = (1 - \rho_{12} - \rho_{23} + \rho_{31})/8$$

$$P_{+--} = P_{--+} = (1 + \rho_{12} + \rho_{23} - \rho_{31})/8$$

où les ρ_{ij} est le coefficient de corrélation partiel entre les variables x_i et x_j .

- dans le cas où les variables sont indépendantes mais dissymétriques (Rosenblueth, 1983) :

$$P_{++} = P_+(x_1) \cdot P_+(x_2)$$

$$P_{--} = P_-(x_1) \cdot P_-(x_2)$$

$$P_{+-} = P_+(x_1) \cdot P_-(x_2)$$

$$P_{-+} = P_-(x_1) \cdot P_+(x_2)$$

où les valeurs de $P_+(x_1)$, $P_-(x_1)$, $P_+(x_2)$, $P_-(x_2)$ sont données par les relations (VI).

Chapitre VII

ANALYSE PROBABILISTE DE LA STABILITÉ DES OUVRAGES

Chapitre VII

ANALYSE PROBABILISTE DE LA STABILITE DES OUVRAGES

VII.1 INTRODUCTION

L'approche probabiliste ou semi- probabiliste de la sécurité des ouvrages est à la mode. On l'a introduite dans les règlements de calcul des ouvrages en béton, béton armé, béton précontraint et métal et certains responsables la verraient bien pénétrer en force dans le domaine des fondations de ces ouvrages et même dans le dimensionnement de tous les ouvrages de génie civil. A l'heure actuelle, une telle évolution des méthodes de la mécanique des sols est certainement prématurée, mais elle est probable à plus ou moins long terme. Il était, pour cette raison indispensable de consacrer un chapitre du présent document aux fondements de l'approche probabiliste de la sécurité des ouvrages en mécanique des sols. L'élaboration des méthodes de calcul probabilistes est encore loin de son terme, malgré le nombre élevé des publications consacrées à ce sujet.

- Mais, il est déjà possible de poser les bases de ce que doivent être les méthodes d'études probabiliste de la stabilité des ouvrages et aussi d'indiquer ce qu'elle ne peuvent pas être. Les thèmes développés dans ce chapitre ont trait à quelques calculs d'ouvrages par des approches autres que déterministes. Il y sera développé en particulier une analyse des notions de coefficient de sécurité « classique » et de probabilité de rupture, une introduction à l'analyse probabiliste de la stabilité de pentes et enfin une étude consacrée à la consolidation probabiliste bidimensionnelle.

VII.2 COEFFICIENT DE SECURITE ET PROBABILITE DE RUPTURE

VII.2.1 Coefficient de sécurité

En mécanique des sols, le dimensionnement des ouvrages comporte traditionnellement deux étapes : une étude de stabilité, objet du présent chapitre et un calcul en déformation, dans les cas où c'est nécessaire.

Pour l'étude de stabilité, la pratique en vigueur depuis plusieurs dizaines d'années comporte le calcul d'un état limite (rupture rotationnelle pour les remblais sur les sols mous, poinçonnement pour les fondations profondes ou superficielle, renversement, glissement,

poinçonnement et rotation d'ensembles pour les ouvrages de soutènement, etc.) et de calcul d'un état « de service » défini par référence à cet état limite à l'aide d'un *coefficient de sécurité* **F**.

Le coefficient de sécurité est défini différemment selon le problème traité :

- dans les études de stabilité de pente, il est généralement égal au rapport des efforts résistants aux efforts moteurs qui s'exercent sur une masse de sol limitée par une surface de rupture de forme de données. Pour les calculs en rupture circulaire, pour chaque cercle de rupture possible, on pose ainsi

$$F = \frac{\text{moments des forces résistantes par rapport au centre du cercle}}{\text{moments des forces motrices par rapport au centre du cercle}}$$

Et le coefficient de sécurité de la pente est pris égal au minimum des valeurs de **F** calculées sur l'ensemble des cercles de rupture envisagés, tandis que, pour les calculs en rupture non circulaire on utilise plutôt le rapport des forces résistantes aux forces motrices. La stabilité est assurée par un processus itératif dans lequel on modifie les hypothèses géométriques du problème jusqu'à obtenir la valeur désirée du coefficient de sécurité ;

- pour les fondations superficielles, on calcule la capacité portante du sol en plasticité parfaite (à l'aide d'abaques) puis on calcule la charge de service en divisant la capacité portante par un coefficient de sécurité ;
- pour les fondations profondes, on calcule séparément la contribution de la pointe et de la surface des pieux et l'on détermine la charge de service en affectant chacun de ces termes d'un coefficient de sécurité particulier ;
- pour les ouvrages de soutènement, on évalue les forces qui s'exercent sur l'ouvrage et on calcule les rapports « effort résistant/effort moteur » correspondant aux différents modes de rupture possibles et l'on modifie la géométrie de l'ouvrage jusqu'à ce que ces coefficients de sécurité partiels aient tous des valeurs convenables.

En pratique, cette diversité des définitions du coefficient de sécurité ne pose pas de problème tant que l'on étudie des ouvrages isolés : on sait qu'il faut prendre $F = 1.5$ pour que les remblais sur sols compressibles, $F = 3$ pour les fondations superficielles, $F = 2$ ou pour les frottements latéral et l'effort de pointe, respectivement, dans les calculs de pieux, etc.

La notion de sécurité est pourtant contestée, pour deux grandes raisons :

- On lui a reproché de ne pas être un « coefficient de sécurité » au sens strict du terme, parce qu'il n'est pas par rapport à un état limite correct [différentes recherches ont été réalisées ou sont en cours sur ce point, notamment au laboratoire de mécanique des solides de Palaiseau et au laboratoire central des ponts et chaussées_ on pourra se reporter pour plus de détails au

publications de Salençon (1974), Coussy (1978), Tristan- Lopez (1981) et Gennoui (1982) sur ce sujet] ;

- On se heurte fréquemment au problème de dimensionnement d'ouvrages complexes, comportant par exemple des fondations superficielles, des ouvrages de soutènement et des déblais (figure 34), pour lesquels la définition d'un coefficient de sécurité unique est ardue, sinon impossible.

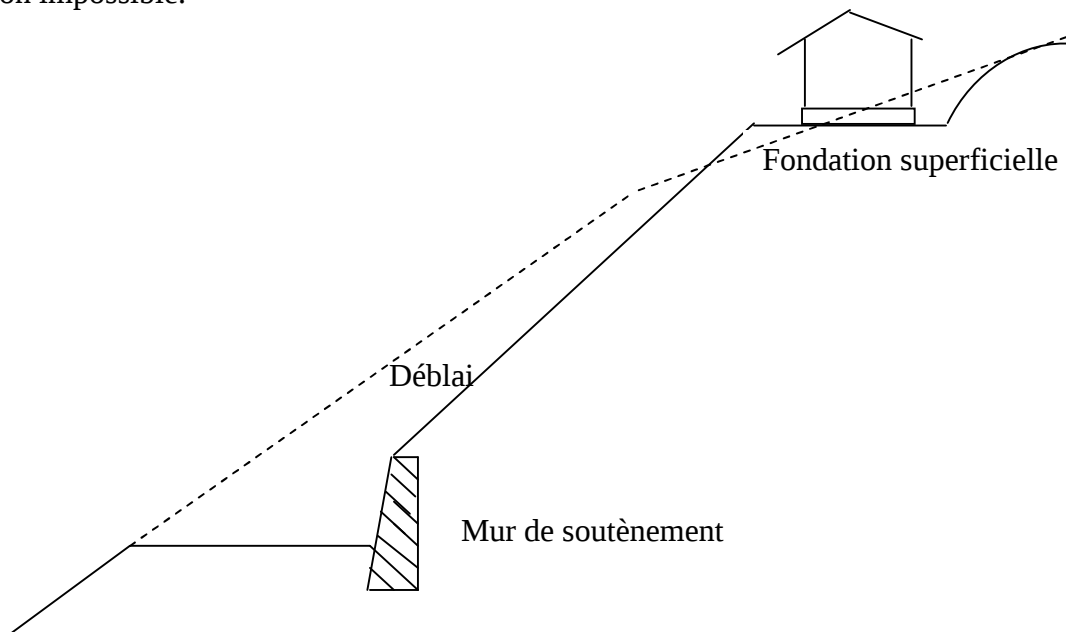


Figure VII.1.1 Exemple de situation complexe pour le calcul du coefficient de sécurité

Si le premier reproche est justifié sur le plan théorique, il ne trouble pas la conscience des projeteurs dans la mesure où la pratique de l'utilisation des coefficients de sécurité est considérée globalement comme satisfaisante par les mécaniciens des sols.

Le problème des ouvrages complexes était beaucoup plus gênant en pratique et l'on a cherché très tôt à unifier la définition et le calcul des coefficients de sécurité, sans grand succès, il est vrai. Les espoirs de solution de ce problème délicat reposent aujourd'hui sur la notion de probabilité de rupture.

VII.2.2 Probabilité de rupture

A l'heure où ces lignes sont écrites, personne ne s'offusque de lire le terme de probabilité de rupture pour un ouvrage de génie civil, mais cela n'a pas toujours été le cas. Lors de débuts de l'application des méthodes statistiques et probabilistes en mécanique des sols, il parut tout à fait incongru à beaucoup de spécialistes que l'on applique la théorie du hasard à l'étude de comportement d'un matériau dont les propriétés sont parfaitement définies en chaque point. On définit d'ailleurs encore le coefficient de sécurité comme une

conséquence de la loi de comportement de matériau, c'est-à-dire comme la mesure d'une limite à ne pas dépasser pour éviter les phénomènes indésirables (déformation plastique, notamment qui se produisent peu avant la rupture.

En réalité, le coefficient de sécurité est plus d'une barrière de protection de nature rhéologique : c'est aussi et peut être surtout, un « coefficient d'incertitude », une précaution contre notre ignorance des propriétés réelles du sol de fondation dans la région de l'ouvrage à construire. Sous cet aspect, le recours au vocabulaire probabiliste devient naturelle : quand on dimensionne un ouvrage avec un coefficient de sécurité de 1,3 par exemple, une rupture peut tout à fait se produire si les essais réalisés lors de la reconnaissance géotechnique ont donné une image erronée des propriétés mécaniques du sol de fondation : la variabilité naturelle des propriétés des sols et l'échantillonnage (au sens statique) réalisé lors de la reconnaissance laissent subsister le risque d'une estimation biaisée des paramètres, c'est-à-dire qu'à tout dimensionnement est associée une probabilité de rupture, faible en général, mais jamais nulle. Le principe de calcul de la probabilité de rupture des ouvrages est simple : on compare de façon tout à fait habituelle la résultante des efforts moteurs C (charge) et la résultante des efforts résistantes R et l'on dit qu'il y a rupture lorsque C devient supérieur à R . Mais l'on ajoute que compte tenu des informations disponibles sur le sol de fondation, on ne connaît pas que de façon certaine les valeurs de C et R : on peut seulement calculer la distribution des probabilités des valeurs de ces deux paramètres (Figure VII.2)

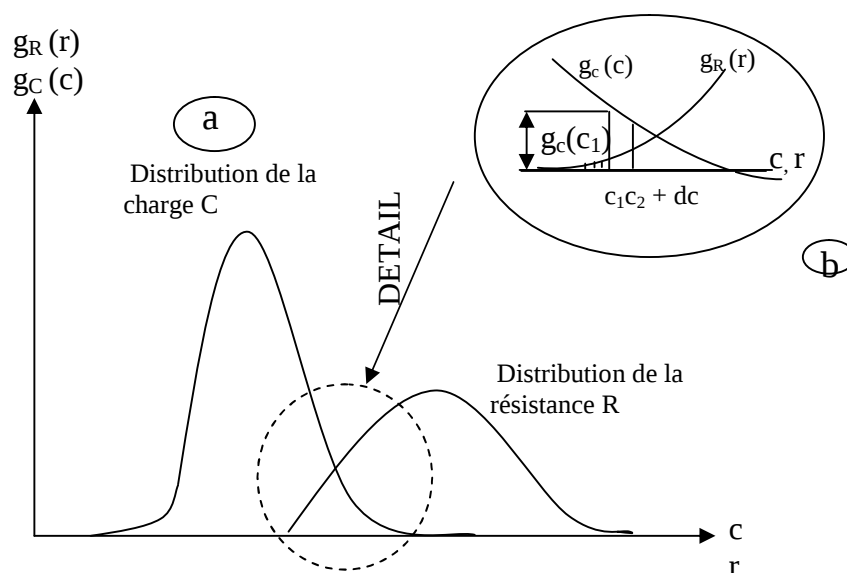


Figure VII.2 Distribution des probabilités de C et R

La probabilité que la résistance R soit inférieure à une valeur déterminée de charge C_1 est égale à :

$$\text{Prob} \{R < c_1\} = \int_{-\infty}^{c_1} g_R(r) dr = \text{aire hachurée de la figure VIII.2b.} \quad (\text{VII.1})$$

Comme la probabilité de la valeur c_1 de la charge est elle-même égale à $g_c(c_1) dc$, on peut par intégration calculer la probabilité de rupture

$$P_r = \int_{-\infty}^{+\infty} G_c(c) \left[\int_{-\infty}^c g_R(r) dr \right] dc \quad (\text{VII.2})$$

$$P_r = \int_{-\infty}^{+\infty} G_R(r) g_c(c) dc \quad (\text{VII.3})$$

en désignant par $g_r(r)$ la fonction de répartition de la variable aléatoire R .

En réalité, comme on le verra dans la suite de ce chapitre, le calcul de la probabilité de rupture est une opération complexe qui nécessite une bonne connaissance des méthodes de calcul probabiliste mais aussi des méthodes de calcul traditionnelles des ouvrages concernés.

VII.2.3 Coefficient de sécurité ou probabilité de rupture

On pourrait penser que la notion de probabilité de rupture va remplacer à plus ou moins long terme celle de coefficient de sécurité pour dimensionnement des ouvrages. En fait, comme on l'a indiqué précédemment, le coefficient de sécurité recouvre deux précautions différentes :

- on cherche, par son intermédiaire, à compenser le caractère très partiel des informations rassemblées lors de la reconnaissance géotechnique du sol de fondation ;
- on cherche également, dans le cas par exemple des remblais sur sols compressibles, à éviter l'apparition des grandes déformations plastiques qui précèdent la rupture ;
- la probabilité de rupture se substitue parfaitement au coefficient de sécurité et constitue même un progrès vis-à-vis du premier objectif. Pour ce qui concerne le second, elle est totalement inopérante et il est difficile d'imaginer ce qui pourrait remplacer le coefficient de sécurité dans ce second rôle.

VII.3 METHODE DE CALCUL PROBABILISTE DU COMPORTEMENT DES OUVRAGES

Les calculs mécaniques des sols nécessaires au dimensionnement des ouvrages (ouvrages de soutènement, ouvrages en terre, fondations, etc.) ont tous pour objectif de déterminer la valeur de paramètres (coefficient de sécurité, temps de consolidation, amplitude des tassement, etc.) qui dépendent des propriétés physiques et mécaniques des sols, de la géométrie des problèmes et des conditions initiales et aux limites imposées.

Si l'on traite les variations dans l'espace des propriétés des sols, les variations dans le temps des conditions aux limites et les fluctuations de la géométrie des sols et des ouvrages comme des phénomènes aléatoires dont chacun peut être représenté par une variable aléatoire X_i , les résultats des calculs nécessaires au dimensionnement des ouvrages sont eux-mêmes des variables aléatoires Y , fonctions des variables aléatoires X_i .

Chacune des variables aléatoires Y a sa propre fonction de répartition $G(y)$ et sa densité de probabilité $g(y)$, qu'il s'agit d'évaluer.

Le calcul probabiliste des ouvrages de mécanique des sols donc « simplement » un problème de calcul de fonction de variables aléatoires : $Y = f(X_i)$.

Pour deux variables indépendantes X_1 et X_2 de densités de probabilité g_1 et $g_2(x)$, la forme exacte de la densité de probabilité de la somme $Y = X_1 + X_2$, produit $Z = X_1 \cdot X_2$ ou du quotient $W = X_1/X_2$ est la suivante (Lumb, 1974) :

$$g(y) = \int_{-\infty}^{+\infty} g_1(y - \alpha) g_2(\alpha) d\alpha,$$

$$g(z) = \int_{-\infty}^{+\infty} g_1(z/\alpha) |\alpha|^{-1} \cdot g_2(\alpha) d\alpha$$

$$g(w) = \int_{-\infty}^{+\infty} g_1(z \cdot \alpha) |\alpha| \cdot g_2(\alpha) d\alpha$$

Et la combinaison de ces équations permet en principe de calculer la densité de probabilité de toute fonction comportant uniquement des sommes, des produits et des quotients. Mais cette méthode peut être très laborieuse en pratique, notamment parce qu'elle nécessite l'évaluation d'intégrales parfois complexes.

Par ailleurs, il n'existe pas toujours de relation explicite entre les paramètres physiques, mécaniques et géométriques des ouvrages et les résultats qu'il s'agit d'évaluer.

Enfin, la plupart des méthodes de calcul utilisées pour le dimensionnement des ouvrages ne sont qu'approchées, ce qui introduit à la fois un biais et une incertitude supplémentaire qu'il faut incorporer dans l'analyse de la fonction $Y = f(X_i)$.

Différentes méthodes peuvent être utilisées pour déterminer de façon approchée la densité de probabilité $g(Y)$ à partir des lois de distribution des paramètres X_i du modèle de calcul. Nous décrivons ci-après quatre d'entre elles :

- Approximation par les séries de Taylor
- Approximation par intégration numérique
- Approximation par une loi normale ou lognormale
- Simulation par la méthode de Monte-Carlo.

VII.4 ANALYSE PROBABILISTE DE LA STABILITE DES PENTES

VII.4.1 Introduction

Une dizaine de méthodes de calcul probabiliste ont été publiées jusqu'à présent pour l'analyse de la stabilité des pentes. Pourtant, ces méthodes pour l'instant des produits de recherche sans applications pratiques dans les bureaux d'étude, où les coefficients de sécurité traditionnels restent le seul outil du projeteur.

Cette situation, a priori étonnante, est certainement due pour partie à la lenteur habituelle du transfert des innovations des chercheurs aux praticiens. Elle a toutefois aussi une autre origine : l'insuffisance des méthodes testées jusqu'à présent. En effet, dans la plus part des articles disponibles dans la littérature, on se rend compte que l'utilisation des méthodes de calcul probabiliste conduirait à une augmentation parfois sensible mais toujours coûteuse des coefficients de sécurité habituellement admis.

Le caractère pessimiste des méthodes de dimensionnement probabiliste n'est pourtant pas inéluctable :

On peut raisonnablement penser que l'amélioration de la description des variations naturelles des propriétés des sols, l'amélioration des méthodes de calcul probabiliste et l'amélioration de la description des mécanismes de rupture devraient conduire à des méthodes de calcul utilisables pour les projets.

L'exemple du paragraphe précédent montre comment l'on peut diminuer la probabilité de rupture par une simple amélioration de la description du sol (et une complexification simultanée de la méthode de calcul). Le résultat obtenu dans cet exemple n'est pas encore suffisant, mais il montre sans nul doute la voie à suivre.

VII.4.2 Comment améliorer les méthodes de calcul probabiliste : un exemple

Considérons un massif de sol compressible d'épaisseur H sur lequel on veut construire un remblai de sable d'épaisseur H_r , de poids volumique γ_r et d'angle de frottement interne $\phi_r=35^\circ$. La reconnaissance géotechnique du site a comporté cinq sondages scissométriques avec un essai tous les mètres de profondeur. Calculons la probabilité de rupture pour les valeurs des paramètres indiqués sur la figure VII.3.

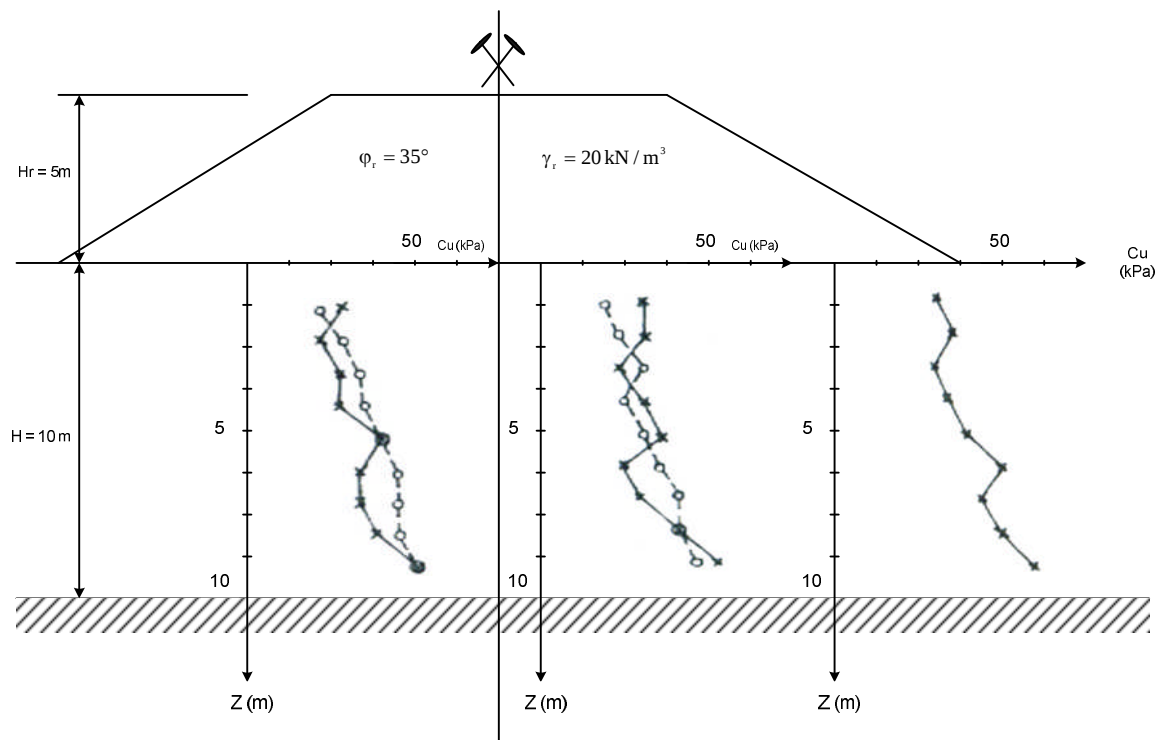


Figure VII.3 Hypothèses géométriques et mécaniques de l'analyse de stabilité

VII.4.3 Analyse de la stabilité par abaque (rupture circulaire, sol homogène)

La plus simple des analyses probabilistes consiste à calculer la stabilité du remblai à l'aide des abaques de Pilot et Moreau (1973), en supposant que la cohésion non drainée c_u a une valeur uniforme incertaine.

L'analyse probabiliste va comporter deux phases :

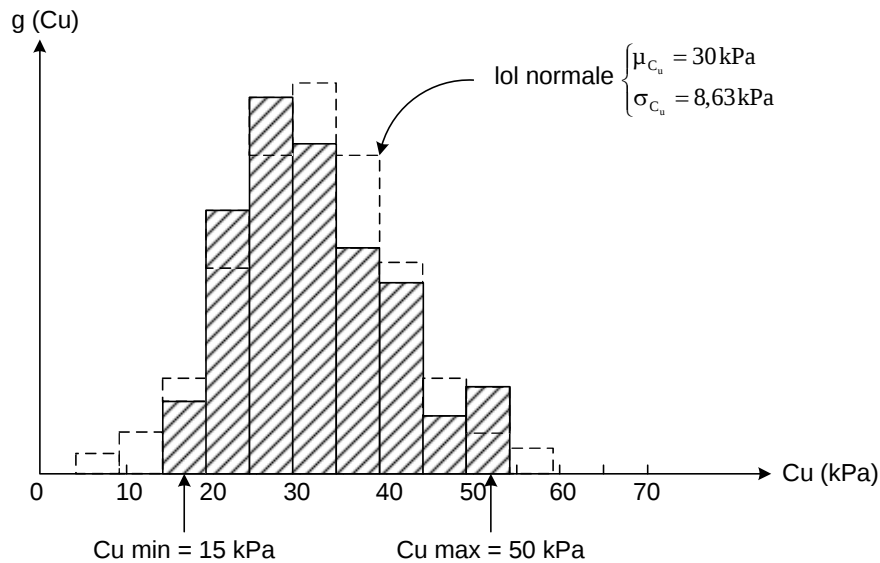
- Il faut d'abord estimer la loi de distribution des valeurs mesurées de la cohésion non drainée.
- Puis, il faut déduire de la loi de distribution de c_u la distribution des valeurs du coefficient de sécurité F . la probabilité de rupture sera alors égale à

$$P_r = \text{Prob} \{F < 1\} \quad (\text{VII.4})$$

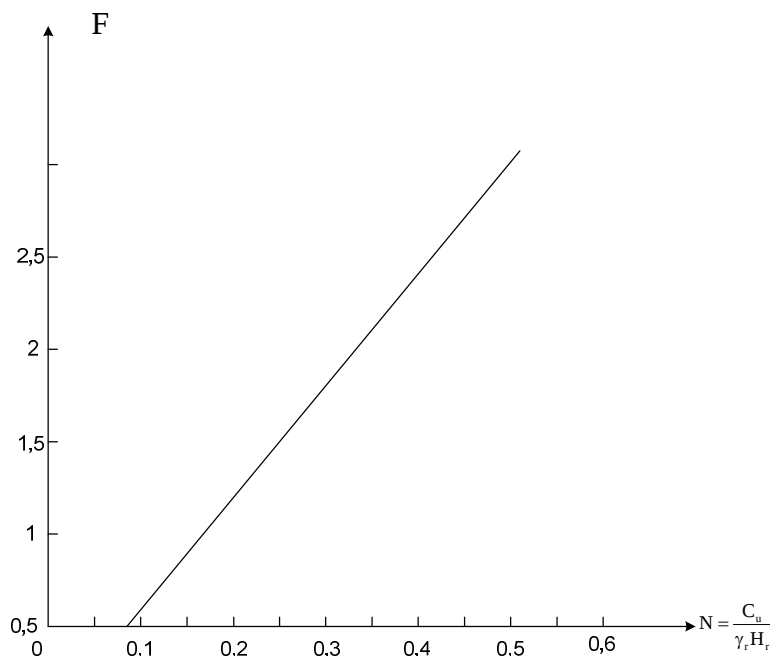
Les 45 valeurs de la cohésion non drainée indiquées sur la figure VIII.4 ont pour moyenne $\bar{c}_u = 30,11 \text{ kPa}$ pour écart type $\sigma_{c_u} = 8,63 \text{ kPa}$ et pour valeurs extrêmes

$$C_{u_{\min}} = 15 \text{ kPa} \text{ et } C_{u_{\max}} = 50 \text{ kPa}.$$

Si l'on admet que les valeurs de c_u suivent une loi normale, on trouve la loi représentée sur la figure VIII.5a



a - Loi normale estimée d'après les 45 valeurs mesurées de C_u



b - Abaque pour le calcul de stabilité ($\varphi_2 = 35^\circ$; $H_r / H = 0,5$; pente 1/2)

Figure VII.5 Etapes du calcul de stabilité à l'aide d'abaques

Pour faire les calculs de stabilité on va utiliser l'abaque de la figure VIII.5, qui a été établi d'après les abaques de Pilot et Moreau (1973) pour $\phi_r = 35^\circ$, des talus de remblai de pente 1/2 et un rapport $H/H_r = 2$. On voit que F est lié à c_u par une relation à peu près linéaire

$$F \approx 0,06 + 5,2 N = 0,06 + 0,052 c_u$$

Par conséquent, le coefficient de sécurité a une loi de distribution de même forme que c_u , mais avec des paramètres différents. On obtient : $\mu_F = 1,62$

$$\sigma_F = 0,052 \quad \sigma_{c_u} = 0,449.$$

Dans le cas d'une loi normale, on peut calculer aisément la probabilité de rupture

$$P_r = \Pr od \{F < 1\} = \Pr od \left\{ \frac{F - \mu_F}{\sigma_F} < \frac{1 - \mu_F}{\sigma_F} \right\} = \Pr od \left\{ \frac{F - \mu_F}{\sigma_F} < -1,381 \right\}$$

soit

$$P_r = 0,0837$$

Ce calcul serait exact si l'on était certain de la validité du schéma de calcul et de la représentativité des valeurs de c_u utilisées pour représenter la résistance au cisaillement du sol. Or chacun sait que les modèles de calcul utilisés pour les analyses de stabilité ne sont qu'approchés et que, d'autre part, les valeurs de c_u déterminées au scissomètre conduisent à des valeurs de F trop fortes dans les sols plastiques (indices de plasticité $I_p > 20$) et trop faible dans les sols peu plastiques ($I_p < 20$) ; à tel point que l'on corrige systématiquement les résultats en fonction de l'indice de plasticité [correction $\mu(I_p)$ de Bjerrum].

En pratique cela implique que la rupture n'est pas obtenue pour $F < 1$ mais pour F inférieur à une valeur F^* dépendant de l'indice de plasticité du sol et de l'incertitude sur le modèle de calcul. On peut admettre que F^* est une variable aléatoire dont la loi de distribution est normale (faute d'informations plus précise) avec une moyenne μ_{F^*} et un écart type σ_{F^*} . Si l'on admet que, pour l'exemple traité, $\mu_{F^*} = 1,2$ et $\sigma_{F^*} = 0,1$, on trouve

$$P_r = \Pr od \{F < F^*\} = \Pr od \{F - F^* < 0\} = \Pr od \{Y < 0\} \quad (\text{VII.5})$$

En introduisant la variable aléatoire $Y = F - F^*$ dont la densité de probabilité est

$$g(y) = \int_{-\infty}^{+\infty} \frac{1}{2\pi \cdot \sigma_{F^*}} \exp \left[-\frac{1}{2} \left(\frac{F - \mu_F + \alpha}{\sigma_F} \right)^2 - \frac{1}{2} \left(\frac{\alpha - \mu_{F^*}}{\sigma_{F^*}} \right)^2 \right] d\alpha$$

Pour éviter le calcul de cette intégrale, on utilisera une autre forme de la probabilité de rupture P_r :

$$P_r = \int_{-\infty}^{+\infty} G_F(\alpha) g_F^*(\alpha) d\alpha,$$

qui consiste à intégrer le produit de la probabilité que $F^* = \alpha \pm \frac{d\alpha}{2}$ par la probabilité $G_F(\alpha)$

que F soit inférieur à α . Un calcul numérique approché conduit à (tableau VII.1).

$$P_r \approx \sum_{i=1}^{30} G_F(0,1) g_F^*(0,1i) \times 0,1 = 0,1806.$$

Tableau VII.1 Détail du calcul de P_R

i	$g_f\left(\frac{i}{10}\right)$	$g_F^*\left(\frac{i}{10}\right)$	$G_F \cdot g_F^* \cdot 0,1$	i	$G_F\left(\frac{i}{10}\right)$	$G_F \cdot g_F^* \cdot 0,1$	$G_F \cdot g_F^* \cdot 0,1$	i	$G_F\left(\frac{i}{10}\right)$	$g_f^*\left(\frac{i}{10}\right)$	G f.g *0 ,1
1		3.10^{-26}	€	11	0,1234	2,420	$2,99.10^{-2}$	21	0.8575	1.10^{-17}	€
2		8.10^{-22}	€	12	0,1748	3,989	$6,97.10^{-2}$	22	0.9018	8.10^{-22}	€
3		1.10^{-17}	€	13	0,2380	2,420	$5,76.10^{-2}$	23	0.9350	3.10^{-26}	€
4		5.10^{-14}	€	14	0,3121	$5,4.10^{-1}$	$1,69.10^{-2}$	24			
5		9.10^{-11}	€	15	0,3946	$4,4.10^{-2}$	$1,7.10^{-3}$	25			
6		6.10^{-8}	€	16	0,4822	$1,3.10^{-3}$	$6,3.10^{-7}$	26			
7		1.10^{-5}	$4,42.10^{-3}$	17	0,5707	1.10^{-5}	$5,7.10^{-7}$	27			
8	$3,4.10^{-2}$	$1,3.10^{-3}$	$2,38.10^{-4}$	18	0,6557	6.10^{-8}	€	28			
9	$5,4.10^{-2}$	$4,4.10^{-2}$	$4,54.10^{-3}$	19	0,7336	9.10^{-11}	€	29			
10	$8,4.10^{-2}$	$4,4.10^{-1}$		20	0,8013	5.10^{-14}	€	30			

VII.4.4 Analyse de stabilité dans l'hypothèse d'un sol tricouche (rupture circulaire)

On peut améliorer la description de la résistance au cisaillement du sol, qui augmente avec la profondeur, en subdivisant la couche compressible en sous couches. Dans ce qui suit, on a pris trois couches d'épaisseur $H/3$.

L'analyse des données doit être recommencées à l'intérieur de chacune des sous couches : on obtient les résultats représentés sur la figure VII.6.

L'analyse probabiliste a été effectuée par la méthode de Monte-Carlo en supposant que les valeurs de la cohésion non drainée étaient indépendantes d'une sous couche à l'autre et qu'elles suivaient une loi normale.

Dans chacune des sous couches, vingt calculs ont été réalisés en rupture circulaire. Les données des calculs (valeurs de c_{u1} , c_{u2} , c_{u3} , dans les trois sous couches) et les coefficients de sécurité calculés sont rassemblés dans le tableau VIII.5.

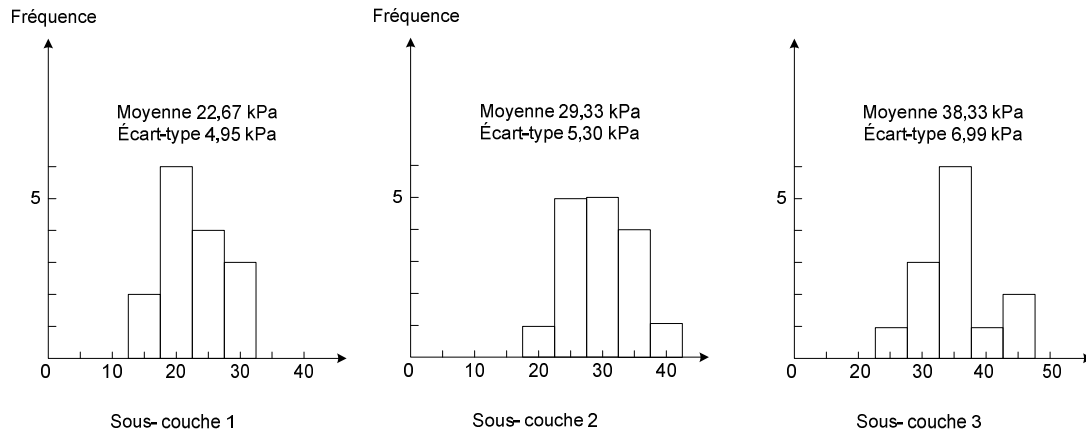


Figure VII.6 Valeurs de la cohésion non drainée mesurées dans chaque couche

Tableau VII.2 Calculs de stabilité (méthode de Monte-Carlo)

i	C_{u1}	C_{u2}	C_{u3}	F
1	21,85	36,08	27,093	1,29
2	21,34	28,61	49,28	1,26
3	25,42	27,67	36,15	1,43
4	20,30	32,41	38,56	1,22
5	28,69	31,96	27,32	1,51
6	23,83	37,55	47,54	1,40
7	24,61	38,19	32,69	1,42
8	23,26	28,07	32,31	1,37
9	19,25	31,21	28,62	1,18
10	23,76	21,30	31,14	1,18
11	28,98	20,82	37,16	1,24
12	22,00	37,62	36,45	1,31
13	21,13	27,68	43,22	1,27
14	25,54	29,51	34,98	1,48
15	25,13	21,00	46,85	1,19
16	30,35	36,30	39,98	1,72
17	30,94	25,06	41,07	1,42
18	21,49	24,77	39,17	1,27
19	24,42	21,22	33,50	1,19
20	15,17	23,13	39,88	0,98

L'histogramme des valeurs de F ainsi trouvées est représenté sur la figure VII.7.

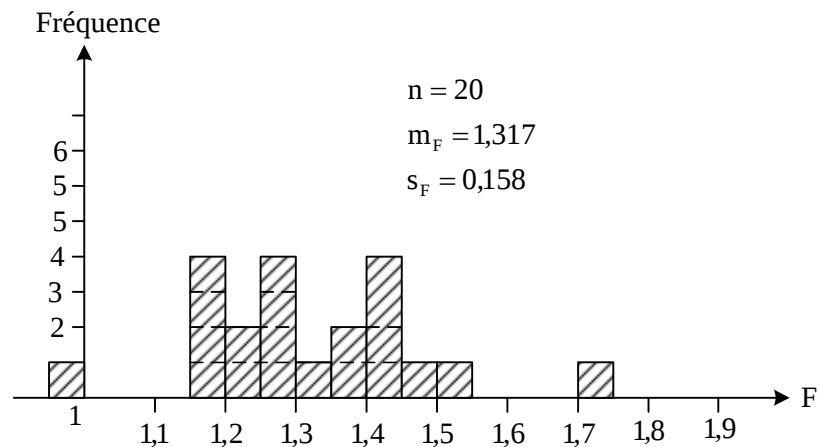


Figure VII.7 Histogramme des résultats des calculs

On voit que, bien que le coefficient de sécurité moyen soit plus faible que dans l'analyse de la section précédente, l'écart type σ_F est beaucoup plus faible :

$$\begin{aligned}\mu_F &= 1,32 \\ \sigma_F &\approx 0,16,\end{aligned}$$

De telle sorte que, si l'on admet que F suit une loi normale, la probabilité de rupture

$$P_r = \text{Pr ob}\{F < 1\} = \text{Pr ob}\left\{\frac{F - \mu_F}{\sigma_F} < \frac{1 - \mu_F}{\sigma_F}\right\} = \text{Pr ob}\left\{\frac{F - \mu_F}{\sigma_F} < -2\right\}$$

vaut $P_r = 0,02275$, c'est-à-dire qu'elle est quatre fois plus faible que dans l'analyse précédente.

On peut avoir une idée de la probabilité de rupture par rapport à $F^*(\mu_{F^*} = 1,2; \sigma_{F^*} = 0,1)$ en calculant :

$$P_{r_1} = \text{Pr ob}\{F < 1,2\} = 0,227$$

Un calcul plus précis, analogue à celui de l'analyse précédente, conduit à $P_{r1} \approx 0,235$.

VII.4.5 Conclusion

L'exemple traité ci-dessus appelle quelques commentaires. Il montre tout d'abord que si l'on admet que le modèle de calcul est exact, l'amélioration de la description des variations de la cohésion non drainée diminue sensiblement l'incertitude sur le coefficient de sécurité calculé et par conséquent aussi la probabilité de rupture.

Toutefois, l'incertitude relative au modèle de calcul (représentée par F^* dans le cas considéré) joue un rôle très important : pour les hypothèses adoptées, elle réduit à néant l'effet

positif d'une amélioration de la description des propriétés du sol, comme on ne dispose en général que de peu d'informations sur l'exactitude des modèles de calcul, cette incertitude pénalise très souvent les calculs probabilistes et peut expliquer leur caractère souvent pessimiste (par rapport aux pratiques traditionnelles).

Enfin, dans l'exemple présenté, on a considéré comme variable aléatoire comme cohésion non drainée du sol. D'autres paramètres peuvent être également traités comme variables aléatoires, poids volumique et angle de frottement interne du remblai géométrie de la surface du terrain naturel et du remblai géométrie du substratum, poids volumiques du sol de fondation.

Les probabilités de rupture calculées sont pour cette raison un peu plus faibles que les probabilités de rupture réelles (toutefois, les variations naturelles de c_u et l'incertitude sur le modèle de calcul représentent les causes principales de l'incertitude sur les résultats).

VII.4.6 Une méthode d'analyse probabiliste de la stabilité des pentes (Alonso, 1976)

La méthode générale d'analyse de la stabilité des pentes développée par Alonso (1976) est un exemple intéressant de méthode de calcul probabiliste.

Cette méthode consiste en une analyse probabiliste de la méthode des tranches. Elle tient compte de la variabilité de la cohésion, de la pression interstitielle, de l'angle de frottement interne, du poids volumique du sol, de la hauteur des tranches et du paramètre qui décrit le degré mobilisation de la résistance au cisaillement disponible.

Parmi tous ces paramètres ce sont les variations de la cohésion et de la pression interstitielle et la méthode de calcul qui sont déterminantes.

Bishop (1955) a établi pour le moment moteur M_a et le moment résistant M_r les expressions suivantes :

$$M_a = r \cdot \Delta \times \sum_{i=1}^{N_s} \gamma_i h_i \sin \sigma_i \quad \text{VII.6)}$$

$$M_r = r \cdot \Delta \sum_{i=1}^{N_s} \frac{c_i + \gamma_i \tan \Phi_i - \mu_i \tan \phi_i}{\cos \vartheta_i + R_i \sin \vartheta_i \tan \phi_i}, \quad \text{(VII.7)}$$

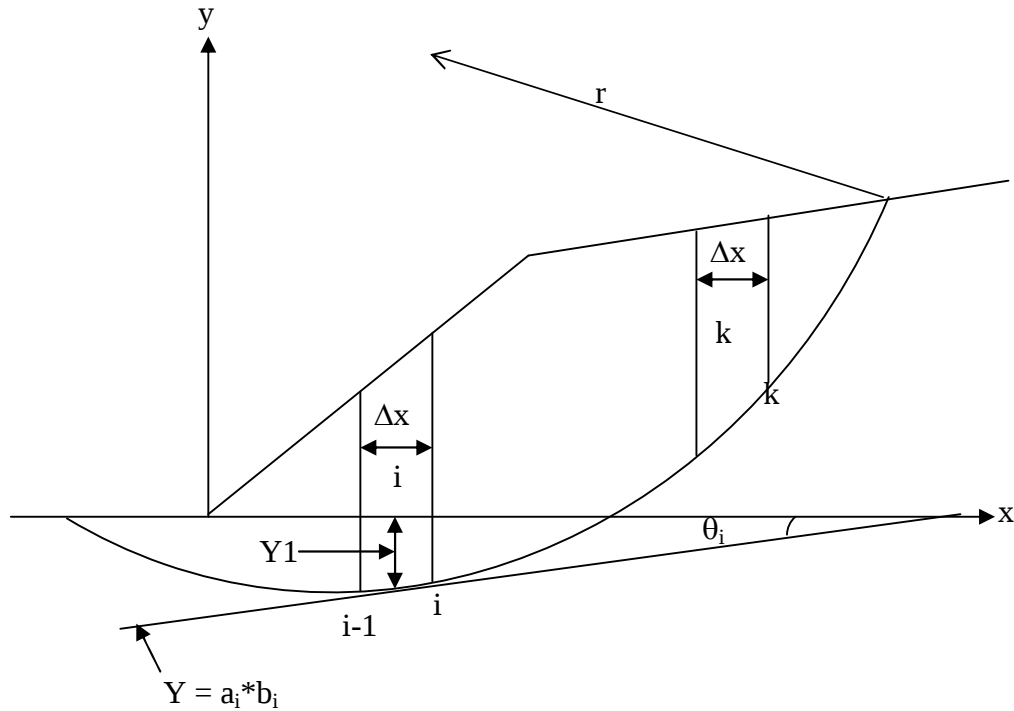


Figure VII.8 Géométrie de la surface de glissement

Avec les notations (figure VII.8) :

R - rayon du cercle de la rupture

Δ_x - largeur constante des tranches (égale pour toutes les tranches)

N_s - nombre de tranches

γ_i - poids volumique du sol de la i^{e} tranche

H_i - hauteur de la i^{e} tranche

θ_i - angle de la tangente à la surface de rupture avec l'horizontale

C_i - cohésion du sol à la base de la i^{e} tranche

ϕ_i - angle de frottement du sol à la base de la i^{e} tranche

U_i - pression interstitielle moyenne à la base de la i^{e} tranche

R_i - degré de mobilisation de la résistance au cisaillement à la base de la i^{e} tranche

La moyenne et la variance du coefficient de sécurité $F = M_r/M_a$ sont égales à :

$$\mu_F = E[F] = \frac{E[M_r]}{E[M_a]} \quad (\text{VII.8})$$

$$V[F] = \frac{V[M_r]}{E[M_a]^2} - \frac{2 E[M_r] \text{cov}[M_r, M_a]}{E[M_a]^2} + \frac{E[M_r]^2 V[M_a]}{E[M_a]^4}. \quad (\text{VII.9})$$

Les valeurs de $E[M_r]\mu_{Mr}$ et $E[M_a]\mu_{Ma}$ s'obtiennent en remplaçant chaque variable par sa valeur moyenne dans les expressions données plus haut. On a d'autre part :

$$V[M_a] = \sum_{i=1}^N \sum_{j=1}^N \frac{\partial M_a}{\partial \gamma_i} \frac{\partial M_a}{\partial \gamma_j} \bigg|_m \text{cov}[\gamma_i, \gamma_j] + \sum_{i=1}^N \sum_{j=1}^N \frac{\partial M_a}{\partial h_i} \frac{\partial M_a}{\partial h_j} \bigg|_m \text{cov}[h_i, h_j]$$

$$V[M_r] = W_c + W_{tg} + W_r + W_u + W_h + W_a$$

$$W_c = \sum_{i=1}^N \sum_{j=1}^N \frac{r^2 \Delta x^2}{\cos \theta_1 + R_1 \sin \theta_1 \text{tg} \phi_1} \bigg|_m \frac{1}{\cos \theta_1 + R_1 \sin \theta_1 \text{tg} \phi_1} \bigg|_m \text{cov}[c_1, c_2]$$

$$W_{tg\theta} = \sum_{i=1}^N \sum_{j=1}^N \frac{\Gamma^2 \Delta x^2 (-c_1 R_1 \sin \theta_1 + \cos \theta_1 \gamma_1 h_1 - \mu \cos \theta_1) (-c_1 R_1 \sin \theta_1 + \cos \theta_1 \gamma_1 h_1 - \mu \cos \theta_1)}{[(\cos \theta_1 + R_1 \sin \theta_1 \text{tg} \Phi_1)(\cos \Phi_1 + R_1 \sin \theta_1 \text{tg} \Phi_1)]^2} \bigg|_m \text{cov}[\text{tg} \phi_1 \text{tg} \phi_1]$$

$$W_r = \sum_{i=1}^N \sum_{j=1}^N \frac{r^2 \Delta x^2 h_1 \text{tg} \theta_1 h_1 \text{tg} \phi_1}{\cos \theta_1 + R_1 \sin \theta_1 \text{tg} \Phi_1} \frac{1}{\cos \theta_1 + R_1 \sin \theta_1 \text{tg} \Phi_1} \bigg|_m \text{cov}[\gamma_1, \gamma_2]$$

$$W_u = V[u] \sum_{i=1}^N \sum_{j=1}^N \frac{r^2 \Delta x \theta_1 \text{tg} \phi_1}{(\cos \theta_1 + R_1 \sin \theta_1 \text{tg} \Phi_1)(\cos \theta_1 + R_1 \sin \theta_1 \text{tg} \Phi_1)} \bigg|_m$$

$$W_h = V[h] \sum_{i=1}^N \left(\frac{r^2 \Delta x \gamma_1 \text{tg} \phi_1}{(\cos \theta_1 + R_1 \sin \theta_1 \text{tg} \Phi_1)} \right)^2$$

$$W_a = V[R]V[h] \sum_{i=1}^N \frac{r^2 \Delta x \sin \phi_1 (u_i \text{tg} \phi_1 - \lambda_1 h_1 \text{tg} \phi_1 - c_1)}{(\cos \theta_1 + R_1 \sin \theta_1 \text{tg} \Phi_1)} \bigg|_m \bigg)^2$$

$$\text{cov}[M_r, M] = \sum_{i=1}^N \sum_{j=1}^N \frac{\partial M_a}{\partial \gamma_i} \frac{\partial M_r}{\partial \gamma_j} \bigg|_m \text{cov}[\gamma_i, \gamma_j] + V[h] \sum_{i=1}^N \sum_{j=1}^N \frac{\partial M_a}{\partial \gamma_i} \frac{\partial M_r}{\partial \gamma_j} \bigg|_m$$

Il faut donc déterminer les covariances et autocovariances des paramètres.

VII.4.7 Propriétés du sol

On peut traiter les propriétés du sol comme des fonctions de variables aléatoires. Par exemple, la cohésion c peut être représentée par la somme d'une dérive $f(\cdot)$ et d'une variable aléatoire centrée $c'(c, y)$:

$$c = C(x, y) = f(x, y, \alpha_{\wedge_1}, \dots, \alpha_{\wedge_n}) + C'(x, y) \quad (\text{VII.10})$$

Les variables aléatoires α_j , étant les estimateurs des paramètres de la dérive.

Sur le segment $(i-1, i)$ c'est à dire sur la base de la i^{e} tranche, la cohésion moyenne est égale à :

$$C_i = \frac{1}{\Delta x} \int_{x_{i-1}}^{x_i} [f(x, y, \alpha_1, \dots, \alpha_n + C'(x, y))] dx \quad (\text{VII.11})$$

si l'on remplace l'arc de cercle (i-1,i) par un segment de droite d'équation $y = a_i x + b_i$ (figure VI.8), on obtient :

$$E[C_i] = \frac{1}{\Delta x} \int_{x_{i-1}}^{x_i} E[f(x, y, \alpha_1, \dots, \alpha_n)] dx. \quad (\text{VII.12})$$

La covariance des C_i pour sa part pour expression :

$$\begin{aligned} \text{cov}[C_K, C_M] &= \\ &= \text{COV} \left[\frac{1}{\Delta x} \int_{x_{i-1}}^{x_i} [f(x, y, \hat{\alpha}_1, \dots, \hat{\alpha}_n) + c'(x, y)] dx, \right. \\ &\quad \left. \frac{1}{\Delta x_m} \int_{x_{i-1}}^{x_i} [f(x, y, \hat{\alpha}_1, \dots, \hat{\alpha}_n) + c'(x, y)] dx \right] \\ &= \text{COV} \left[\frac{1}{\Delta x} \int_{x_{i-1}}^{x_i} [f(x, y, \hat{\alpha}_1, \dots, \hat{\alpha}_n) + c'(x, y)] dx, \frac{1}{\Delta x} \int_{x_{i-1}}^{x_i} [f(x, y, \hat{\alpha}_1, \dots, \hat{\alpha}_n) + c'(x, y)] dx \right] \\ &+ \text{COV} \left[\frac{1}{\Delta x} \int_{x_{i-1}}^{x_i} c'(x, y) dx, \frac{1}{\Delta x_m} \int_{x_{i-1}}^{x_i} [f(x, y, \hat{\alpha}_1, \dots, \hat{\alpha}_n) + c'(x, y)] dx \right] \\ &+ \text{COV} \left[\frac{1}{\Delta x} \int_{x_{i-1}}^{x_i} [f(x, y, \hat{\alpha}_1, \dots, \hat{\alpha}_n) + c'(x, y)] dx, \frac{1}{\Delta x} \int_{x_{i-1}}^{x_i} c'(x, y) dx \right] \\ &+ \text{COV} \left[\frac{1}{\Delta x} \int_{x_{i-1}}^{x_i} c'(x, y) dx, \frac{1}{\Delta x_m} \int_{x_{i-1}}^{x_i} c'(x, y) dx \right] \\ &= C_{11} + C_{21} + C_{12} + C_{22} \end{aligned} \quad (\text{VII.13})$$

Si l'on suppose qu'il existe régression linéaire entre C et la profondeur y , c'est-à-dire si :

$$f(x, y, \hat{\alpha}_1, \dots, \hat{\alpha}_n) = \hat{\alpha}_1 + \hat{\alpha}_2 y, \quad (\text{VII.14})$$

ce qui est souvent vrai, on a :

$$\begin{aligned} E[C_K] &= \\ \hat{\alpha}_1 + \hat{\alpha}_2 (a_k \frac{x_k + x_{k-1}}{2} + b_k) &= \hat{\alpha}_1 + \hat{\alpha}_2 y_k^* \end{aligned}$$

Où Y_K désigne l'ordonnée du centre du segment (k-1). Par conséquent le terme C_{11} se calcule à partir de la relation suivante :

$$C_{11} = V[\hat{\alpha}_1] + \text{cov}[\hat{\alpha}_1, \hat{\alpha}_2](y_m + y_k) + V[\hat{\alpha}_2] y_k y_m \quad (\text{VII.15})$$

α_1 et α_2 sont calculées à partir d'un échantillon de n valeurs par la méthode des moindres carrées.

$$\hat{\alpha}_1 = \sum_{i=1}^n d_i C(y_i); \quad \hat{\alpha}_2 = \sum_{i=1}^n a_i C(y_i)$$

avec :

$$d_i = \frac{1}{n} - \bar{y} a_i ;$$

$$a_i = y_i - \bar{y} / \left(\sum_{i=1}^n y_i^2 - n \bar{y}^2 \right)$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i.$$

La figure VIII.9 montre la variation de la cohésion c en fonction de y ainsi que la dérive obtenue par régression linéaire et $c'(x,y)$.

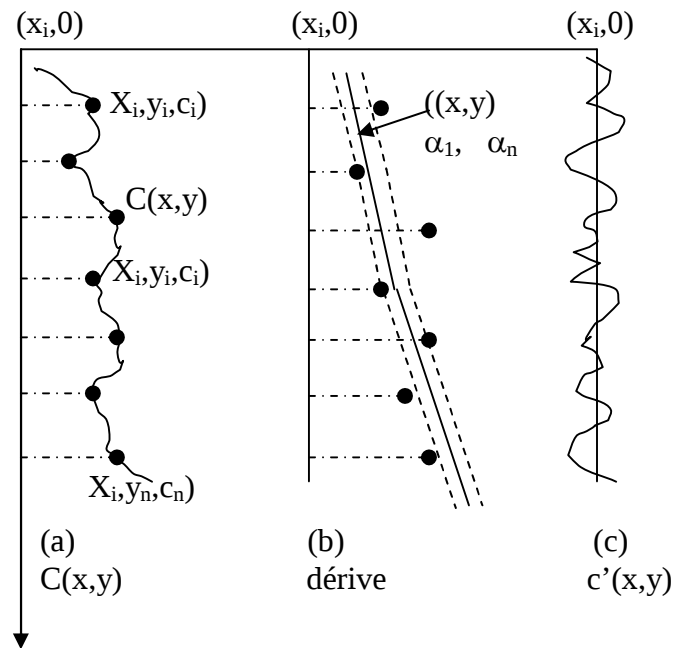


Figure VII.09 Variation de la cohésion du sol

Pour obtenir l'expression de C_{21} , on utilise le fait que :

$$\text{Cov} [X,Y] = E [X,Y] - [X] \cdot [Y]$$

Et que, comme $E [C'(\cdot)] = 0$, on a :

$$\text{Cov} [f(\cdot), C'(\cdot)] = E [f(\cdot) * C'(\cdot)].$$

Comme $f(\cdot) = \alpha_1 + \alpha_2 y$, on obtient en conséquence:

$$\text{Cov} [f(\cdot), C'(\cdot)] = E [\alpha_1 \cdot C'(\cdot)] + [\alpha_2 y_m \cdot C'(\cdot)].$$

Cela permet d'obtenir pour C_{21} l'expression finale :

$$C_{21} = \frac{1}{(\Delta x)^2} \int_{X_{K-1}}^{X_K} \int_{X_{K-1}}^{X_M} \sum_{i=1}^N d_i E[C'(x, a_K x + b_K) \cdot C'(x_0, y_i)] d_x d_y +$$

$$\frac{1}{(\Delta x)^2} \int_{X_{K-1}}^{X_K} \int_{X_{K-1}}^{X_M} (a_m x + b_m) \sum_{i=1}^N a_i E[C'(x, a_K x + b_K) \cdot C'(x_0, y_i)] d_x d_y.$$
VII.16)

Si la fonction d'auto corrélation de C' a une forme exponentielle, on peut écrire (Alonso, 1975 ; Lumb, 1975) :

$$E [C'(x, a_K x + b_K) \cdot C'(x_0, y_i)] = B_c e^{-\alpha_c(\tau)} \quad \text{(VII.17)}$$

avec :

$$\tau = [(x_0 - x)^2 + (y_i - a_K x + b_K)^2]^{1/2}$$

B_c - Variance du paramètre C'

α_c - Degré d'auto corrélation, d'autant plus petit que l'auto corrélation est forte

Dans ce cas, on peut effectuer une intégration partielle et on trouve :

$$C_{21} = \frac{B_c}{\Delta x} \sum_{i=1}^n (d_i + a_i y_m^x) I_i^K.$$

$$I_i^K = \int_{X_{K-1}}^{X_K} \exp \left[-\alpha_c [(x_0 - x)^2 + (y_i - a_K x + b_K)^2]^{1/2} \right] dx. \quad \text{(VII.18)}$$

C_{12} se calcule par la même formule que C_{21} , en inversant m et k .

C_{22} représente la corrélation entre les deux variables centrées C'_k et C'_m .

Si l'on suppose que la droite d'équation $y = ax + b$ passe par les milieux des segments k et m l'expression de C_{22} est la suivante :

$$C_{22} = \frac{1}{(\Delta x)^2} \int_{X_{K-1}}^{X_K} \int_{X_{K-1}}^{X_M} B_c e^{-\alpha_c |(x-y)\sqrt{(1+a^2)}|} dx dy \quad \text{(VII.19)}$$

Pour le paramètre $\text{tg}\phi$, l'analyse est identique à celle qui vient d'être exposé pour c

Pour le poids volumique, on peut supposer que γ est une fonction aléatoire de la profondeur, ce qui permet d'adopter la même démarche que pour la cohésion c .

On suppose en outre que la corrélation entre c et $\text{tg}\phi$ est nulle. Une corrélation existe entre le poids volumique et les paramètres de résistance du sol, mais vu la distance de corrélation (environ 2 m) et les dimensions des tranches analysées, cette corrélation est négligeable. On fait la même hypothèse sur toutes les corrélations des paramètres deux à deux.

VII.4.8 Pression interstitielle

L'expérience prouve qu'il existe une forte autocorrélation entre les pressions interstitielle des différentes tranches. On prend donc $\rho_{ij} = 1$ dans l'expression :

$$\text{cov}[u_i, u_j] = \sigma_u^2 \quad (\text{VII.20})$$

VII.4.9 Géométrie du talus

La précision de relevés topographiques n'ayant pas la raison varier d'une tranche à l'autre, on considère la hauteur des tranches comme une variable aléatoire non autocorrélée, de variance σ_h^2 constante :

$$\text{cov}[h_i, h_j] = \delta_{ij} \sigma_h^2 \quad (\text{VII.21})$$

avec :

$$\delta_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

VII.5 Processus de mobilisation de la résistance au cisaillement

Les valeurs de R_j pour les différentes tranches sont considérées de variance σ_R^2 . D'où

$$\text{cov}[R_i, R_j] = \delta_{ij} \sigma_R^2. \quad (\text{VII.22})$$

VII.5.1 Application de la méthode

Le schéma suivant (figure VIII.11) montre la procédure utilisée pour la recherche du cercle correspondant à la probabilité maximale de rupture : on se fixe à priori un ensemble de cercles à tester. Pour chacun des cercles on calcule la valeur moyenne et la variance du coefficient de sécurité F , ce qui permet de déterminer la probabilité de rupture correspondante, compte tenu de la forme supposée de la loi de distribution de F (loi normale ou loi lognormale).

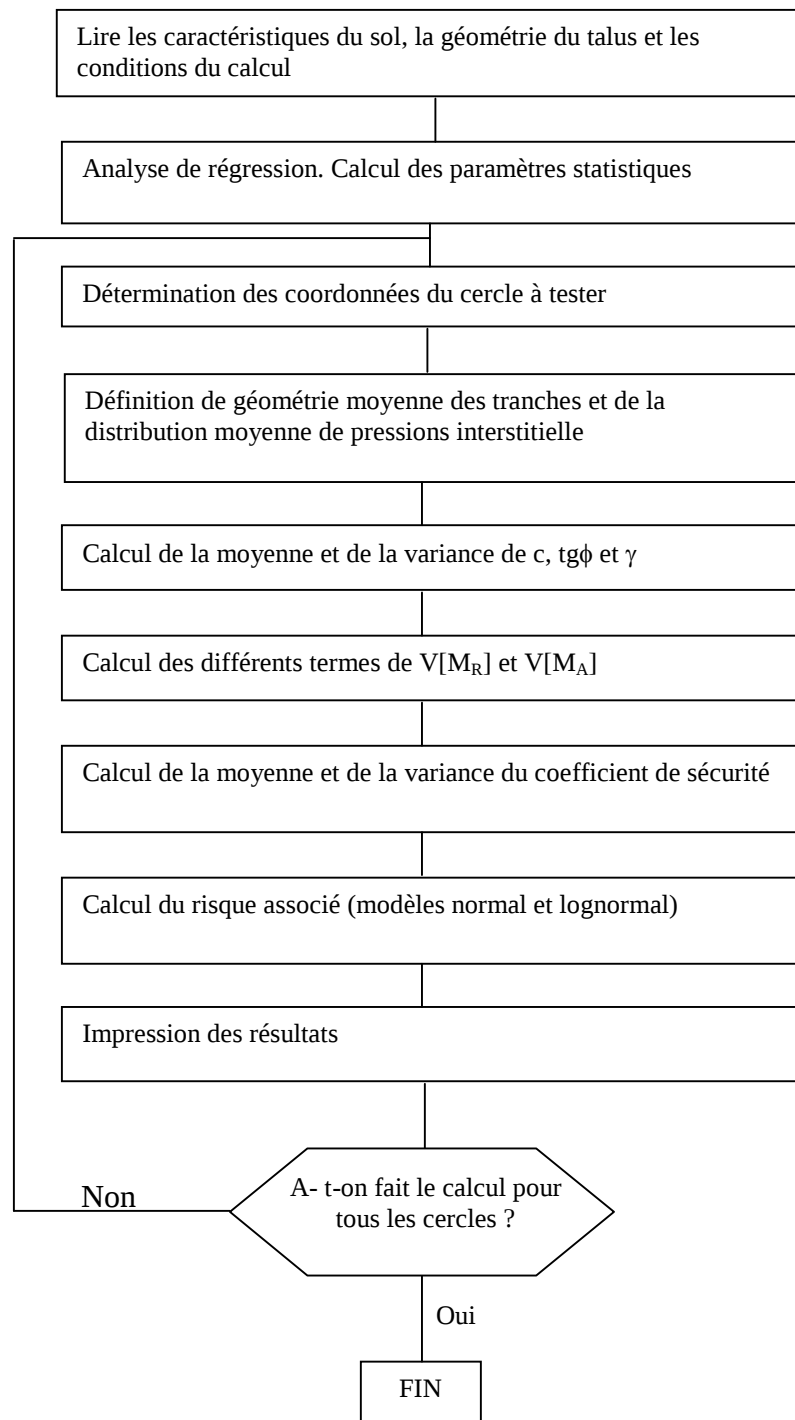


Figure VII.11 Procédure de détermination du cercle de rupture le plus défavorable

Alonso a appliqué cette procédure à l'étude de la stabilité d'une pente dans l'argile sensible d'Ottawa. La géométrie et les paramètres du problème étudié sont indiqués sur la figure VII.12

Les paramètres statistiques de la pression interstitielle ont été calculés à partir des mesures effectuées sur différents piézomètres installés aux points indiqués sur la figure VII.13

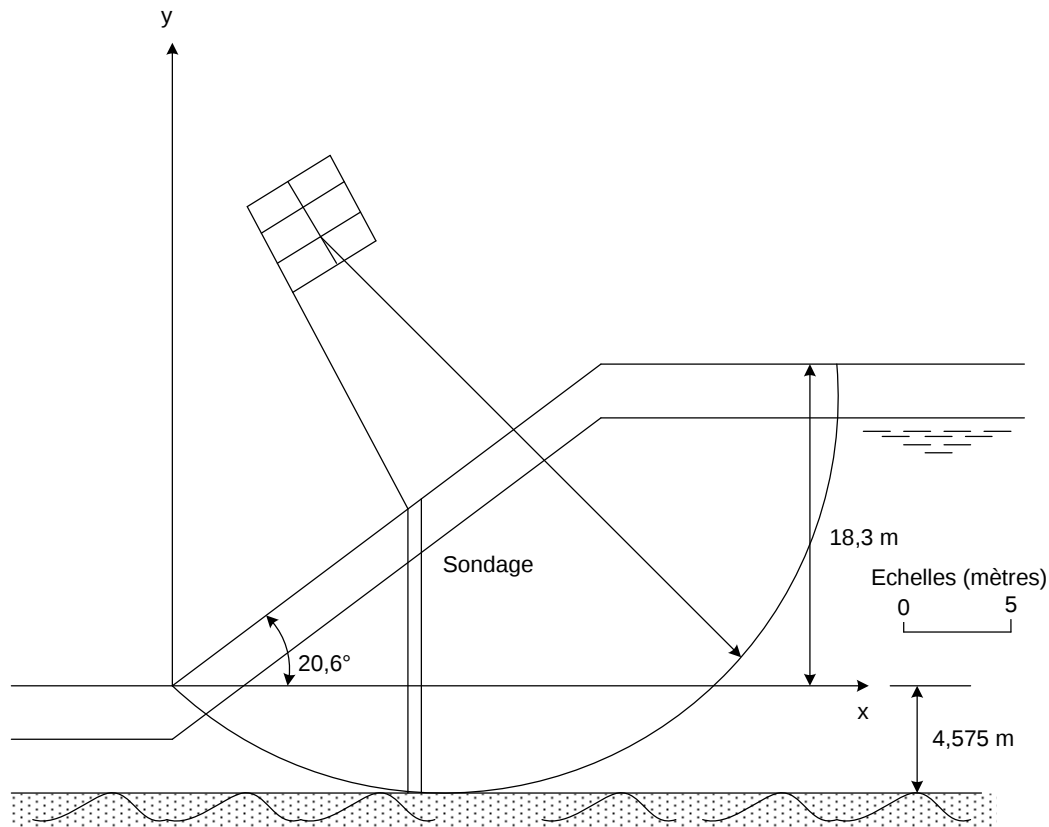


Figure VII.12 Géométrie et paramètre du cas étudié par Alonso (1976)

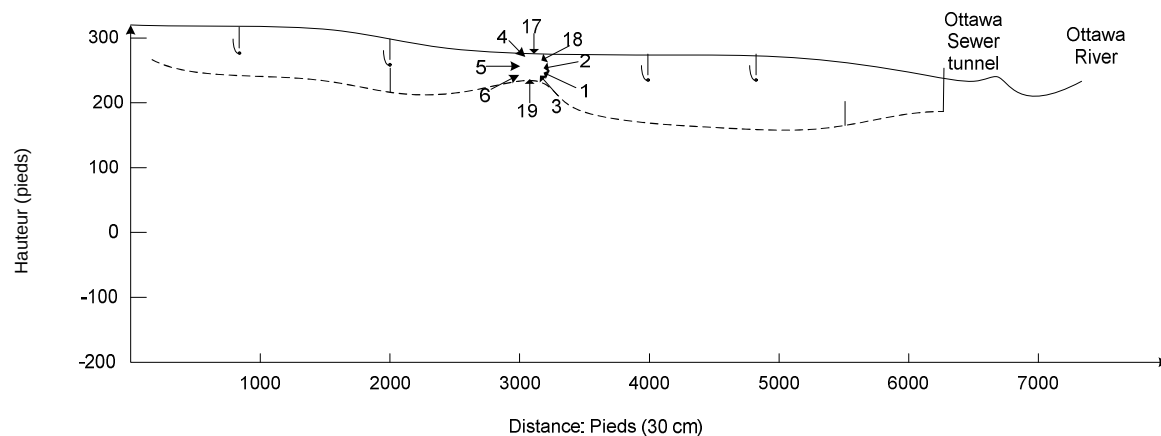


Figure VII.13 Position des piézomètres et conclusions des mesures

Comme le calcul pour chaque cercle de la probabilité de rupture augmente sensiblement la durée des calculs, on peut profiter du fait empirique (sur la base des résultats complets d'Alonso) que le minimum des coefficient de sécurité et le maximum des probabilités de rupture coïncident pour déterminer d'abord le cercle de coefficient de sécurité minimal puis effectuer le calcul de probabilité sur ce cercle unique. Alonso a effectué le calcul de

probabilité pour chaque cercle calculé et obtenu les résultats représentés sur la figure VII.46 sous forme de courbes d'égale probabilité de rupture et d'égale coefficient de sécurité. On notera que les courbes d'égale probabilité de rupture diffèrent suivant la loi de probabilité adoptée pour F (loi normale ou loi lognormale), mais que, dans les deux cas, il semble que le cercle critique soit le même pour l'approche déterministe classique et pour l'approche probabiliste.

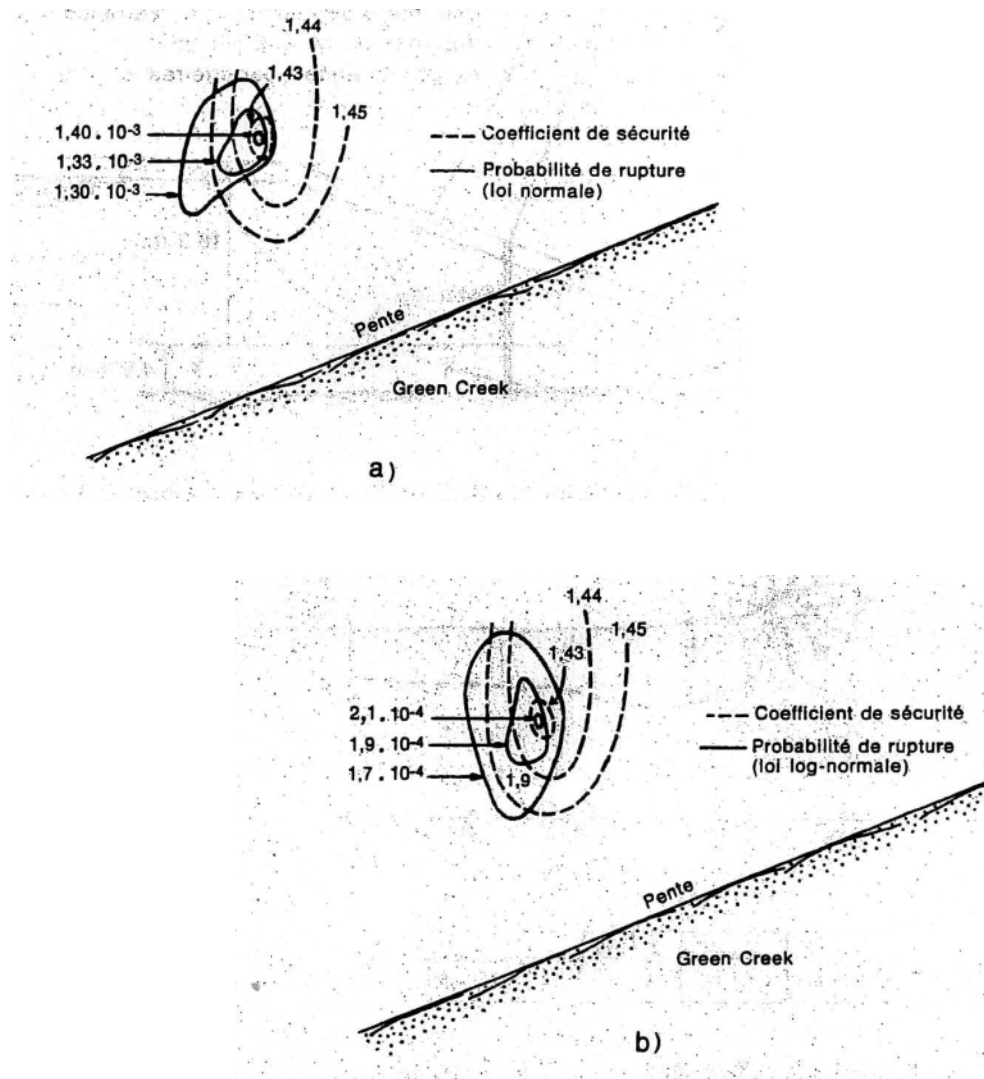


Figure VII.14 Courbes d'égale probabilité de rupture et d'égale coefficient de sécurité

Alonso a également étudié la contribution des différents facteurs à la variance des moments actifs et résistants dans cet exemple de Green Creek.

Il a trouvé que, l'incertitude sur le résultat provient essentiellement de la pression interstitielle u et de la cohésion c (figure VIII.15), l'incertitude étant beaucoup plus, grande pour le moment résistant que pour le moment des forces motrices.

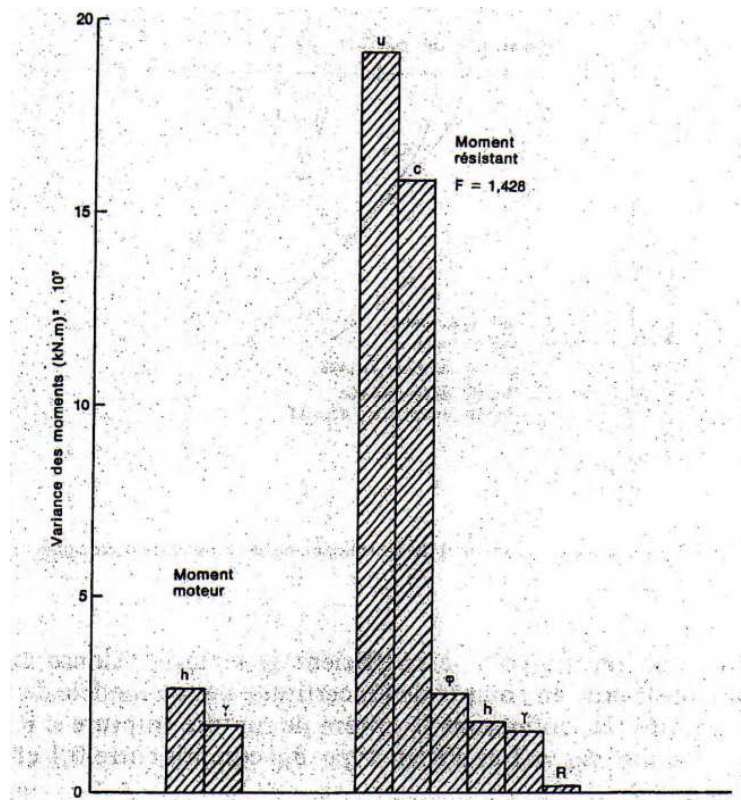


Figure VII.15 Contributions de différents facteurs à la variance des moments (cercle critique)

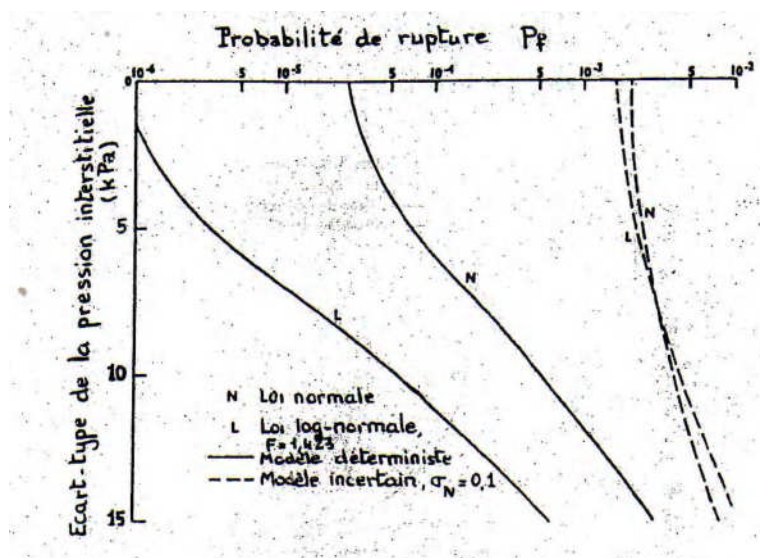


Figure VII.16 Influence de la pression interstitielle sur la probabilité de rupture

Les figures VII.16 et VII.17 montrent l'influence des variations de u et de c et ϕ sur la probabilité de rupture, suivant la forme normale ou log-normale que l'on retient pour le coefficient de sécurité.

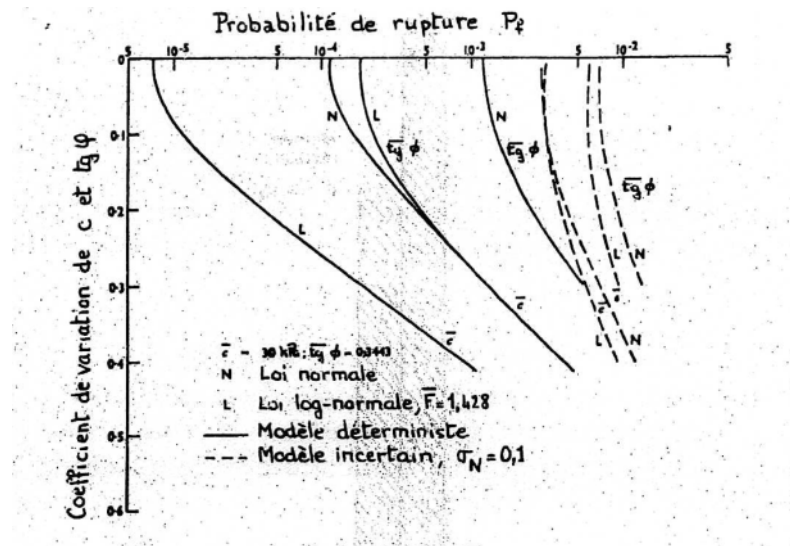


Figure VII.17 Influence de la cohésion et du frottement sur la probabilité de rupture

Alonso indique à ce propos que la prise en compte de l'incertitude liée au modèle de calcul (variable aléatoire N définissant le critère de rupture (rupture si $F < N$), de valeur moyenne $\mu_N = 1$ et d'écart type σ_N compris entre 0,1 et 0,15) a trois effets :

- le risque de rupture se trouve augmenté,
- la probabilité de rupture varie moins,
- les résultats deviennent peu sensibles à la distribution adoptée pour F .

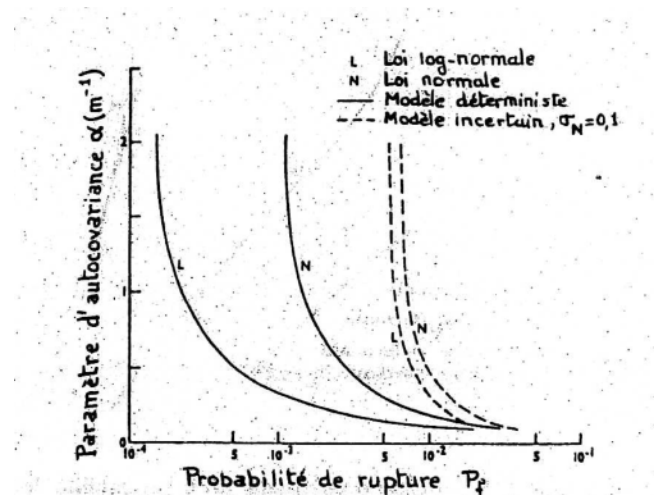


Figure VII.18 Influence du degré d'autocorrélation des propriétés du sol sur la probabilité de rupture.

Alonso a également étudié l'influence du degré d'autocorrélation des propriétés du sol (α) sur la probabilité de rupture. Les résultats montrent (figure VII.18) que, plus α est faible

c'est-à-dire plus les propriétés du sol sont liées dans l'espace, plus la probabilité de rupture est grande. D'autre part, on voit qu'au-delà d'un certain seuil ($\alpha > 0,75 \text{ m}^{-1}$, par exemple), la variation du paramètre devient si rapide qu'elle n'a plus d'influence sur le coefficient de sécurité, qui résulte d'une intégration sur une surface assez étendue.

Une conclusion importante de l'étude d'Alonso est qu'il n'existe pas de relation unique entre le coefficient de sécurité et la probabilité de rupture, puisque le coefficient de sécurité dépend essentiellement des valeurs moyennes, tandis que la probabilité de rupture dépend également des coefficients de variation. La figure VIII.19 présente les relations obtenues dans l'hypothèse où seules les valeurs moyennes des paramètres varient, les coefficients de variation restant constants. Sur la figure VIII.19 ont été ajoutées les courbes proposées par Meyerhof (1970), sur la base d'observations empirique et de son expérience. Cette figure peut être utilisée en première approximation pour l'analyse des glissements dans l'argile Champlain.

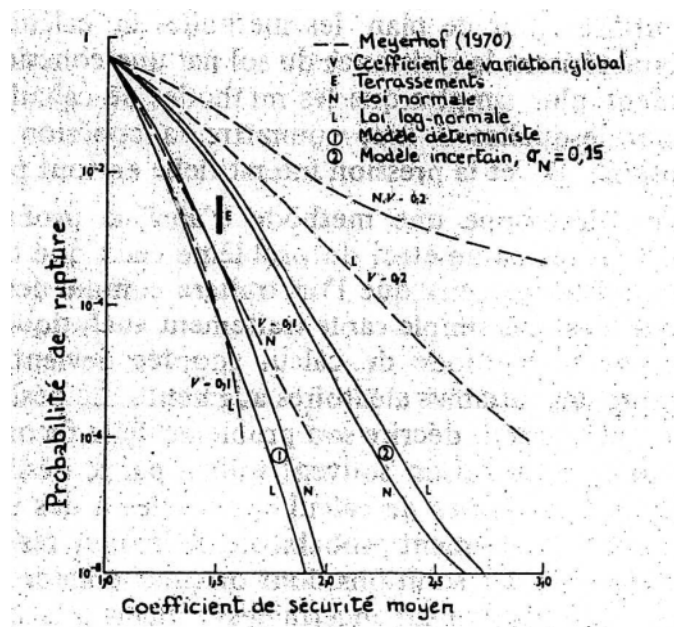


Figure VII.19 Relation entre le coefficient de sécurité et la probabilité de rupture.

Baghery (1980) a appliqué la méthode développée par Alonso à l'étude de la stabilité du remblai A du site expérimental de remblais sur sols compressibles de Cubzac-les-Ponts (Gironde)

VII.6 PRINCIPES GENERAUX DES METHODES D'ANALYSE PROBABILISTE DE LA STABILITE DES PENTES

Comme on a pu le constater dans l'exemple précédent, l'analyse probabiliste de la stabilité d'une pente fait intervenir trois éléments d'importance égale :

- La description des variations naturelles des propriétés physiques et mécaniques des sols et des conditions géométriques du problème traité,
- une méthode de calcul déterministe de la stabilité des pentes
- et une méthode de traitement analytique ou numérique des fonctions de variables aléatoires.

Le choix des paramètres de calcul est lié à celui de la méthode de calcul déterministe utilisée. Sur ce plan, les méthodes de calcul en contraintes totales, qui caractérisent la résistance du sol par une cohésion non drainée, sont évidemment plus simples que les méthodes de calcul en contraintes effectives, pour lesquelles il faut connaître la cohésion c , l'angle de frottement interne ϕ et la pression interstitielle en tout point.

Lorsque l'on développe une méthode d'analyse probabiliste, il faut choisir parmi tous les paramètres du problème ceux que l'on considérera comme déterministes et ceux que l'on traitera comme des variables aléatoires. Ce choix n'est pas simple car le traitement analytique ou numérique des équations de la méthode de calcul adoptée devient très complexe quand le nombre des variables aléatoires augmente. Le désir d'exhaustivité du chercheur, qui voudrait décrire son problème de la façon la plus précise possible, est pour cette raison souvent vaincu par le désir contradictoire de développer des méthodes de calcul applicables à des problèmes réels.

A l'autre extrême, le traitement probabiliste de la seule résistance au cisaillement du sol fournit des solutions dont on peut craindre qu'elles soient optimistes (ou «insuffisamment incertaines») quant à la stabilité réelle de l'ouvrage.

Seules des analyses détaillées, telles que celle de la figure VIII.15 précédente, peuvent permettre de faire des choix raisonnables. On voit sur cette figure que, pour les calculs en contraintes effectives, l'incertitude sur la stabilité réelle de la pente dépend principalement de l'incertitude sur les pressions interstitielles et de l'incertitude sur la cohésion du sol. Pour simplifier le traitement numérique de la stabilité des pentes, on pourrait donc limiter à deux le nombre des variables aléatoires, en intégrant dans l'incertitude sur le modèle de calcul l'influence des variations des autres paramètres. Le bien-fondé d'un tel choix doit toutefois être vérifié dans chaque cas, en l'état actuel de nos connaissances.

Une fois que l'on a choisi les variables aléatoires du calcul, il faut en déterminer les caractéristiques statistiques, ce que l'on peut faire d'après des résultats de mesures ou d'essais ou en utilisant des valeurs publiées dans la littérature. Outre la moyenne et l'écart type, il est indispensable de déterminer la dérive éventuelle des variables aléatoires dans l'espace et leur fonction d'autocorrélation, ainsi que les corrélations multiples entre ces variables. Faute de tenir compte de ces caractéristiques de la variabilité naturelle des sols, on risque d'obtenir des résultats d'intérêt pratique à peu près nul.

La méthode de calcul déterministe qui constitue l'intermédiaire obligatoire, entre les hypothèses du calcul et la probabilité de rupture que l'on cherche à déterminer doit être choisie avec soin. Les méthodes de calcul «explicites», dans lesquelles le résultat est une fonction explicite des hypothèses du calcul, sont en général plus simples à traiter que les méthodes de calcul dans lesquelles le résultat provient d'un calcul itératif.

Deux critères peuvent être utilisés pour caractériser la stabilité (ou l'instabilité) des pentes :

- la marge de sécurité M_S , égale à la différence entre le moment résistant et le moment moteur, peut être comparée à 0 (ou à une valeur aléatoire représentant l'incertitude sur le modèle de calcul) ;
- le coefficient de sécurité F , égal au rapport du moment résistant au moment moteur, peut être comparé à 1 (ou à une valeur aléatoire représentant l'incertitude sur le modèle de calcul).

Les deux méthodes peuvent être utilisées et sont en pratique équivalentes. Il semble toutefois que le coefficient de sécurité F soit d'un emploi plus fréquent.

Pour déterminer la distribution statistique du résultat du calcul déterministe, qu'il s'agisse de la marge de sécurité ou du coefficient de sécurité, on peut utiliser l'une des quatre méthodes décrites au paragraphe précédent. La méthode de Monte-Carlo est fréquemment utilisée parce qu'elle est simple, mais elle demande un volume de calculs très important et nécessite une procédure adaptée pour la génération des ensembles de valeurs aléatoires des paramètres. Son grand mérite est d'être utilisable même quand le problème analysé n'admet pas de solution explicite. On a toutefois toujours intérêt à examiner s'il n'est pas possible d'utiliser en priorité les autres méthodes.

VII.7 PERSPECTIVES

Les exemples d'analyse probabilistes décrits ci-dessus reposent sur une schématisation très simple (unidimensionnelle) du comportement des sols sous les ouvrages. Si l'on admet que les analyses probabilistes ne peuvent avoir beaucoup plus de valeur que les schémas déterministes sous-jacents, la simplicité des modèles utilisés pour ces études limite la confiance que l'on peut leur accorder en pratique.

Dans le cas analysé par Magnan et Baghery (1982), il est frappant de constater par exemple que le tassement observé sous le remblai est très nettement supérieur aux valeurs calculées lors de l'analyse probabiliste

Figure VIII.20.

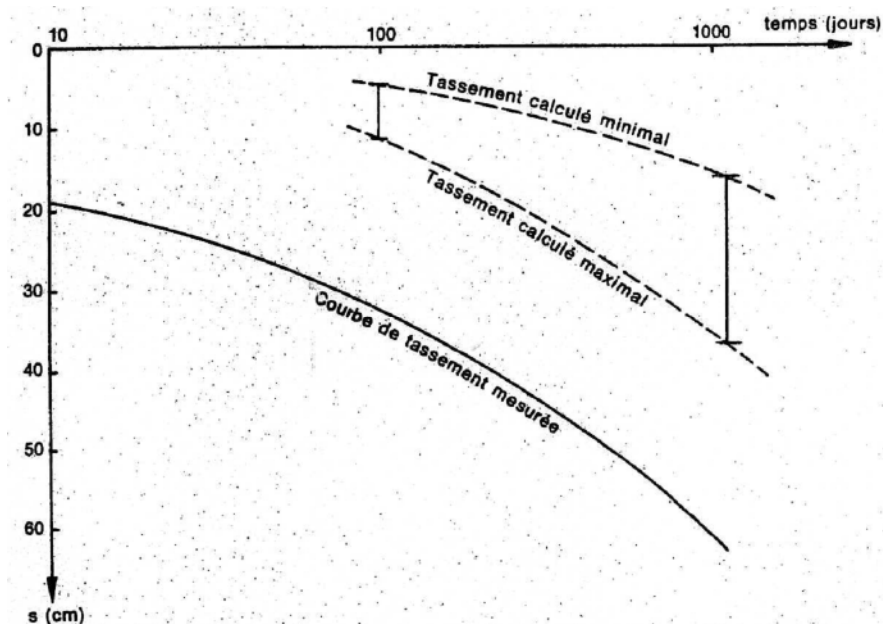


Figure VIII.20 Comparaison des mesures et du résultat de l'analyse probabiliste
(Remblai B de Cubzac-les-Ponts, Magnan et Baghery, 1982).

Cette divergence est certainement due pour partie à l'imperfection du schéma de calcul unidimensionnel adopté, mais elle a aussi une autre origine : l'ordre de grandeur du coefficient de consolidation utilisé pour le calcul est faux, par suite des imperfections des méthodes utilisées pour le déterminer en laboratoire. Il est d'ailleurs plutôt rassurant de constater que l'introduction de techniques statistiques et probabilistes dans des secteurs où les méthodes de calcul déterministes traditionnelles sont globalement considérées comme satisfaisantes ne rend pas inutile le perfectionnement des méthodes de mesure des propriétés des sols et de calcul des ouvrages.

VII.8 ANALYSE DE LA CONSOLIDATION BIDIMENSIONNELLE DES SOLS

VII.8.1 Introduction

Malgré le développement d'outils de calculs (méthodes numériques) de plus en plus performants et une description de sol de plus en plus proche de la réalité, il existe toujours un écart entre les résultats de calcul et les valeurs des tassements et des surpressions interstitielles réels à la fois en laboratoire et in situ.

C'est dans l'espoir de réduire ce biais existant, ou, tout au moins de progresser dans la compréhension de ces divergences, que les géotechniciens poussent leurs investigations dans l'application des statistiques et probabilités, avec comme objectifs de prendre en compte l'effet de la variabilité des propriétés physiques et mécaniques de sol dans l'analyse probabiliste de la consolidation qui reste encore un grand sujet d'intérêt.

Les études traitant de la consolidation des sols d'un point de vue probabiliste sont toutefois rares.

Il faut noter que pour la plupart de ces études, seul l'aspect unidimensionnel de la consolidation a été pris en compte. Dans beaucoup de cas, la théorie de la consolidation unidimensionnelle sous-estime la vitesse de tassement des sols fins. Cette sous-estimation provient, dans notre cas, du caractère unidimensionnel de la déformation.

VII.8.2 Maillages et hypothèses de calcul

Dans l'évaluation de l'influence de l'effet de la variabilité des propriétés de consolidation, la déformabilité E et la perméabilité k , sur les résultats de calculs, en prenant en compte la déformation bidimensionnelle dans l'analyse probabiliste de la consolidation, nous nous proposons de calculer les tassements et les surpressions interstitielles à chaque pas de temps et en des points préalablement choisis d'un massif de sol soumis à une charge uniforme en surface.

Les deux paramètres E et k ont été supposés indépendants. Leurs valeurs ont été créées au moyen d'un générateur de nombres aléatoires, sur la base d'une loi lognormale, plus représentative que la loi normale (littérature et réalité physique), avec des valeurs moyennes et coefficients statistiques théoriques déduits des études d'Haghighou (1983) sur l'argile molle de Cubzac-Les-Ponts et consignés dans le tableau VIII.3.

Tableau VII.3 Coefficients statistiques des paramètres de calcul

	E	K
<i>Moyenne</i>	520 kPa	$11,4048.10^{-5}$ m/j
<i>Variance</i>	10816 (kPa)^2	$6,37.10^{-9} \text{ (m/j)}^2$
<i>Coeff. De variation</i>	0,20	0,70

Le type de maillage utilisé dans les calculs représente la moitié du sol de fondation à cause de la symétrie du problème.

La figure **VII.21** illustre bien le caractère bidimensionnel de la déformation. Elle montre :

- un massif de sol soumis à une charge uniforme en surface. La charge est appliquée progressivement pendant un jour et est maintenue constante jusqu'à la fin de la consolidation, supposée achevée au bout de 20.000 jours ;
- des propriétés (module de Young E, coefficient de perméabilité k) variables,
 - verticalement, d'une couche à une autre, mais constants dans une même couche (figure VII.21a),
 - aléatoirement dans le massif de sol, d'une couche à une autre et dans une même couche horizontale, d'un élément à un autre (figure VII.21b).

Le maillage initial est constitué de 576 éléments quadrilatères à 8 nœuds et de 1833 nœuds. Il est limité horizontalement, à $x=0$ m (axe du maillage) et $x = +40$ m. Le substratum, imperméable, est à -10 m et le terrain naturel à 0 m.

Les conditions limites suivantes ont été adoptées :

- surpressions interstitielles nulles le long de la surface de la nappe ($x = 0$),
- déplacements horizontaux u nuls sur le bord vertical du maillage, situé à $x = 40$ m,
- déplacements horizontaux et verticaux nuls sur le bord inférieur du maillage (surface non drainante).

Toutes les autres caractéristiques du sol ont été supposées déterministes, de même que la géométrie et la charge appliquée.

VII.8.3 Méthode d'analyse et outils de calculs

Comme pour l'analyse probabiliste de la consolidation unidimensionnelle, analysée précédemment par les mêmes auteurs, la méthode des éléments finis et la technique de Monte Carlo sont appliquées à un sol de fondation dont la moitié du maillage est représenté dans la figure VII.21, pour la résolution de la consolidation bidimensionnelle dans les deux cas suivants :

- variabilité unidimensionnelle des paramètres E et k : dix valeurs sont nécessaires pour chaque calcul. Les valeurs des paramètres attribuées à chaque calcul sont variables d'une couche à une autre mais constantes dans une même couche ;

- variabilité bidimensionnelle de ces paramètres : les paramètres de calcul varient aléatoirement dans le massif de sol, d'une couche à une autre et dans une même couche horizontale, d'un élément à un autre. 360 valeurs sont nécessaires à l'exécution d'un calcul, soit 36000 valeurs pour l'ensemble des 100 calculs.

La méthode numérique est ici la méthode des éléments finis ; nous avons utilisé le programme Rosalie Groupe 09 du LCPC, développé pour la résolution des problèmes de consolidation uni- et bidimensionnelle.

La méthode statistique est la technique de simulation de Monte Carlo, technique la plus performante à l'heure actuelle.

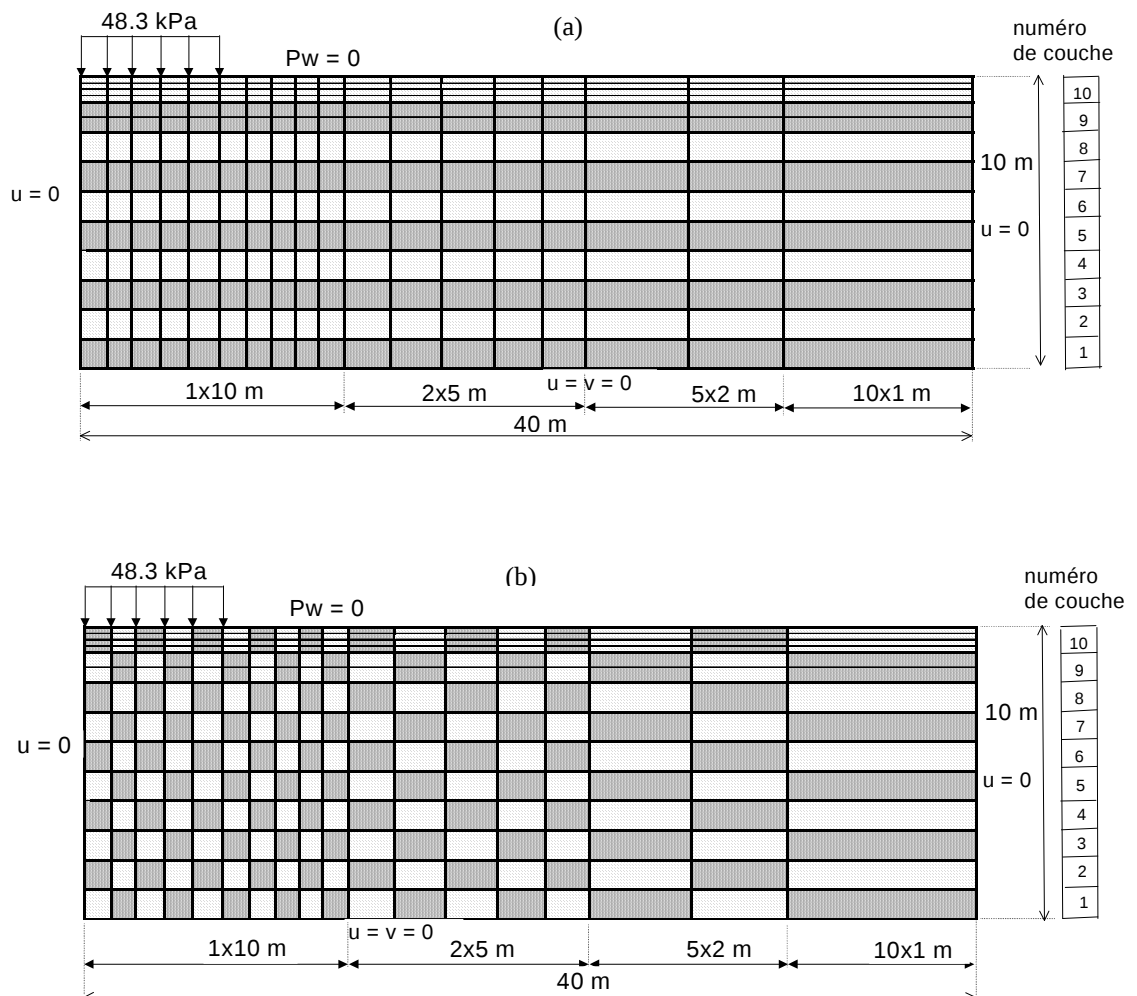


Figure VII.21 Maillages utilisés pour les calculs en éléments finis : maillage (a) pour le calcul de la variabilité uniquement verticale ; maillage (b) pour le cas de la variabilité verticale et horizontale.

▪ L'outil de calcul déterministe

Pour pouvoir prendre en compte le caractère hétérogène du sol au travers de ses propriétés physiques et mécaniques, et donc pouvoir décomposer le massif en éléments plus petits avec des valeurs de la déformabilité et de la perméabilité, nous utilisons la méthode des éléments finis au travers du Groupe 09, Rosalie qui permet le calcul du comportement au cours du temps des massifs de sol saturé d'un liquide interstitiel compressible ou incompressible.

▪ La démarche probabiliste : méthode de Monte Carlo

Cette méthode repose sur la simulation numérique répétée de l'écoulement à l'intérieur d'un arrangement déterministe d'éléments de sol en forme de couches, rectangles ou blocs. Chacun de ces éléments est considéré comme un massif homogène auquel est affectée une valeur particulière de la propriété de sol, fixée en utilisant la fonction de densité de probabilité de la propriété considérée et sa loi de variabilité spatiale.

Une solution déterministe par méthode numérique est obtenue pour chaque ensemble de valeurs des paramètres du calcul. La répétition de ce processus de calcul permet d'établir la loi de distribution de la solution du problème.

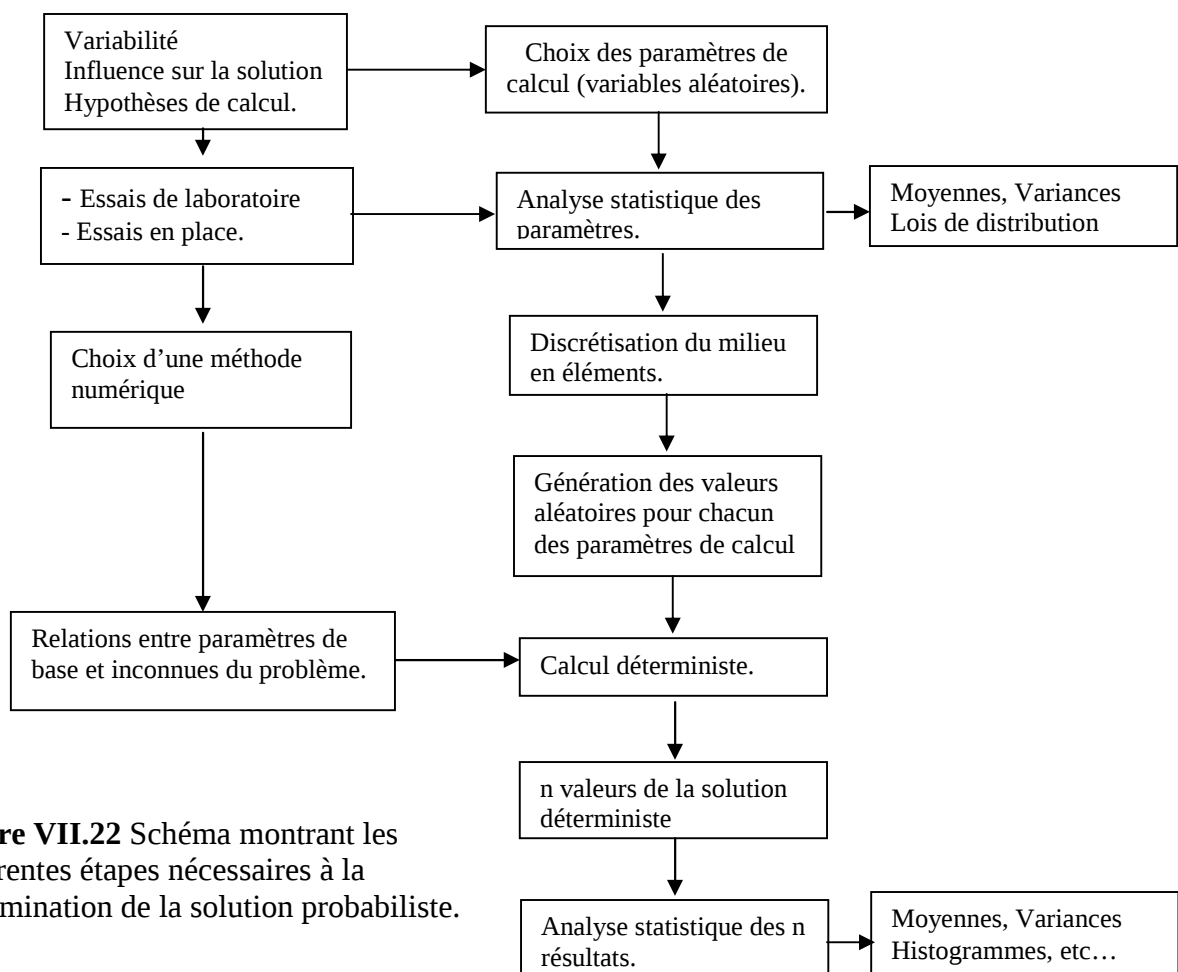


Figure VII.22 Schéma montrant les différentes étapes nécessaires à la détermination de la solution probabiliste.

VII.8.4 Résultats des calculs

1. Démarche

Nous avons calculé les tassements, les déplacements horizontaux et les surpressions interstitielles ainsi que les coefficients statistiques correspondants, à chaque pas de temps dans un massif homogène de sol soumis à une charge uniformément répartie et reposant sur un substratum imperméable, en tous des points du massif préalablement choisis, représentés dans les figures VII.23 et VII.24.

Par ailleurs, trois calculs déterministes sont effectués, pour comparer leurs résultats à ceux obtenus dans le cas de l'analyse probabiliste : le premier utilise la moyenne arithmétique, le second la moyenne géométrique et le troisième la moyenne harmonique des coefficients de perméabilité, tandis que la moyenne arithmétique du module de Young est utilisée dans les trois cas.

Deux séries de cent calculs par éléments finis ont été effectuées, en utilisant les maillages représentés sur la figure VII.21.

D'abord la variabilité des propriétés des sols est supposée unidimensionnelle (valeurs uniformes dans chaque plan horizontal et les surpressions interstitielles se dissipent verticalement). Ensuite, nous analyserons le cas d'une variation bidimensionnelle des propriétés du sol. Dans ce cas, le massif de sol tasse et se déforme latéralement alors que l'eau interstitielle emprunte des chemins d'écoulement préférentiels dans tout le volume du sol.

La méthode de Monte Carlo a été utilisée pour estimer la distribution des résultats.

2. Variabilité unidimensionnelle

Dans ce cas les déplacements sont bidimensionnels (tassements et déplacements horizontaux) et les surpressions interstitielles se dissipent verticalement. Dix valeurs de E et dix valeurs de k sont nécessaires pour chaque calcul et 1000 valeurs de chacun des paramètres pour l'ensemble des 100 calculs.

▪ Résultats

Les résultats, en termes d'évolution du tassement, des surpressions interstitielles, des déplacements horizontaux et des coefficients statistiques correspondants (écarts-types et coefficients de variation), en fonction du temps ou en fonction de la profondeur, sont donnés sous forme de courbes, montrées dans les figures VII.23 et VII.24.

On observe les résultats suivants :

- L'allure des courbes obtenues, traduisant l'évolution, au cours du temps ou en fonction de la profondeur, du tassement, des surpressions interstitielles ou des déplacements horizontaux, est harmonieuse et est semblable à celle des courbes du même type, déterminées dans les études classiques de la consolidation ;

- La valeur instantanée du tassement est élevée ;

Si les courbes de tassement, en fonction de la profondeur ou au cours du temps, ont un mode de variation identique, les isochrones des surpressions interstitielles prises à des valeurs de temps faibles (10 et 100 jours) diffèrent du reste des courbes du même type. Cela semble traduire qu'en début de consolidation l'eau interstitielle « cherche » ses chemins préférentiels dans le massif ;

- L'incertitude sur le tassement varie peu avec la profondeur et en tout point du massif en début de consolidation et jusqu'à 5000 jours. Le coefficient de variation, peu variable durant la consolidation, est de l'ordre de 6% ;

- A la fin de la consolidation, on note une grande dispersion des valeurs calculées des surpressions interstitielles. Dans ce cas, le coefficient de variation correspondant peut dépasser 100% ;

- Les déplacements horizontaux ne sont pas négligeables et la valeur de l'incertitude relative correspondante est assez faible.

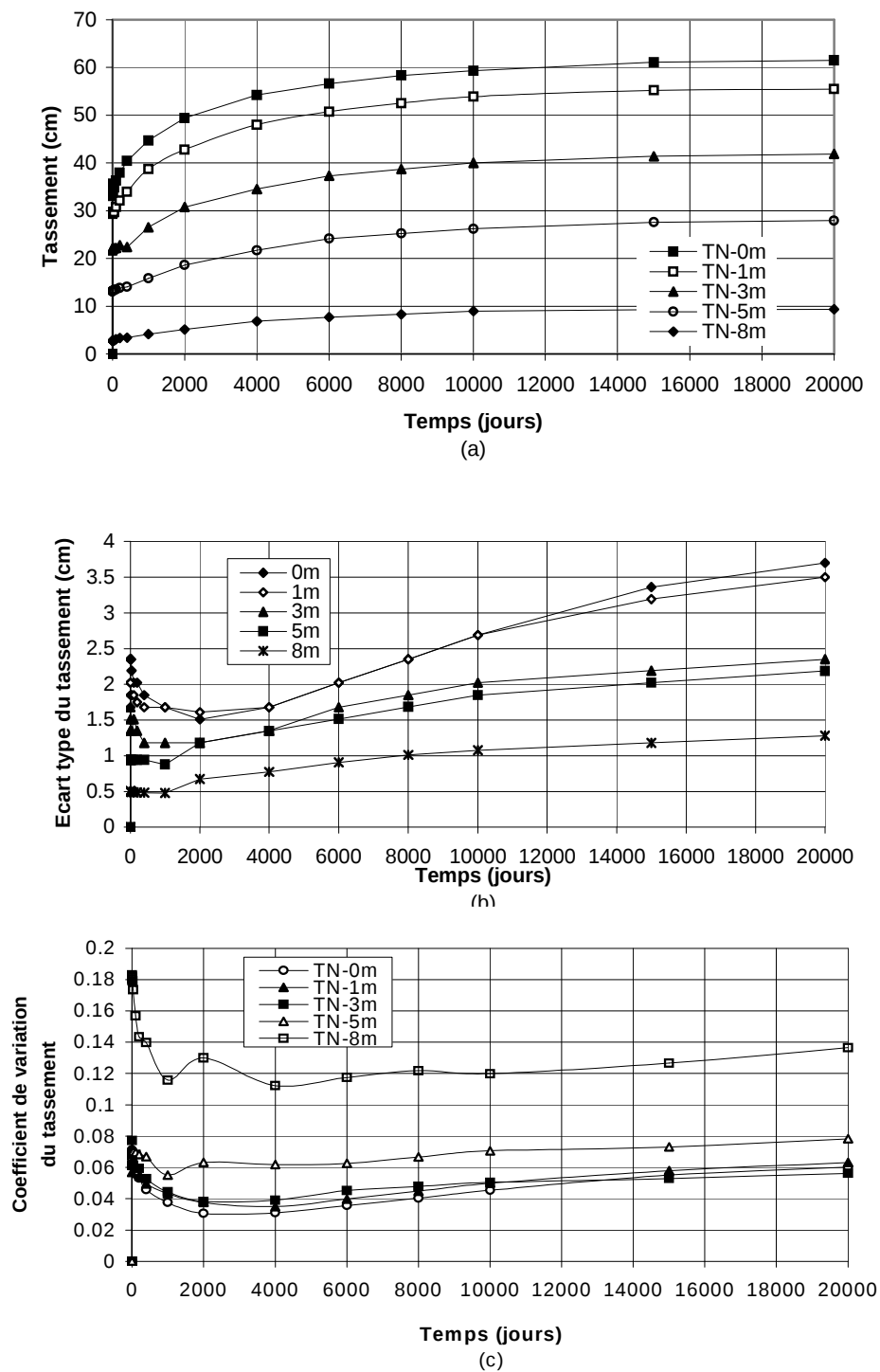


Figure VII.23 Evolution au cours du temps du tassement de plusieurs points du massif (a), des écarts-types (b) et des coefficients de variation (c) (variabilité unidimensionnelle).

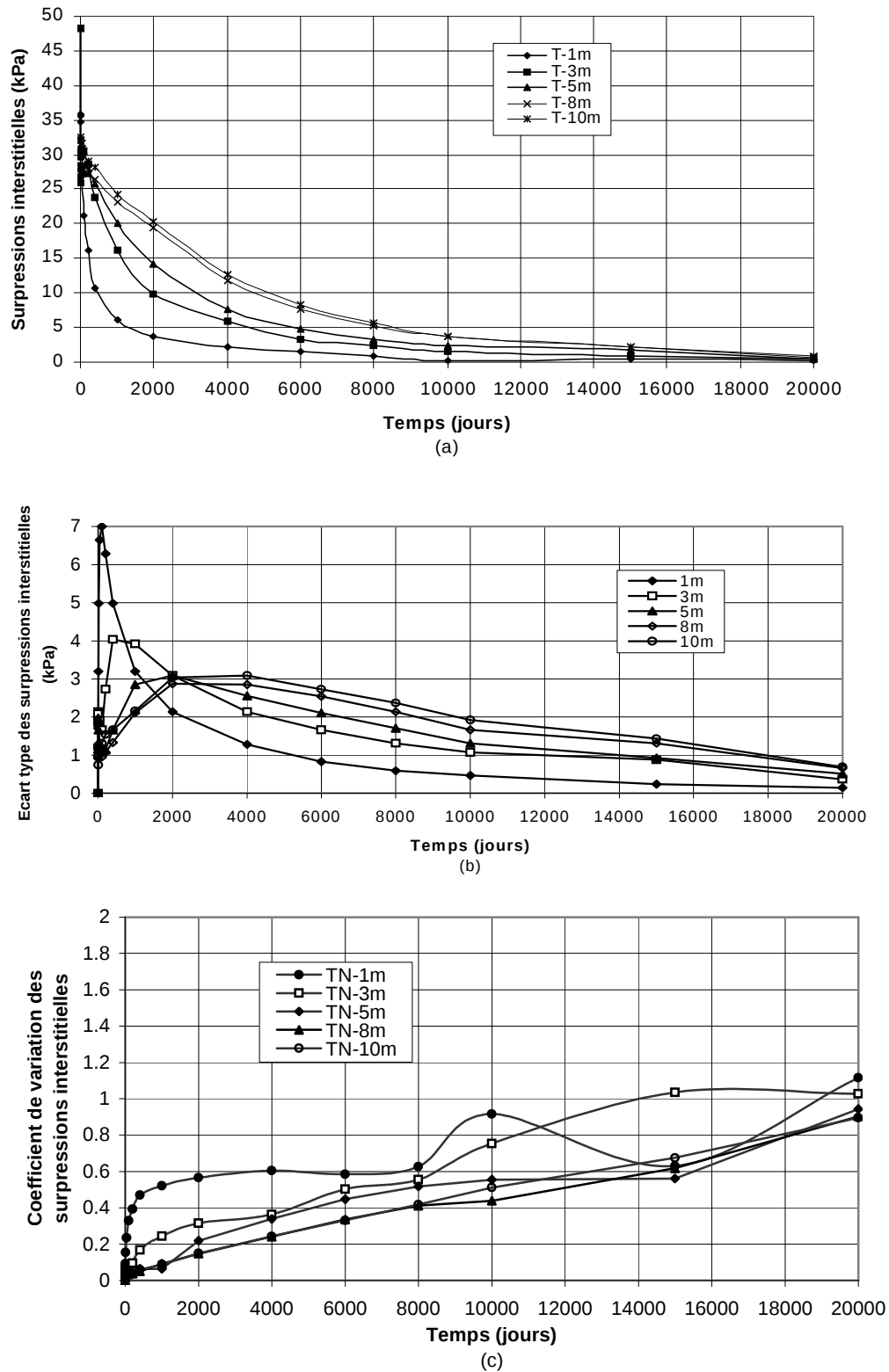
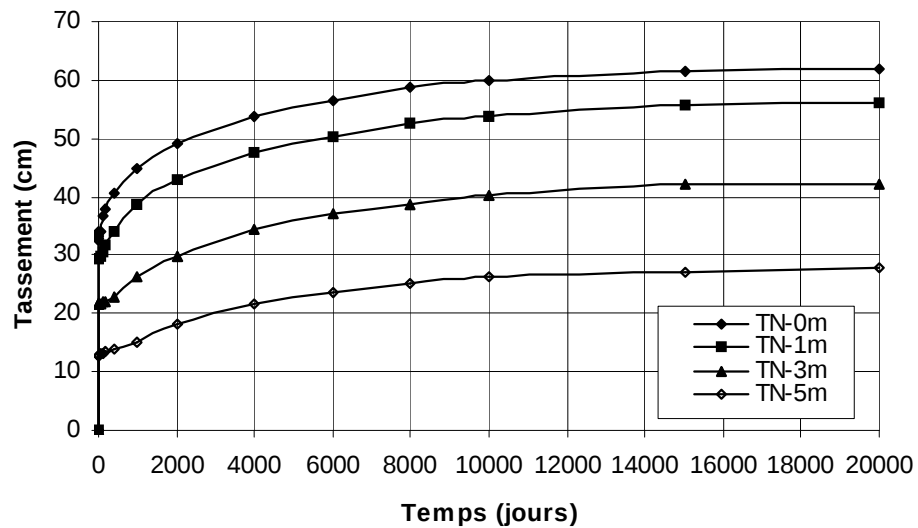
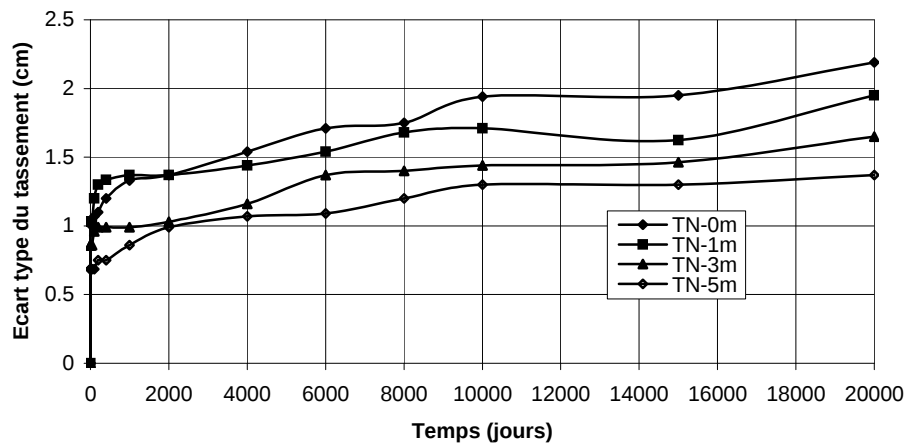


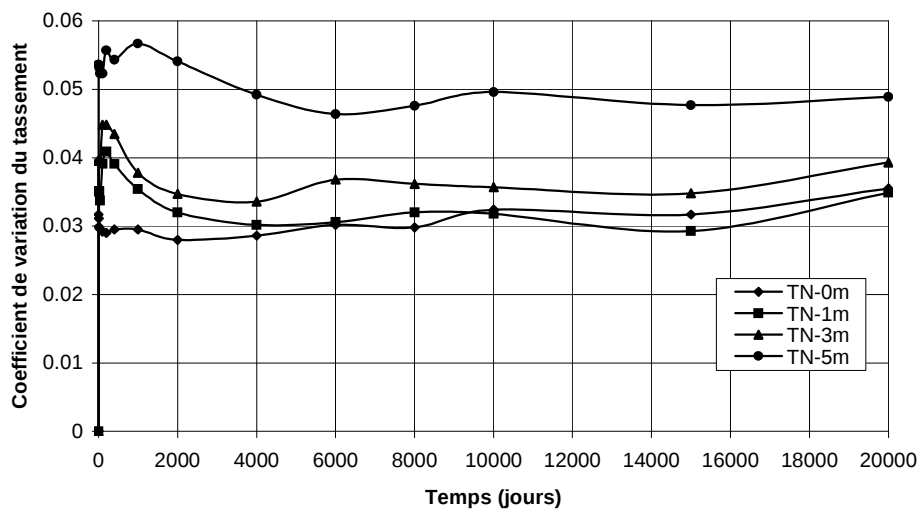
Figure VII.24 Evolution au cours du temps des surpressions interstitielles moyennes en plusieurs points du massif (a), des écarts-types (b) et des coefficients de variation (c) (variabilité unidimensionnelle).



(a)



(b)



(c)

Figure VII.25 Evolution au cours du temps du tassement de plusieurs points du massif (a), des écarts-types (b) et des coefficients de variation correspondants (c) (variabilité bidimensionnelle).

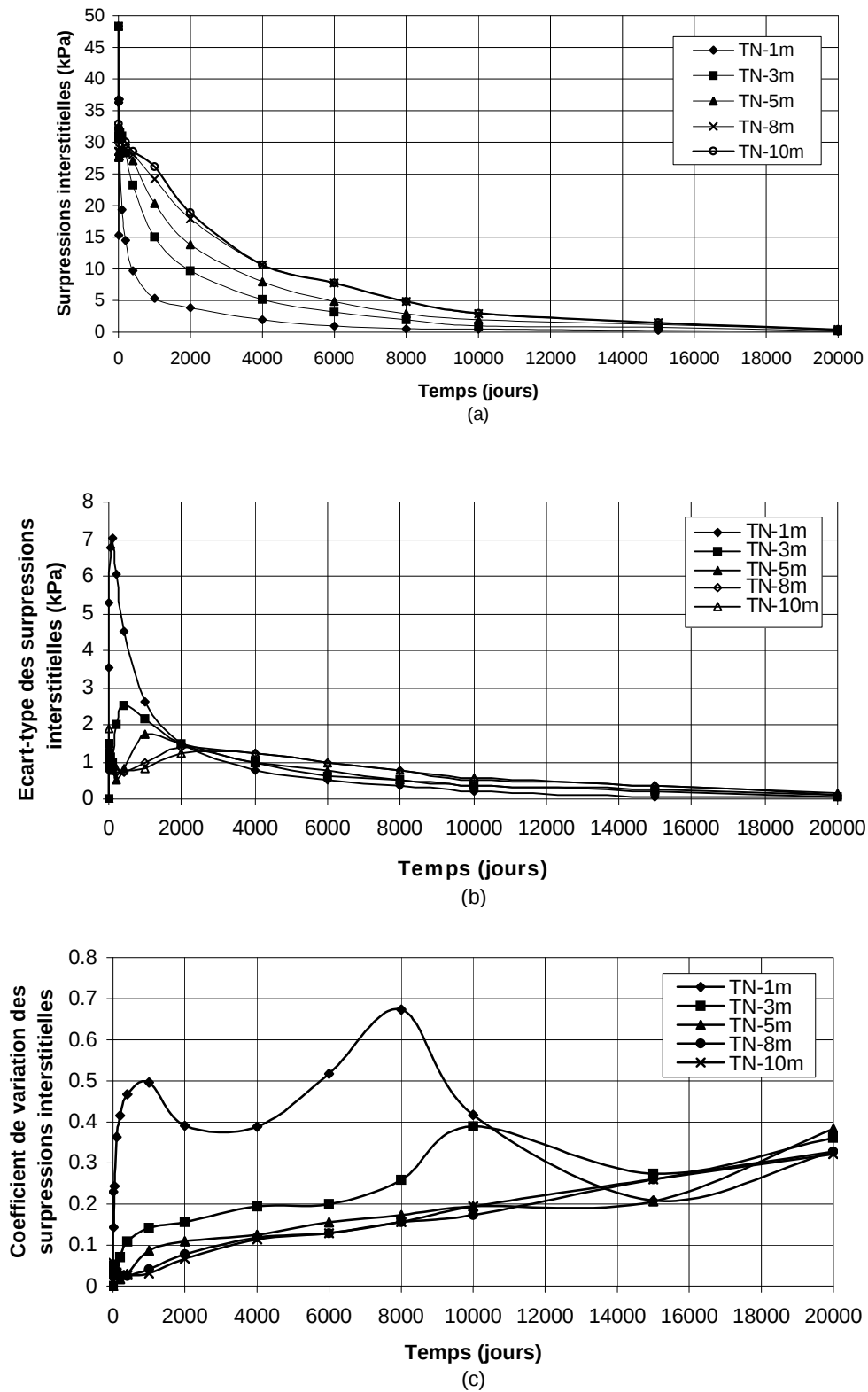


Figure VII.26 Evolution au cours du temps des surpressions interstitielles moyennes en plusieurs points du massif (a), des écarts-types (b) et des coefficients de variation correspondants (c) (variabilité bidimensionnelle).

3. Variabilité bidimensionnelle

Dans ce cas de l'analyse probabiliste de la consolidation bidimensionnelle à variabilité bidimensionnelle, les paramètres de calcul, le module d'élasticité E et la déformabilité k , varient aléatoirement dans le massif de sol, d'une couche à une autre et dans une même couche horizontale, d'un élément à un autre.

Le massif de sol étant divisé en 10 couches d'épaisseur égale, 360 valeurs de chacun des paramètres sont nécessaires à l'exécution d'un calcul, soit 36000 valeurs à l'ensemble des 100 calculs.

L'étude du comportement du sol hétérogène, comme dans les études précédentes, va consister à suivre l'évolution de la moyenne et des coefficients statistiques des valeurs du tassement, des déplacements horizontaux et des surpressions interstitielles, à différentes profondeurs. Les résultats de la série des 100 calculs, symbolisées par un ensemble de courbes, rassemblées sur les figures VIII.25 et VIII.26, nous amènent globalement aux mêmes remarques que précédemment. Ainsi nous constatons la même allure régulière des courbes obtenues, semblables à celles obtenues dans des études classiques de la consolidation. Comme précédemment, nous relevons les quelques aspects suivants :

- la valeur du tassement initial est importante ;
- les surpressions interstitielles sont négligeables en fin de consolidation ;
- les valeurs des déplacements horizontaux ne sont pas négligeables ; l'erreur commise dans l'évaluation de ces valeurs n'est pas élevée ;
- le degré d'incertitude absolue du tassement est faible et est constant à la surface du sol et pour toute valeur du temps de consolidation. Il est de l'ordre de 3% ;
- le tassement, à tout instant, est approximativement une fonction linéaire de la profondeur. Il en est de même de l'incertitude relative correspondante ;
- en début de consolidation et jusqu'à 5000 jours, l'incertitude sur le tassement varie très peu ; la valeur maximale est observée au point de la surface libre, sous l'axe du massif ;
- l'incertitude relative sur les surpressions interstitielles, maximales et élevées à environ -1 m de profondeur, sont faibles en début et surtout en fin de consolidation.

4. Comparaison des calculs probabilistes et déterministes

Les résultats des deux analyses de la consolidation bidimensionnelle et des calculs déterministes correspondants sont comparés entre elles sur les figures VII.27 et VII.28.

On remarque que :

- l'allure des courbes de tassement est identique pour tous les calculs ;

- le mode de variation de l'incertitude sur le tassement est identique dans les deux cas d'analyse probabilistes. Par exemple, l'incertitude maximale est observée pour la même valeur du temps (20.000 jours) et au même point du massif (à la surface du sol, sous l'axe du massif) ;
- l'incertitude est plus grande dans le cas de la variabilité unidimensionnelle. Le rapport des valeurs des incertitudes à certaines valeurs du temps (par exemple à la fin de la consolidation) et en différents points du massif peut atteindre la valeur 2. Dans les deux cas, néanmoins, la valeur ne dépasse pas 5 cm ;
- la valeur moyenne du tassement est sensiblement la même ;
- la solution probabiliste est assez bien approchée par la solution déterministe à moyenne arithmétique des deux paramètres ;
- on observe les valeurs du tassement les plus faibles dans le cas de la méthode déterministe à moyenne harmonique ;

La comparaison des valeurs des surpressions interstitielles pour tous les calculs est illustrée dans la figure VII.28. En début de consolidation, la dissipation des surpressions interstitielles est forte dans le cas de la variabilité unidimensionnelle. En fin de consolidation, ou pour des valeurs de temps élevées, on observe le phénomène inverse. Pour ces valeurs de temps, la vitesse de dissipation interstitielle est la plus grande dans le cas du calcul déterministe à moyenne arithmétique.

En ce qui concerne l'incertitude sur les surpressions interstitielles, on relève les points suivants :

- l'incertitude évolue de la même façon dans les deux analyses probabilistes. Ainsi la valeur du temps pour laquelle on observe le maximum de l'incertitude dépend du point considéré du massif. Et cette valeur du temps est d'autant plus faible que le point du massif est proche de la surface libre ;
- l'incertitude est plus grande dans le cas de la variabilité unidimensionnelle. L'écart le plus grand, entre les deux valeurs de l'incertitude, est observé, en chaque point du massif, pour la valeur du temps correspondant au maximum de l'incertitude, et il est d'autant plus grand que le point considéré du massif est proche de la surface non drainante. Le rapport des deux incertitudes, à -10 m de profondeur, est supérieur à 2 ;
- la solution probabiliste et la solution déterministe, en dépit des observations énoncées plus haut, sont assez proches l'une de l'autre.

5. Conclusion

L'étude comparaison des résultats des deux analyses de la consolidation bidimensionnelle a montré l'effet réducteur de la variabilité des paramètres de sol, sur la solution.

En effet, pour toute valeur du temps de consolidation et en tout point du massif, l'incertitude sur la solution, même si le mode variation est identique dans les deux cas, est considérablement réduite dans le cas de la variabilité bidimensionnelle.

On y observe aussi parfois (incertitude sur les surpressions interstitielles) une variation plus régulière et plus douce, dans le temps et avec la profondeur.

La solution probabiliste (en particulier dans le cas de la variabilité unidimensionnelle), dans le cas du tassement, est mieux approchée par la solution déterministe à moyenne arithmétique des paramètres, alors qu'en termes de surpressions interstitielles, c'est la solution déterministe à moyenne géométrique qui est la plus représentative.

En ce qui concerne les valeurs des déplacements horizontaux obtenus dans les deux analyses probabilistes effectuées, on n'observe aucune différence.

6. Comparaison des différentes analyses effectuées : effet de la dimension de la variabilité spatiale

Les tableaux VIII.4 et VIII.5 comparent les valeurs des coefficients statistiques des valeurs du tassement et des surpressions interstitielles respectivement, à 100, 1000, 5000 et 20.000 jours, pour différentes profondeurs.

L'étude comparative des résultats des quatre études probabilistes de la consolidation a révélé le rôle important de la dimension de l'analyse.

Il apparaît clairement que, quel que soit le caractère imposé à l'écoulement d'eau interstitielle, la valeur moyenne du tassement est plus grande dans le cas de l'analyse bidimensionnelle, sauf en fin de consolidation, à des profondeurs voisines du substratum imperméable. L'écart entre les valeurs du tassement à l'instant initial, dans les cas uni- et bidimensionnel, est important : à la surface du sol, il vaut plus de 30 cm. Cet écart diminue au cours du temps, par suite d'une vitesse de tassement plus grande dans le cas de la consolidation unidimensionnelle. D'autre part, l'effet de l'écoulement d'eau sur la solution est plus significatif dans le cas de l'analyse de la consolidation unidimensionnelle. Quelle que soit la dimension de l'analyse, l'aspect unidimensionnel de la variabilité exagère l'incertitude sur la solution, ce qui est bien mis en évidence par les valeurs du coefficient de variation montré dans le tableau VII.4.

Tableau VII.4 Comparaison des coefficients statistiques des valeurs du tassement à 0, -3 et -8 m de profondeur obtenus des quatre analyses effectuées.

		Calculs probabilistes (Monte Carlo, éléments finis)											
		Consolidation unidimensionnelle						Consolidation bidimensionnelle					
		Variabilité 1-D			Variabilité 2-D			Variabilité 1-D			Variabilité 2-D		
Z (m)	T (j)	m _s (cm)	σ _s (cm)	CV	m _s (cm)	σ _s (cm)	CV	m _s (cm)	σ _s (cm)	CV	M _s (cm)	σ _s (cm)	CV
0	100	5.8	1.06	0.18	6.09	0.37	0.06	36.34	2.00	0.055	36.7	1.07	0.03
	1000	17.6	2.26	0.13	18.8	0.74	0.038	44.7	1.60	0.036	45.03	1.33	0.03
	5000	38.2	3.36	0.088	40.73	1.01	0.025	55.33	1.89	0.034	55.9	1.65	0.03
	20.000	57.13	2.30	0.040	58.4	0.89	0.015	61.45	3.70	0.060	61.74	2.19	0.035
1	100	0.068	0.094	0.73	0.07	0.031	0.44	22.1	1.49	0.067	22.2	0.86	0.039
	1000	5.4	2.14	0.40	6.03	0.73	0.12	26.5	1.25	0.048	26.6	0.96	0.036
	5000	22.6	4.13	0.18	24.9	1.27	0.26	36.1	1.65	0.046	36.6	1.21	0.033
	20.000	39.40	2.55	0.064	40.64	0.83	0.020	41.8	2.74	0.065	42.1	1.65	0.039
8	100	0.017	0.035	17	0.29	0.35	12	3.05	0.29	0.095	3	0.26	0.086
	1000	0.32	0.30	0.93	0.37	0.10	0.27	4.2	0.36	0.086	4.1	0.36	0.09
	5000	5.35	1.63	0.30	6.04	0.50	0.08	7.4	0.94	0.13	7.4	0.68	0.09
	20.000	11.3	1.66	0.15	11.52	0.52	0.045	9.5	1.26	0.13	9.4	0.82	0.087

Tableau VII.5 Comparaison des coefficients statistiques des surpressions interstitielles.

Comparaison des résultats obtenus des quatre analyses effectuées.

		Calculs probabilistes (Monte Carlo, éléments finis)											
		Consolidation unidimensionnelle						Consolidation bidimensionnelle					
		Variabilité 1-D			Variabilité 2-D			Variabilité 1-D			Variabilité 2-D		
Z (m)	T (j)	m _s (kPa)	σ _s (cm)	CV	m _s (kPa)	σ _s (kPa)	CV	m _s (kPa)	σ _s (kPa)	CV	M _s (kPa)	σ _s (kPa)	CV
1	100	28.50	9.75	0.34	26.2	3.55	0.14	21.3	7.33	0.34	24.2	6.77	0.28
	1000	10.76	5.69	0.53	9.03	1.72	0.19	6.2	3.4	0.55	5.53	2.44	.44
	5000	4.38	2.75	0.63	3.45	0.75	0.22	1.63	1.1	0.67	1.39	0.62	0.45
	20.000	0.56	0.60	1.07	0.24	0.095	0.40	0.13	0.145	1.12	0.06	.03	0.50
3	100	48.28	0.05	0.001	48.28	0.015	3E-4	28.3	1.1	0.039	28.5	0.62	0.021
	1000	39.45	2.27	0.06	38.47	1.44	0.04	20.6	2.9	0.14	21	1.60	0.076
	5000	19.06	5.85	0.31	17	1.75	0.10	1.9	2.35	1.24	6.44	0.98	0.15
	20.000	2.44	2.27	0.93	1.66	0.39	0.33	0.54	0.51	0.94	0.29	0.1	0.34
10	100	48.3	0	0	48.3	0.006	E-05	30.6	1.31	0.43	30.5	0.94	0.03
	1000	47.32	1.12	0.024	47.15	0.38	0.008	30.6	2.16	0.089	30.5	1.81	.07
	5000	27.36	6.66	0.24	24.2	2.08	0.085	9.6	2.97	0.31	9	1.1	.12
	20.000	3.45	2.95	0.85	1.66	0.53	0.32	0.77	0.69	0.90	0.41	0.13	0.32

Ce degré d'incertitude est faible et constant (de l'ordre de 4%) dans le cas de la consolidation bidimensionnelle à variabilité bidimensionnelle.

On compare aussi les résultats des quatre séries de calculs relatifs aux dissipations des surpressions interstitielles sur la figure VII.28 et dans le tableau VII.5.

On relève ce qui suit :

- les surpressions interstitielles se dissipent plus vite dans le cas de l'analyse bidimensionnelle ; dans ce cas, les dissipations instantanées sont importantes (plus de 15 kPA) ;
- quel que soit la dimension de la déformation dans l'analyse probabiliste, c'est l'écoulement bidimensionnel d'eau interstitielle qui favorise le plus les dissipations des surpressions interstitielles.

On notera que l'incertitude évolue de la même façon pour tous les calculs, aussi bien en fonction du temps qu'en fonction de la profondeur. Elle est la plus faible dans le cas de la consolidation bidimensionnelle, sauf pour des temps faibles, à des profondeurs voisines du substratum ; dans chaque cas d'analyse de la consolidation, la variabilité unidimensionnelle surestime l'incertitude ; les valeurs les plus fortes sont celles de l'analyse unidimensionnelle à variabilité unidimensionnelle. Dans ce cas, elles valent, à -1 m de profondeur et à 100 jours, 10kPa, soit 20% de la valeur initiale.

Généralement, la différence entre la solution probabiliste et la solution déterministe est assez faible ; néanmoins l'hypothèse d'un sol « homogène équivalent » est plus acceptable dans le cas déterministe à moyenne géométrique.

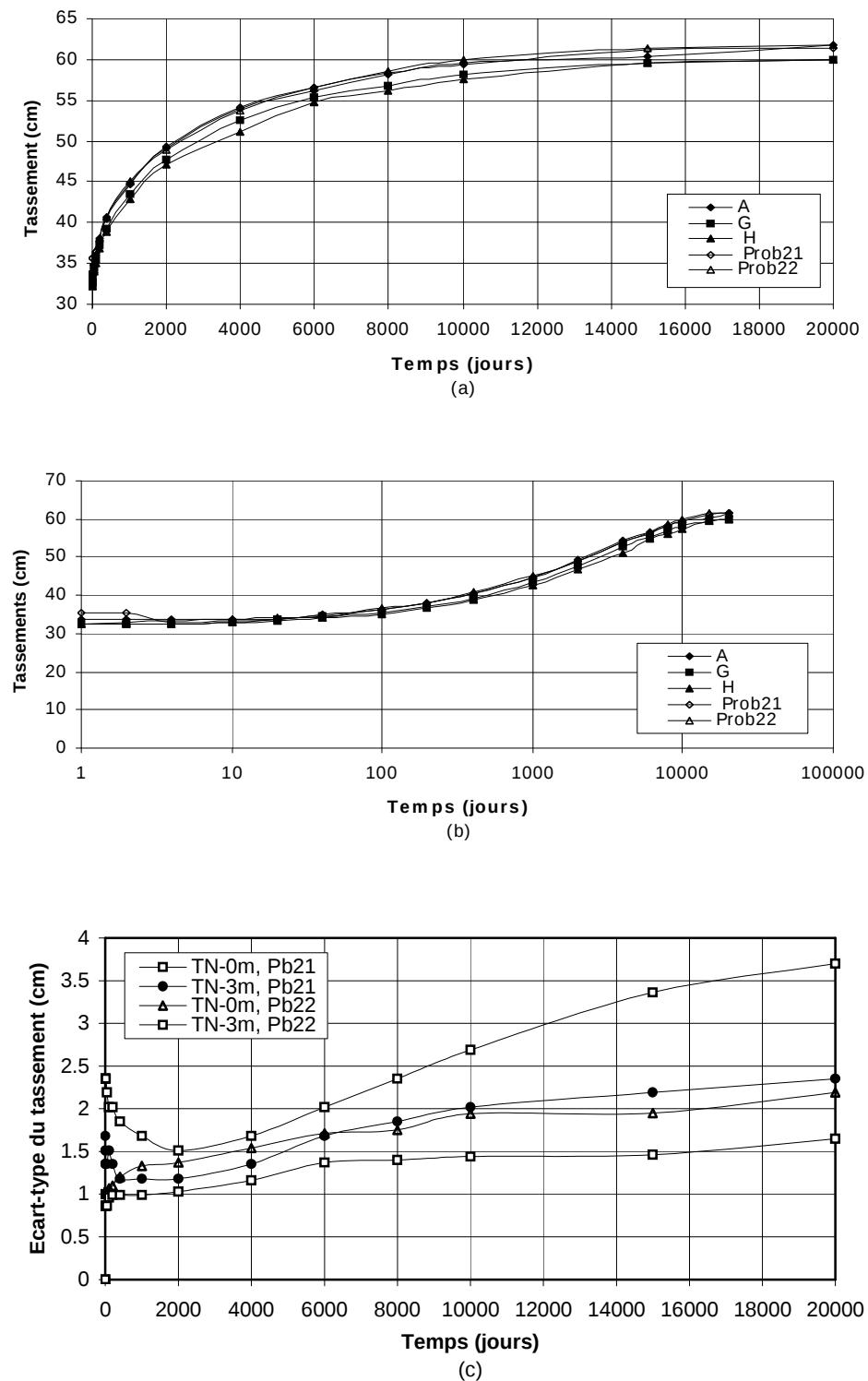


Figure VII.27 Comparaison des calculs probabilistes et déterministes :

(a) Comparaison de l'évolution des tassements de la surface du sol selon les calculs (échelle linéaire du temps logarithmique) ; (b) Comparaison de l'évolution des tassements de la surface du sol selon les calculs (échelle du temps logarithmique) ; (c) Comparaison de l'évolution des écarts-types selon les calculs (valeurs à 0 m et 3m de profondeur).

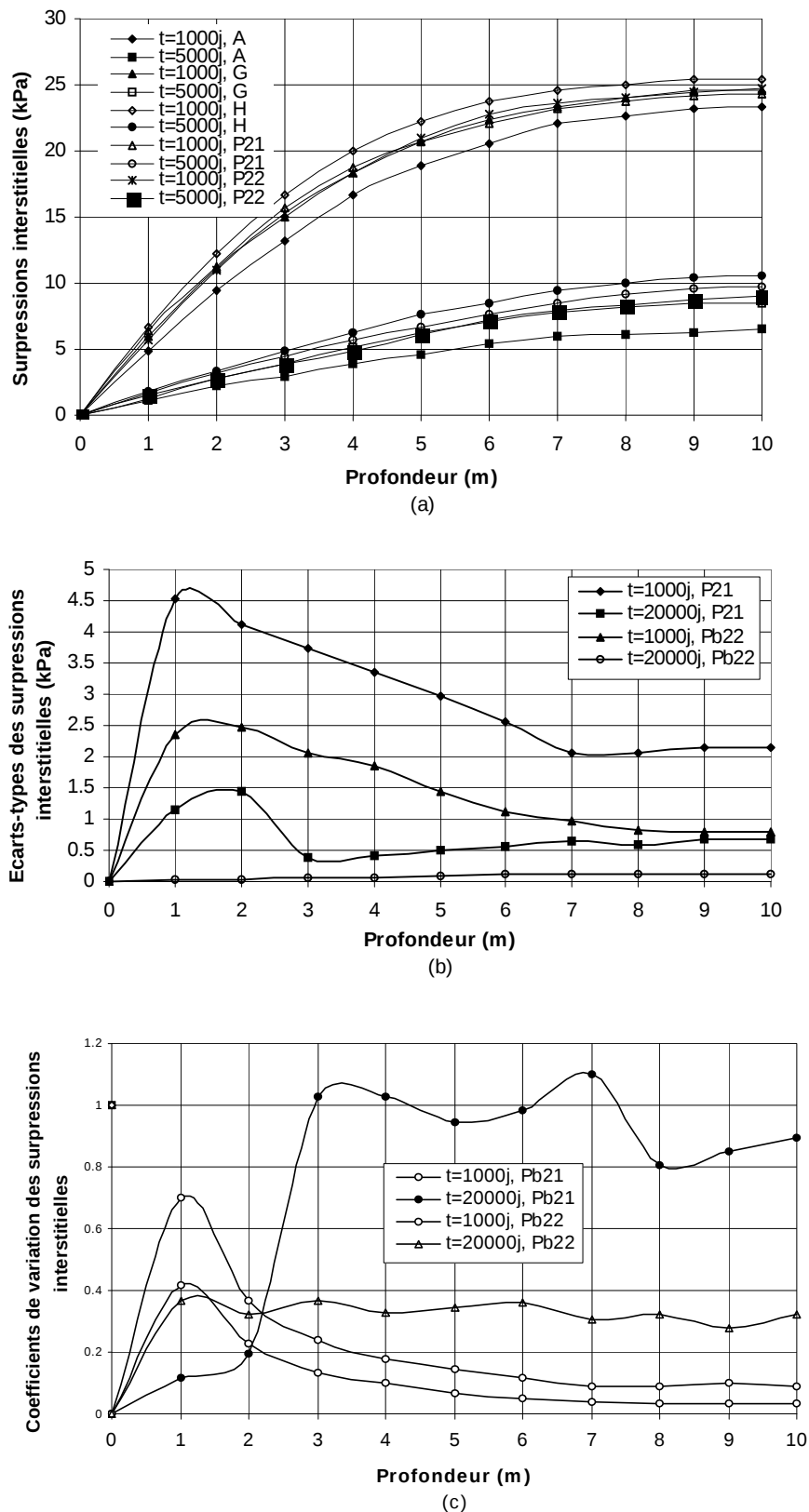


Figure VII.28 Comparaison des calculs probabilistes et déterministes :

(a) Comparaison de l'évolution des surpressions interstitielles en fonction de la profondeur aux temps 1000 et 20000 jours ; (b) Comparaison de l'évolution des écarts types des surpressions interstitielles ; (c) et des coefficients de variation correspondants.

VII.9 CAPACITE PORTANTE

L'influence de la variabilité spatiale des paramètres de sol sur la dispersion de la capacité portante des fondations a été étudié par Genevois (1984) et par Favre et al (1995a) en utilisant le modèle de comportement de Druker-Prager et la méthode des éléments finis stochastique. Dans cette étude, la variabilité spatiale de l'angle de frottement, du module d'élasticité, de leurs corrélations et autocorrélations pour un sol non cohérent ont été analysées. Les résultats obtenus ont montré que la dispersion de la capacité portante des fondations est d'abord sensible au nombre de variables aléatoires du sol ; c'est la dispersion de ces variables de base qui joue un rôle et non leur valeur moyenne. Aussi, l'autocorrélation sur une variable joue un rôle analogue à la corrélation entre deux variables.

Par ailleurs, l'effet sur la capacité portante des sols spatialement aléatoires a été étudié par Fenton et al (2000) en combinant la théorie des champs aléatoires avec la méthode des éléments finis élastoplastique. Une première investigation a été menée en examinant l'effet d'un effort de cisaillement spatialement aléatoire sur la capacité portante. L'objectif principal est de déterminer le point où la variabilité spatiale des propriétés du sol affecte la distribution de la capacité portante.

Pour la dispersion du coefficient de variation (CV) de l'effort de cisaillement du sol, la valeur prévue de la capacité portante tend vers la solution de Prandtl. Pour l'augmentation des valeurs de (CV), la valeur prévue de la capacité portante chute rapidement.

L'évaluation de la capacité portante pour une fondation peu profonde par la moyenne des méthodes probabilistes a été récemment menée par Cherubini C. et al (2000). L'utilisation de ces méthodes était pour déterminer l'indice de fiabilité de ce type de fondation, cela sur la base des valeurs proposées des coefficients de variations de l'angle de frottement et de la cohésion effective. L'auteur a effectué une comparaison entre les résultats de la méthode déterministe et la méthode de Monte Carlo, résultats jugés parfois très proches.

Griffiths et al (2002) ont présenté les résultats d'une étude probabiliste de la capacité portante d'une fondation filante, rigide et rugueuse posée sur un sol cohérent. L'analyse a concerné l'influence de la cohésion non drainée dont la distribution est aléatoire. L'auteur a combiné la théorie des champs aléatoires avec les éléments finis non linéaires et la méthode de Monte Carlo. Dans une étude paramétrée, la moyenne de la cohésion non drainée est tenue constante tandis que le coefficient de variation et la distance de corrélation spatiale de la cohésion varient systématiquement. L'influence de la variation spatiale de la cohésion sur la moyenne de la capacité portante a été discutée. Aussi, une comparaison entre une fondation rugueuse et une autre lisse a été réalisée.

VII.10 AUTRES CALCULS

- Dans la seconde partie du travail de Baghry S et al. (1983), l'étude a consisté à déterminer la probabilité de rupture associée aux valeurs du coefficient de sécurité calculées par la méthode de Bishop simplifiée sur un des remblais du site Cubzac-Les-Ponts.

L'étude de la variabilité du compactage des sols sur la stabilité des pentes des barrages en remblais a été menée par Breysse et al (2000). Les caractéristiques mécaniques des remblais utilisés dans la construction de ces ouvrages ont une variabilité importante, ce qui influe sur leur stabilité. De ce fait, l'auteur a mené ses analyses en considérant la cohésion non drainée (C_u) comme variable aléatoire dans une étude bidimensionnelle développée pour la sécurité de ce type d'ouvrages. Cette variable a été déterminée à travers les mesures in situ, puis modélisés et simulés numériquement par la suite. Cette modélisation et cette simulation mènent à définir une valeur caractéristique pour ce paramètre (C_u), sachant que cette valeur est liée à un niveau assumé de risque. A la fin, l'auteur a comparé entre l'ampleur de cette variabilité et les autres sources d'erreurs sur la détermination de stabilité de ces ouvrages.

Ahmed M. Hassen et al. (1999) ont effectué une analyse probabiliste de la stabilité d'une pente. Ils ont mis l'accent sur les similitudes et les différences des surfaces de glissement correspondant à des coefficients de sécurité minimum calculés par la méthode déterministe et la méthode probabiliste ainsi que sur l'indice de fiabilité de ces surfaces. Comme résultat important, ils ont mis en évidence la coïncidence des surfaces critiques probabilistes pour des petites valeurs dominantes des propriétés du sol.

- Les analyses de la stabilité de pentes par approche stochastique ont été aussi étudiées par Chiasson, (1997).

En utilisant une approche stochastique d'analyse de stabilité des pentes, Chiasson (1997) cherche à quantifier les risques de rupture d'un ouvrage. Ceci peut se faire en considérant les variables d'entrée comme des variables aléatoires, ou des fonctions aléatoires. Ainsi, le résultat final, i.e. le coefficient de sécurité, correspond à une variable aléatoire (V.A) dépendante. Cette V.A a une fonction de distribution statistique avec une espérance mathématique et une variance. La probabilité de rupture de l'ouvrage quant à elle représente celle d'obtenir un coefficient inférieur à 1.0. La nature du résultat final dépend de la façon dont sont traitées les variables d'entrée. La plupart des approches probabilistes utilisent une V.A unique par paramètre de résistance. Chez ces approches, les réalisations «résistances» ne sont pas des fonctions spatiales mais sont plutôt identiques en tous points du dépôt. Un tel modèle ne correspond pas à la variabilité spatiale observée chez les sols. Un modèle plus approprié fait usage de la notion

d'autocorrelation spatiale. Ceci garantit que des points rapprochés ont en moyenne des réalisations similaires (forte corrélation) tandis que des points éloignés sont indépendants ou non corrélés. La plupart des approches probabilistes existantes sont donc un cas limite, celui des espaces aléatoires parfaitement autocorrectifs.

Les coefficients de sécurité présentés résultent de calcul sur des dépôts virtuels, i.e. des réalisations d'un espace aléatoire. A chaque nouveau dépôt virtuel correspond une surface de glissement critique (la plus susceptible à glisser). Cette surface n'est pas préalablement imposée. Elie adopte plutôt une forme prenant avantage des lentilles de faibles résistances et évitant celles de fortes résistances générées, Cet article s'attarde surtout sur une description du modèle d'analyse adopte accompagnée de quelques résultats.

Le modèle adopté représente la résistance non drainée d'un sol comme une variable régionalisée (Materon, 1971). Soulier et al. (1990) et Chiasso et al. (1995) démontre la justesse de ce modèle chez les sols argileux. Dans ces études, les mesures fluctuent respectivement autour d'une constante et d'une dérive linéaire. Le modèle utilise donc :

$$S_u(x,z) = m(z) + R(x,z) \quad (\text{VII.23})$$

Où S_u = la résistance non drainée aux coordonnées (x, z) , m = la dérive selon la profondeur z , et R = la composante fluctuation aléatoire de la résistance. La dérive m est simplement d'ordre linéaire et s'écrit:

$$m = A_0 + A_1 z \quad (\text{VII.24})$$

où A_0 = la résistance non drainée à la profondeur zéro (kPa), A_1 = le gradient de la résistance au cisaillement (kPa/m). Ces deux coefficients sont définis par l'utilisateur. Lorsque $A_1 = 0$, la résistance au cisaillement du matériau fluctue autour d'une résistance moyenne constante. Sinon, elle fluctue autour d'une droite d'accroissement selon la profondeur.

Une fonction aléatoire stationnaire représente la composante fluctuation R . Sa structure de variabilité se caractérise par un covariogramme $C(h)$. Quatre modèles mathématiques sont admis. Ces modèles sont l'exponentiel, le sphérique, le gaussien et l'effet de pépite pure. Les trois premiers sont à deux paramètres: variance a priori $C(0)$ et portée a . Le dernier correspond à une absence totale de structure, i.e. $a = 0$ et un seul paramètre est à définir (C_0). La méthode permet aussi de considérer une anisotropie géométrique. Par exemple, la portée a peut être différente selon la direction verticale et horizontale. Ainsi, il est possible de représenter un dépôt dont la variabilité est plus forte dans une direction donnée. Ceci s'observe typiquement chez les dépôts d'origine sédimentaires, où la variabilité est généralement plus forte selon l'axe vertical, et où elle est souvent isotrope dans le plan horizontal. Soulier et al. et Chiasso et al. ont trouvé une portée verticale de 2.0 m et des valeurs s'étendant entre 10 et 30 m dans le plan horizontal.

La méthode d'analyse stochastique de la stabilité se compose de deux étapes, La première génère un champ de résistance au cisaillement non drainé. Celui-ci représente une coupe verticale d'un

dépôt virtuel. L'utilisateur définit d'abord une grille de N_2 noeud. La méthode engendre ensuite à chacun de ces noeuds des réalisations de résistances au cisaillement. Ces réalisations ou simulations s'obtiennent à l'aide d'une décomposition de Cholesky ou LV (Alabert, 1987). Pour une simulation non conditionnelle, la matrice covariance C des N_2 noeuds se décompose de façon à avoir :

$$C = L - U \quad (\text{VII.25})$$

et où $U = LT$. La composante fluctuation $R(x,z)$ de l'équation (1) s'obtient alors par un produit matriciel d'un vecteur w composé de N_2 déviations indépendantes de loi normale avec la matrice triangulaire inférieure L :

$$R = L \cdot w \quad (\text{VII.26})$$

Le résultat est une fluctuation obéissant à la loi normale. Une simulation non conditionnelle de résistances au cisaillement s'obtient en ajoutant la dérive m . Il est possible aussi de générer des simulations dites conditionnelles. Celles-ci permettent de créer aux points de mesures des réalisations identiques à celles enregistrées, tout en demeurant aléatoires partout ailleurs sur la grille. Le lecteur désireux d'en savoir plus peut consulter Chiasson et Djebbari, (1997).

La décomposition de Cholesky est une étape très gourmande en temps de calcul. Par contre, la grille et les points de mesures ne changent pas pendant l'analyse d'une structure spatiale et d'une géométrie d'excavation donnée. Ainsi, la décomposition se fait qu'une fois

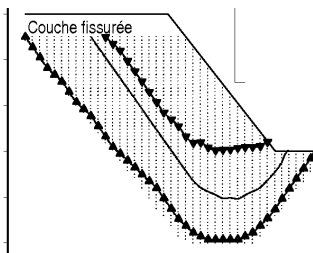


Figure VII.29 Surfaces stochastiques de rupture.

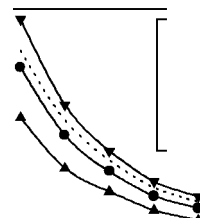


Figure VII.30 Coefficients stochastiques de sécurité.

et le temps de calcul apparaît donc uniquement à la première simulation. Seul un nouveau vecteur de déviations normales doit être généré pour chaque nouvelle simulation. Toutes les simulations subséquentes sont donc extrêmement rapides. Ceci rend la méthode très bien adaptée au problème, puisqu'un grand nombre de dépôts virtuels doivent être générés dans un court délai de calcul.

L'étape deux est un calcul du coefficient de sécurité critique (F_s) d'une pente. Cette

pente est excavée dans le dépôt virtuel de l'étape précédente et selon une géométrie définie par l'utilisateur. Le F_s se calcule à l'aide d'une méthode d'équilibre limite en contrainte totale identique à celle de Janbu (1973). La forme de la surface potentielle de glissement n'est pas définie a priori. Elle s'obtient plutôt par une optimisation par analyse dynamique (Baker, 1980 ; Cordier, 1980) d'un ensemble continu de segments joignant les nœuds de la grille.

La Figure VIII.29 présente une grille typique et trois surfaces critiques: une obtenue par approche déterministe et les deux autres correspondant aux coefficients de sécurité minima et maxima d'une série de $N = 700$ réalisations. Cette figure montre que la grille semble suffisante pour contenir toutes les surfaces critiques de glissement. Une grille trop petite risque biaiser l'histogramme du F_s . Chez plusieurs dépôts virtuels, la surface critique peut se retrouver à l'extérieur de la grille. Par contre, une grille trop grande produit des temps de calcul excessivement long en plus de problèmes d'ordre numérique. La relation obtenue entre la profondeur d'une excavation et l'histogramme du coefficient de sécurité est illustrée à la figure VIII.30. Quelques remarques s'imposent. D'abord, le coefficient de sécurité diminue en fonction de la profondeur d'excavation, en accord avec les résultats d'analyses déterministes. Ensuite, l'étendue entre le premier et le dernier décile diminue en fonction de la profondeur d'excavation (même résultat pour l'écart-type), Ainsi, plus l'excavation est profonde, plus le coefficient de sécurité est en moyenne faible, mais moins il est incertain. Chez la probabilité de rupture, c'est la tendance de la moyenne qui domine puisqu'elle ne cesse d'augmenter avec la profondeur. Elle passe d'une valeur négligeable à 2 mètres d'excavation, à un peu plus de 10% à 8 mètres de profondeur. La baisse de dispersion avec z s'explique quant à elle par la longueur de la surface critique. Plus l'excavation est profonde, plus la surface critique tend à s'allonger. Un phénomène de régularisation prend ainsi place et augmente en intensité avec la longueur de la surface. D'autres résultats peuvent être trouvés dans Chiasson et Djebbari (1996, 1997) et Djebbari (1996).

Cette méthode permet donc de représenter de façon fidèle la structure de la variabilité spatiale des sols. Les résultats montrent entre autre comment le risque de rupture augmente avec la profondeur, et comment la distribution statistique du F_s se resserre autour de la moyenne lorsque la profondeur d'une excavation augmente. Cette étude fait aussi ressortir les lacunes des approches déterministes chez les sols. En effet, le coefficient déterministe surestime systématiquement le coefficient probabiliste moyen. Enfin, la surface critique de glissement déterministe n'a aucune correspondance à l'ensemble infini de surface critique probabiliste. Ceci explique les nombreuses observations de surface réelle de glissement qui ne correspondent pas aux prédictions déterministes. Par contre, les analyses déterministes ne sont pas à rejeter du revers de la main puisqu'elles peuvent s'avérer suffisantes lorsque la dispersion des mesures est faible.

CONCLSION GENERALE

CONCLUSION

La nécessité de placer la mécanique dans l'incertain est aujourd'hui un fait reconnu tant il est vrai que la conception des objets industriels « *se place dans le royaume de la réalité physique où il n'y a pas d'absolu et où la connaissance est imparfaite* » (A.M. Freudenthal, 1947).

Les dernières décennies ont vu ainsi le développement de méthodes et d'outils, l'introduction de l'approche probabiliste mais également d'autres approches, apportant à chacun dans sa recherche ou son métier « *une précision supplémentaire* ».

La description et l'analyse de l'incertain en géotechnique, avec pour champs d'application la sécurité des ouvrages et les risques, sont des questions encore à la mode en ce début du 21^{ème} siècle

Par exemple, tous les calculs et toutes les prévisions sur les coefficients de sécurité devraient en définitive être remplacés par des estimations raisonnables des probabilités de ruine.

La prise en compte dans le calcul de structures diverses des incertitudes inhérentes au modèle, pouvant porter sur les propriétés du matériau, les chargements ou la géométrie, semble aujourd'hui incontournable si l'on cherche à obtenir des prédictions numériques "fiables", exploitables dans un processus de conception ou une prise de décision

Les domaines d'application des approches probabilistes sont très variés : matériaux composites, céramiques, verres, bétons, bois, métalliques, milieux granulaires, milieux poreux, polycristaux...

Afin de prendre en compte l'effet de la variabilité des propriétés physiques et mécaniques de sol dans l'analyse de problèmes de mécanique des sols, les géotechniciens poussent et testent leurs investigations dans l'application de techniques d'analyse diverses.

L'apport en géotechnique des applications des méthodes statistiques et probabilistes, entre autres, est considérable. Et malgré le développement d'outils de calculs (méthodes numériques) de plus en plus performants et une description de sol de plus en plus proche de la réalité, il existe toujours un écart entre les résultats de calcul et les valeurs solutions réelles.

Ce présent rapport se veut une petite contribution à la vulgarisation d'un certain nombre de techniques qui permettent la caractérisation des propriétés physiques et mécaniques des sols et qui prennent en compte de l'effet de cette variabilité sur la solution. Il en est ainsi des

méthodes classiques (lois de distribution, paramètres statistiques,...) ou des techniques de géostatistique pour décrire la variation spatiale.

Par ailleurs, nous avons passé en revue toutes les notions utiles à la résolution d'un problème de géotechnique en prenant en considération l'hétérogénéité des sols. Il va sans dire que la simulation de Monte Carlo est très utile dans ce cas. Il en va de même des corrélations entre propriétés de sol, refusées par les uns et admises par les autres, mais qui permettent souvent des descriptions plus rapides et plus justes.

L'état de l'art des méthodes et autres techniques d'analyse des sols hétérogènes ainsi que les divers exemples de calcul d'ouvrages de géotechnique nous ont, néanmoins, conduit à quelques remarques :

- les réalisateurs introduisent déjà le calcul des probabilités dans leurs études de fiabilité, mais il est utile de remarquer que dans les applications courantes de calcul d'ouvrages divers, l'introduction des produits de recherches dans le domaine des statistiques et probabilités est encore assez timide,
- il existe de nombreux facteurs d'incertitude qu'il est difficile d'évaluer il s'avère difficile, pour de nombreuses raisons, de les prendre en compte dans les calculs. Il s'agit notamment de toutes les sources de variabilité de l'information géotechnique : il y a celles dues à la variabilité naturelle des terrains (qui n'est pas probabilisée a priori et n'est peut-être pas probabilisable à l'échelle de nos ouvrages) mais il y a aussi celles qui proviennent de toutes les opérations nécessaires à l'acquisition de ces données (forage, prélèvement, transport, traitement des données, etc.).
- la description des comportements d'ouvrages de géotechniques passe par des modélisations pas toujours fidèles. Il est difficile dans ce cas de rendre compte de la réalité et d'aboutir à des solutions justes, quelque soit la technique utilisée.

Toutes ces raisons incitent à la prudence et de ce fait, les recherches sur les techniques d'analyse de la variabilité des propriétés des sols et de leur influence sur la solution est toujours d'actualité et elle le sera tant que persistent des zones d'ombres dans la description de la réalité.

Références bibliographiques

1. **Auvinet G., Mellah R., Masrouri F., Rodriguez J.F.**, (2000), "La méthode des éléments finis stochastiques en géotechnique", *Revue Française de Géotechnique* N° 93, 4e trimestre 2000, PP, 67-79.
2. **Baguelin F., Kovarik J.-B.**, (2000), "Une méthode de détermination des valeurs caractéristiques des paramètres géotechniques", *Revue Française de Géotechnique* N° 93, 4e trimestre 2000, PP, 35-41.
3. **Baghery S., Magnan J.-P.**, (1983), "Analyse probabiliste de la stabilité et des tassements des remblais du site expérimental de Cubzac -les-ponts", Rapport de recherche tiré de la thèse de Docteur Ingénieur soutenue par Baghery Shohrab (1980), "Probabilités et statistiques en mécanique des sols – Analyse probabiliste de la stabilité et des tassements des remblais sur sols compressible (site expérimental de Cubzac-les-ponts) " ECPC, Paris, France.
4. **Bekkouche A.**, (1987), "Sécurité des grands barrages en terre. Approche probabiliste des problèmes d'écoulement liés aux reconnaissances et contrôles" Thèse de Docteur, Ecole Centrale de Paris France.
5. **Bekkouche A., Aissa Mamoune S.M. & Houmadi Y.** (2003) "Variabilité des paramètres de gonflement sous une fondation superficielle" (05-07 Novembre 2003) ; Symposium International sur les fondations superficielles FONDSUP ; Paris, France.
6. **Bekkouche A., Houmadi Y. & Aissa Mamoune S.M.** (2004) "Variabilité bidimensionnelle des paramètres de gonflement" (05-07 Novembre 2003) ; Conférence Internationale de Géotechnique – Beyrouth 2004 ; Paris, France.
7. **Benjamin, J.R. and Cornell C.A.**, (1970) "Probability, statistics and decision for civil engineers" Université de Standford, McGraw-Hill (Eds), USA.
8. **Bolle A.**, (1988) "Approche probabiliste en mécanique des sols avec prise en compte de la variabilité spatiale", Thèse de Docteur, Ecole EPFL de Lausanne Suisse.
9. **Bolle A.**, (2000), "Etude et prise en compte de variabilité spatiale", *Revue Française de Géotechnique* N° 93, 4e trimestre 2000, PP, 55-66.
10. **Bolle A.**, (2001) "Approche probabiliste en reconnaissance et en simulation", Module A3, Faculté des sciences appliquées Université de Liège, Belgique.
11. **Boulefkhad T.** (1986), "Influence de la variabilité des sols sur les calculs des tassements des fondations superficielles" Thèse de Docteur Ingénieur. Université de Clermont II. UFR de recherche scientifique et technique, France.
12. **Breysse D., Kouassi P., Poulain. D.**, (2000), "Influence of the variability of compacted soils on slope stability of embankments, Applications of statistics and probability", Melchers & Stewart (eds), ICASP7, 2000 Balkema, Rotterdam, PP, 367-373.
13. **Brzakala W.**, (1995), "Stochastic finite layers method, Probabilistic analysis of shallow foundations based on field data, Applications of statistics and probability", Lemaire, Favre & Mébarki (eds), ICASP7, 1995 Balkema, Rotterdam, PP, 47-54.
14. **Brzakala W., Pula W., Mahmoud El-Meligy**, (1995), "Settlement analysis by non- linear finite element reliability method, Applications of statistics and probability", Lemaire, Favre & Mébarki (eds), ICASP7, 1995 Balkema, Rotterdam, PP, 55-53.

15. **Cafaro F. Cherubini C. & Cotechia E** (2000), "Use of the scale of fluctuation to describe the geotechnical variability of an Italian clay, Applications of statistics and probability", Melchers & Stewart (eds), ICASP8, 2000 Balkema, Rotterdam, PP, 481-486.
16. **Cambou B.** (1979) "Influence de l'hétérogénéité spatiale à faible échelle dans une analyse probabiliste en mécanique des sols",
17. **Cassan M.**, (2000), "Utilisation de la statistique descriptive en géotechnique", Revue Française de Géotechnique N° 93, 4e trimestre 2000, PP, 20-34.
18. **Chiasson, P. et Djebbari, Z.** 1997. Stochastic slope stability analysis of temporary cuts in clay. Comptes rendus du xrv-: Congres International de Mecanique des Sols et de Fondations, Hambourg, septembre 1997.
19. **Chiasson, P. et Djebbari, Z.** 1996. Slope Stability of Temporary Cuts by Conditional and Unconditional Simulations. Comptes rendus de la 49 "me conference canadienne de géotechnique, 23-25 septembre, St-John's, Bitech pub. Vol. 2, pp. 903-910.
20. **Chiasson, P., J. Lafleur, M. Soulie, and K.T. Law** 1995. Characterizing spatial variability of a clay by geostatistics. Revue canadienne de géotechnique, **32** : 1-10.
21. **Chiasson P.**, (1997) "Analyse de la stabilité de pentes par approche stochastique" Acte du 13eme Congrès français de mécanique, association universitaire de mécanique, poitiers- Futuroscope, 1-5 Septembre 1997, Vol.2, PP, 57-60.
22. **Cherubini C.**, (1994), "The variability of geotechnical parameters, Probabilities and Materials", NATO.ASI, D.P.Breyse (eds), PP, 69-80.
23. **Cherubini C. Di Nauta P., Tesoro G.** (2000), "Probabilistic bearing capacity of strip shallow foundation on c, ϕ soils, Applications of statistics and probability", Melchers & Stewart (eds), ICASP8, 2000 Balkema, Rotterdam, PP, 487-490.
24. **Djebbari, Z.** 1996. Simulation géostatistique et calcul de stabilité de pentes par programmation dynamique. These de M.Sc.A., Université de Moncton, Moncton.
25. **Favre J.-L.**, (2000), "Les différents types d'erreurs et leur prise en compte dans les calculs géotechniques", Revue Française de Géotechnique N° 93, 4e trimestre 2000, PP, 11-20.
26. **Favre J.-L. et Magnan J.-P.**, (2000) "Incertitudes en géotechnique et dimensionnement des ouvrages" Revue Française de Géotechnique N° 93, 4e trimestre 2000.
27. **Favre J. L., Genevois B.**, (1995) "Analyse de la sécurité des fondations superficielles vis-à-vis d'un défaut de portance: effet de la variabilité spatiale des paramètres du sol", Revue française de géotechnique.
28. **Favre J. L. & Touati H.**, (1995), "Probabilistic estimation of the errors in finite element calculation code for shallow foundation, Applications of statistics and probability", Lemaire, Favre & Mébarki (eds), ICASP7, 1995 Balkema, Rotterdam, PP, 1309-1315.
29. **Favre J. L.**, (2004), " Sécurité des ouvrages. Risques. Modélisation de l'incertain, fiabilité, analyse des risques ", Ellipses.
30. **Fenton G. A.**, (1990), Simulation and analysis of random field, Thèse Ph.D, Université de Princeton, New Jersey, USA.
31. **Fenton G. A.**, (1997a), "Random field representation of CPT data: Parameter estimation", Norwegian Geotechnical Institute, Report 514170-4, Oslo, Norway, 1997.
32. **Fenton G. A. , D. V. Griffiths** , (1997b) "Extreme Hydraulic Gradient Statistics in a Stochastic Earth Dam", ASCE Journal of Geotechnical and Geoenvironmental Engineering, 123(11), PP, 995-1000,

33. **Fenton G. A., Griffiths D. V.** (2000) "Bearing capacity of spatially random soils", 8th ASCE Specialty Conference on Probabilistic Mechanics and Structural Reliability PMC2000-097.
34. **Fenton G. A., Griffiths D. V.** (2002) "Probabilistic foundation settlement on spatially random soil", Journal of geotechnical and geoenvironmental engineering. PP, 381-390.
35. **Gaouar E.M., Deplagne F. & Bacconnet C.**, (1995), "Multidimensional simulations of physical or mechanical parameters inside materials, Applications of statistics and probability", Lemaire, Favre & Mébarki (eds), ICASP7, 1995 Balkema, Rotterdam, PP,1245-1308.
36. **Genevois B.**, (1984), "Rôle de la variabilité spatiale des paramètres de sols sur la capacité portante des fondations superficielles", Thèse de Docteur Ingénieur, Ecole Centrale des arts et manufactures, France.
37. **Griffiths D. V., Fenton G. A.** (1997) "Three-Dimensional seepage through spatially random soil", Journal of geotechnical and geoenvironmental engineering Février 1997, PP, 153-160.
38. **Griffiths D. V., Fenton G. A.** (1998) "Probabilistic analysis of exist gradients due to steady seepage", Journal of geotechnical and geoenvironmental engineering Septembre 1998, PP, 779-787.
39. **Griffiths D. V., Fenton G. A.** (2002) "Bearing capacity of rough rigid strip footing on cohesive soil : Probabilistic study", Journal of geotechnical and geoenvironmental engineering Septembre 2002,. PP, 743-755.
40. **Hadj-Hamou T., Favre J. L. & Yoon S.H.**, (1995), "Probabilistic analysis of shallow foundations based on field data, Applications of statistics and probability", Lemaire, Favre & Mébarki (eds), ICASP7, 1995 Balkema, Rotterdam, PP, 15-21.
41. **Jaksa M.B.**, (1995) "The influence of spatiale variability on the geotechnical design properties of a stiff, overconsolidated clay" Ph.D thesis, University of Adelaide, Australia.
42. **Jaksa M.B., Brooker P.I., Kaggwa.W.S.**, (1997), "Modeling the spatial variability of the undrained shear strength of clay soils using geostatistics", E.Y. Baafi N.A. Schofields (eds.), Géostatistics Wollong'96, volume 2, 1284-1295.
43. **Magnan J.-P.**, (1982), "Les méthodes statistiques et probabilistes en mécanique des sols", Presse de l'ENPC France
44. **Magnan J.-P.**, (2000), "Quelques spécificités du problème des incertitudes en géotechnique", Revue Française de Géotechnique N° 93, 4e trimestre 2000, PP, 3-10.
45. **Magnan J.-P., Abdulbaki A., Pouget P.** (1994), "Analyse probabiliste de la stabilité de deux remblais édifiés sur versant instable", Revue française de géotechnique N°69, PP, 3-10.
46. **Magnan J.-P., Bouheraoua A.** (1995) "Stochastic analysis of soil consolidation in theory and practice, Applications of statistics and probability", Favre & Mébarki (eds), ICASP7, Balkema, Rotterdam, PP, 77-83.
47. **Magnan J.-P., Bouheraoua A.** (1996) "Analyse probabiliste de la consolidation unidimensionnelle des sols". Revue française de géotechnique 4eme Trimestre 1996 . PP,19-30.
48. **Magnan J.-P., Bouheraoua A.** (2001) "Analyse probabiliste de la consolidation bidimensionnelle des sols". Revue française de géotechnique. 2eme Trimestre 2001 . PP,10-25.
49. Matheron, G. 1971. The theory of regionalized variables and its applications. Centre de géostatistique, Ecole des Mines de Paris, Fontainebleau.
50. **Mellah R., Auvinet G., Masrouri F.** (2000a), "Stochastic finite element method

- applied to non-linear analysis of embankments", ELSEVEIR, Probabilistic engineering Mécanic 15 (2000), PP, 251-259.
51. **Mellah R., Auvinet G., Masrouri F.** (2000b), "Stochastic finite element method applied to non-linear analyses in geomechanics, Applications of statistics and probability", Melchers & Stewart (eds), ICASP8, 2000 Balkema, Rotterdam, PP, 79-85.
 52. **Nishimura S., Fujii H., Shimada K.**, (1995), "FEM consolidation analysis considering variability of soil parameters", Settelement analysis by non-linear finite element reliability method, Aplications of statistics and probability, Lemaire, Favre & Mébarki (eds), ICASP7, 1995 Balkema, Rotterdam, PP, 85-92.
 53. **Pouget P., Cadou N., Magnan J.-P.**, (2000), "Deux applications de l'analyse factorielles à des données géotechniques", Revue Française de Géotechnique N° 93, 4e trimestre 2000, PP, 43-54.
 54. **Smith M., Marche C., Robert B.**, (2004), "La nature de l'incertitude dans un calcul de risque barrage (Canada)", Colloque les risques en Génie Civil – ISET de Sfax-AIT-18-19 Mars 2004 – Tunisie.
 55. **Soulié M., Masengo. E.**, (1994), "Small scale variability in clay samples, Probabilities and Materials", NATO.ASI, D.P. Breyse (eds), PP, 81-94.
 56. **Vanmarcke, E. H.**, 1984 "Random fields: Analysis and synthesis", The Massachusetts Institute of Technology (MIT) Press. Cambridge, Massachusetts: USA.