

# STATISTIQUE DE LA DÉCISION APPLIQUÉE À L'ESTIMATION DES PARAMÈTRES DE CISAILLEMENT DES SOLS

## I INTRODUCTION

L'analyse de la performance des structures géotechniques est basée d'une part, sur l'analyse de la variabilité physique des paramètres du sol et d'autre part, ces paramètres ne sont pas toujours mesurés directement, sinon en petit nombre mais font l'objet d'un modèle qui tient compte des paramètres contrôlés. En plus de ces incertitudes, plusieurs erreurs sont commises sur les réalisations elles-mêmes de ces paramètres, faisant naître ainsi des incertitudes physiques.

La prédiction de la performance, avant la construction des ouvrages, dépend de la qualité et la quantité de l'information obtenue à partir des essais réalisés. Si la valeur de l'information est connue en fonction des coûts de l'observation, une étude économique est donc possible. La théorie de la statistique de décision, appliquée à l'estimation, s'adapte bien à ce genre d'analyse.

Les reconnaissances des matériaux, qui ont servi à la construction du noyau du barrage de Grand'Maison (France), ont été effectuées en 1977. En 1978, l'information concernant les paramètres de cisaillement, a été jugée insuffisante et par conséquent une nouvelle campagne d'essais a été programmée. L'étude proposer va tenter d'une part, d'analyser cette décision ainsi que la valeur de l'information nouvelle et, d'autre part, d'analyser l'opportunité d'une troisième campagne d'essais; et ceci en utilisant les techniques d'estimation des paramètres à partir de la théorie de la statistique de décision.

## 2 THÉORIE DE LA STATISTIQUE DE DÉCISION

### 2.1 Concept de bases

#### 2.1.1 Espace des états de la nature

L'espace des états de la nature est un espace dont les éléments sont des distributions de probabilité et dont chacune est censée correspondre à un état de la nature est un ensemble de description conjecturale de l'Univers à un moment donné. L'ensemble de ces descriptions imaginaires est appelé ensemble des états pos-

sibles et noté par  $\Omega$ . Si de plus, on définit une distribution de probabilité  $\mathbf{P}$  sur l'ensemble des hypothèses  $\mathbf{A}$ , l'espace des états de la nature est ainsi un triplet  $(\Omega, \mathbf{P}, \mathbf{A})$ , c'est à dire un espace de probabilité. Il faudra bien distinguer la distribution de probabilité  $\mathbf{P}$  des éléments  $W$  de l'espace  $\Omega$  qui sont aussi des distributions de probabilité.

Pratiquement, on admettra souvent que la forme analytique des éléments  $w \in \Omega$  est connue et que chaque  $W$  est entièrement déterminé par les valeurs d'un ensemble fini de paramètres (par exemple la moyenne et la variance si les distributions sont normales). Si ces paramètres sont au nombre de  $p$ , l'espace des états de la nature prendra la forme d'un espace euclidien à  $p$  dimensions.

#### 2.1.2 Espace des observations

Le statisticien dispose d'un ensemble d'observations, prenant la forme d'un ensemble à  $n$  variables aléatoires (ou éventuellement une infinité) dont les distributions de probabilité dépendent du paramètre  $w$  sinon leurs observations ne fournissent aucune information sur l'espace  $\Omega$ . Nous avons à tenir compte:

- De l'ensemble des valeurs possibles du multiplet des  $n$  variables aléatoires observées; soit  $\mathbf{S}$  cet espace;
- Des événements possibles à l'occasion de leur observation; soit  $\mathbf{E}$  l'ensemble de ces événements;
- D'une distribution de probabilité  $\mathbf{P}_w$  dépendant de  $w$  et attribuant une probabilité à chaque événement de  $\mathbf{E}$ .

Il s'agit en fait d'une famille d'espace de probabilité  $(\mathbf{S}, \mathbf{E}, \mathbf{P}_w)$ , indexé par le paramètre  $W$ .

Il s'agit en fait d'une famille d'espace de probabilité  $(\mathbf{S}, \mathbf{E}, \mathbf{P}_w)$ , indexé par le paramètre  $W$ .

#### 2.1.3 Espace des décisions $\Delta$

C'est un ensemble de décisions possibles. L'ensemble des événements décisionnels peut être représenté par des sous-ensembles de  $\Delta$  noté  $D$ . Le couple  $(\Delta, D)$  définit un espace probabilisable (quand on aura ajouté une probabilité, définie pour tout événement décisionnel, ce sera un espace de probabilité).

A. BEKKOUCHE  
Université Abou Bekr  
Belkaid de Tlemcen  
A. DJEDID  
Université Abou Bekr  
Belkaid de Tlemcen

## RÉSUMÉ

La prédiction de la performance avant la construction des ouvrages dépend de la qualité et de la quantité de l'information obtenue à partir des essais réalisés. Si la valeur de l'information est connue en fonction des coûts de l'observation, une étude économique est donc possible. La théorie de la statistique de décision, appliquée à l'estimation, s'adapte bien à ce genre d'analyse.

## MOTS-CLÉS

Optimisation, Information, Géotechnique, Cisaillement.

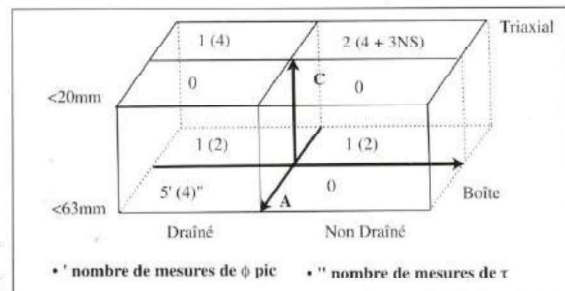


Figure 1. Plan d'expérience pour l'analyse de variance

#### 2.1.4 Espace des récompenses R

C'est l'ensemble des récompenses possibles qui peuvent être reçues par le statisticien comme le résultat d'une décision  $d \in \Delta$  et  $w \in \Omega$  étant le résultat de l'expérimentation. Donc  $\sigma(w, d)$ , note la récompense dans R quand le statisticien prendra une décision  $d$ , le résultat de l'expérimentation étant  $w$ . Soit B, l'ensemble des événements possibles de  $R, |\sigma(w, d): \Omega \times \Delta \rightarrow R|$

#### 2.2 Fonction utilité

Il est supposé que le statisticien aura à définir une préférence sur l'ensemble des récompenses et donc donné un ordre total sur R. Pour chaque distribution  $P \in \mathbf{P}$  une famille de distribution de probabilité et une fonction U à valeur dans R définie sur R, l'espérance de U si elle existe,  $E(U/P)$  quand P est une distribution sur R, est donnée par:  $E[U/P] = \int_{\mathbf{R}} U(r) dP(r)$

La fonction réelle U définie sur R est appelée fonction d'utilité si elle a les propriétés suivantes:

- Pour  $P_1$  et  $P_2 \in \mathbf{P}$  deux distributions de probabilités telles que  $E[U/P_1]$  et  $E[U/P_2]$  existent;
- $P_1 \leq P_2$  si et seulement si  $E[U/P_1] \leq E[U/P_2]$

Pour chaque récompense  $r \in \mathbf{R}$ ,  $U(r)$  est appelée utilité de r et pour chaque distribution de probabilité  $P \in \mathbf{P}$ ,  $E[U/P]$  est appelée utilité de P. Dans les problèmes de l'ingénieur, il convient de parler du côté négatif de l'utilité qui est appelé perte.

#### 3 ANALYSE DES FACTEURS D'INFLUENCE

L'influence des procédures d'essais et de la densité initiale a été testée par l'analyse de variance-covariance (Lebart et al). Cette analyse permettra de dégager un modèle cohérent de l'angle de frottement de pie à partir des résultats de laboratoire. Un modèle linéaire, sans interaction, a été choisi pour tester les différentes hypothèses nulles (non influence de facteur ou de tel groupe de facteurs).

$$\Phi_{ijk}^I = \Phi_0 + A_i + B_j + C_k + a\gamma + \varepsilon_{ijk}^I$$

Le plan d'expérience pour l'analyse de variance est donné par la figure (1).

Malgré le déséquilibre des cellules, l'analyse est possible sans toutefois pouvoir tester le modèle complet en une seule analyse nombre insuffisant de données. Il a été procédé à des analyses sur un ou deux facteurs à la fois. Les résultats sont donnés par le tableau 1.

L'analyse de la covariance sur la densité initiale a montré que ce facteur n'avait pas d'influence significative. On notera que les densités varient sur une faible plage.

| Facteurs | Probabilité critique | Degré de signification |
|----------|----------------------|------------------------|
| A        | 3°                   | Très                   |
| B        | 6°                   | Peu                    |
| C        | 17°                  | Non                    |
| A+B      | 8°                   | Peu                    |
| A+C      | 11°                  | Non                    |

Tableau 1: Résultats d'analyse de variance

Les mêmes analyses ont été menées sur la contrainte de cisaillement de rupture et ont confirmé ces résultats. Aussi, le modèle suivant a été adopté:

$$\begin{aligned} \Phi_i^I &= \Phi_0 + A_i + \varepsilon_i^I \\ \Phi_0 &= 40,2^\circ \\ A_i &= -A_2 = 2,7 \\ \forall i, i \quad E[\varepsilon_i^I] &= 0 \quad V[\varepsilon_i^I] = 10,8 \quad \sigma_\varepsilon = 3,28 \end{aligned}$$

Il a été estimé que la boîte de cisaillement à l'encontre du triaxial procurait un biais systématique par rapport à la réalité. Aussi, les valeurs ont été débiaisées, soit  $A_1 = 5,4^\circ$ ,  $A_2 = 0^\circ$  et  $\Phi_0 = 37,5^\circ$ . Le modèle est donc:

$$\begin{aligned} \Phi^I &= 37,5 + \varepsilon^I \\ \Phi_0 &= 37,5^\circ \\ \forall i \quad E[\varepsilon^I] &= 0 \quad V[\varepsilon^I] = 9,54 \quad \sigma_\varepsilon = 3,09 \end{aligned}$$

Les essais indiquent une forte dispersion, due à celle donnée par la boîte de cisaillement. Ce désavantage, pour la prédiction de l'angle de frottement interne de pie, est discuté dans le calcul d'optimisation du nombre d'essais.

|                  |                                                                                                                                                                                                           |
|------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Analyse Variance | Facteurs d'influence<br>Type d'appareil $A_{ji} = 1, 2$ (Boîte et Triaxial)<br>Condition d'essai $B_{jk} = 1, 2$ (drainé et non drainé)<br>Nature des matériaux $C_{kk} = 1, 2$ (Ecrêté à 63mm et à 20mm) |
| Covariance       | Densité sèche initiale $\gamma_d$ (Variable continue)                                                                                                                                                     |



#### 4 OPTIMISATION DU NOMBRE D'ESSAIS DE CISAILLEMENT

En 1977, lors des premières reconnaissances des matériaux servant pour la construction du barrage de Grand'Maison, on disposait de cinq essais à la boîte de cisaillement de moyenne  $\phi_0 = 42.5^\circ$  et de variance  $S^2_{\phi_0} = 14.64$ . Si l'on part de l'hypothèse que  $\phi()$  suit une loi normale de moyenne  $W$  et variance connue  $\sigma^2_{\phi_0}$ . Soit :

$$f(\Phi) = \int (\Phi, w) = \frac{1}{\sqrt{2\pi\sigma_\phi^2}} \exp\left\{-\frac{1}{2} \frac{(\Phi - W)^2}{\sigma_\phi^2}\right\}$$

Le paramètre  $W$  de la loi de  $\phi$  est lui-même considéré comme variable autoritaire. Ne sachant que peu de choses sur cette variable, on peut choisir sa loi dans la famille des conjuguées naturelles de  $f(\phi)$ , c'est-à-dire, dans notre cas, une loi normale. Soit  $\xi(w)$ , la distribution à priori des paramètres  $W$ :

$$\xi(w) = \frac{1}{\sqrt{2\pi\sigma_\phi^2}} \exp\left\{-\frac{1}{2} \frac{(w - \mu_0)^2}{\sigma_0^2}\right\}$$

La moyenne  $\mu_0$  et l'écart type  $\sigma_0$  sont estimés à partir des essais de 1977, qui servent alors d'information à priori:

$$\mu_0 = \frac{1}{5} \sum_{i=1}^5 \sigma_{i0} = 42.9^\circ$$

$$\sigma_0 = \frac{s_\phi}{\sqrt{5}} = 1.71^\circ \quad \sigma_0^2 = 2.93$$

Quel est le risque attaché à cette première information à priori ? L'étude de la sécurité et l'étude économique devraient nous donner une fonction de coût dépendant des trois paramètres aléatoires du problème (les états de la nature, ici le paramètre  $W$ ; les réalisations de la variable de calcul de l'ouvrage, ici c'est le paramètre  $\phi$  et les paramètres de décision, ici la décision portera sur la valeur à adopter pour  $w$  soit  $w_d$ ).

La fonction de coût est alors  $L(w, w_d)$ . En l'absence des études économiques et de sécurité, on peut directement adopter la fonction objective sous une forme très simple et propre au calcul analytique:

$$L(w, w_d) = R_0 (w - w_d)^2$$

Le risque à priori serait alors:

$$\int_{-\infty}^{+\infty} R_0 (w - w_d)^2 \xi(w) dw$$

Ce risque est minimum pour la meilleure décision à priori  $W^* = \mu_0$ . Le risque à priori devient  $R^*(w^*, d) = R_0 \sigma_0^2$ .

Quel bénéfice apporteraient les essais de 1978 ? La fonction de vraisemblance des cinq essais triaxiaux est une loi normale:

$$f(\Phi / w) = \frac{1}{(2\pi\sigma_\phi^2)^{n/2}} \exp\left\{-\frac{1}{2} \sum_{i=1}^n \frac{(\Phi_i - w)^2}{\sigma_\phi^2}\right\} \text{ avec } n = 5$$

$$\text{et } \bar{\phi} = 1/5 \sum_{i=1}^5 \phi_i = 37.5^\circ$$

La distribution à posteriori de  $w$  est une loi normale (conjuguée naturelle):

$$f(w / \Phi_i) = \frac{1}{(2\pi\sigma_n^2)^{1/2}} \exp\left\{-\frac{1}{2} \frac{(w - \mu_n)^2}{\sigma_n^2}\right\}$$

Dont les paramètres  $\mu_n$  et  $\sigma_n$  se calculent explicitement:

$$\mu_n = \frac{n \bar{\phi} / \sigma_\phi^2 + \mu_0 / \sigma_0^2}{n / \sigma_\phi^2 + 1 / \sigma_0^2}$$

$$\frac{1}{\sigma_n^2} = \frac{1}{\sigma_\phi^2} + \frac{1}{\sigma_0^2}$$

Le risque à posteriori serait égal à:

$$\int_{-\infty}^{+\infty} R_0 (w - w_d)^2 f(w, \phi_i) dw$$

qui est minimum et égal à  $R_0 \sigma_n^2$  pour la meilleure décision égale à  $\mu_n$ .

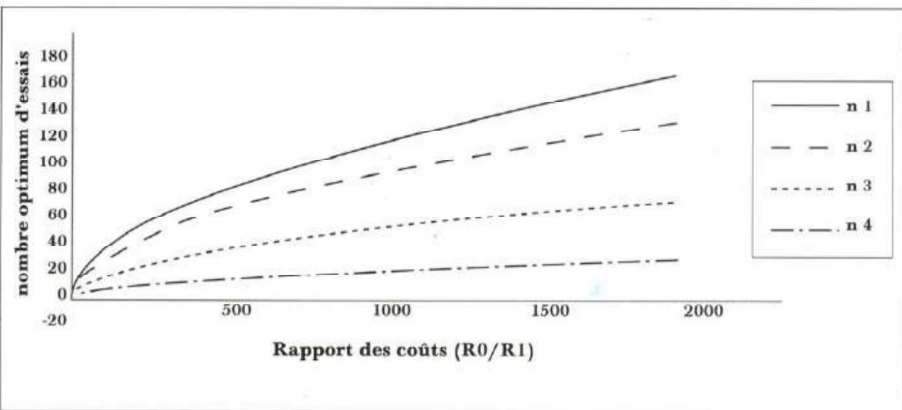
Le gain que procure la nouvelle information est égal à la diminution du risque moins le coût de l'information; c'est-à-dire le coût des essais triaxiaux et que l'on suppose proportionnel au nombre d'essais, soit  $R_1 n$ .

$$G = R_0 (\sigma_0^2 - \sigma_n^2) - R_1 n = R_0 \sigma_0^2 \frac{n}{n + \sigma_0^2 / \sigma_\phi^2}$$

Ce gain est maximum pour:

$$n^* = \sigma_0 \sqrt{\frac{R_0}{R_1} - \frac{\sigma_\phi^2}{\sigma_0^2}}$$

figure 2- Nombre optimal d'essais en fonction des coûts et des variances



# BIBLIOGRAPHIE

- [1] A. BEKKOUCHE (1987) «Sécurité des grands barrages en remblais. Approche probabiliste des problèmes d'écoulements liés aux reconnaissances et contrôles»; Thèse de doctorat en mécanique des sols et structures, Ecole Centrale des Arts et Manufactures de Paris.
- [2] M. H. DEGROOT (1970) «Optimal statistical decisions», Mc Graw-Hill book Company
- [3] H. RAIFA / S. Schlaifer (1961) «Applied statistical decision theory», Harvard University
- [4] A. P. Page J6L MELSA (1971) «Estimation theory with application to communications and control», Mc Graw-Hill book Company
- [5] M. TRIBUS (1972) «Décisions rationnelles dans l'incertain», Traduit de l'américain par J. Pelzer, Massons et Cie éditeurs
- [6] E. VIOLA (1983) «On the bayesian approach to estimate compressive strenght of concret in situ»; 4th ICASP in soil and structural engineering Italy.
- [7] RN. WITHMAN (1984) «Evaluating calculated risk in geotechnical engineering», Journal of geotechnical engineering, ASCE, Vol 110 N°2 - February.

## 5 APPLICATIONS NUMÉRIQUES

**Campagne d'essai 1977:** La question essentielle est celle du choix de  $\sigma_\phi^2$  qui est supposée connue. Après les essais à la boîte de cisaillement, on n'a en principe d'autre choix que de prendre la variance estimée de ces essais comme estimation de  $\sigma_\phi^2$  puisqu'on ne sait rien des essais de l'année suivante; soit:

$$\sigma_\phi^2 = 14,64 \quad \sigma_0^2 = 2,93 \quad \frac{\sigma_\phi^2}{\sigma_0^2} = 5$$

EDF en a réalisé 5 essais alors que le nombre optimal qu'il aurait fallu réaliser en 1978 (Figure 2) est:

$$\begin{aligned} n^*_1 &= 7 && \text{pour } R_0/R_1 = 10 \\ n^*_1 &= 33 && \text{pour } R_0/R_1 = 100 \end{aligned}$$

**Campagne d'essais triaxiaux 1978:** La campagne de 1977, que l'on a pris pour information a priori, a fournie des données très désavantageuses à cause du biais des essais à la boîte et de la dispersion beaucoup plus grande que ceux-ci. On est alors en droit de prendre une valeur inférieure de la variance, estimée à partir de l'échantillon complet (1977-1978) débaisé, soit (Figure 2):

$$\sigma_\phi^2 = 9,03 \quad \sigma_0^2 = 2,93 \quad \frac{\sigma_\phi^2}{\sigma_0^2} = 3,1$$

$$\begin{aligned} n^*_{21} &= 6 && \text{pour } R_0/R_1 = 10 \\ n^*_{22} &= 27 && \text{pour } R_0/R_1 = 100 \end{aligned}$$

Si l'on juge que la variance est toujours très grande, on pourrait l'estimer à partir des essais triaxiaux seulement. Soit (Figure 2):

$$\sigma_\phi^2 = 2,6 \quad \sigma_0^2 = 2,93 \quad \frac{\sigma_\phi^2}{\sigma_0^2} = 0,88$$

$$\begin{aligned} n^*_{31} &= 4 && \text{pour } R_0/R_1 = 10 \\ n^*_{32} &= 15 && \text{pour } R_0/R_1 = 100 \end{aligned}$$

On pourrait dire que l'EDF s'est implicitement placé dans le cas  $R_0/R_1 = 16$ , si la fonction objective a quelque réalisme.

**Nouvelle campagne:** Les essais réalisés en 1977 et 1978 sont à même de servir d'information a priori pour le paramètre  $w$ , dont les caractéristiques de la loi seront maintenant:

$$\mu_0 = \phi_0 = \bar{\phi} = 37,5^\circ$$

$$\mu_n = 37,5^\circ \quad \sigma_n^2 = 0,4$$

La variance à posteriori  $\sigma_n^2$  devient le nouveau  $\sigma_0^2$  et le nombre optimal d'essais d'une troisième campagne, aurait été (Figure 2):

$$\begin{aligned} n^*_4 &= 10 && \text{pour } R_0/R_1 = 100 \\ n^*_4 &= 45 && \text{pour } R_0/R_1 = 1000 \end{aligned}$$

## 6 CONCLUSION

Bien que le nombre de données ne permette pas de tester les quatre facteurs à la fois, l'analyse a permis de montrer que le facteur le plus influent sur l'angle de frottement est le type d'appareil utilisé.

Si la fonction objective choisie ait quelque réalisme, l'analyse a montré que le nombre d'essai a été insuffisant. On peut penser à l'inverse, que la sécurité vis-à-vis d'instabilité n'est absolument pas contrôlée par la moyenne l'angle de frottement interne, ni par sa variance ponctuelle, mais par d'autres paramètres prenant en compte la variabilité spatiale de  $\phi_{pic}$ , comme son échelle de fluctuation; les ruptures ayant lieu en grande masse.

Ceci montre la nécessité d'utiliser les champs stochastiques, pour mieux connaître les aléas des propriétés de la digue et mieux les prendre en compte dans des études d'états critique globaux. Cette question semble d'importance particulière pour les états critiques d'action interne de l'eau.

Enfin, l'illustration d'un calcul d'optimisation du nombre d'essai montre la nécessité d'une analyse économique couplée à l'analyse de sécurité et sans laquelle aucune analyse de risque et donc aucune optimisation des décisions n'est possible ■