

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

Faculté de technologie
Département de Génie Civil



Mémoire pour l'obtention du diplôme de Master en Génie Civil

Option : Ingénierie des Structures

Thème :

**ANALYSE DES DONNEES EN GEOTECHNIQUE POUR LE
DIMENSIONNEMENT DES OUVRAGES CAS DE L'AEROPORT
D'ALGER**

Présenté le 23 juin 2013 par :

Mlle MEGNOUNIF Imène

Devant le jury composé de:

Z. BENYELLES
D. ZENDAGUI
M. MATALLAH
A. ZADJAOUI

président
examineur
examineur
encadreur

Remerciements

En premier lieu, j'exprime mes profonds remerciements à ma petite famille sur tout à mes parents et mes amis et tous ceux qui m'ont apporté leurs aides et leurs soutiens pour la réalisation de ce mémoire.

J'exprime notamment ma profonde gratitude et mes remerciements à mon encadreur Mr Abdeldjlil ZADJAOUI pour son aide et son encadrement.

J'exprime également mes remerciements à :

- ❖ Mr Zouheir BENYELLES enseignant au département de génie civil d'avoir accepté d'examiner mon travail et de présider le jury.
- ❖ Mr Djawad ZENDAGUI et Mr Mouhammed MATALLAH enseignants au département de génie civil d'avoir accepté d'être examinateurs de ce travail de projet de fin d'étude.
- ❖ Aussi à tous les enseignants du département de génie civil, qui m'ont apporté d'une manière directe ou indirecte des idées qui m'ont aidé à la réalisation de ce mémoire de fin d'études.

RESUME

Actuellement, nombre de techniques classiques d'analyse statistiques ont été étudiées dans le domaine de la mécanique des sols. Toutefois, l'analyse de données, quant à elle, a rarement été appliquée à des données géotechniques. L'objectif de cette étude est donc de rechercher l'usage qui peut être fait de ce type d'analyse en mécanique des sols.

Une première partie est consacrée d'une manière générale aux techniques d'analyse statistiques avant de détailler l'analyse de données. L'étude porte plus particulièrement sur l'analyse en Composante Principale (ACP) et l'analyse Factorielle de Correspondances (AFC), méthodes dites de représentation dans un plan factoriel.

Dans une deuxième partie, des données pénétrométriques provenant d'un site connu hétérogène sont traitées au moyen de l'ACP. Cette étude confirme la validité de la méthode pour déterminer les couches de sols et les définir ensuite statistiquement. Enfin, la représentation factorielle est également utilisée pour faciliter l'établissement de cartes d'isométriques de résistance de pointe.

Mots-clés : Mécanique des sols – Statistiques – Analyse de données – ACP – AFC – Plan factoriel – Couches de sols.

ABSTRACT

At present, many conventional statistical analysis techniques have been studied in the field of soil mechanics.

However, analysis of data, temporarily, has rarely been applied to geotechnical data. The objective of this study is to investigate the use which may be made of this type of analysis in soil mechanics.

The first part is devoted to general statistical analysis technique before detailed data analysis. The study focuses on the Principal Component Analysis (ACP) and Factor of correspond (AFC) analysis, called representation in factorial design methods.

In the second part, the penetrometer data are processed using the ACP. An interpretation of the contribution of this technique in this area is the result of this study.

Keywords : Soil Mechanics - Statistics - Data Analysis - ACP - AFC - Factorial - Layers of soil.

ملخص

حالياً قد تم دراسة العديد من التقنيات التحليلية الإحصائية التقليدية في مجال ميكانيكا التربة. ومع ذلك، تحليل البيانات وفي الوقت نفسه، نادراً ما تم تطبيقها على البيانات الجيوتقنية. الهدف من هذه الدراسة هو البحث عن الاستخدام الذي يمكن تطبيقه من هذا النوع من التحليل في ميكانيكا التربة. ويخصص الجزء الأول تقنيات التحليل الإحصائية العامة قبل تحليل البيانات التفصيلية. وتركيز الدراسة على تحليل المكونات الأساسية (ACP) و تحليل عامل من المقابلة (AFC)، طرق مسماة بالبيانات في معلم ذو عوامل خاصة.

في الجزء الثاني، تتم معالجة البيانات غير المتجانسة من اختراق موقع معروف باستخدام ACP. تؤكد هذه الدراسة على صلاحية طريقة لتحديد طبقات التربة ثم يقام بتعيين الطبقات إحصائياً. وأخيراً يتم استخدام التمثيل ذو عوامل خاصة أيضاً إلى تيسير إنشاء خرائط متساوي القياس المقاومة للإكراميات.

كلمات البحث : ميكانيكا التربة - إحصائيات - تحليل البيانات - AFC - ACP - المستوى ذو عوامل - طبقات من التربة.

SOMMAIRE

REMERCIEMENTS.....	i
RESUME	ii
ABSTRACTE.....	iii
ملخص.....	iv
LISTE DES FIGURES.....	viii
LISTE DES TABLEAUX.....	ix
LISTE DES NOTATIONS.....	vii
INTRODUCTION GENERALE.....	1
Chapitre 1 : ETUDE BIBLIOGRAPHIQUE.....	3
1. VARIABILITE DES PROPRIETES DES SOLS	3
1.1. HETEROGENEITE DES SOLS	3
1.2. LES STATISTIQUES ET PROBABILITES EN MECANIQUE DES SOLS... ..	4
1.3. APPORT DES STATISTIQUES ET PROBABILITES EN MECANIQUE DES SOLS.....	6
1.3.1. Méthodes d'analyse.....	6
1.3.1.1. Statistique descriptive.....	6
1.3.1.2. Statistique pour caractère bivariée (régression et corrélation).....	15
1.3.1.3. Statistique pour caractère multi varies.....	18
1.3.1.4. Analyse factorielle des données.....	19
1.3.1.5. L'analyse factorielle des correspondances (AFC).....	20
1.3.1.6. Variations spatiales.....	26
CONCLUSION.....	26
Chapitre 2: ANALYSE DE DONNEES : ANALYSE FACTORIELLE.....	27
2.1. PRESENTATION GENERALE.....	27
2.2. CARACTERISTIQUES DES DONNEES.....	27
2.2.1. Tableau de données.....	27
2.2.2. Données actives et supplémentaires.....	29
2.2.3. Caractéristiques des points.....	29
2.2.3.1. Poids.....	29
2.2.3.2. Centre de gravité (point d'origine).....	30
2.3. ESPACE DES INDIVIDUS.....	31

2.3.1. Rôle de la métrique.....	31
2.4. L'ESPACE DES VARIABLES.....	31
2.4.1. La métrique des poids.....	31
2.5. MOMENT D'INERTIE.....	32
2.5.1. Inertie du nuage des individus par rapport à un axe passant par G.....	32
2.5.2. Inertie du nuage des individus par rapport à un sous-espace vectoriel V passant par G.....	33
2.5.3. Décomposition de l'inertie totale.....	33
2.6. RECHERCHE DE 1 ^{ER} AXE Δ_1 PASSANT PAR G D'INERTIE MINIMUM..	34
2.6.1. Expressions algébriques de I_{Δ_1} et de $\overline{\ G\alpha_1\ ^2}$	34
2.6.2. Recherche du maximum.....	34
2.7. RECHERCHE DES AXES SUIVANTS.....	35
2.8. PARTICIPATIONS DES AXES A L'INERTIE TOTALE.....	35
2.9. REPRESENTATION DES INDIVIDUS DANS LES NOUVEAUX AXES...36	
2.9.1. Qualité de la représentation des individus.....	37
2.9.2. Interprétation des nouveaux axes en fonction des individus.....	38
2.9.2.1. Contribution absolue d'un individu à un axe.....	38
2.9.2.2. Contribution relative d'un individu à un axe.....	38
2.10. REPRESENTATION DES VARIABLES.....	39
2.10.1. Interprétation des axes en fonction des anciennes variables.....	41
2.10.2. Qualité de la représentation des variables.....	41
2.10.3. Étude des liaisons entre les variables.....	41
2.11. ANALYSE EN COMPOSANTES PRINCIPALES NORMEE.....	41
CONCLUSION.....	42

Chapitre 3 : ANALYSE DE DONNEES PENETROMETRIQUES: CAS DE L'AEROPORT HOUARI BOUMEDIENE.....44

3.1. PRESENTATION DU SITE.....	44
-----------------------------------	----

3.1.1. L'Aéroport Houari Boumediene d'Alger.....	44
3.1.2. Caractérisation du sol.....	46
3.1.2.1. Constitution du sol.....	46
3.1.3. Caractéristiques mécaniques.....	47
3.2. PRESENTATION DES DONNEES.....	48
3.2.1. Origine des données.....	48
3.2.2. Présentation du tableau de données.....	48
3.3. ETUDE SONDAGE PAR SONDAGE (SENS VERTICALE).....	49
3.3.1. Objectif de l'étude.....	49
3.3.1.1. Dans le contexte du problème de drainage.....	49
3.3.1.2. Dans le contexte de l'étude.....	50
3.3.2. Présentation de l'étude.....	50
3.3.2.1. Zonage sur plans factoriels.....	50
3.3.3. Caractéristiques statistiques des couches.....	56
3.3.4. Exploitation des résultats.....	56
3.3.4.1. Carte des profondeurs des sols compressibles.....	56
3.3.5. Etude dans les plans horizontaux.....	59
3.3.5.1. Méthodologie.....	59
3.3.5.2. Courbes isométrique de la résistance de pointe.....	60
Conclusion.....	64
CONCLUSION GENERALE	66
ANNEXE A : ESSAI EN PLACE	
ANNEXE B : DONNEES PENETROMETRIQUES DE L'AEROPORT D'ALGER	
REFERENCE BIBLIOGRAPHIQUE	

LISTE DES FIGURE

Tableau 1 1: présentation des effectifs ni (nombre de sondage par pénétromètre dynamique) et fréquences fi (le pourcentage des sondages par pénétration) associés aux nombre de cavités.	7
Tableau 1 2: des effectifs ni (nombre de sondage carottés) et fréquences fi.....	7
(le pourcentage des sondages carottés) associés aux intervalles.....	7
de classes de la teneur en eau w (%).	7
Figure 1 1: Diagramme en bâton.....	8
Figure 1 2: Histogramme.	9
Figure 1 3: Diagramme en bâton tiré d'un histogramme.	9
Figure 1 4: Courbe cumulative.....	10
Tableau 1 3: Tableau de contingence.....	20
Figure 1 5: Représentation graphique des relations entre X et Y (Laffly, 2009).	22
Figure 1 6: Courbe représentant les valeurs propre en fonction du nombre de dimension (Carpentier, 2005).	24
Figure 2 4: Représentations du nuage de point sur le plan formé par les deux premiers axes principaux.....	37
Figure 2 5: Représentation des variables sur le cercle de corrélation	40
Figure 3 1: Localisation de l'aéroport d'Alger.....	44
Figure 3 2: Plan de l'aéroport HOUARI Boumediene.	45
Figure 3 3: Réseau hydrographique de la plaine de la Mitidja.....	45
Figure 3 4: Aménagement de l'Oued Smar.....	46
Figure 3 5 : Plan d'implantation des sondages pénétrométriques.	49
Figure 3 6.a: Biplot factoriel 1-2 de l'ACP des données du sondage 3. Pour Rp.	51
Figure 3 7.b: Biplot factoriel 1-2 de l'ACP des données du sondage 3. Pour Fl.	51
Figure 3 8.a: Biplot factoriel 1-2 de l'ACP des données du sondage 25. Pour Rp.	52
Figure 3 9.b: Biplot factoriel 1-2 de l'ACP des données du sondage 25. Pour Fl.	53
Figure 3 10.a: Biplot factoriel 1-2 de l'ACP des données du sondage 28. Pour Rp.	53
Figure 3 11.b: Biplot factoriel 1-2 de l'ACP des données du sondage 28. Pour Fl.	54
Figure 3 12.a: Biplot factoriel 1-2 de l'ACP des données du sondage 65. Pour Rp.	54
Figure 3 13.b: Biplot factoriel 1-2 de l'ACP des données du sondage 65. Pour Fl.	54
Figure 3 14.a: Biplot factoriel 1-2 de l'ACP des données du sondage 1041. Pour Rp.	55
Figure 3 15.a: Biplot factoriel 1-2 de l'ACP des données du sondage 1041. Pour Fl.....	55
Figure 3 16: Carte de profondeur de la base de la base des sols compressibles.....	58
Figure 3.11 : Carte des profondeurs de la base de la base des sols compressibles établie graphiquement par Magnan.....	58
Figure 3 172: Carte de courbes isométriques de Rp à z=5 m.....	61
Figure 3 183: Carte de courbes isométriques de Rp à z=7,5 m.....	62
Figure 1.14: Carte de courbes isométriques de Rp à z=10 m.....	62
Figure 3 5: Carte de courbes isométriques de Rp à z=12,5 m.....	63
Figure 3 16 : Carte de courbes isométriques de Rp à z=15 m.....	63
Figure 3 17: Carte de courbes isométriques de Rp à z=17,5 m.....	63
Figure 3 198: Carte de courbes isométriques de Rp à z=20 m.....	64

LISTE DES TABLEAUX

Tableau 1 1: présentation des effectifs n_i (nombre de sondage par pénétromètre dynamique) et fréquences f_i (le pourcentage des sondages par pénétration) associés aux nombre de cavités.	7
Tableau 1 2: des effectifs n_i (nombre de sondage carottés) et fréquences f_i	7
Tableau 1 3: Tableau de contingence.....	20
Tableau 3 1 : Profondeur de la base des sols les plus compressibles.	57

INTRODUCTION GENERALE

Le dimensionnement d'un ouvrage de génie civil quel qu'il soit, nécessite en premier lieu une bonne connaissance du sol de fondation. La première étape d'un projet est donc la campagne de reconnaissance géotechnique, elle permet à l'ingénieur de choisir des valeurs représentatives des caractéristiques du sol nécessaire aux calculs. Toutefois, le choix de ces valeurs est une tâche délicate qui requiert chez l'ingénieur praticien une bonne dose d'expérience et de savoir-faire. En effet, il est à l'évidence impossible de définir en tout point d'un site les propriétés du sol ; par conséquent, la détermination des valeurs représentatives des paramètres s'effectue en général sur la base de quelques résultats d'essais sur éprouvettes prélevées pratiquement au hasard et d'essais en place réalisés suivant un maillage plus ou moins large.

La mécanique des sols classique repose sur des hypothèses simplificatrices qui ignorent les caractéristiques des sols ; le sol est généralement défini par l'hétérogénéité, la diversité la variabilité dans l'espace et l'inconnue.

Les informations retirées des propriétés physiques est mécaniques définissent les valeurs représentatives des paramètres de calcul, ces valeurs sont en globale incertaine ce qui se traduit par des incertitudes sur la prévision du comportement des ouvrages par la suite.

De plus, il est généralement admis que les modèles que nous utilisons pour le calcul du comportement des sols de fondations ne sont pas très satisfaisants.

Beaucoup d'entre eux semblent inadaptés dès qu'on les applique à des massifs de sols de dimensions réelles et qu'on compare leurs prévisions aux résultats des mesures réalisées sur un site réel. Dans une optique « déterministe » classique, cette comparaison défavorable doit conduire à rejeter le modèle rhéologique ou à chercher de meilleures valeurs des paramètres.

Une nouvelle approche de la mécanique des sols s'imposait donc pour tenter de résoudre ces problèmes d'hétérogénéité des sols et de modélisation de leur comportement ou du moins pour mieux les appréhender. Les méthodes statistiques et probabilistes sont donc venues compléter les méthodes déterministes classiques.

C'est pourquoi, depuis environ une quarantaine d'années, une approche statistique et probabiliste de la mécanique des sols s'est développée et a pris une certaine importance dans de nombreuses branches de la géotechnique. L'application des techniques d'analyse statistique et probabiliste a pour objectif de traiter de façon quantitative l'influence de la variabilité naturelle des sols sur le comportement des ouvrages et d'éliminer toute subjectivité dans les choix de l'ingénieur.

A notre connaissance très peu d'études ont été menées dans le domaine de la géotechnique sur l'analyse factorielle de données : rarement vu se développer, par exemple, des recherches sur l'analyse en composante principale ou l'analyse factorielle des correspondances en géotechnique.

Notre travail de projet de fin d'études se veut comme une contribution dans ce sens. Des données pénétrométriques provenant d'un site connu hétérogène sont traitées par la technique dite « Analyse en Composante Principale (ACP) ». On vise la validité de cette

méthode pour déterminer les couches de sols et les définir ensuite statistiquement, puis établir des cartes d'isométriques de résistance de pointe à l'aide de la présentation factorielle

Actuellement, nombre de techniques classiques d'analyse statistiques ont été étudiées dans le domaine de la mécanique des sols. Toutefois, l'analyse de données, quant à elle, a rarement été appliquée à des données géotechniques. L'objectif de cette étude est donc de rechercher l'usage qui peut être fait de ce type d'analyse en mécanique des sols.

Une première partie est consacrée pour exposer l'intérêt et les objectifs des techniques d'analyse statistiques et d'analyse factorielle des données en mécanique des sols.

Dans une deuxième partie, des données pénétrométriques sont traitées au moyen de l'ACP. Le résultat du travail consiste à l'élaboration de courbes isométriques de R_p (résistance à la pointe), élément indispensable pour la vérification de la stabilité de notre ouvrage (structure de l'aéroport).

Une interprétation de l'apport de cette technique dans ce domaine est le fruit de cette étude. Nous achèverons par une conclusion relative à ce travail.

Vu le nombre de pages limité dans ce rapport, la partie des essais en place et la méthodologie de vérification de l'ouvrage n'ont pas été rédigée dans ce rapport. L'annexe A donne un résumé sur l'essai de pénétration dynamique qui est le sujet de notre étude.

Chapitre 1 : ETUDE BIBLIOGRAPHIQUE

Ce chapitre de caractère bibliographique présente une synthèse des méthodes statistiques et probabilistes utilisées en géotechnique surtout dans le domaine des structures en génie civil. Il se veut comme des notions de base et éléments nécessaires pour résoudre la problématique posée en introduction générale de ce mémoire.

2. VARIABILITE DES PROPRIETES DES SOLS

1.4. Hétérogénéité des sols :

L'étude des sols et des roches sert de cadre aux travaux de génie civil. Ces matériaux constituent les couches superficielles de l'écorce terrestre. Leurs caractéristiques les plus appropriés sont (Magnan, 1993):

- la diversité,
- l'hétérogénéité,
- la variabilité dans l'espace
- l'inconnu.

Diversité, car on trouve aussi bien des matériaux massifs et très résistants, comme certaines roches granitiques, que des bancs fissurés de calcaires, des couches solubles de gypse ou de sels, des cailloux et graviers aux arêtes arrondies ou vives, des sables fins, des argiles, ou des tourbes. Certains matériaux sont perméables, d'autres non. Certains sont fissurés, d'autres constituent des massifs homogènes...

Hétérogénéité, car il est rare qu'une couche de sol ou de roche soit constituée d'un matériau unique : les massifs rocheux sont souvent stratifiés et très souvent traversés de fissures remplies d'éléments fins ou de concrétions, les couches de sables ou graviers contiennent souvent des blocs de plus grandes dimensions et des particules plus fines d'argile ou de limon, les couches d'argile peuvent contenir des grains de sable, des coquillages, des particules organiques, les sols tourbeux ont des fractions minérales parfois importantes,...

Variabilité dans l'espace, qui résulte des conditions de formation et d'évolution des matériaux de l'écorce terrestre : formation par soulèvement de matériaux des couches internes du globe terrestre (volcanisme, orogénèse) ou par plissement, formation par désagrégation des roches, par transport gravitaire, glaciaire, fluvial ou éolien, ou par sédimentation en eau douce ou en eau salée... La formation de la plupart des roches sédimentaires a duré pendant des centaines, voire des milliers d'années et on y retrouve les traces des variations des courants, des saisons, du niveau des mers, des températures, de l'origine des matériaux transportés,...

Ignorance, car on ne connaît jamais en tout point d'un massif de sols ou de roches la nature et les propriétés physiques et mécaniques du matériau : ces observations et ces mesures sont effectuées sur des carottes prélevées à l'intérieur du massif à étudier ou sur des affleurements et ne concernent dans le meilleur des cas qu'une fraction très faible du volume du massif (de l'ordre de 10^{-6} à 10^{-8} du volume total. A titre de comparaison, un sondage d'opinion courant teste de l'ordre de 3.10^{-5} de la population). Un tel

échantillonnage serait totalement insuffisant pour définir des données réalistes pour le calcul des ouvrages de génie civil si l'on ne pouvait compléter la structure du massif en s'appuyant sur les règles de la géologie et sur les données des mesures géophysiques.

La mécanique des sols et des roches doit donc découvrir et décrire le matériau avant d'en modéliser le comportement mécanique ou physique. Une fois le matériau identifié et ses frontières définies, on peut alors utiliser des modèles de calcul, dont certains sont issus de la mécanique des milieux continus, d'autres de la mécanique des solides ou de la mécanique des fluides, avant d'en interpréter les résultats pour répondre aux questions posées. Une des difficultés de la mécanique des sols et des roches réside justement dans l'interaction de la reconnaissance des sites et de la modélisation, la qualité des analyses et des prévisions dépendant autant de l'une que de l'autre.

Pour le mécanicien des sols, dont l'objectif final est de dimensionner des ouvrages, cette variabilité naturelle ne présente d'intérêt que par ses conséquences sur les propriétés physiques et mécaniques qu'il utilise dans ses calculs et éventuellement par ses conséquences sur les modèles de calcul eux-mêmes. Du point de vue de la mécanique des sols, la principale question qui se pose est la suivante : les sols d'un site ont-ils des propriétés variables d'un point à un autre et, si c'est le cas, comment peut-on en tenir compte dans les calculs ?

En tout état de cause, connaissant cette variabilité naturelle des sols ainsi que le nombre restreint de sondages et d'essais effectués, on peut facilement imaginer la part de subjectivité qui entre dans la détermination des valeurs représentatives des caractéristiques physiques et mécaniques des sols. En effet, le choix de ces valeurs relève en grande partie du savoir-faire de l'ingénieur praticien et, selon son expérience, il présentera le résultat de son calcul comme une valeur précise et fiable ou comme une valeur prévisionnelle précise à 10 ou 20% près.

De plus, il est généralement admis que les modèles que nous utilisons pour le calcul du comportement des sols de fondations ne sont pas très satisfaisants. Beaucoup d'entre eux semblent inadaptés dès qu'on les applique à des massifs de sols de dimensions réelles et qu'on compare leurs prévisions aux résultats des mesures réalisées sur un site réel. Dans une optique « déterministe » classique, cette comparaison défavorable doit conduire à rejeter le modèle rhéologique ou à chercher de meilleures valeurs des paramètres.

Une nouvelle approche de la mécanique des sols s'imposait donc pour tenter de résoudre ces problèmes d'hétérogénéité des sols et de modélisation de leur comportement ou du moins pour mieux les appréhender. Les méthodes statistiques et probabilistes sont donc venues compléter les méthodes déterministes classiques

1.5. LES STATISTIQUES ET PROBABILITES EN MECANIQUE DES SOLS :

Le développement des méthodes statistiques et probabilistes en mécanique des sols est donc dû :

- ❖ D'une part au besoin ressenti par les géotechniciens de résoudre les problèmes posés par l'influence de la dispersion des caractéristiques des sols sur le calcul des ouvrages ;
- ❖ D'autre part au désir des spécialistes des statistiques et de probabilités de trouver de nouvelles applications concrètes à leurs travaux théoriques.

A l'origine l'utilisation des méthodes statistiques a voulu satisfaire aux deux objectifs suivants :

Caractériser la variabilité naturelle des sols pour l'introduire dans les schémas de calcul classiques des fondations des ouvrages. Comme il est impossible de déterminer en tout point les propriétés d'un sol sur le site d'un projet, l'ingénieur doit choisir des valeurs (moyennes par exemple) représentatives de ces propriétés. Le choix de ces valeurs conditionnant fortement la confiance que l'on peut accorder aux calculs, le premier but a été donc de contrôler la fiabilité des données et d'estimer la précision des résultats (Houmadi, 2004).

Augmenter l'efficacité de la reconnaissance géotechnique. Celle-ci étant budgétairement limitée, il paraît important de l'optimiser en jouant par exemple sur le type, le nombre et la position des sondages. Pour cela on a cherché à utiliser au mieux les corrélations existant entre les différentes propriétés des sols. Cette démarche est particulièrement intéressante car les méthodes de détermination de ces propriétés sont plus ou moins rapides que peu d'essais longs et complexes (Abdul Baki et al., 1993)

Ces études statistiques présentent donc un intérêt pratique certains, mais elles ne doivent pas constituer une fin en soi, leur emploi doit avoir comme but de mieux définir les hypothèses de calcul déterministes. Des méthodes de calcul probabiliste ont également été développées. Ces méthodes, en tenant compte de l'influence de la variabilité des paramètres, permettent de présenter les résultats des calculs sous forme d'une valeur moyenne et d'une incertitude ou mieux encore d'une distribution de probabilité des déplacements et d'une probabilité de rupture.

Le domaine d'application actuel des méthodes statistiques et probabilistes est donc assez vaste mais il reste encore des techniques d'analyse inexploitées ou en cours d'étude. Cependant nous donnerons en guise de conclusion une liste, déjà assez longue ; établie par Magnan en 1983, des principaux secteurs de l'activité du géotechnicien pouvant bénéficier de l'apport de ces méthodes :

A. Au stade de la reconnaissance géotechnique

- i. Choix du nombre et de l'implantation des sondages ;
- ii. Choix des types de sondage et d'essais (utilisation de corrélations) ;
- iii. Mise en forme des résultats et préparation de coupes géotechniques pour les calculs.

B. Au stade de calcul des ouvrages

- i. Calcul de la stabilité de l'ouvrage en termes de probabilité de rupture ;
- ii. Calcul de la distribution de la probabilité (ou de la valeur moyenne de l'incertitude) pour les déformations des sols de fondations au cours du temps.

- iii. Détection des points faibles dans les ouvrages complexes (zones de plus grand risque dans un barrage, ouvrage le plus critique dans un ensemble d'ouvrages de soutènement en zone instable.

C. Au stade de contrôle sur l'ouvrage réalisé

- i. Détermination des zones où l'incertitude est maximale et où il faut placer l'instrumentation ;
- ii. Contrôle de la qualité des travaux, etc.

L'auteur de la même référence précise toutefois que si de nombreux travaux ont déjà été consacrés à ces différents problèmes, on doit constater qu'aucun d'entre eux n'a encore été réglé de façon définitive même si d'énormes ont été faits.

La variabilité naturelle des sols constituant le phénomène d'origine d'un grand nombre d'erreurs et d'incertitudes dans les calculs, elle possède une importance capitale. Nous consacrons la suite de ce rapport bibliographique aux techniques statistiques et probabilistes employées pour son analyse.

1.3. APPORT DES STATISTIQUES ET PROBABILITES EN MECANIQUE DES SOLS

1.3.1. Méthodes d'analyse:

L'étude d'un sol passe par la collecte de données numériques. La statistique permet de traduire ces valeurs numériques en informations utiles pour la réalisation du projet investi. La méthode statistique comporte les étapes suivantes:

- ❖ La statistique descriptive qui est un ensemble des méthodes permettant d'analyser les données à partir de paramètres et de graphes.
- ❖ Les modèles statistiques qui permettent d'ajuster les résultats recueillis sur un échantillon à des lois de probabilité afin de faire des prévisions et des interpolations sur la population. Cette statistique est basée sur la recherche d'un échantillon qui représente les données géotechniques. C'est-à-dire, il représente le mieux possible la diversité de la population entière.

1.3.1.1. Statistique descriptive :

a) Notion de base :

L'ensemble d'individus sur lequel on effectue une analyse statistique est dit population. Le nombre d'individus est l'effectif de la population. L'échantillon est l'ensemble d'individus prélevés dans une population déterminée pour l'étude demandée. Une situation possible, M_i du caractère définie une modalité.

La caractéristique étudiée est dite caractère ou variable statistique. On distingue les caractères qualitatifs et les caractères quantitatifs qui peuvent être discrets ou continus.

Les informations recueillies peuvent être résumés dans des tableaux statistiques. Pour un caractère discret, à chaque valeur observée x_i est affecté un effectif n_i ou une fréquence f_i qui désignent respectivement le nombre ou le pourcentage d'apparition de la valeur observée (exemple tableau 1 1).

Tableau 1 1: présentation des effectifs n_i (nombre de sondage par pénétromètre dynamique) et fréquences f_i (le pourcentage des sondages par pénétration) associés aux nombre de cavités.

Nombre de cavités	Nombre de sondage	Fréquence
0	1	0,018
1	4	0,070
2	7	0,123
3	11	0,193
4	5	0,088
5	8	0,140
6	7	0,123
7	3	0,053
8	3	0,053
9	1	0,018
10	7	0,123

Pour un caractère continu, l'effectif n_i ou une fréquence f_i sont associés à des classes d'intervalles (exemple tableau 1 2).

Tableau 1 2: des effectifs n_i (nombre de sondage carottés) et fréquences f_i (le pourcentage des sondages carottés) associés aux intervalles de classes de la teneur en eau w (%).

Teneur en eau w (%)	Nombre de sondage	Fréquence
[25 - 29[2	0,0322
[29 - 33[6	0,0967
[33 - 37[6	0,0967
[37 - 41[11	0,1774
[41 - 45[10	0,1612
[45 - 49[8	0,1290
[49 - 53[7	0,1129
[53 - 57[3	0,0484
[57 - 61[7	0,1129
[61 - 65[1	0,0161
[65 - 69[1	0,0161

Un tableau peut être présenté par des graphes. Ces derniers permettent une lecture simple et rapide des données recueillies. On présente les graphes les plus utilisés et qui serviront pour la représentation de ce mémoire.

- **Caractère discret :**

Diagramme en bâtons: il représente les effectifs ou fréquences associées aux valeurs observées. La répartition du nombre de cavités observées sur les sondages par pénétromètre résumé dans le tableau 1 peut être représentée par le diagramme en bâtons suivant (figure. 1.1).

On peut lire dans sept sondages il-y- a eu dix cavités. Les sondages ou on a observé trois cavités sont en nombre de onze. Trois cavités correspondent à l'effectif maximal, c'est le mode.

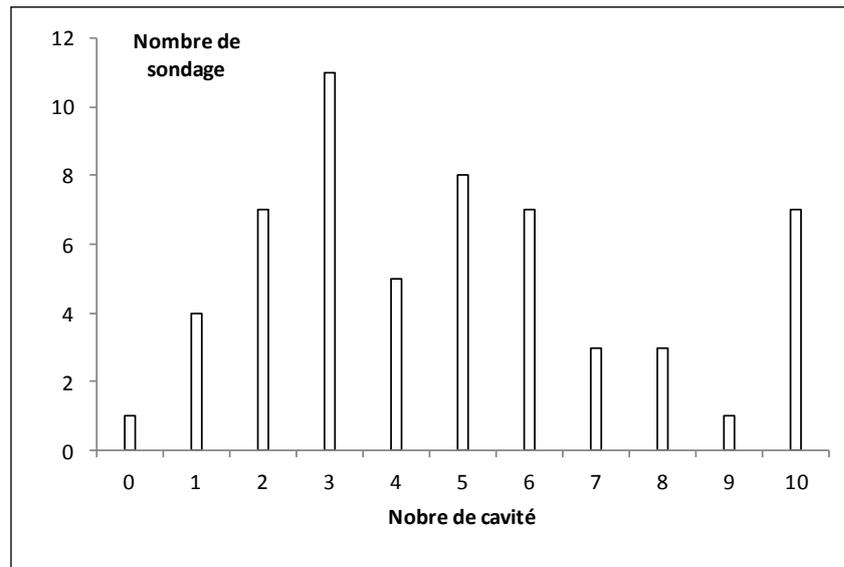


Figure 1 1: Diagramme en bâton.

- **Caractère continu :**

Histogramme: il représente le nombre d'effectif pour des intervalles de paramètres observés. Au tableau 02 correspond l'histogramme suivant (figure 1.2).

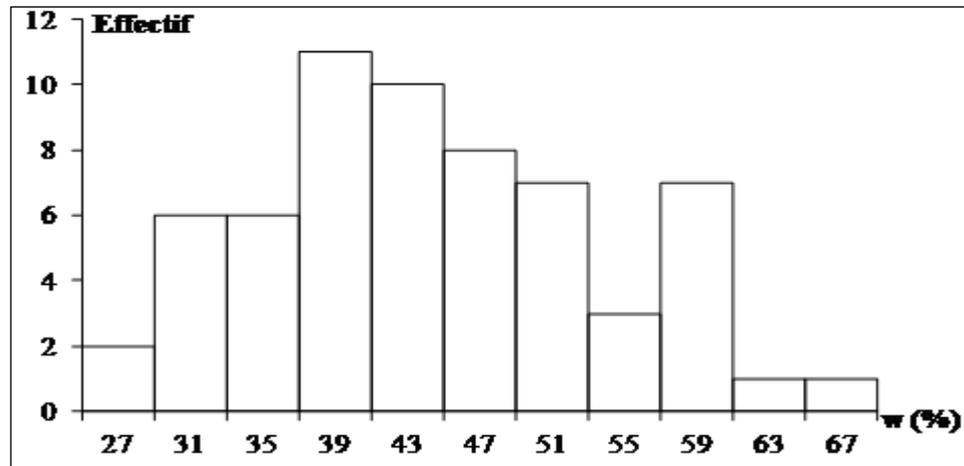


Figure 1 2: Histogramme.

A des intervalles de valeurs de la teneur en eau est associé le nombre de sondages carottés. On peut lire que deux sondages carottés ont une teneur en eau entre 25 et 29%. Alors que onze sondages carottés ont une teneur en eau entre 37 et 41 %.

A partir d'un histogramme on peut tracer un diagramme en bâton, exemple (figure 1.3). Où à chaque centre de classe on affecte la valeur de l'effectif associé.

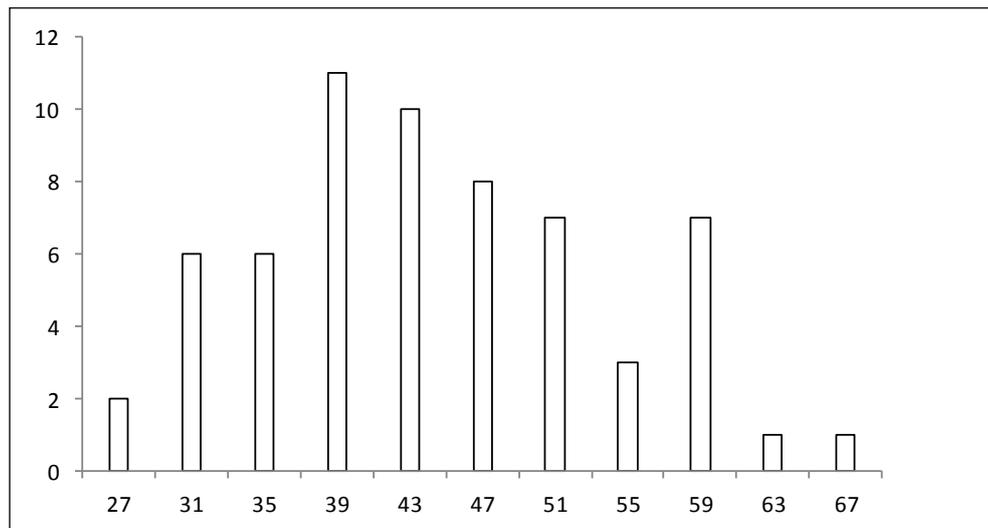


Figure 1 3: Diagramme en bâton tiré d'un histogramme.

Le diagramme intégral correspond à une représentation des effectifs ou fréquences cumulés soit par des courbes en escaliers ou des courbes cumulatives.

La courbe cumulative est la fonction qui représente les fréquences cumulées en fonction du paramètre étudié. Du tableau 1 2, on a (figure 1.4).

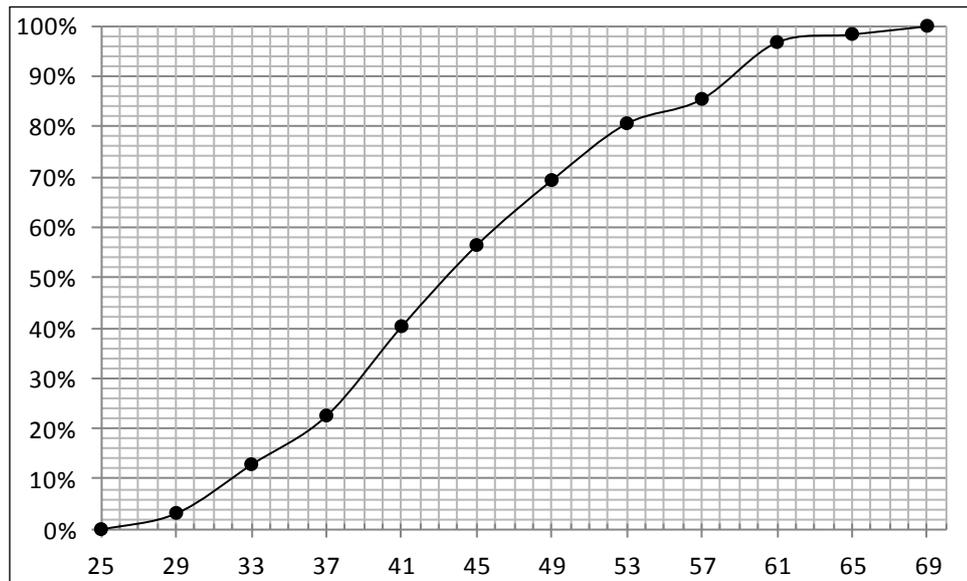


Figure 1 4: Courbe cumulative.

En plus des représentations graphiques l'étude peut être complétée par des paramètres statistiques:

- Les paramètres de position rendent compte de l'ordre de grandeur de l'ensemble des observations. Parmi ces derniers, on a la moyenne (dite aussi espérance mathématique), la médiane et le mode. Ils permettent aussi de localiser la (ou les) zone(s) des fréquences maximums.
- Les paramètres de dispersion se rapportent aux écarts entre les valeurs observées comme la variance, l'écart type (dit aussi moyenne quadratique) et les moments statistiques. L'histogramme, ou le diagramme, des fréquences donnent une idée sur la dispersion des données autour de la moyenne.
- Les paramètres de forme comme les coefficients de variation (C_v), d'asymétrie et d'aplatissement (β_2) donnent une idée sur la régularité des observations ainsi que sur les zones de forte ou faible fréquence (Immediato, 2010).

b) Lois de distribution statistique:

Une loi de probabilité est donnée par une fdp (fonction de densité de probabilité), $f(x)$, et une fonction de répartition $F(t)$.

La fonction f doit être positive ou nulle et d'intégral égale à 1 ($f(x) \geq 0$ et $\int_{-\infty}^{+\infty} f(x) dx = 1$).

La probabilité de non dépassement est donnée par:

$$F(t) = \text{prob}(X \leq t) = \int_{-\infty}^t f(x) dx$$

(1-1)

Parmi les lois de probabilité on distingue celles associées aux caractères discrets tels que:

- ❖ Loi uniforme,
- ❖ Loi de Bernoulli,
- ❖ Loi binomial,
- ❖ Loi binomial négative,
- ❖ Loi de poisson,
- ❖ Loi géométrique,

Les caractères continus suivent les lois suivantes:

- ❖ Loi uniforme,
- ❖ Loi normale ou loi de Laplace-Gauss
- ❖ Loi normale réduite,

Les lois déduites de la loi normale sont:

- ❖ Loi du χ^2 de Pearson
- ❖ Loi de Student
- ❖ Loi de Fisher-Snédecor

Selon Abdul Baki et al. (1993), les lois les plus utilisées en mécanique des sols sont la loi normale et la loi béta. On les applique aussi au logarithme et inverse du paramètre étudié.

- **Loi normale**

Elle est dite aussi loi de Laplace-Gauss. Elle est définie par deux paramètres m et σ qui désignent, l'espérance mathématique (moyenne) et l'écart type. La Fdp est donnée par :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1(x-m)^2}{2\sigma^2}}$$

(1-2)

La fonction est symétrique et se présente sous forme de cloche ayant pour coefficient d'asymétrie et d'aplatissement $\beta_1=0$ et $\beta_2=3$. C'est une loi de référence dite loi centrale.

- **Loi Log-Normale:**

Une variable aléatoire X suit une loi log-normale quand son logarithme suit une loi normale. C'est à dire que: $Y = \ln X$ suit une loi Normale. La fdp s'écrit:

$$f(x) = \frac{1}{x} \cdot \frac{1}{\beta\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{\ln(x)-\alpha}{\beta}\right)^2}$$

(1-3)

α et β sont respectivement l'espérance et l'écart type de la variable $Y = \ln(X)$

- **Loi bêta:**

C'est une loi pour les caractères continues, bornée par deux paramètres de forme notés α et β , et définies sur $[0,1]$. C'est une loi avec deux paramètres seulement.

Sa fonction de densité $f(x; \alpha; \beta)$ de probabilité a plusieurs formes pour des valeurs variables de α et β :

$$\alpha = m_x \left(\frac{m_x(1 - m_x)}{V} - 1 \right)$$

(1-5)

$$\beta = (1 - m_x) \left(\frac{m_x(1 - m_x)}{V} - 1 \right)$$

(1-6)

$$f(x; \alpha; \beta) = \begin{cases} \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\int_0^1 u^{\alpha-1}(1-u)^{\beta-1} du} & x \in [0, 1] \\ 0 & \text{sinon} \end{cases}$$

(1-7)

$$f(x; \alpha; \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \mathbb{1}_{[0,1]}(x)$$

(1-8)

$$f(x; \alpha; \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1} \mathbb{1}_{[0,1]}(x)$$

(1-9)

Où Γ est la fonction gamma définie par:

$$\Gamma(\alpha) = \int_0^{+\infty} e^{-t} \cdot t^{\alpha-1} dt$$

(1-10)

$[0,1]$ (x) : est la fonction caractéristique de $[0,1]$,

B: est la fonction bêta constante de normalisation permettant à la densité de s'intégrer à l'unité,

Dans cette loi espérance mathématique :

$$E(x) = \frac{\alpha}{\alpha+\beta} \quad (1-11)$$

et la variance

$$V(x) = \frac{\alpha\beta}{(\alpha+\beta)^2 \cdot (\alpha+\beta+1)} \quad (1-12)$$

Rappelle sur les principales formules statistiques:

$$\text{- Moyenne: } \bar{X} = m_x = \frac{1}{n} \sum_{i=1}^n x_i \quad (1-13)$$

$$\text{- Variance: } V(x) = \frac{1}{n-1} \sum_{i=1}^n (x_i - m_x)^2 \quad (1-14)$$

$$\text{- Ecart type: } \sigma = \sqrt{V(x)} \quad (1-15)$$

$$\text{- Coefficient de variation: } Cv(x) \text{ ou } Cx = \frac{\sigma_x}{m_x} \quad (1-16)$$

$$\text{- Coefficient d'asymétrie: } \sqrt{\beta_1} = \frac{\mu_3}{\sigma^3} \quad (1-17)$$

$$\text{- Coefficient d'aplatissement: } \beta_2 = \frac{\mu_4}{\sigma^4} \quad (1-18)$$

$$\mu_3 = \frac{n}{(n-1)(n-2)} \sum_{i=1}^n (x_i - m_x)^3 \quad (1-19)$$

$$\begin{aligned} \mu_4 = & \frac{n}{(n-1)(n-2)(n-3)} \\ & \times \left\{ (n+1) \sum_{i=1}^n (x_i - m_x)^4 - 3 \frac{(n-1)}{n} \left[\sum_{i=1}^n (x_i - m_x)^2 \right]^2 \right\} \\ & + 3 \left[\frac{1}{n-1} \sum_{i=1}^n (x_i - m_x)^2 \right]^2 \end{aligned} \quad (1-20)$$

μ_3 Est proche de zéro lorsque la distribution des valeurs du paramètre

c) Testes statistiques:

Lors des tests statistiques paramétriques, nous nous référerons aux mêmes lois de distribution pour trouver la probabilité que les données soient conformes à une certaine hypothèse.

Les tests statistiques sont faits pour assurer une comparaison entre les observations à partir d'une étude statistique et aux valeurs, aux quelles on peut s'attendre d'une hypothèse quelconque, Il existe plusieurs teste entre eux, les tests d'adéquation et comparaison de distributions, qu'on utilise dans ce chapitre.

- **Test d'adéquation de Kolmogorov:**

Propose de construire une statistique de test basée sur la distance entre la fonction de répartition F et une estimation de la fonction de répartition de la variable X qui est supposée suivre une loi quelconque obtenue à partir des observations, on pose les hypothèses de tests suivantes :

H_0 : X suit la loi F

H_1 : X suit une autre loi.

L'équation

$$f_n(x) = \frac{\text{card}(i/x_i \leq x)}{n} \quad x \in \mathbb{R} \quad (1-21)$$

Kolmogorov propose la statistique de teste suivante

$$D_n = \sup |f_n(x) - f(x)|$$

- **Test d'adéquation du chi 2 : loi discrète**

La loi de distribution se rapproche de la distribution du χ^2 à $(k-r-1)$ degrés de liberté, quand la loi théorique est valide (Monbet, 2009)

On utilise une statistique de test du chi2 donnée par :

$$T = \sum_{k=1}^K \frac{(n_i - e_i)^2}{e_i}$$

(1-22)

K : le nombre de classe dans les quelle on a regroupé les n valeurs,

n_i : Nombre d'observation qui tombent dans la $i^{\text{ème}}$ classe,

e_i : Nombre théorique des éléments de la classe i ;

r : Nombre de paramètres qu'il a fallu estimer pour définir la loi tester,

On suppose que la loi de distribution théorique, est la même loi de distribution réelle de la population.

Le teste par rapport à χ_α^2 qui est une valeur particulière de T , ou α représente le niveau de signification qui se déroule de la façon suivante:

- Diviser l'étendue de l'échantillon en K classe, La classe est limitée par les valeurs de variable, et on compte le nombre des observations qui tombent en chaque classe.
- Calculer les effectifs e_i des classes pour la loi testée;

$$e_i = n. [g(x) - g(x_{i-1})]$$

(1-23)

$g(x)$, est une fonction de répartition

- Calculer T_1 ,
- Chercher dans une table des valeurs de χ_α^2 correspondant au nombre de degré de liberté $(k-r)$, puis fixer une valeur du niveau de signification α qui est liée à χ_α^2 , et dire si la loi testée est valide ($T_1 \leq \chi_\alpha^2$) (ABDUL BAKI et al., 1993).

- **Test de Shapiro-Wilk:**

Ce teste se fait que pour les variable qui suivent une loi normale, il est basé sur une comparaison de deux estimateurs liés a la variance, cette comparaison est faite par un rapport (Monbet, 2009).

La procédure tu teste est la suivante:

- Les observations s sont classées par ordre de grandeurs croissantes:
- Calculer la valeur de W qui est donné par:

$$W = \frac{(\sum_{i=1}^n a_i x_{(i)})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

(1-24)

$a_{(i)}$: constantes données par:

$$(a_1, \dots, a_n) = \frac{E^T V^{-1}}{(E^T V^{-1} V^{-1} E)^{1/2}}$$

(1-25)

- Tester la signification de W par la table de Shapiro-Wilk

1.2.1.2. Statistique pour caractère bivariée (régression et corrélation):

Les paramètres étudiés peuvent être mixtes. Dans le cas de caractère bivariée, la variable statistique à deux dimensions est une variable composée de deux variables statistiques à une dimension, dont l'une est explicative et l'autre est expliquée et qui se présente par un tableau à double entrée appelé tableau de contingence.

Lorsqu'on a deux grandeurs statistiques X et Y, on a pour objectif:

- ❖ D'analyser la liaison et voir s'il existe une relation entre X et Y. l'analyse se fait en général par un nuage de n points définie par $\{(x_i, y_i); 1 \leq i \leq n\}$, cette analyse graphique permet de comprendre les différentes caractéristiques énumérées ci-dessous:

- ❖ Situer les proximités entre les individus,
- ❖ Etudier la forme globale des points,
- ❖ Voir, s'il existe une forme de liaison ou de régularité,
- ❖ Détecter visuellement les points qui s'écartent des autres.

- ❖ Définir la forme de la relation entre X et Y. La relation peut être Linéaire positive, linéaire négative comme on peut avoir des liaisons monotones non-linéaire, liaisons non-linéaires et non-monotones ou absence de liaison.

- ❖ Quantifier l'intensité de la liaison.

a) Droite de corrélation :

Lorsqu'on a un nuage de point qui présente une tendance linéaire, on cherche la meilleure droite qui passe le plus proche possible des points expérimentaux. La droite est représentée de telle manière à minimiser les erreurs e_i qui présentent les distances entre les points et la droite, cette dernière est nommée droite de corrélation ou de régression et qui a pour équation:

$$Y = aX + b. \quad (1-26)$$

- **Critère des moindres carrés**

La droite est tracée en déterminant les coefficients a et b, de telle sorte que $\sum e_i \rightarrow 0$ avec :

$$e_i = y_i - (a x_i - b) \quad (1-27)$$

Le principe de la méthode des moindres carrés, est que la somme des carrés des distances e_i soit minimale. Les paramètres a et b du modèle doivent donc vérifier le minimum de la fonction

$$\varphi(a, b) = \left(\sum e_i \right)^2 = (y_i - (a x_i - b))^2 \quad (1-28)$$

Le développement de cette équation peut être arrangé suivant l'équation suivante :

$$\varphi(a, b) = \left(\sum_{i=1, \dots, n} x_i^2 \right) a^2 + nb^2 + \left(2 \sum_{i=1, \dots, n} x_i \right) ab - \left(2 \sum_{i=1, \dots, n} x_i y_i \right) a - \left(2 \sum_{i=1, \dots, n} y_i \right) b + \left(\sum_{i=1, \dots, n} y_i^2 \right) \quad (1-29)$$

$\varphi(a, b)$ Est une fonction à deux variables. Le tracé de la fonction $\varphi(a, b)$ représente une concavité vers le haut. On déduit alors que le minimum existe et il est unique. Donc le minimum est obtenu là où les dérivées premières s'annulent.

On obtient:

$$\begin{cases} \frac{\partial \varphi}{\partial a}(a, b) = \frac{\partial}{\partial a} \left(\sum_{i=1, \dots, n} (y_i - ax_i - b)^2 \right) = 0 \\ \frac{\partial \varphi}{\partial b}(a, b) = \frac{\partial}{\partial b} \left(\sum_{i=1, \dots, n} (y_i - ax_i - b)^2 \right) = 0 \end{cases} \quad (1-30)$$

On divise tout les termes du second membre du système par n, on obtient:

$$\begin{cases} a \frac{1}{n} \sum_{i=1; n} x_i^2 + b \frac{1}{n} \sum_{i=1; n} x_i = \frac{1}{n} \sum_{i=1; n} x_i y_i \\ a \frac{1}{n} \sum_{i=1; n} x_i + b = \frac{1}{n} \sum_{i=1; n} y_i \end{cases} \quad (1-31)$$

Le système peut être écrit par:

$$\begin{cases} a \overline{X^2} + b \overline{X} = \overline{XY} \\ a \overline{X} + b = \overline{Y} \end{cases} \quad (1-32)$$

En résolvant le système, on utilise le déterminant on obtient:

$$a = \frac{1}{V(X)} \begin{vmatrix} \overline{XY} & \overline{X} \\ \overline{Y} & 1 \end{vmatrix} = \frac{\overline{XY} - \overline{X} \overline{Y}}{V(X)} \quad (1-33)$$

et:

$$b = \overline{Y} - \overline{X} \frac{\overline{XY} - \overline{X} \overline{Y}}{V(X)} \quad (1-34)$$

- **Qualité de l'ajustement linéaire entre les caractères X et Y**

Le coefficient R ne mesure que le caractère linéaire d'une liaison. Son usage doit être réservé à des nuages où les points répartis suivant une tendance linéaire. La qualité de la corrélation peut être mesurée par un coefficient de corrélation R.

$$R = \frac{\sum_{i=1; n} (x_i - \overline{X})(y_i - \overline{Y})}{\sqrt{\sum_{i=1; n} (x_i - \overline{X})^2 \sum_{i=1; n} (y_i - \overline{Y})^2}} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} \quad (1-35)$$

- **Interprétation du coefficient de corrélation**

On considère le vecteur \vec{Y} représentant les valeurs observées du caractère Y et \vec{Y}^* représentant les valeurs estimées par la droite de corrélation :

$$\vec{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad \text{et} \quad \vec{Y}^* = \begin{pmatrix} ax_1 + b \\ ax_2 + b \\ \vdots \\ ax_n + b \end{pmatrix} \quad (1-36)$$

La corrélation est d'autant satisfaisante que l'angle θ formé entre les deux vecteurs soit faible. Ce pendant une relation entre l'angle θ et les vecteurs \vec{Y} et \vec{Y}^* est donnée par le produit scalaire :

$$\langle \vec{Y}; \vec{Y}^* \rangle = \|\vec{Y}\| \|\vec{Y}^*\| \cos \theta \quad (1-37)$$

Après arrangement, on trouve :

$$\cos(\theta) = \frac{\langle \vec{Y}; \vec{Y}^* \rangle}{\|\vec{Y}\| \|\vec{Y}^*\|} = \frac{\text{Cov}(X; Y)}{\sigma_X \sigma_Y} = R \quad (1-38)$$

1.2.1.3. Statistique pour caractère multi varies:

Les paramètres étudiés peuvent être mixtes. La variable statistique à plusieurs dimensions est une variable composée de nombreuses variables statistiques à une dimension, dont l'une est expliquée et les autres sont explicatives et qui se présente comme suite.

U	V	W	X	M
u_1	v_1	w_1	x_1			m_1
u_2	v_2	w_2				
.						.
.
.						.
u_N	v_N	w_N	x_N			m_N

On suppose que la fonction $f(U, V, W, \dots)$ soit la meilleure courbe qui se rapproche le maximum du nuage de point. Cette tendance est formée par la variable M qui est expliquée par l'ensemble des autres variables. Avec:

$$f(U, V, W, \dots) = m_i = a \cdot u_i + b \cdot v_i + c \cdot w_i + \dots \quad i \in [1, N] \quad (1-39)$$

On suppose le modèle suivant:

$$\begin{aligned}
 m_1 &= a \cdot u_1 + b \cdot v_1 + c \cdot w_1 + \dots + cst \\
 m_2 &= a \cdot u_2 + b \cdot v_2 + c \cdot w_2 + \dots + cst \\
 m_3 &= a \cdot u_3 + b \cdot v_3 + c \cdot w_3 + \dots + cst \\
 m_N &= a \cdot u_N + b \cdot v_N + c \cdot w_N + \dots + cst
 \end{aligned}
 \tag{1-40}$$

De la on tire la relation suivante:

$$\begin{pmatrix} m_1 \\ \vdots \\ m_N \end{pmatrix} = \begin{pmatrix} u_1 & v_1 & w_1 & x_1 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_N & v_N & w_N & x_N & \dots & 1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

(1-41)

Soit:

$$Y = \begin{pmatrix} m_1 \\ \vdots \\ m_N \end{pmatrix}, \quad X = \begin{pmatrix} u_1 & v_1 & w_1 & x_1 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_N & v_N & w_N & x_N & \dots & 1 \end{pmatrix} \text{ et } \theta = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

(1-42)

Donc:

$$Y = X \cdot \theta \tag{1-43}$$

On cherche le vecteur $\vec{\theta}$ Pour pouvoir avoir la fonction f(x):

$$[X]^T \cdot [Y] = [X]^T [X] [\theta] \tag{1-44}$$

$$([X]^T \cdot [X])^{-1} \cdot [X]^T [Y] = ([X]^T \cdot [X])^{-1} \cdot ([X]^T \cdot [X]) \cdot [\theta]$$

(1-45)

Donc:

$$([X]^T \cdot [X])^{-1} \cdot [X]^T [Y] = [\theta] \tag{1-46}$$

1.2.1.4. Analyse factorielle des données :

L'analyse factorielle est un procédé de la famille de la statistique multi variée, utilisée pour décrire la variabilité entre des variables observées, au moyen de variables latentes. Cette méthode est utilisée dans toute discipline faisant face à de grandes quantités de données entre autre la mécanique des sols.

L'analyse factorielle donne des résultats souvent proches de l'Analyse en composantes principales (ACP). Cependant, les deux méthodes sont formellement différentes (les matrices de variance-covariance analysées sont différentes).

Analyse factorielle désigne souvent de façon générale les méthodes d'analyse des données, dont fait partie l'ACP (analyse en composante principale) et l'AFC (l'analyse factorielle des correspondances).

1.2.1.5. L'analyse factorielle des correspondances (AFC):

a) Généralité :

L'AFC est une analyse en composante principale (ACP). Elle a pour but de lire l'information contenue dans un repère multidimensionnel en diminuant la dimension de l'espace tout en conservant un maximum d'information englobé dans l'espace de départ.

L'analyse factorielle est une technique utilisée pour exploiter des observations à partir d'un tableau de contingence: elle permet de donner une représentation graphique qui permet de voir les rapprochements et les oppositions entre les caractéristiques des individus (Cibois, 2006).

Soient deux variables X et Y, comportant respectivement a et b modalités. Les observations représentées dans un tableau (tableau 4) de contingence à a lignes et b colonnes donnent les effectifs conjoints c'est-à-dire les effectifs observés pour chaque combinaison d'une modalité i de X et d'une modalité j de Y (Laffly, 2009.).

	X1	X2	X3	...	Xa	mains
Y1	n_{11}	n_{12}	n_{13}	...	n_{1b}	$=\sum_{j=1}^b n_{1j}$
Y2	n_{21}	n_{22}	n_{23}	...	n_{2b}	$=\sum_{j=1}^b n_{2j}$
Y3	n_{31}	n_{32}	n_{33}	...	n_{3b}	\vdots
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
Yb	n_{a1}	n_{a2}	n_{a3}	...	n_{ab}	\vdots
pieds	$=\sum_{i=1}^a n_{i1}$	$=\sum_{i=1}^a n_{i2}$	N

Tableau 1 3: Tableau de contingence.

Ce tableau devient une matrice de la forme:

	X1	X2	X3	...	Xa	Somme ligne
Y1	n_{11}	n_{12}	n_{13}	...	n_{1b}	$=\sum_{j=1}^b n_{1j}$
Y2	n_{21}	n_{22}	n_{23}	...	n_{2b}	$=\sum_{j=1}^b n_{2j}$
Y3	n_{31}	n_{32}	n_{33}	...	n_{3b}	\vdots
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
Yb	n_{a1}	n_{a2}	n_{a3}	...	n_{ab}	\vdots
Somme colonne	$=\sum_{i=1}^a n_{i1}$	$=\sum_{i=1}^a n_{i2}$	N

n_{ij} : Les valeurs du tableau correspondant au effectifs conjoints entre X et Y, avec :

$\{i, j \in \mathbb{N} \text{ et } i \in [1, a] \mid j \in [1, b]\}$

n_i : Effectif total de la ligne i ,

n_j : effectif total de la colonne j

$N_{..}$ somme des effectifs de la matrice.

Avec:

$$N = \sum_{i=1}^a \sum_{j=1}^b n_{ij} \quad (1-47)$$

f_{ij} : fréquence de la case $i^{\text{ème}}$ colonne et $j^{\text{ème}}$ ligne,

$$\text{Avec} \quad f_{ij} = \frac{n_{ij}}{N_{..}} \quad (1-48)$$

Les fréquences lignes (ou coordonnées des profils lignes) sont calculées par :

$$f_{i.} = n_{ij} / n_i = f_{ij} / f_i = \text{Effectif de la cellule } (i, j) / \text{Effectif de la ligne } i$$

Les coordonnées du profil ligne moyen sont calculées par :

$$f_{.j} = n_{.j} / N_{..} = \text{Effectif de la colonne } j / \text{Effectif total}$$

Les fréquences colonnes (ou coordonnées des profils colonnes) sont calculées par :

$$f_{c_{ij}} = n_{ij} / n_{.j} = f_{ij} / f_{.j} = \text{Effectif de la cellule } (i, j) / \text{Effectif de la colonne } j$$

Les coordonnées du profil colonne moyen sont calculées par :

$$f_{i.} = n_i / N_{..} = \text{Effectif de la ligne } i / \text{Effectif total}$$

Si on veut observer la valeur relative des réponses par rapport aux colonnes, c'est-à-dire comment réagit les modalités X1, X2, X3, X4 selon les modalités Y1, Y2, Y3, par exemple.

Pour ce on transforme le tableau réduit en divisant la valeur de la fréquence lignes, et faire une représentation graphique dans un espace à trois dimensions. Il est même possible d'associer à tous les points un cercle dont la surface est proportionnelle à la valeur des sommes des lignes (figure 1.5) (Laffly, 2009.).

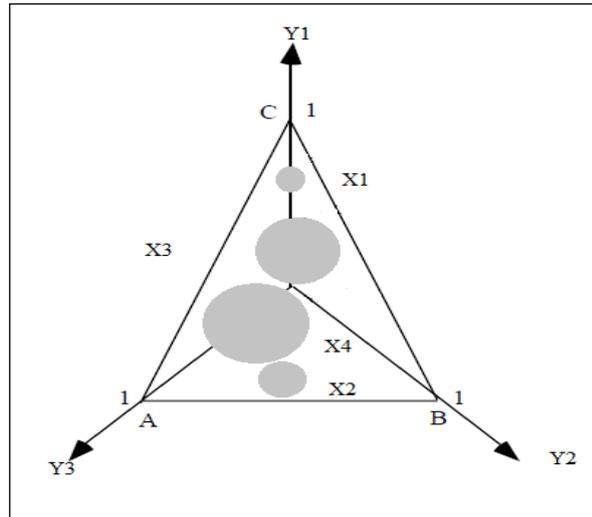


Figure 1 5: Représentation graphique des relations entre X et Y (Laffly, 2009).

b) Distances entre profils. Métrique du Φ^2

Chaque ligne du tableau contenant les fréquences lignes est une liste des coordonnées d'un point dans un espace à (a) dimensions. Les points trouvés forment un nuage des individus-lignes.

On définit de même le nuage des individus-colonnes à partir du tableau des fréquences colonnes.

On s'intéresse aux directions de "plus grande dispersion" de chacun des nuages de points. Cela revient à mesurer la "distance du Φ^2 " entre deux points présentés par la ligne i et la ligne i' définie par :(Carpentier, 2005)

$$d_{\Phi^2}^2(L_i, L_{i'}) = \sum_j \frac{(fl_{ij} - fl_{i'j})^2}{f_j} \quad (1-49)$$

Avec:

$d_{\Phi^2}^2(L_i, L_{i'})$: Distance entre les points caractérisés par les lignes i et i',

fl_{ij} : fréquence ligne de la ligne i

$fl_{i'j}$: fréquence ligne de la ligne i'

f_j : Les coordonnées du profil ligne moyen

c) Taux de liaison et Phi-2

Les taux de liaison sont donnés par:

$$t_{ij} = \frac{f_{ij} - f_i \cdot f_j}{f_i \cdot f_j} \quad (1-50)$$

t_{ij} : taux de liaison,

f_{ij} : fréquence de la cellule correspondant a la ligne i et à la colonne j ,

$f_{i.}$: Les coordonnées du profil colonne moyen,

La moyenne de la série des taux de liaison divisée par les coefficients $f_{i.} \cdot f_{.j}$ est nulle.

La variance de cette série (avec la même division) est le coefficient Φ^2 :

$$\Phi^2 = \sum_{i,j} f_{i.} \cdot f_{.j} \cdot t_{ij}^2 = \sum_{i,j} \frac{(f_{ij} - f_{i.} \cdot f_{.j})^2}{f_{i.} \cdot f_{.j}} \quad (1-51)$$

La méthode d'AFC peut être vue comme une décomposition adéquate du Φ^2 selon plusieurs axes factoriels.

d) L'analyse factorielle des correspondances:

L'application de la méthode donne comme résultats :

- ❖ Construit des images des nuages d'individus-lignes ou colonnes, de manière que les distances entre images soient des distances euclidiennes et non plus des distances calculées selon la métrique du Φ^2 ,
- ❖ Rechercher les directions de plus grande dispersion dans ces nuages de points images.

La matrice dont on recherche les valeurs et vecteurs propres est un concept mathématique difficile, qui ne possède pas de signification immédiatement intuitive. De ce fait, on calcule des produits scalaires entre lignes (ou entre colonnes) de cette matrice (Carpentier, 2005).

e) Valeurs propres

Le nombre de valeurs propres est le minimum du nombre de lignes ou de colonnes dans tableau de contingence. La première valeur propre est égale à 1, et n'est pas mentionnée dans les résultats. Les autres valeurs propres sont des nombres positifs inférieurs à 1 et leur somme est égale à Φ^2 .

En traçant une courbe représentant les valeurs propre en fonction du nombre de dimension, on peut choisir les axes factoriels à conserver. La figure 1.6 présente un exemple cité par (Carpentier, 2005).

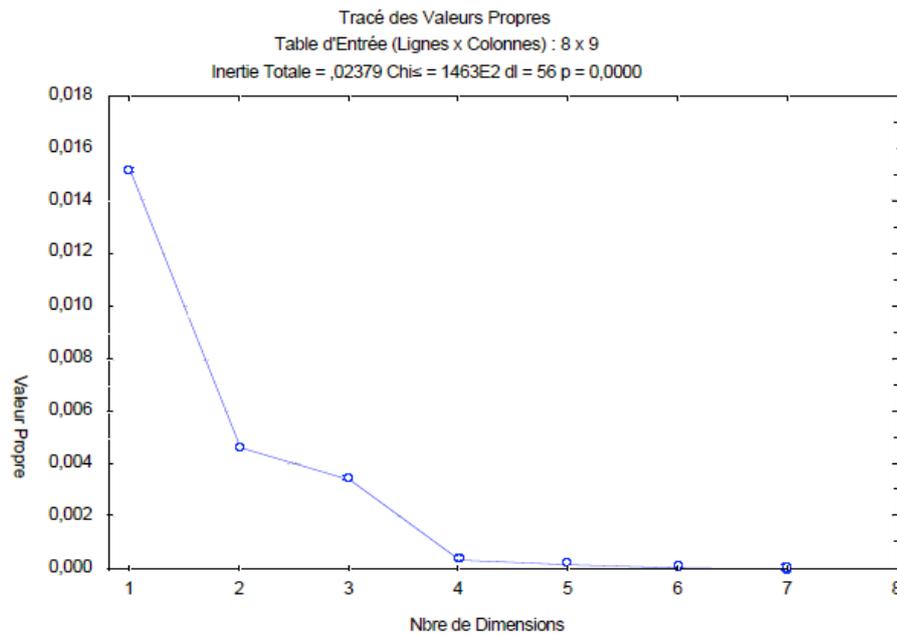


Figure 1 6: Courbe représentant les valeurs propre en fonction du nombre de dimension (Carpentier, 2005).

Dans cette exemple il-y-a une brusque décroissance des valeurs propres entre la 3^{ème} et la 4^{ème} valeur propre. On retient donc les 3 premiers axes factoriels.

f) Résultats relatifs aux individus-lignes

La colonne indique les qualités de représentation des individus ligne par les axes factoriels. Ces qualités sont calculées par la formule suivante (Carpentier, 2005)

$$Qlt(L_i, F_K) = \frac{\sum_{K=1}^n (\text{coordonnées de la ligne } i \text{ selon } F_K)^2}{\sum_j (\text{coordonnées de la ligne } j \text{ selon } F_j)^2}$$

(1-52)

L_i : La ligne i ,

F_j : le facteur principal j ,

K : le numéro de l'axe factoriel,

n : nombre des axes factoriels,

La colonne ou Inertie relative est calculée de la manière suivante :

L'inertie relative d'un individu ligne est obtenue en divisant l'inertie absolue (la somme des inerties des cellules de la ligne) de l'individu par la somme de toutes les inerties, c'est-à-dire par Φ^2 (Carpentier, 2005).

Pour chacun des axes factoriels, le tableau donne les coordonnées factorielles de l'individu-ligne selon cet axe.

Ces coordonnées ont les propriétés suivantes :

- ❖ Selon chaque axe, la moyenne des coordonnées des individus-lignes divisée par les masses, est nulle.
- ❖ Selon chaque axe, la moyenne des carrés des coordonnées des individus-lignes pondérées par les masses, est la valeur propre correspondante.
- ❖ Les coordonnées selon deux axes différents, divisées par les masses, forment deux séries statistiques indépendantes (covariance nulle)

Les résultats relatifs aux individus-colonnes s'expliquent de la même manière que les résultats relatifs aux individus-lignes (Carpentier, 2005).

g) Résultats graphiques

Les transformations rendent les valeurs obtenues (pour les individus ligne ou colonne) comparables. Les graphiques factoriels pourront être représentés en faisant paraître sur un même graphique les individus lignes et les individus colonnes.

On peut réaliser et essayer d'interpréter des graphiques :

- en dimension 1 : on place les individus le long d'un axe factoriel,
- en dimension 2 : on place les individus dans un plan défini à partir de deux axes factoriels,
- en dimension 3 : on place les individus dans une représentation en perspective d'un espace à 3 dimensions (Carpentier, 2005).

h) Interprétation géométrique

Les distances entre deux individus-lignes, ou entre un individu-ligne et l'origine des axes, peuvent être facilement interprétées.

En effet: la distance euclidienne entre deux points-lignes, représentés par leurs coordonnées factorielles est égale à la distance du Φ^2 entre les profils-lignes initiaux. La proximité entre un point-ligne L et un point-colonne C ne possède pas d'interprétation géométrique immédiate. En revanche, l'angle de sommet O dont les côtés passent par L et C a la propriété suivante :

- si l'angle (OL, OC) est aigu, la modalité-ligne L et la modalité colonne C s'attirent
- si l'angle (OL, OC) est obtus, la modalité-ligne L et la modalité colonne C se repoussent
- si l'angle (OL, OC) est droit, la modalité-ligne L et la modalité colonne C n'interagissent pas (taux de liaison voisin de 0) (Carpentier, 2005).

1.2.1.6. VARIATIONS SPATIALES

Lorsqu'on mesure une caractéristique en un point, on peut considérer la valeur obtenue comme la réalisation d'une variable aléatoire en ce point.

Il en est de même pour tous les points d'un échantillon donné. On a une infinité de variables aléatoire représentant conjointement l'échantillon.

Il y a deux étapes principales dans une étude lors de variation spatiale:

- a) Le variogramme a pour concept que deux observations situées l'une près de l'autre devraient, en moyenne, se ressembler plus que deux observations éloignées. Il est basé sur une fonction mathématique, utilisée pour l'identification des caractéristiques des variables aléatoires.
- b) Le krigeage est une méthode qui utilise les valeurs connues des caractéristiques des variables aléatoire pour une estimation optimale aux points non mesurés (Chauvet, 1999).

CONCLUSION :

Ces méthodes d'analyse statistique appliquée à la mécanique des sols apparaissent désormais couramment dans les travaux de l'ingénieur en génie civil. Nous pourrions dire qu'elles ont fait leurs preuves et sont donc par conséquent devenues des méthodes classiques.

Toutefois, les méthodes d'analyse de donnée et plus particulièrement d'analyse factorielle n'ont pas été autant développées en mécanique des sols que les méthodes déjà plus classiques d'analyse de régression multiple ou de variations spatiales. La deuxième partie de ce mémoire de fin d'études est consacrée à une étude de l'analyse factorielle appliquée à une banque de données pénétrométriques. C'est pourquoi nous allons nous attacher dans le chapitre suivant à présenter ces méthodes plus en détail.

Chapitre 2: ANALYSE DE DONNEES : ANALYSE FACTORIELLE

2.1. PRESENTATION GENERALE :

En engineering, notamment en mécanique des sols, la compréhension d'un phénomène nécessite de suivre l'évolution de plusieurs variables, en nombres souvent élevé et voir les interactions et les répulsions entre ces variables. En application, cela se traduit par une étude d'un tableau $X(n ; p)$, de p variables et n observations qui représentent l'évolution de chaque variable. Ainsi, les informations sont regroupées dans une matrice $X(n ; p)$.

Pour les descriptions multidimensionnelles, des méthodes factorielles sont apparues au début des années 30. Vers les années 60, Benzecri J.P. a développé les méthodes descriptives appelée ACP (Analyse en composantes principales). Cette méthode est basée sur les aspects géométriques et représentations graphiques. (Duby C. et al., 2006).

A partir d'un tableau rectangulaire de p variables quantitative et n individus, l'ACP est un outil particulièrement puissant pour explorer la structure de telles données. En fait l'ACP consiste à analyser des données sur des espaces de dimensions supérieures à 2. Cependant, au-delà de la dimension 3, il est difficile, voir impossible de visualiser la forme du nuage de point et d'en déduire les relations existantes entre ces individus et variables.

La méthode ACP est un outil permet de déterminer des sous-espaces de dimension $q < p$ de tel sorte à comprimer les informations retirées du tableau et voir les relations entre variables et ressemblances entre les individus. L'idée consiste à trouver un espace de dimension réduite tel que la projection du nuage des points sur ce nouveau espace déforme le moins possible l'image des points et conserve le maximum d'informations sur les proximités entre les points. Il est à rappeler qu'une projection des points sur un espace quelconque déforme la géométrie des points et fait perdre l'information. Duby. et al., 2006.

La représentation des individus sur des espaces réduits, en particulier sur le plan (dimension 2) permet de voir s'il existe une structure ou organisation non connues au par avant. Ainsi, cherche à distinguer des groupes d'individus qui se ressemblent ou qui se différencient des autres. Quant à la représentation des variables, elle permet d'étudier les structures des liaisons linéaires, celles qui sont corrélées entre elles et celles qui ne le sont pas. (Duby. et al., 2006).

2.2. CARACTERISTIQUES DES DONNEES :

2.2.1. Tableau de données :

Un tableau $X(n, p)$ est une matrice de n lignes représentant les n individus et p colonnes représentant les p variables.

X=

	V ₁	V ₂	V ₃	V _p	
	X ₁₁	X ₁₂	X ₁₃	X _{1n}	u ₁
	X ₂₁	X ₂₂	X ₂₃	X _{2n}	u ₂
	:						:
	:						:
	X _{n1}	X _{np}	u _n

Chaque individu (appelé aussi unité) est représenté par un vecteur qui donne les valeurs des P variables,

$$U_i^T = (x_{i1} \ x_{i2} \ \dots \ x_{ip}) \quad \text{Donc,} \quad U_i = \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{ip} \end{pmatrix}$$

Les individus forment des nuages de 'n' points dans un espace vectoriel dans \mathbb{R}^p dit espace des variables. Afin de visualiser l'ensemble des individus, on place tout les points qui représente les unités ou individus dans un espace affine dont l'origine est un vecteur de \mathbb{R}^p .

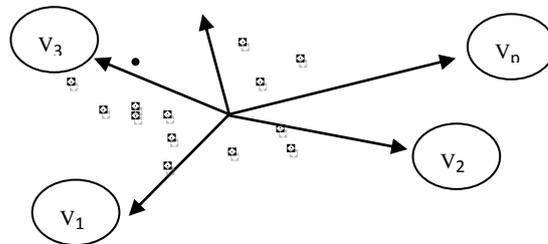


Figure 2 1: Nuage des individus dans l'espace de dimension p défini par les variables

De même, chaque variable $V_i = (x_{i1} \ x_{i2} \ \dots \ x_{in})$ est représentée par un vecteur qui donne les valeurs des n individus. Ainsi, les variables forment un nuage de 'p' points dans l'espace vectoriel \mathbb{R}^n dit espace des individus. Pour visualiser les variables, on place tout les points qui représentent les variables dans un espace affine correspondant de \mathbb{R}^n .

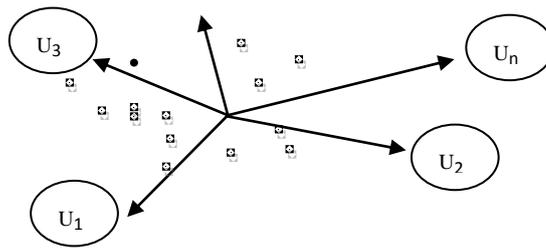


Figure 2 2: Nuage des variables dans l'espace de dimensions n défini par les individus.

2.2.2. Données actives et supplémentaires :

Les variables qualitatives sont exclues de cette analyse. Les variables quantitatives disponibles pour l'analyse ACP se divisent en deux ensembles.

- Les variables actives qui servent pour le calcul des axes principaux,
- Les variables supplémentaires qui sont des variables illustratives.

Dans le tableau de contingence on sépare les individus actifs et supplémentaires comme indiqué si dessous. (Saporta, 2006).

Tableau 2 3: Tableau des données indiquant les valeurs actives et les valeurs supplémentaire

	Variables actives	Variables sup
Individus actifs	X	S
Individus sup		0 ⋮ 0

2.2.3. Caractéristiques des points :

2.2.3.1. Poids :

Les poids P_i des individus désignent l'importance des données dans le tableau, ces poids sont comparables a des fréquences, se sont des nombres positifs dont la somme est égale à 1.

Les poids des individus permettent de voir si les échantillons sont regroupés ou rectifiés. On peut les rassembler dans une matrice 'D' diagonale de taille 'n' qui est définie comme suit:

$$D = \begin{bmatrix} p_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & p_n \end{bmatrix}$$

Dans le cas usuel les individus ont la même importance égale à 1/n, c'est-à-dire les données sont recueillies de façon aléatoire à probabilités égales (SAPORTA, 2006). Alors

$$D = \frac{1}{n} \times I \tag{2-1}$$

2.2.3.2. Centre de gravité (point d'origine) :

L'origine d'un espace vectoriel, c'est-à-dire le vecteur nul V(0, 0,...,0) n'est certainement pas une origine acceptable. Il est plus judicieux de choisir le centre de gravité du nuage comme origine, qu'on note par G. Il est donné par l'équation:

$$\sum_{i=1}^n p_i \overrightarrow{GU_i} = 0 \tag{2-2}$$

p_i : est le poids de l'individu $u_i \quad \forall i = 1, \dots, n$

Le centre de gravité G du nuage des individus est alors le point dont les coordonnées sont les valeurs moyennes des variables donné comme suit:

$$G = \begin{pmatrix} \frac{1}{n} \sum_{i=1}^n x_{i1} \\ \vdots \\ \frac{1}{n} \sum_{i=1}^n x_{ij} \\ \vdots \\ \frac{1}{n} \sum_{i=1}^n x_{ip} \end{pmatrix} = \begin{pmatrix} x_{.1} \\ \vdots \\ x_{.j} \\ \vdots \\ x_{.p} \end{pmatrix} \tag{2-3}$$

Prendre G pour origine, revient à travailler sur le tableau des données centrées suivant:

$$X_{ct} = \begin{pmatrix} x_{11} - x_{.1} & \cdots & x_{1j} - x_{.j} & \cdots & x_{1p} - x_{.p} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{j1} - x_{.1} & & x_{jj} - x_{.j} & & x_{jp} - x_{.p} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{p1} - x_{.p} & & x_{jp} - x_{.j} & & x_{pp} - x_{.p} \end{pmatrix}$$

Le vecteur des coordonnées centrées de l'individu U_{ci} :

$$U_{ci} = \begin{pmatrix} x_{i1} - x_{.1} \\ \vdots \\ x_{ij} - x_{.j} \\ \vdots \\ x_{ip} - x_{.p} \end{pmatrix} \quad (2-4)$$

Le vecteur des coordonnées centrées de la variable V_{cj} :

$$V_{cj} = (x_{j1} - x_{.j} \quad \dots \quad x_{ji} - x_{.j} \quad \dots \quad x_{jp} - x_{.j}) \quad (2-5)$$

2.3. ESPACE DES INDIVIDUS:

2.3.1. Rôle de la métrique :

L'ACP, basée sur des aspects géométriques, s'intéresse en particulier aux distances entre les individus afin d'interpréter les condensations et dispersions contenus dans un nuage de points. Cela revient à mesurer la distance entre chaque couple d'individus (Saporta, 2006).

Pour faire une représentation géométrique, il faut choisir une distance entre deux points de l'espace. Cependant, en physique on peut facilement tirer la distance entre deux points dans l'espace dont les axes sont perpendiculaires par la formule de Pythagore. Les dimensions sont de même nature, c'est des longueurs qu'on mesure avec la même unité. Par contre, en statistique chaque dimension correspond à un individu dont les composantes ne portent pas des unités homogènes. Dans ce cas la formule de Pythagore n'est pas recommandée. Afin de remédier à cela on donne la distance par la formule suivante (Buby. et al. , 2006) et (Saporta, 2006).

$$d^2(U_i, U_i') = \sum_{j=1}^p (x_{ij} - x_{i'j})^2. \quad (2-6)$$

2.4. L'ESPACE DES VARIABLES :

Chaque variables x_j est une liste de n valeurs, donc on la considère un vecteur x_j d'un espace E à n dimensions appelé l'espace des variables.

2.4.1. La métrique des poids :

Dans l'espace des variables on s'intéresse aux angles entre les vecteurs en raison de la propriété suivante le coefficient de corrélation linéaire est tout simplement le cosinus de l'angle entre deux variables centrées.

Ainsi, entre deux vecteurs $\overrightarrow{oU_1}$ et $\overrightarrow{oU_{i'}}$, on associe le produit scalaire:

$$\langle \overrightarrow{oU_1}; \overrightarrow{oU_{i'}} \rangle = \sum_{j=1}^p x_{1j} x_{i'j} = U_1^T U_{i'} \quad (2-7)$$

La norme du vecteur $\overrightarrow{oU_1}$ est :

$$\|\overrightarrow{oU_1}\|^2 = \sum_{j=1}^p x_{1j}^2 = U_1^T U_1 \quad (2-8)$$

L'angle entre deux vecteurs est donné par son cosinus, par l'équation:

$$\cos(\alpha) = \frac{\langle \overline{ou_1}; \overline{ou'_1} \rangle}{\|\overline{ou_1}\| \|\overline{ou'_1}\|} = \frac{\sum_{j=1}^p x_{ij} x'_{1j}}{\sqrt{\sum_{j=1}^p x_{ij}^2} \sqrt{\sum_{j=1}^p x'_{1j}{}^2}} = \frac{u_1^T u'_1}{\sqrt{u_1^T u_1} \sqrt{u'^T_1 u'_1}} \quad (2-9)$$

2.5. MOMENT D'INERTIE :

En physique, le moment d'inertie définit la géométrie des masses, donc, la répartition de la matière. L'inertie totale du nuage de points, est la moyenne équilibrée des distances au carrée des points au centre de gravité. Elle est donnée par la formule suivante:

$$I_g = \frac{1}{n} \sum_{i=1}^n d^2(G, u_i) = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^p (x_{ij} - x_{.j})^2 = \frac{1}{n} \sum_{i=1}^n u_{ci}^T u_{ci} \quad (2-10)$$

Le moment d'inertie sert à mesurer la dispersion du nuage des individus par rapport au centre de gravité.

Plus le moment d'inertie est grand et plus le nuage des points est dispersé. Contrairement à cela, un nuage de point est concentré autour du centre de gravité se traduit par un faible moment d'inertie.

En un point b, l'inertie peut être calculée par la relation donnée par le théorème de Huygens:

$$I_b = I_g + \|g - b\|^2 \quad (2-11)$$

On remarque que I_G peut s'écrire sous la forme suivante :

$$I_G = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^p (x_{ij} - x_{.j})^2 = \sum_{j=1}^p \left[\frac{1}{n} \sum_{i=1}^n (x_{ij} - x_{.j})^2 \right] = \sum_{j=1}^p \text{Var}(v_j) \quad (2-12)$$

$\text{Var}(v_j)$ est la variance expérimentale de la variable v_j .

Sous cette forme, on constate que l'inertie totale est égale à la trace de la matrice de covariance Σ des p variables v_j (Duby. et al., 2006).

$$I_G = \text{trace}(\Sigma). \quad (2-13)$$

2.5.1. Inertie du nuage des individus par rapport à un axe Δ passant par G :

Soit $h_{\Delta i}$ la projection orthogonale des individus u_i sur l'axe Δ . L'inertie du nuage des individus par rapport à un axe Δ mesure la distance entre Δ et les points du nuage des individus.

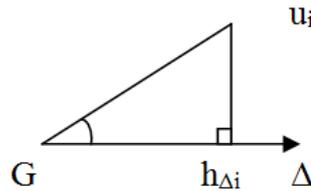


Figure 2.3 : Projection orthogonale des individus sur l'axe Δ

Cette inertie est égale, par définition à (Duby. et al. , 2006):

$$I_{\Delta} = \frac{1}{n} \sum_{i=1}^n d^2(h_{\Delta i}, u_i) \quad (2-14)$$

2.5.2. Inertie du nuage des individus par rapport à un sous-espace vectoriel V passant par G :

Cette inertie est, par définition, égale à :

$$I_V = \frac{1}{n} \sum_{i=1}^n d^2(h_{V i}, u_i) \quad (2-15)$$

Où $h_{V i}$ est la projection orthogonale de u_i sur le sous-espace V (Duby C ; et al., 2006).

2.5.3. Décomposition de l'inertie totale :

Si on note V^* le complémentaire orthogonal de V dans \mathbb{R}^p et $h_{V^* i}$ la projection orthogonale de u_i sur V^* , en appliquant le théorème de Pythagore, on peut écrire :

$$d^2(h_{V i}, u_i) + d^2(h_{V^* i}, u_i) = d^2(G, u_i) = d^2(G, h_{V i}) + d^2(G, h_{V^* i}) \quad (2-16)$$

On en déduit, que c'est le théorème de Huygens :

$$I_V + I_{V^*} = I_G \quad (2-17)$$

Dans le cas où le sous-espace est de dimension 1 (un axe), I_{V^*} est une mesure de l'allongement du nuage selon cet axe.

On emploie pour I_{V^*} les expressions « inertie portée par l'axe » ou « l'inertie expliquée par l'axe »

En projetant le nuage des individus sur un sous-espace V , on perd l'inertie mesurée par I_V (En effet, en projection les distances ne peuvent que diminuer), on ne garde que celle mesurée par I_{V^*} .

De plus, si on décompose l'espace \mathbb{R}^p en plusieurs axes orthogonaux entre eux: (Duby. et al., 2006) et (Saporta, 2006)

$$\Delta_1 \oplus \Delta_2 \oplus \dots \oplus \Delta_p$$

On peut écrire:

$$I_G = I_{\Delta_1^*} + I_{\Delta_2^*} + \dots + I_{\Delta_p^*} \quad (2-18)$$

2.6. RECHERCHE DE 1^{ER} AXE Δ_1 PASSANT PAR G D'INERTIE MINIMUM :

Sa consiste à chercher l'axe le plus Proche de l'ensemble des points, autrement dit si l'on doit projeter ce nuage sur cet axe, c'est lui qui donnera l'image la moins déformée du nuage, donc c'est l'axe Δ_1 passant par G d'inertie I_{Δ_1} minimum.

En appliquant la relation entre les inerties données au paragraphe précédent, rechercher Δ_1 tel que: I_{Δ_1} est minimum, est équivalent à chercher Δ_1 tel que $I_{\Delta_1^*}$ est maximum.

I_{Δ_1} est minimum $\leftrightarrow I_{\Delta_1^*}$ est maximum

On définit l'axe Δ_1 par son vecteur directeur unitaire $\overrightarrow{G\alpha_1}$.

Il faut donc trouver $\overrightarrow{G\alpha_1}$ tel que $I_{\Delta_1^*}$ est maximum sous la contrainte que $\|\overrightarrow{G\alpha_1}\|^2 = 1$.

2.6.1. Expressions algébriques de $I_{\Delta_1^*}$ et de $\|\overrightarrow{G\alpha_1}\|^2$:

$$d^2(G, h_{\Delta_1^*}) = \langle Gu_i, G\alpha_1 \rangle^2 = \alpha_1^T U_{cti} U_{cti}^T \alpha_1 \quad (2-19)$$

En utilisant la symétrie du produit scalaire. On en déduit:

$$I_{\Delta_1^*} = \frac{1}{n} \sum_{i=1}^n \alpha_1^T U_{cti} U_{cti}^T \alpha_1 = \alpha_1^T \left[\frac{1}{n} \sum_{i=1}^n U_{ci} U_{ci}^T \right] \alpha_1 \quad (2-20)$$

Entre crochets on reconnaît la matrice de covariance empirique Σ des p variables.

$$I_{\Delta_1^*} = \alpha_1^T \Sigma \alpha_1 \quad (2-21)$$

$$\|\overrightarrow{G\alpha_1}\|^2 = \alpha_1^T \alpha_1 \quad (2-22)$$

2.6.2. Recherche du maximum:

Il faut trouver α_1 tel que $\alpha_1^T \Sigma \alpha_1$ soit maximum avec la contrainte $\alpha_1^T \alpha_1 = 1$. On cherche un optimum d'une fonction de plusieurs variables liées par une contrainte (les inconnues sont les composantes de α_1).

On utilise dans ce cas, la méthode des multiplicateurs de Lagrange. Dans le cas de la recherche de α_1 , il faut calculer les dérivées partielles de :

$$g(\alpha_1) = g(\alpha_{11}, \alpha_{12}, \dots, \alpha_{1p}) = \alpha_1^T \Sigma \alpha_1 - \lambda_1 (\alpha_1^T \alpha_1 - 1). \quad (2-23)$$

En utilisant la dérivée matricielle, on obtient :

$$\frac{\partial(g(\alpha_1))}{\partial \alpha_1} = 2\Sigma \alpha_1 - 2\lambda_1 \alpha_1 = 0 \quad (2-24)$$

Le système à résoudre est :

$$\begin{cases} \Sigma \alpha_1 - \lambda_1 \alpha_1 = 0 \dots (1) \\ \alpha_1^T \alpha_1 - 1 = 0 \dots (2) \end{cases} \quad (2-25)$$

De l'équation matricielle (1) de ce système on déduit que α_1 est vecteur propre de la matrice Σ associé à la valeur propre λ_1 .

En multipliant à gauche par α_1^T les deux membres de l'équation (1) on obtient :

$$\alpha_1^T \Sigma \alpha_1 - \lambda_1 \alpha_1^T \alpha_1 = 0 \quad (2-26)$$

Et en utilisant l'équation (2) on trouve que :

$$\alpha_1^T \Sigma \alpha_1 = \lambda_1 \quad (2-27)$$

$\alpha_1^T \Sigma \alpha_1$ Est égale à l'inertie $I_{\Delta_1^*}$ qui doit être maximum. Cela signifie que la valeur propre λ_1 est la plus grande valeur propre de la matrice de covariance Σ et que cette valeur propre est égale à l'inertie portée par l'axe Δ_1 .

L'axe Δ_1 pour lequel le nuage des individus à l'inertie minimum a comme vecteur directeur unitaire le premier vecteur propre associé à la plus grande valeur propre de la matrice de covariance Σ .

2.7. RECHERCHE DES AXES SUIVANTS:

2.8.

Cela consiste à chercher un deuxième axe Δ_2 orthogonal au premier et d'inertie minimum. Comme précédemment on cherche Δ_2 passant par G, qui est définie par son vecteur directeur unitaire α_2 . L'inertie du nuage des individus par rapport à son complémentaire orthogonal est égale à :

$$I_{\Delta_2^*} = \alpha_2^T \Sigma \alpha_2 \quad (2-28)$$

Elle doit être maximum avec les deux contraintes suivantes :

$$\alpha_2^T \alpha_2 = 1 \quad \text{Et} \quad \alpha_2^T \alpha_1 = 0. \quad (2-29)$$

La relation $\alpha_2^T \alpha_1 = 0$ exprime que le deuxième axe doit être orthogonal au premier (le produit scalaire des deux vecteurs directeurs est nul). En appliquant la méthode des multiplicateurs de Lagrange, on trouve que α_2 est le vecteur propre de Σ correspondant à la deuxième plus grande valeur propre. On peut montrer que le plan défini par les axes Δ_1 et Δ_2 est le sous-espace de dimension 2 qui porte l'inertie maximum.

De la même manière précédente on peut trouver d'autres axes. On appelle les nouveaux axes, « axes principaux » (Duby C. et al., 2006).

2.8. PARTICIPATIONS DES AXES A L'INERTIE TOTALE :

En utilisant le théorème de Huygens, on peut décomposer l'inertie totale du nuage des individus comme suit:

$$I_G = I_{\Delta_1^*} + I_{\Delta_2^*} + \dots + I_{\Delta_p^*} = \lambda_1 + \lambda_2 + \dots + \lambda_p \quad (2-30)$$

La contribution absolue de l'axe Δ_k à l'inertie totale du nuage des individus est égale à :

$$ca\left(\frac{\Delta_k}{I_G}\right) = \lambda_k \quad (2-31)$$

λ_k est la valeur propre qui lui est associée.
Sa contribution relative est égale à :

$$cr\left(\frac{\Delta_k}{I_G}\right) = \frac{\lambda_k}{\lambda_1 + \lambda_2 + \dots + \lambda_p} \quad (2-32)$$

Ces définitions peut être étalées à tous les sous-espaces produits par les nouveaux axes. Ainsi, le pourcentage d'inertie expliqué par le plan engendré par les deux premiers axes Δ_1 et Δ_2 est égal à :

$$cr\left(\frac{\Delta_1 \oplus \Delta_2}{I_G}\right) = \frac{\lambda_1 + \lambda_2}{\lambda_1 + \lambda_2 + \dots + \lambda_p} \quad (2-33)$$

Les indicateurs qui sont des pourcentages d'inertie, définissent la variabilité du nuage des individus expliquée par ces sous-espaces.

Si les dernières valeurs propres ont des valeurs faibles, on pourra négliger la variabilité qu'expliquent les axes correspondants.

On se contente souvent de faire des représentations du nuage des individus dans un sous-espace engendré par les deux premiers axes (Duby. et al., 2006).

2.9. REPRESENTATION DES INDIVIDUS DANS LES NOUVEAUX AXES :

La représentation des individus dans les nouveaux plans, admet de calculer les coordonnées des individus dans les nouveaux axes. Pour obtenir y_{ik} , coordonné de l'unité u_i sur l'axe Δ_k , on projette orthogonalement le vecteur $\overrightarrow{Gu_i}$ sur cet axe et on obtient :

$$y_{ik} = \langle \overrightarrow{Gu_i}, \overrightarrow{\alpha_k} \rangle = \alpha_k^T U_{ci} \quad (2-34)$$

$$Y_i = A^T U_{ci} \quad (2-35)$$

Y_i : Le vecteur des coordonnées de l'unité u_i

A : La matrice du changement de base (A matrice des vecteurs propres orthogonaux et de norme 1 est une matrice orthogonale, son inverse est égale à sa transposée) (Duby C. et al., 2006).

La figure suivante illustre une représentation des variables sur axes principaux tiré d'un exemple par le logiciel XLstat.

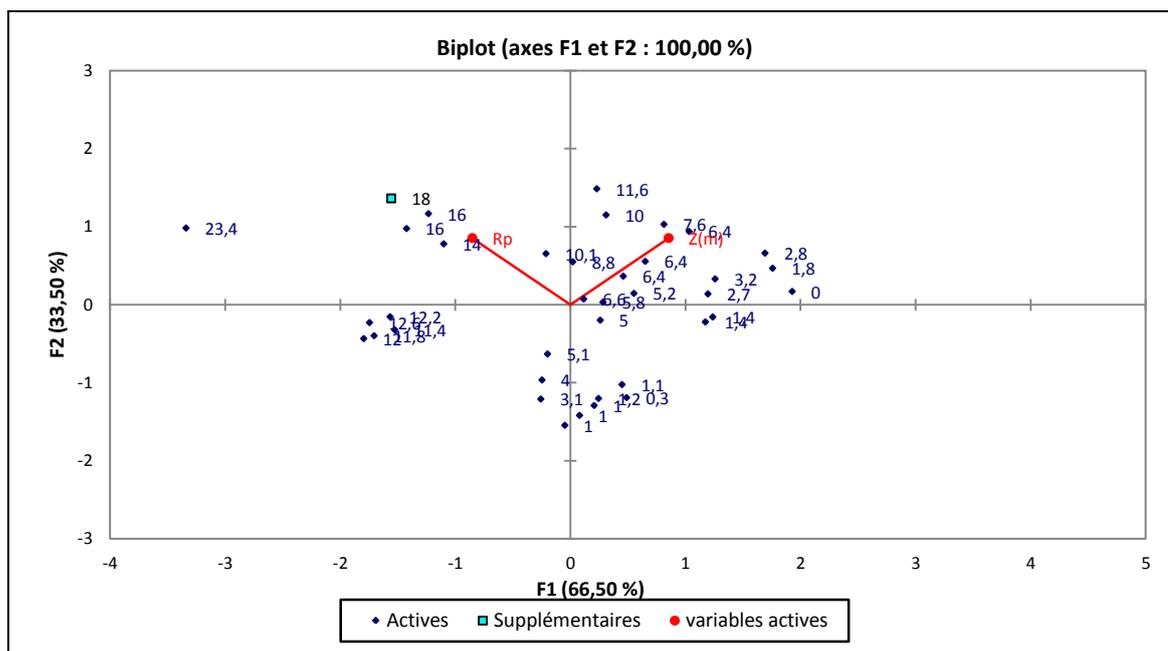


Figure 2 4: Représentations du nuage de point sur le plan formé par les deux premiers axes principaux

Remarque:

L'orientation des axes peut changer d'un logiciel à l'autre. Le signe des coordonnées des individus sur un axe n'a donc pas de signification. En revanche, la comparaison des signes peut s'interpréter. Si deux individus u_i et u_j ont sur un axe Δ , le premier une coordonnée positive et le second une coordonnée négative, cela signifie qu'ils s'opposent sur cet axe.

2.9.1. Qualité de la représentation des individus :

Les points représentant les individus sont projetés sur un axe (ou sur un plan), lorsque la distance qui les sépare est grande on peut garantir que les points représentants ces individus sont éloignés dans l'espace. Mais le contraire n'est pas toujours vrai (deux individus dont les projections sont proches sur un axe (ou sur un plan) peuvent ne pas être proches dans l'espace).

Pour interpréter correctement les proximités entre les points il faut s'assurer de la qualité de la représentation.

Pour assurer une bonne représentation d'un individu u_i sur un sous-espace de dimensions quelconque, il faut que l'angle entre le vecteur $\overrightarrow{Gu_i}$ et le sous-espace soit petit donc le cosinus de l'angle ou le carré du cosinus soit proche de 1.

En utilisant le théorème de Pythagore, on peut montrer que le carré du cosinus de l'angle d'un vecteur avec un plan engendré par deux vecteurs orthogonaux, est égal à

la somme des carrés des cosinus des angles du vecteur avec chacun des deux vecteurs qui génèrent le plan.

Si le carré du cosinus de l'angle est proche de 1, alors on conclut que l'individu u_i est bien représenté par sa projection sur le sous-espace.

Et si deux individus sont bien représentés en projection et ont des projections voisines, alors on pourra dire que ces deux individus sont proches dans l'espace.

Le carré du cosinus de l'angle θ_{ik} entre $\overrightarrow{Gu_i}$ et un axe Δ_k de vecteur directeur unitaire α_k est égal à:

$$\cos(\theta_{ik})^2 = \frac{(\overrightarrow{Gu_i}; \overrightarrow{G\alpha_k})^2}{\|\overrightarrow{Gu_i}\|^2} = \frac{\alpha_k^T U_{cti} U_{cti}^T \alpha_k}{U_{cti}^T U_{cti}} = \frac{[\sum_{j=1}^p (x_{ij} - \bar{x}_j) \alpha_{kj}]^2}{\sum_{j=1}^p (x_{ij} - \bar{x}_j)^2} \quad (2-37)$$

En utilisant le théorème de Pythagore on peut calculer le carré du cosinus de l'angle $\theta_{ikk'}$ entre $\overrightarrow{Gu_i}$ et le plan généré par deux axes $\Delta_k \oplus \Delta_{k'}$:

$$\cos^2(\theta_{ikk'}) = \cos^2(\theta_{ik}) + \cos^2(\theta_{ik'}) \quad (2-38)$$

Remarque :

Si un individu est très proche du centre de gravité dans l'espace, alors le point représentant cet individu sur un sous-espace sera bien représenté.

2.9.2. Interprétation des nouveaux axes en fonction des individus :

Quand on calcule l'inertie $I_{\Delta_k^*}$ expliquée par l'axe Δ_k , on peut voir quelle est la part de cette inertie due à un individu u_i particulier.

2.9.2.1. Contribution absolue d'un individu à un axe :

$I_{\Delta_k^*}$ étant égale à $\frac{1}{n} \sum_{i=1}^n d^2(h_{\Delta_{ki}}, G)$, la contribution absolue de u_i à cette inertie est par définition égale à:

$$ca\left(\frac{u_i}{\Delta_k}\right) = \frac{1}{n} d^2(h_{\Delta_{ki}}, G) \quad (2-39)$$

Puisque tous les individus ont la même importance. Un individu participera d'autant plus à la confection d'un axe, que sa projection sur cet axe sera éloignée du centre de gravité du nuage.

Inversement, un individu dont la projection sur un axe sera proche du centre de gravité participera faiblement à l'inertie portée par cet axe.

On utilise ces contributions pour interpréter les nouveaux axes de l'ACP en fonction des individus.

2.9.2.2. Contribution relative d'un individu à un axe :

Pour un individu particulier u_i , donner sa contribution relative à l'inertie portée par cet axe :

$$cr\left(\frac{u_i}{\Delta_k}\right) = \frac{\frac{1}{n} d^2(h_{\Delta_{ki}}, G)}{I_{\Delta_k^*}} = \frac{\frac{1}{n} (\overrightarrow{Gu_i}; \overrightarrow{G\alpha_k})^2}{\lambda_k} = \frac{\frac{1}{n} \alpha_k^T U_{ci} U_{ci}^T \alpha_k}{\lambda_k} \quad (2-40)$$

L'examen de ces contributions permet d'interpréter les axes principaux avec les individus.

On peut remarquer que (C. Duby et al., 2006):

$$\sum_{i=1}^n \text{cr} \left(\frac{u_i}{\lambda_k} \right) = 1. \quad (2-41)$$

2.10. REPRESENTATION DES VARIABLES :

La représentation des variables et des individus se fait d'une manière similaire il suffit de faire les raisonnements dans \mathbb{R}^n au lieu de \mathbb{R}^p .

Les variables n'ont pas le même sens que les individus ; après avoir représenté les individus dans l'espace des anciennes variables, on fait un changement de base dans cet espace. Les nouveaux axes sont des combinaisons linéaires des anciens axes et peuvent donc être considérés comme de nouvelles variables combinaisons linéaires des anciennes. Ces nouvelles variables notées $C_1, C_2, \dots, C_k, \dots, C_p$ sont les composantes principales.

Soit : C_k la nouvelle variable correspondant à l'axe Δ_k :

$$C_k = \sum_{j=1}^p \alpha_{kj} V_{ctj} = X_{ct} \alpha_k \quad (2-42)$$

De façon générale : $C = [C_1 \ C_2 \dots C_k \dots C_p] = X_{ct} A . \quad (2-43)$

Il est intéressant de voir comment les anciennes variables sont liées aux nouvelles et pour cela on calcule les corrélations des anciennes variables avec les nouvelles.

Le coefficient de corrélation entre les anciennes et les nouvelles variables représente les coordonnées des anciennes variables. Donc la représentation se fait sur un cercle appeler "cercle des corrélations", qui vient du fait qu'un coefficient de corrélation variant entre -1 et +1, les représentations des variables de départ sont des points qui se trouvent à l'intérieur d'un cercle de rayon 1 (Duby. et., al, 2006).

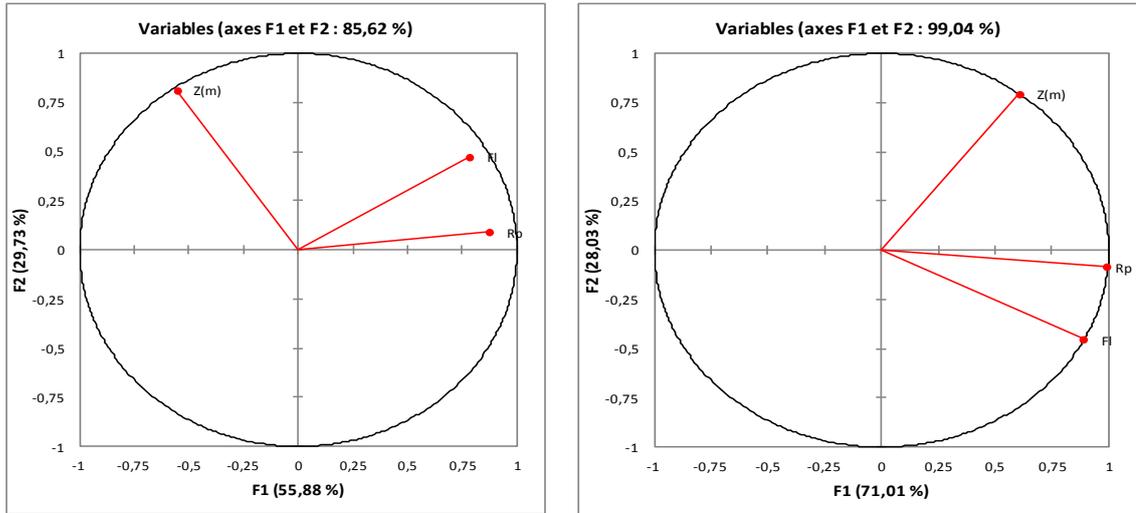


Figure 2 5: Représentation des variables sur le cercle de corrélation

On peut montrer que les variances, covariances et coefficients de corrélation empiriques des composantes principales entre elles ou avec les variables de départ sont (Duby. et., al, 2006).:

$$\text{var}(C_k) = \frac{1}{n} \alpha_k^T X_{ct}^T X_{ct} \alpha_k = \alpha_k^T \Sigma \alpha_k = \lambda_k \quad (2-44)$$

$$\text{cov}(C_k, V_{cj}) = \frac{1}{n} \alpha_k^T X_c^T V_{cj} = \frac{1}{n} \alpha_k^T X_c^T X_c \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \alpha_k^T \Sigma \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \lambda_k \alpha_k^T \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \lambda_k \alpha_{kj} \quad (2-45)$$

Et enfin:

$$R(C_k, V_j) = \sqrt{\lambda_k} \frac{\alpha_{kj}}{\sqrt{\text{var}(V_j)}} \quad (2-46)$$

Ou α_{kj} est la j^{ème} coordonnée du vecteur directeur unitaire α_k de Δ_k .
De façon générale, la matrice de covariance des composantes principales est égale à Σ_C :

$$\Sigma_C = \frac{1}{n} A^T X_{ct}^T X_{ct} A = A^T \Sigma A = \Lambda, \quad (2-47)$$

Ou Λ est la matrice diagonale des valeurs propres de Σ :

$$\Lambda = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_p \end{bmatrix}$$

Et la matrice des covariances entre les composantes principales et les anciennes variables vaut:

$$\text{cov}(C, V) = \frac{1}{n} X_{ct}^T X_{ct} A = \Sigma A = A \Lambda \quad (2-48)$$

selon la même référence, si on remarque que la variance empirique d'une variable est le carré de la norme du vecteur qui la représente dans la géométrie euclidienne choisie et que le coefficient de corrélation empirique de deux variables est égal au produit scalaire des deux vecteurs qui les représentent, on pourra interpréter les angles des vecteurs comme des corrélations.

2.10.1. Interprétation des axes en fonction des anciennes variables :

Une ancienne variable V_j explique d'autant mieux un axe principal qu'elle sera fortement corrélée avec la composante principale correspondant à cet axe.

2.10.2. Qualité de la représentation des variables :

Encore une autre fois selon la même référence, une variable est bien représentée lorsque sa corrélation avec la composante principale correspondante en valeur absolue est proche de 1. De même la variable est bien représentée sur un plan si elle est proche du bord du cercle des corrélations.

2.10.3. Étude des liaisons entre les variables :

Sur le cercle des corrélations, on peut interpréter les positions des anciennes variables les unes par rapport aux autres en termes de corrélations comme le site (Duby. et al., 2006).

Deux points très proches du cercle des corrélations sont très corrélés positivement entre elles. Si elles sont proches du cercle, mais dans des positions symétriques par rapport à l'origine, elles seront très corrélées négativement, et variables proches du cercle des corrélations et dont les vecteurs qui les joignent à l'origine forment un angle droit, ne seront pas corrélées entre elles.

2.11. ANALYSE EN COMPOSANTES PRINCIPALES NORMEE :

Dans l'ACP simple, tous les individus ont le même poids dans l'analyse, mais aussi, toutes les variables jouent le même rôle et les nouveaux axes sont issus de la matrice de covariance empirique des variables. Cela pose parfois des problèmes qu'on site, si les anciennes variables sont hétérogènes, comme par exemple des poids, des tailles et des âges, quel sens peut-on donner aux composantes principales qui sont alors des combinaisons linéaires de variables mixtes ; si on change d'unités sur ces variables, on peut changer complètement les résultats de l'ACP, et une variable contribuera d'autant plus à la confection des premiers axes, que sa variance est forte.

Pour résoudre ces problèmes, on cherchera à normaliser les variables et à travailler sur des variables sans dimension. La méthode la plus utilisée est celle qui admet de

diviser les valeurs des variables par leur écart-type, c'est-à-dire que l'on travaille sur des variables centrées et réduites.

Cela revient à faire la même analyse que pour l'ACP simple, mais on choisit une autre distance euclidienne entre les individus que la distance euclidienne classique. La distance choisie vaut :

$$d^2(u_i, u_{i'}) = \sum_{j=1}^p \frac{1}{\sigma_j^2} (x_{ij} - x_{i'j})^2 \quad (2-49)$$

Cette nouvelle distance permet de faire jouer un rôle à chacune d'entre les variables.

Si on reprend tous les calculs de l'ACP simple, mais en remplaçant les variables de départ par les variables centrées réduites, on voit que ce n'est plus la matrice de covariance, mais la matrice de corrélation R qui intervient pour la recherche des nouveaux axes.

Les propriétés de l'ACP normée par rapport à l'ACP simple découlent du fait que la matrice de corrélation R n'a que des 1 sur sa diagonale principale. Cela entraîne que sa trace est toujours égale à p. Et la trace de la matrice est égale à l'inertie totale du nuage calculée avec la distance euclidienne que l'on a choisie.

Donc l'inertie totale du nuage des individus dans \mathbb{R}^p est toujours égale à p. Cette particularité donne une règle supplémentaire pour choisir le nombre d'axes que l'on va garder pour les interprétations, conçue sur le raisonnement suivant : on a p valeurs propres dont la somme vaut p ; on peut ne considérer comme significatives que les valeurs propres dont la valeur est supérieure à 1, puisque la valeur moyenne des valeurs propres vaut 1 et leur somme vaut p. Sa peut servir de guide pour le choix de la dimension du sous-espace que l'on veut garder.

Une autre particularité de l'ACP normée est que la représentation des variables avec les cercles de corrélation correspond exactement à la représentation des variables dans \mathbb{R}^n .

CONCLUSION :

L'interprétation des représentations factorielles prend en compte plusieurs critères et ne se limite pas systématiquement à une simple description du nuage de points que ce soit en ACP ou en AFC.

D'une manière générale, nous pouvons dire que l'origine des axes G est le centre de gravité du nuage : il représente donc la tendance moyenne des individus. Les individus très loin du centre présentent un écart important par rapport à cette moyenne et influencent beaucoup les statistiques.

L'inertie d'un axe (ou la valeur propre associée à son vecteur directeur) témoigne de la représentativité de cet axe relativement au nuage de points. Plus un plan possède une grande inertie, plus la projection du nuage sur ce plan est réaliste et moins il y a eu de perte d'information.

Les coordonnées des points sur les axes factoriels sont les coefficients de corrélation des variables ou des individus qu'ils représentent avec ses axes. La valeur du produit scalaire des vecteurs joignant deux points à l'origine n'est autre que le coefficient de corrélation entre les deux variables correspondantes : deux points voisins indiquent des variables fortement corrélées.

Il est important de conclure aussi, que ce n'est pas forcément avec les vecteurs propres que l'on explique le mieux les constitutions des axes factoriels, mais plutôt avec les deux indicateurs que sont les contributions des points aux axes et les cosinus carrés.

L'analyse et l'interprétation d'une représentation factorielle nécessite donc du recul et la prise en compte de nombreux éléments. Elle doit reposer sur des critères valides et non uniquement sur une description a priori de son allure même si celle-ci peut apporter des indications intéressante. C'est la difficulté principale rencontré dans ce travail, on bien senti cela dans l'application à une banque de donnée des essais pénétrométriques pour un ouvrage très important.

Chapitre 3 : ANALYSE DE DONNEES PENETROMETRIQUES: CAS DE L'AÉROPORT HOUARI BOUMEDIENE.

3.4. PRESENTATION DU SITE:

3.4.1. L'Aéroport Houari Boumediene d'Alger :

Dans le cadre de la construction d'une nouvelle aérogare et d'une gare de fret dans l'aéroport international d'Alger, une expertise du sous-sol et des fondations a été réalisée en 1983.

Le site étudié (Aéroport d'Alger) se situe à environ 20 km au sud-est de la ville d'Alger, soit dans la grande plaine côtière, « Plaine de la Mtidja », qui s'étend entre les montagnes de l'Atlas tellien et la méditerranée. La figure suivante (Figure 3.1) donne un aperçu de la situation de l'aéroport.



Figure 3 1: Localisation de l'aéroport d'Alger.

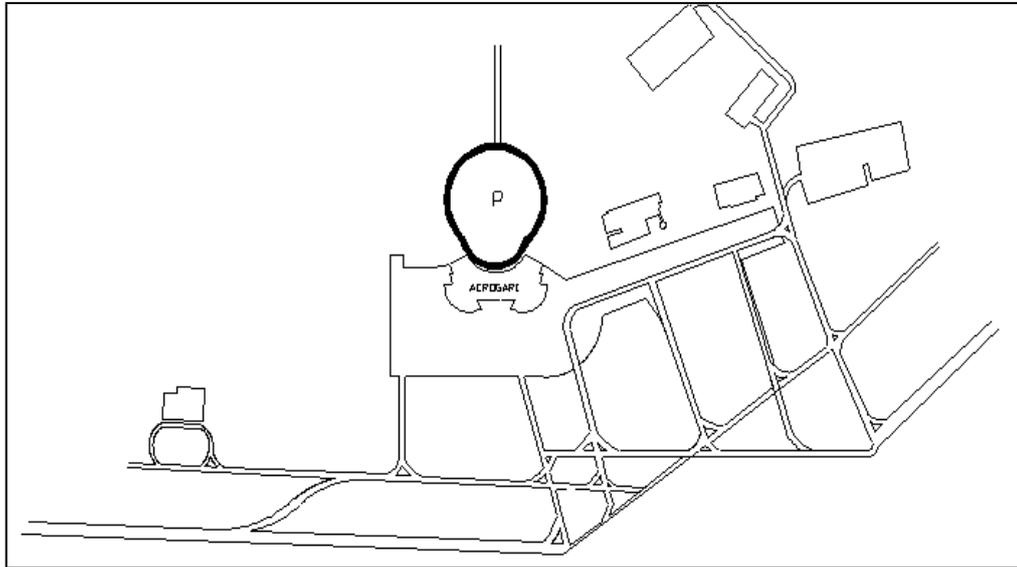


Figure 3 2: Plan de l'aéroport HOUARI Boumediene.

La plaine de la Mitidja est sillonnée par cinq principaux oueds issus de l'Atlas et assurant son drainage vers le Nord avec un sens d'écoulement général Sud Ouest-Sud Est. Le principal cours d'eau intéressant le site est l'Oued Smar qui reçoit à l'origine deux affluents : l'Oued Boutrick et l'Oued Saïd. Ses eaux ont été déviées au Sud et l'Ouest de l'aéroport pour permettre la construction de pistes. La figure 3.3 représente le réseau hydrographique de la région et la figure 3.4 montre les aménagements réalisés sur l'Oued Smar.

Le terrain de l'aéroport est peu dénivelé, soit une cote de +16 m à 23 m, c'est-à-dire seulement quelques mètres au-dessus du niveau de la mer.

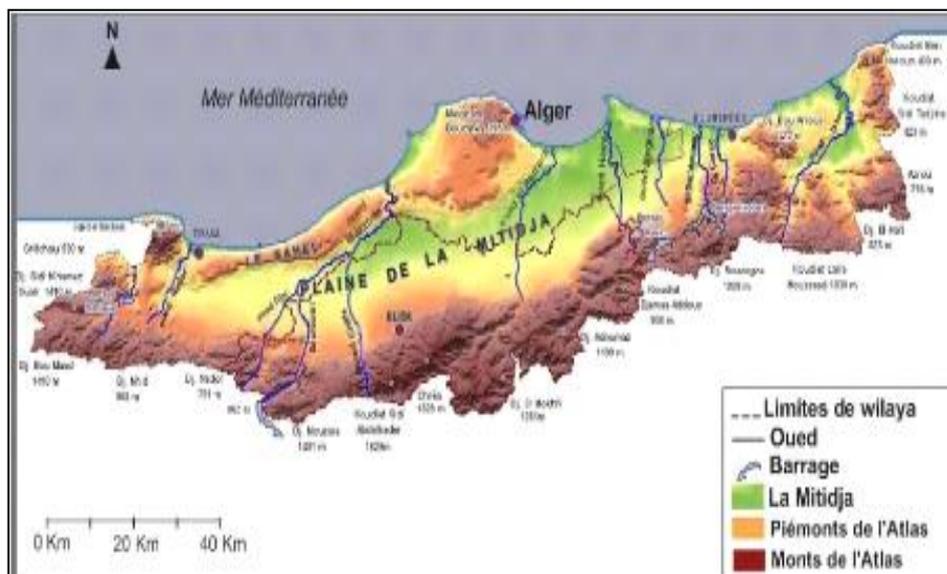


Figure 3 3: Réseau hydrographique de la plaine de la Mitidja.

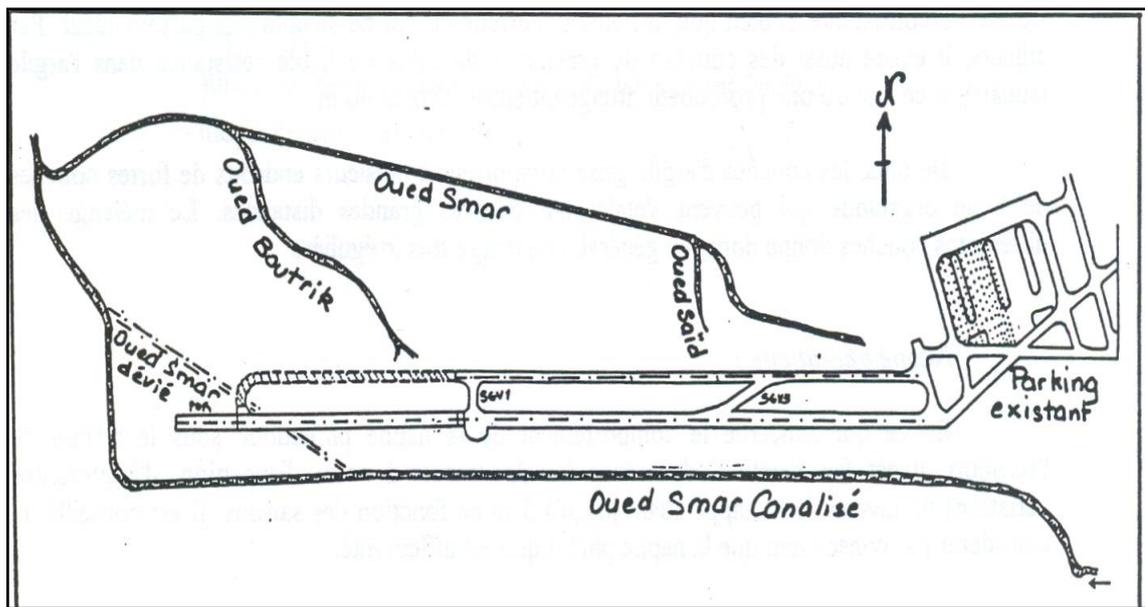


Figure 3 4: Aménagement de l'Oued Smar.

3.4.2. Caractérisation du sol:

3.4.2.1. Constitution du sol :

a) Nature du sol :

Sous une couche de terre végétale ou sous une couche de terrain remblayé atteignant une épaisseur de 0,3 m à 1,5 m, se trouve dans la partie sur du terrain, une série de couches d'argile grise jusqu'à 19 m voire même 22 m de profondeur. Ces couches sont en alternance mi-consistance ou consistance jusqu'à très consistantes.

Jusqu'à une profondeur de 16 m à 19 m, parfois même 22 m, ces couches d'argiles sont interrompues par des couches de graviers, de galets et de sable qui sont au moins moyennement denses.

Sous ses couches géologiquement jeunes d'argiles grise, à partir d'une profondeur de 19 m à 22 m et ce jusqu'à une très grande profondeur, se trouve une couche apparemment beaucoup plus anciennes d'argile jaunâtre et tachetée. Cette couche est très consistante ou même très dure.

Au nord du terrain, la couche jaunâtre d'argile monte peu à peu pour atteindre le niveau du terrain. Sa résistance diminue quelque peu.

Les intercalations de sable ou de gravier sont rares jusqu'à une profondeur de 16 m mais elles deviennent plus fréquentes entre 16 m et 20 m. Elles sont toutefois très irrégulières aussi bien horizontalement que verticalement et sont toutefois très irrégulières aussi bien horizontalement que verticalement et sont très fréquemment interrompues par de l'argile mi-consistante si bien que le pouvoir porteur du sol ne semble pas surtout idéal. Par ailleurs, il existe aussi des couches de gravier et de sable

de faible résistance dans l'argile jaunâtre et ce jusqu'à une profondeur atteignant entre 20 m et 40 m.

De plus, les couches d'argile grise comportent à plusieurs endroits de fortes couches de boue organique qui peuvent s'étaler sur de très grandes distances. Le mélange des différentes couches donne donc, en général, une image très irrégulière.

b) Nappe phréatique :

En ce qui concerne le comportement de la nappe phréatique sous le terrain de l'aéroport, il est fait mention, dans tous les documents à notre disposition, d'importantes variations du niveau de la nappe allant jusqu'à 5 m en fonction des saisons. Il est conseillé de considérer par conséquent que la nappe phréatique est affleurante.

3.4.3. Caractéristiques mécaniques :

Il ressort de la synthèse des différents essais en place et plus particulièrement des pénétrogrammes que les sols en place sont très hétérogènes du point de vue de leur résistance. Ainsi, en général, les couches suivantes apparaissent successivement :

- ❖ En surface, les sols présentent des résistances de pointes élevées témoignant de la présence des sols surconsolidés sujets aux variations de teneurs en eau (fluctuations de la nappe et évaporation). L'épaisseur moyenne est 2,3 m.
- ❖ Au dessous de cette couche de surface, on trouve des sols présentant des faibles résistances de pointe ($R_p < 1$ MPa) reflétant des argiles molles ou des vases.
- ❖ Au-dessous, on rencontre des sols très résistants avec des résistances de pointe élevées à très élevées ($R_p > 3$ MPa à $R_p > 10$ MPa) témoignant de la présence de sols pulvérulents.

On note que les couches de faible résistance sont parfois intercalées de couches résistantes.

La reconnaissance géophysique a mis en évidence les mêmes couches au moyen des mesures de résistivité.

Enfin, les essais en laboratoire ont confirmé les résultats déjà obtenus. Les sols rencontrés sur six sondages différents sont, d'une manière générale, des sols :
Fins ;

De densité très variable (de peu denses à très denses) ; Proches de la saturation à la saturation ; Plastiques ; Sur-consolidés en surface et sous consolidés en profondeur ; Moyennement compressible.

3.5. PRESENTATION DES DONNEES :

3.5.1. Origine des données :

données que nous traiterons dans cette étude proviennent de sondages effectués avec un pénétromètre statique de type Gouda (voir annexe A). Nous disposons d'un grand nombre de courbes pénétrométriques sur lesquelles sont représentées simultanément la résistance de pointe opposée par le sol à l'enfoncement de la pointe conique du pénétromètre et la force de frottement du sol sur la paroi latérale de la tête de mesure.

Etant donnée l'importance du nombre de sondages, nous nous limiterons aux 67 sondages du parking de l'aéroport et à quelques sondages sur les taxiways TW9 et TW10.

L'implantation des essais au pénétromètre est donnée sur la figure 3.5. Pour chaque courbe pénétrométrique, nous avons relevé les valeurs de la résistance de pointe R_p et de la force de frottement latéral F_l tous les 0,5 m jusqu'à 20 m ou moins suivant le refus. Etant donnée la qualité médiocre de certains pénétrogrammes, des valeurs manquent parfois, à certaines profondeurs.

Les coordonnées d'implantation des sondages, quant à elles, ont été définies par rapport à un repère choisi de manière tout à fait subjective.

3.5.2. Présentation du tableau de données :

Le tableau de données voir annexe B comporte 7 colonnes et 2775 lignes et est construit de la manière suivante :

- ❖ Chaque ligne correspond à une mesure de R_p et F_l à une profondeur précise et pour un sondage donné. Les lignes correspondant à un même sondage sont regroupées.
- ❖ Les colonnes contiennent successivement :
 - i. Le numéro de ligne de mesure ou libellé de l'individu ;
 - ii. La mesure qui correspond au numéro du sondage ;
 - iii. La profondeur z (en mètres) de la mesure ;
 - iv. La valeur de résistance de pointe R_p (en bars) ;
 - v. La valeur du frottement latéral F_l (en bars).

Il est important de noter que ce tableau correspond au tableau brut de données, il sera éventuellement transformé ou découpé pour les études qui suivent.

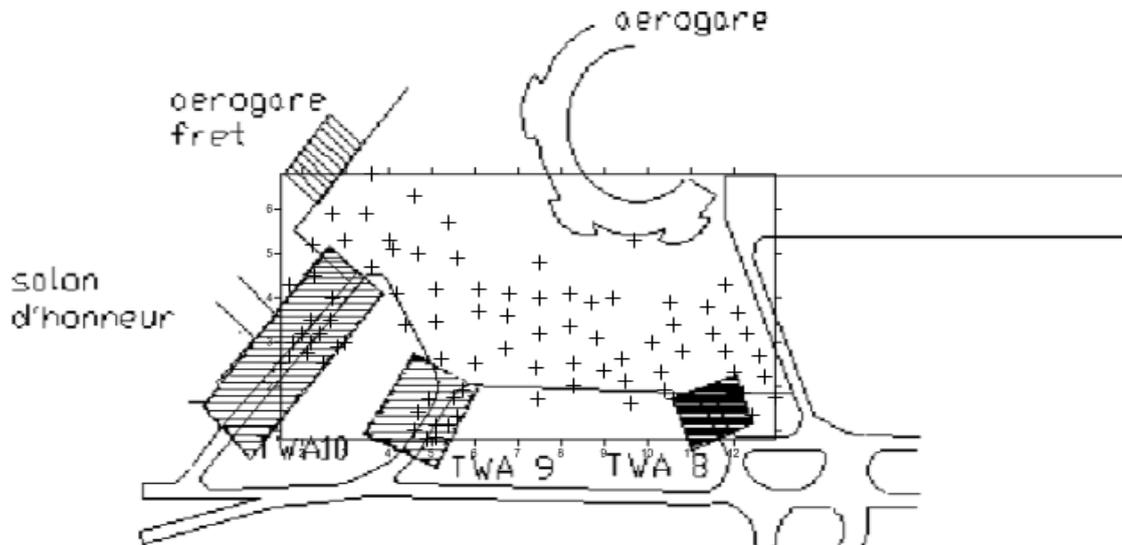


Figure 3 5 : Plan d'implantation des sondages pénétrométriques.

3.6. ETUDE SONDAGE PAR SONDAGE (SENS VERTICALE) :

3.6.1. Objectif de l'étude :

L'étude porte sur deux points complémentaires à savoir le contexte du problème de drainage et le contexte de l'étude.

3.6.1.1. Dans le contexte du problème de drainage :

Cette étude fait suite en quelque sorte à une expertise de l'étude de consolidation menée par Magnan en 1995. Lors d'une expertise précédente, effectuée en 1986, pendant les travaux d'extension de pistes de l'aéroport, il est apparu que le site de l'aéroport Houari Boumediene est parcouru par des vallées fossiles emplies de matériaux plus déformables sur des profondeurs pouvant atteindre une vingtaine de mètres. Ces vallées dont l'emplacement ne correspond pas toujours à celui des cours d'eau actuels (avant même leur dérivation pour l'extension de l'aéroport), nécessitent un traitement par drainage pour contrôler les tassements postérieurs à la mise en service des ouvrages.

Dans le contexte de cette expertise, Magnan s'est attaché à établir une carte des profondeurs des sols qui sont à la fois les moins résistants et les plus compressibles. Cette carte représente des zones en fonction de la profondeur de la base des sols les plus compressibles. Pour déterminer cette base ; Magnan a utilisé une méthode graphique ; il observait sur les courbes pénétrométriques les variations de la résistance de pointe R_p suivant quatre classes de valeurs :

- Classe A : $R_p < 1 \text{ MPa}$;*
- Classe B : $1 \text{ MPa} < R_p < 2 \text{ MPa}$;*
- Classe C : $2 \text{ MPa} < R_p < 3 \text{ MPa}$*
- Classe D : $R_p > 3 \text{ MPa}$.*

Une fois caractérisé chaque sondage, il a établi la carte des épaisseurs des sols les plus déformables en prenant comme valeur critique de la résistance de pointe $R_p = 2$ MPa. Les cinq zones où l'on rencontre les sols de résistance de pointe $R_p < 2$ MPa jusqu'à des profondeurs de plus de 20 m, 10 à 20 m, 5 à 10 m, 0 à 5 m y sont représentées.

En ce qui concerne, nous voulons établir cette carte mais en utilisant le logiciel SURFER pour caractériser chaque sondage et les couches qui le constituent.

3.6.1.2. Dans le contexte de l'étude

Dans cette étude sur les sols de l'aéroport d'Alger, non seulement nous possédant un grand nombre de données qui plus est, la diversité des sols du site ainsi que le contraste entre la vallée fossile et les matériaux qui remplissent, constituent des conditions idéales pour tester la méthode définie précédemment.

Nous allons donc, dans cette étude, suivre la démarche suivante : pour chaque essai pénétrométrique, nous déterminons les différentes couches d'après les plans factoriels puis nous les caractérisant statistiquement.

Il est clair que dans cette étude l'AFC n'a pas d'avantage particulier par rapport à l'ACP dans ce cas précis, nous ne traiterons les données qu'en analyse en composante principale.

3.6.2. Présentation de l'étude :

3.6.2.1. Zonage sur plans factoriels :

Nous suivrons la même démarche pour tous les sondages successivement aussi le tableau de données a dû être découpé en tableau correspondant chacun à un essai.

Dans un premier temps, nous effectuons l'ACP et éditons le plan factoriel qui en résulte. Ensuite chaque individu est libellé par le numéro qui lui correspond. Enfin, en partant du premier individu et en évoluant suivant les libellés croissants, nous définissons chaque zone dans laquelle tous les individus sont considérés comme proches les uns des autres. Chaque groupe peut alors être répertorié au moyen des numéros des premier et dernier individus lui appartenant.

Bien sûr, les problèmes d'échelle et de notion de proximité existent toujours et ne peuvent pas mieux être résolus. Cependant, pour quatre sondages proches les uns des autres, les échelles des représentations factorielles ont été prises identiques pour permettre des comparaisons (figure 3.9). Nous avons pu constater que les zones déterminées pour chaque sondage étaient les mêmes quelle que soit l'échelle mais qu'il était délicat de superposer les plans factoriels étant donnée leur « unicité ». Il n'est donc pas très intéressant de prendre une échelle unique puisque nous ne pouvons pas comparer les zones déterminées et de plus graphiquement l'étalement du nuage de points diminue.

Les ACP sont réalisées sur l'ensemble des individus d'un même sondage et avec les variables z, Rp et Fl. Cette fois, le nombre d'individus étant plus faible (normalement 40 par sondage), nous pouvons utiliser les graphiques biplots factoriels. Nous nous contentons des plans factoriels passant par le premier et le deuxième axes car leur représentativité est satisfaisante. Les figures 3.5 à 3.8) donnent quelques exemples de représentations factorielles avec les zonages effectués.

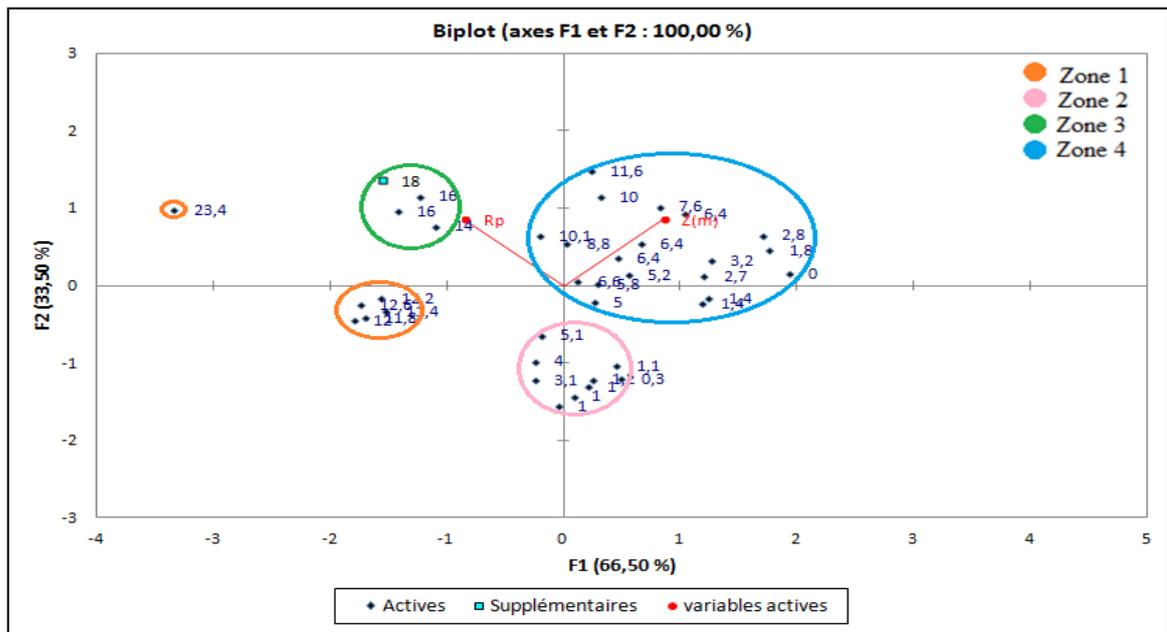


Figure 3 6.a: Biplot factoriel 1-2 de l'ACP des données du sondage 3. Pour Rp.

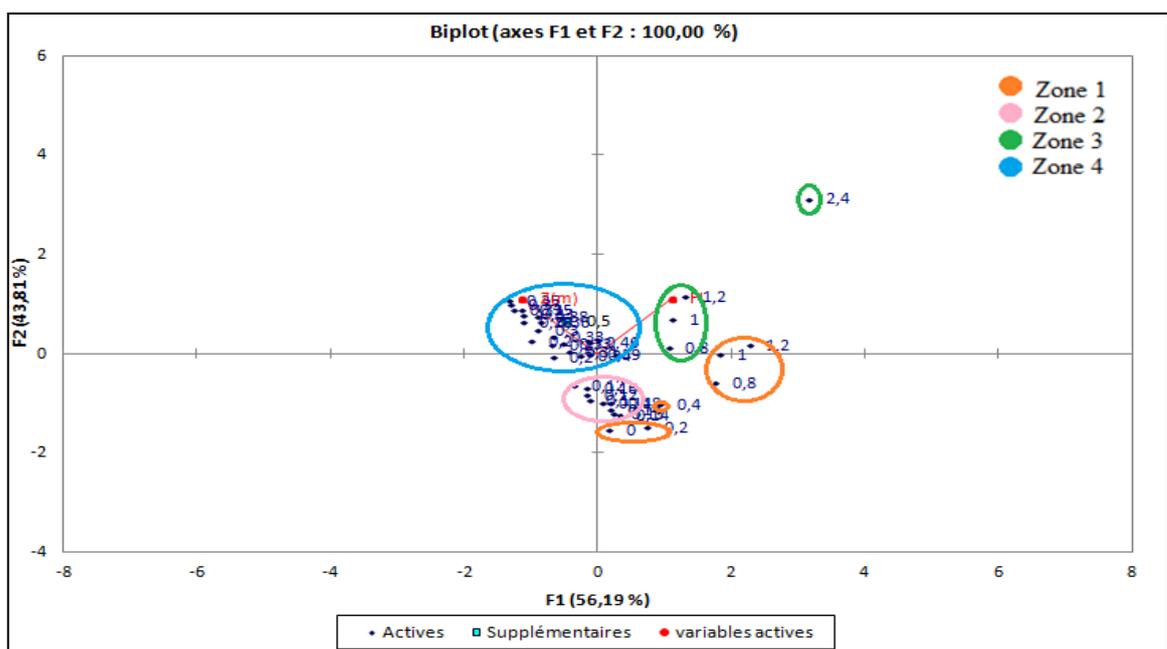


Figure 3 7.b: Biplot factoriel 1-2 de l'ACP des données du sondage 3. Pour Fl.

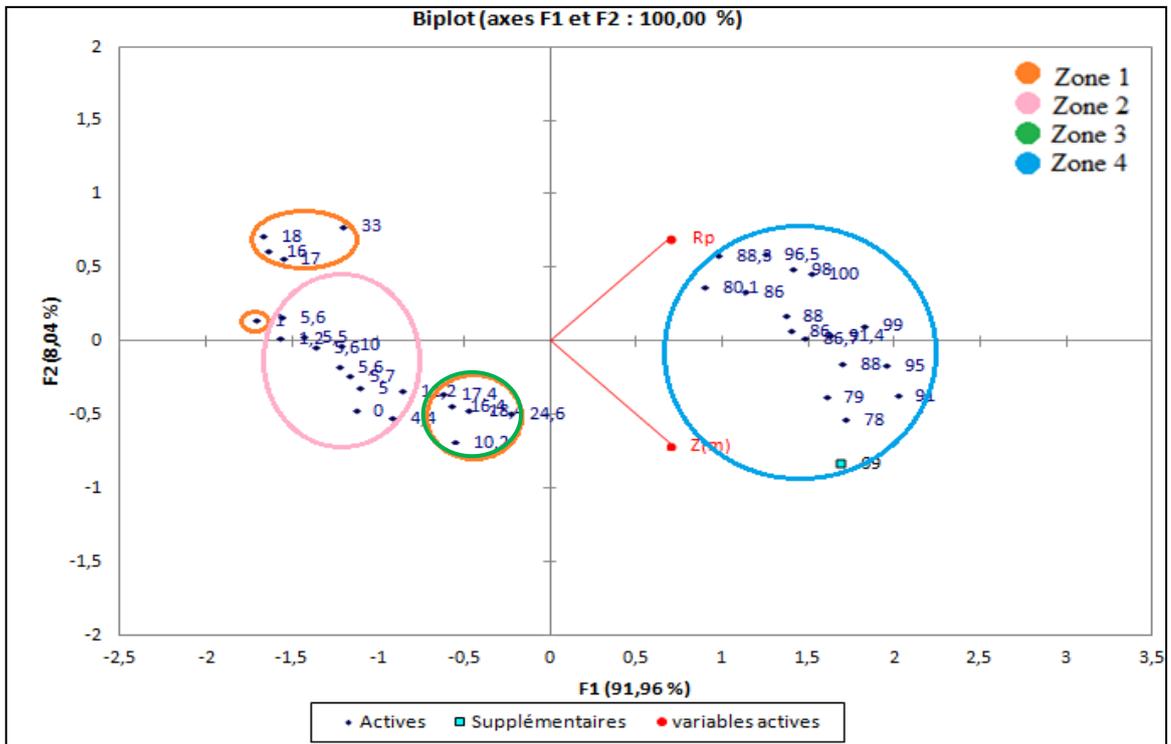


Figure 3 8.a: Biplot factoriel 1-2 de l'ACP des données du sondage 25. Pour Rp.

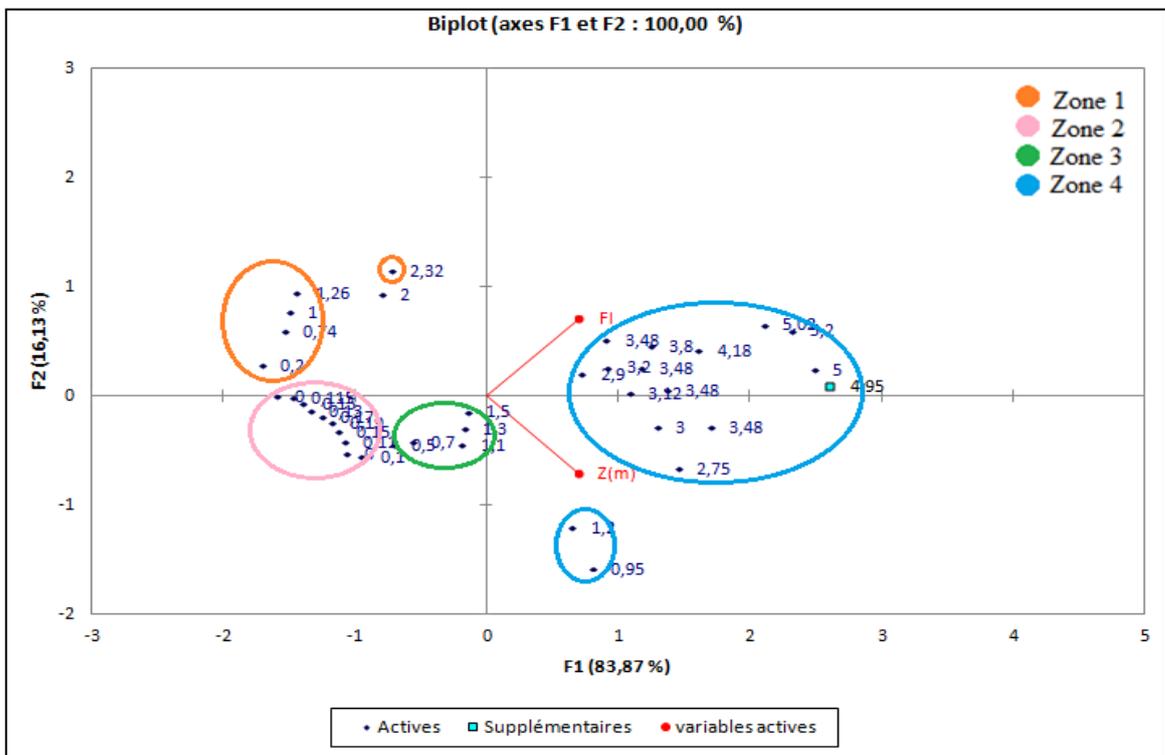


Figure 3 9.b: Biplot factoriel 1-2 de l'ACP des données du sondage 25. Pour F1.

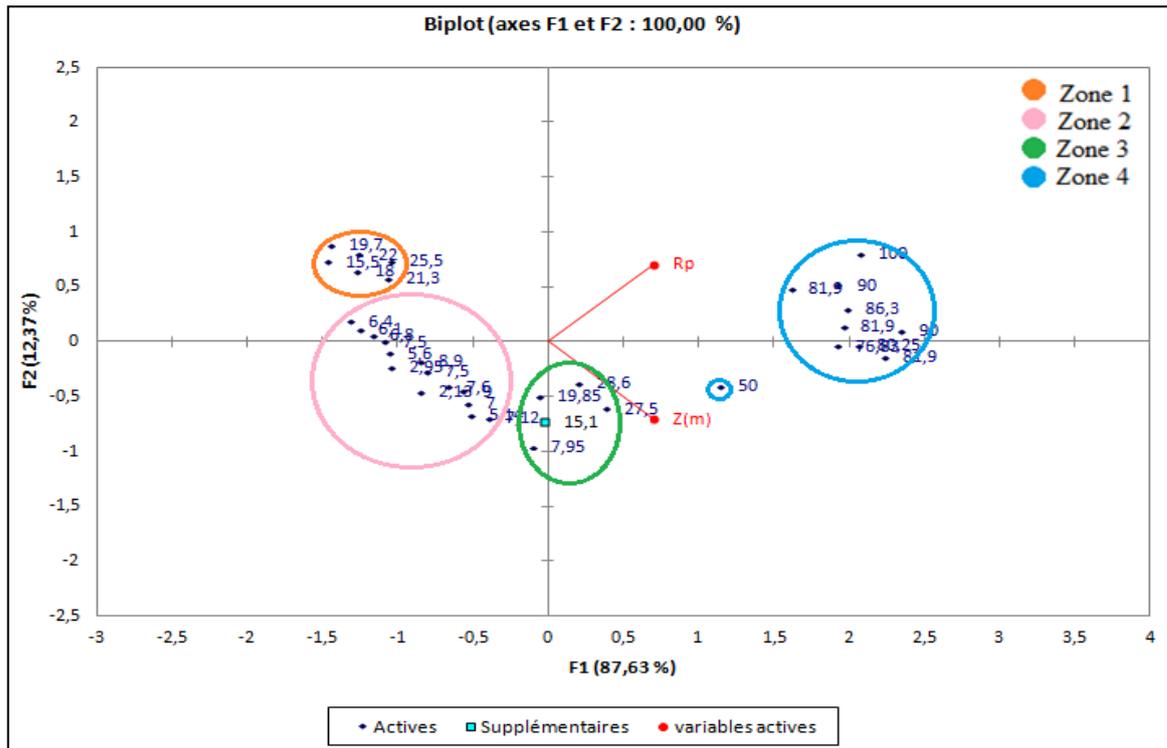
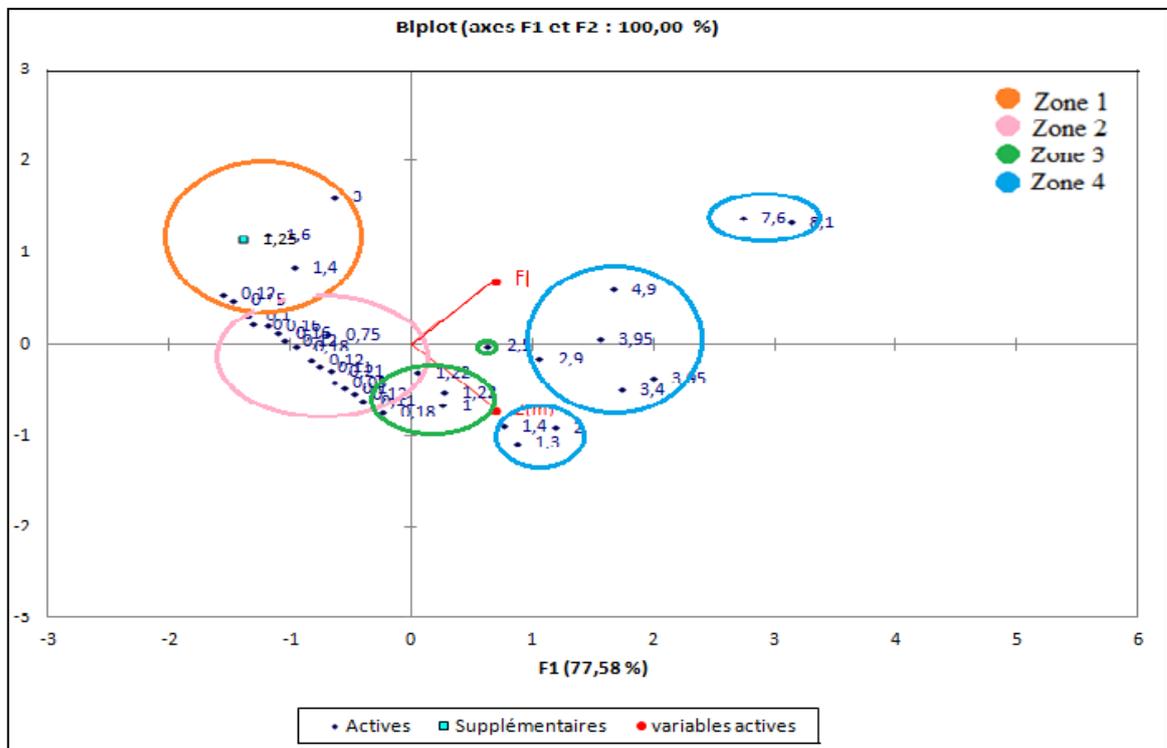


Figure 3 10.a: Biplot factoriel 1-2 de l'ACP des données du sondage 28. Pour Rp.



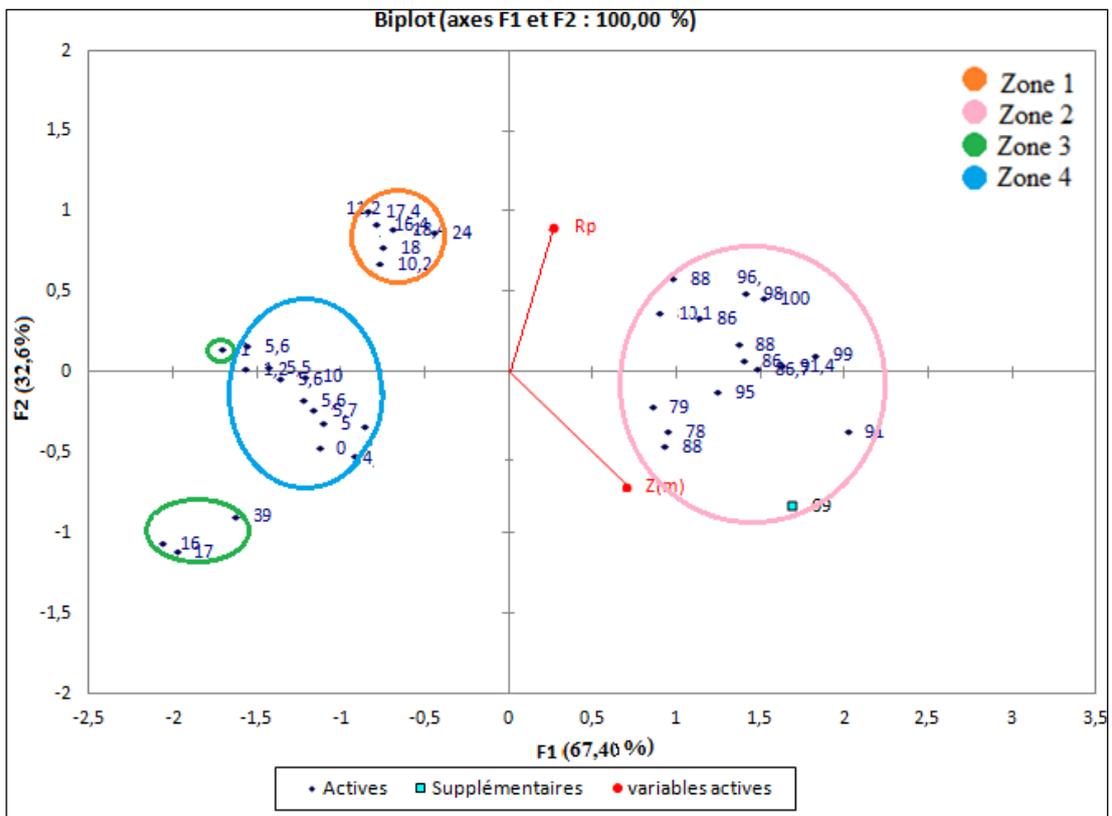


Figure 3 14.a: Biplot factoriel 1-2 de l'ACP des données du sondage 1041. Pour Rp

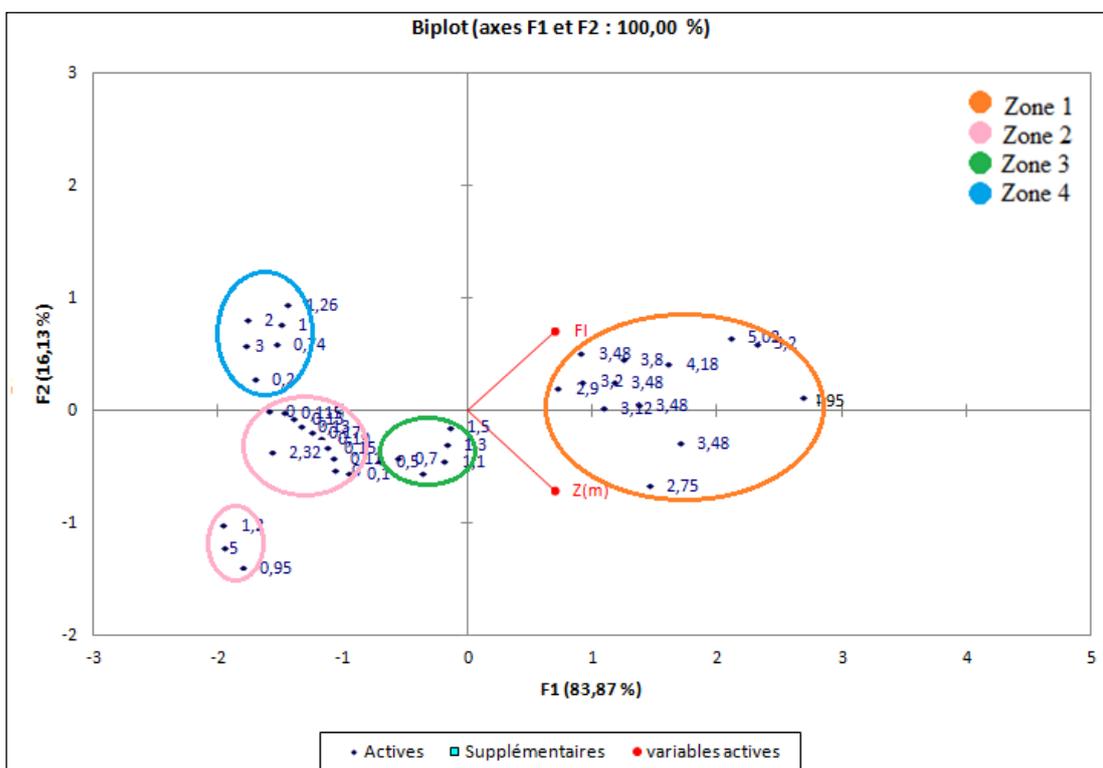


Figure 3 155.a: Biplot factoriel 1-2 de l'ACP des données du sondage 1041. Pour F1

3.6.3. **Caractéristiques statistiques des couches. :**

Pour chaque sondage, nous répertorions les zones déterminées sur les représentations factorielles. Ensuite, nous calculons pour les valeurs R_p et F_l leur moyenne, leur variance et leur coefficient de variation et ce pour chaque couche par ses caractéristiques moyenne. De plus, nous pourrions comparer les coefficients de variations de chacune des couches avec le coefficient de variation global du sol : si nous observons une très nette diminution des coefficients de variation des couches par rapport à celui de tout le sol et si des couches sont relativement faibles, nous pourrions conclure à l'homogénéité des couches.

Le tableau récapitulatif des statistiques, sondage par sondage est donné dans en annexe B.

D'une manière générale, nous constatons bien une nette voire très nette diminution des coefficients de variation entre chaque couche et tout le sol. Toutefois, les valeurs obtenus ne sont pas toujours satisfaisante et notamment pour les couches les plus profondes ; ceci s'explique sans doute par la présence des inclusions irrégulières de sable et de graviers que la reconnaissance géotechnique mis en évidence.

Nous pouvons donc tout de même conclure que la méthode exposée semble valide pour mettre en évidence les couches de sols. Dans la suite, nous allons établir la carte des profondeurs des sols compressibles et la comparer à celle obtenue par Magnan.

3.6.4. **Exploitation des résultats :**

3.6.4.1. **Carte des profondeurs des sols compressibles :**

En premier lieu, pour établir la carte des profondeurs de sols compressibles, il faut pour chaque sondage déterminer la profondeur de la base des sols les plus compressibles ($R_p < 2$ MPa).

Pour cela, nous recherchons dans le tableau établi ultérieurement la couche la plus profonde telle que la valeur moyenne de la résistance de pointe plus ou moins l'écart-type soit inférieure à la valeur critique de 2 MPa :

$$m_{R_p} = \pm \sqrt{\sigma_{R_p}^2 \leq 2 \text{MPa}}$$

(3-1)

On a pris la valeur moyenne plus ou moins l'écart type pour des raisons de sécurité. En effet, la moyenne de la résistance de pointe sur une couche peut être supérieure à 2 MPa mais il peut exister à l'intérieur des inclusions plus ou moins importantes de sol moins résistant. L'écart type permet de tenir compte des valeurs excentrées par rapport à la valeur moyenne.

Lorsque cette couche est repérée, nous en déduisons la profondeur de la base des sols les plus compressibles qui correspond à la profondeur de l'interface inférieure de la

couche. Le tableau suivant (Tableau 3.1) ci-dessous donne pour chaque sondage cette profondeur.

n°	Z(m)	n°	Z(m)	n°	Z(m)	n°	Z(m)	n°	Z(m)
1	3,5	24	12	45	10	65	16	1020	16
2	19	25	10,5	46	19,5	67	12,5	1021	5,5
3	19	26	5	47	19,5	68	9,5	1039	0
4	8	27	14	48	19,5	980	0	1040	6,5
5	13	28	12,5	49	14	981	0	1041	8,5
6	1,5	30	17	50	13	982	6		
7	0	31	20	51	13	983	17,5		
8	2	32	18,5	52	5	984	5		
9	0	33	13,5	53	20	985	0		
10	18	34	10	54	0	986	3		
11	15,5	35	10	55	10	991	13,5		
12	10	36	17,5	56	1,5	992	6,5		
16	15,5	37	19,5	57	1,5	993	6		
17	16	38	10,5	58	0	1010	14,5		
18	7,5	39	10,5	59	17,5	1011	15,5		
19	0	40	0	60	1	1012	15		
20	8	41	0	61	6,5	1013	15		
21	5	42	4,5	62	20	1014	11		
22	10	43	0	63	16	1018	11,5		
23	1,5	44	7,5	64	20	1019	12,5		

Tableau 3 1 : Profondeur de la base des sols les plus compressibles.

Au vu de ce tableau, nous pouvons déduire à quelle zone appartient chaque sondage et réaliser un travail de cartographie. Pour établir la carte des profondeurs des sols compressibles, nous traçons approximativement les limites de zones en nous référant aux valeurs du tableau ci-dessus. C'est-à-dire que suivant la profondeur déterminée pour un sondage, celui-ci se trouve plus ou moins loin des limites qui sont 0 m, 5 m, 10m, 15 m et 20 m.

Nous avons établi une carte (figure 3.10) avec une variante au Nord-Ouest du terrain. Elle semble correcte vis-à-vis des profondeurs des profondeurs déterminées ci-dessus.

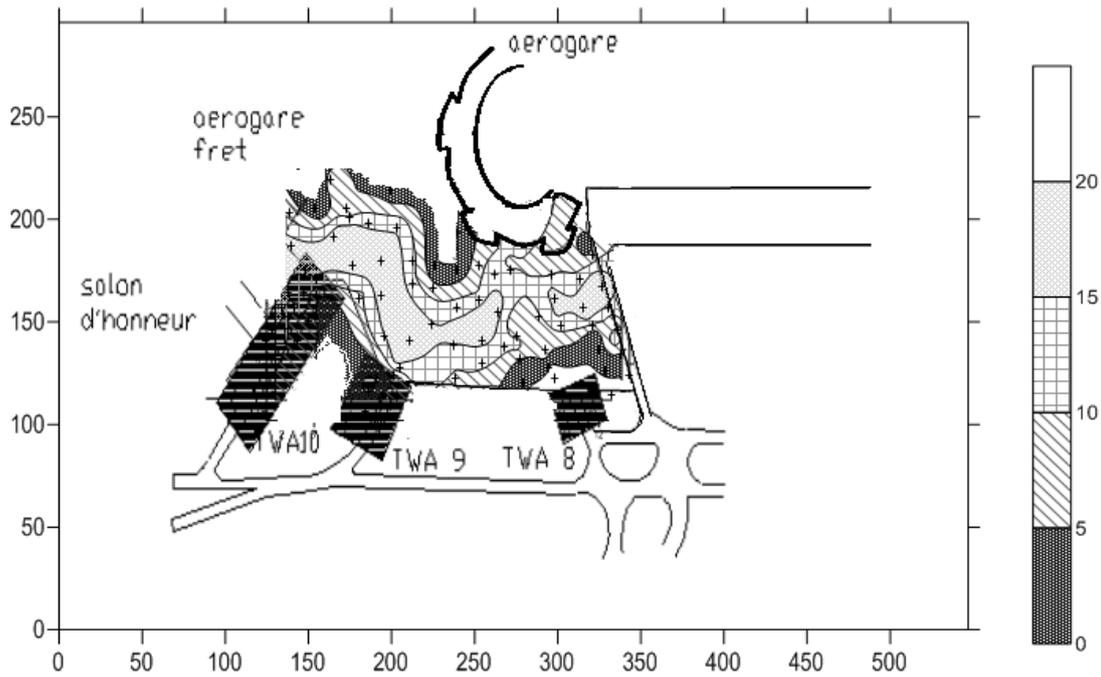


Figure 3 16: Carte de profondeur de la base de la base des sols compressibles.

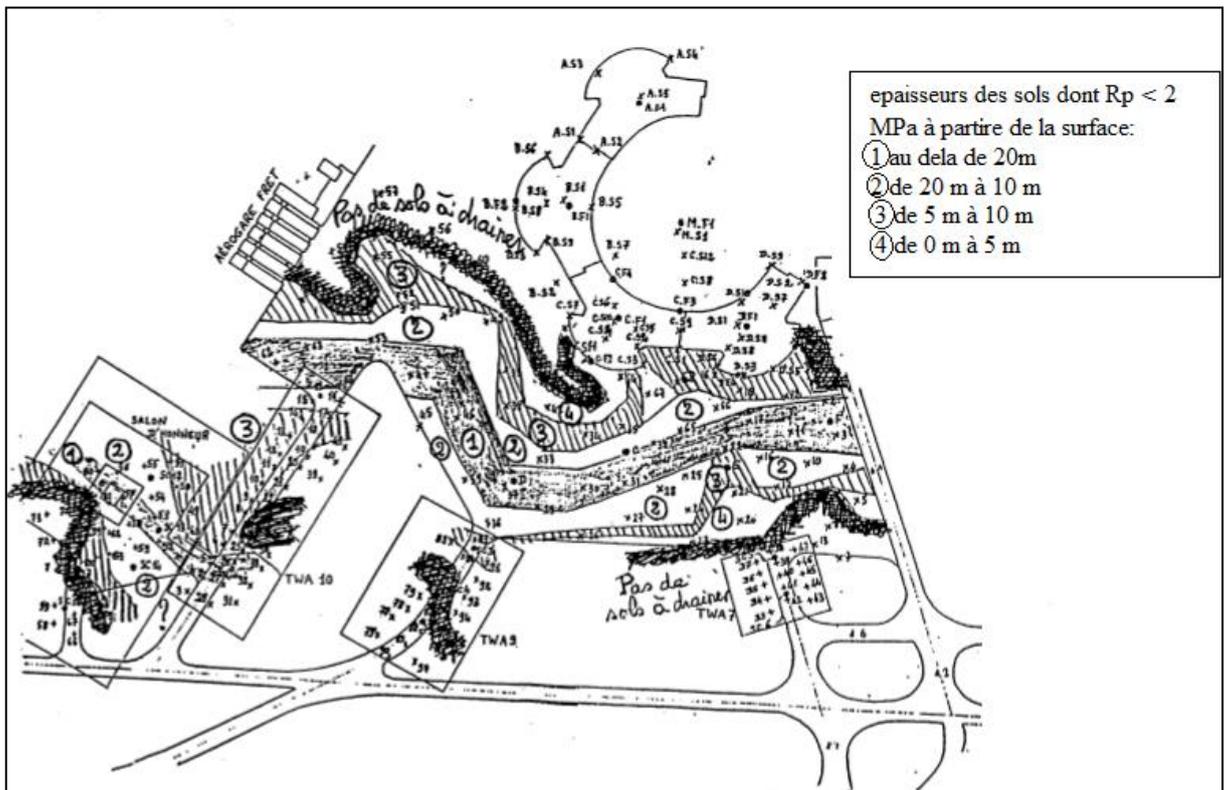


Figure 3.11 : Carte des profondeurs de la base de la base des sols compressibles établie graphiquement par Magnan.

La figure 3.11 représente la carte établie par Magnan. Si l'on compare notre hypothèse avec cette carte, on constate que les deux cartes se ressemblent légèrement dans toute la zone du parking. La similitude est moins évidente dans les zones de s Taxiways.

Il faut noter la limite de notre étude : la méthode graphique de Magnan permet de déterminer la profondeur de sol compressible au mètre près tandis que nous raisonnons en terme de couches compressibles. C'est-à-dire que nous considérons comme base des sols les plus compressibles l'interface inférieure de la couche la plus profonde telle que $R_p < 2 \text{ MPa}$. Dans notre cas, la présence éventuelle de zones plus résistantes en partie inférieure de couche n'est pas prise en compte et peut expliquer des différences de profondeur.

D'une manière générale, nous pouvons dire que notre carte obtenue n'est pas correcte à 100% puisqu'elle ne correspond pas à celle établie par Magnan. Ceci nous permet de conclure que notre méthode de détermination de couches et nécessite des études plus approfondie pour justifier sa fiabilité. Cependant, elle implique un travail important « à la main » pour définir les zones sur les plans factoriels puis pour traiter chaque couche dans un sol nécessiterait in traitement informatique adéquat afin de la rendre moins subjective et plus performante.

3.6.5. Etude dans les plans horizontaux :

La stratigraphie n'est peut être pas la seule utilisation possible de l'ACP en mécanique des sols, c'est pourquoi nous voulons tirer parti de la quantité de données que nous possédant pour approfondir nos recherches. Lorsque nous avons étudié les données sondage par sondage nous n'avons pas tenu compte des coordonnées x et y dans un plan horizontal. L'idée serait donc d'étudier les résultats des essais pénétrométriques dans des plans horizontaux situés à différentes profondeurs.

3.6.5.1. Méthodologie:

Pour cette étude dans des plans horizontaux, nous devons trier les données et découper le tableau non plus suivant le numéro de sondage mais suivant la profondeur. Le refus de certains sondages se présentant avant 20m, nous a conduits à compléter les données de R_p et F_l . Nous avons constaté dans la banque de données qu'à partir du refus les dernières valeurs de R_p et F_l sont arrondies à la demi-dizaine la plus proche.

La résistance de pointe R_p , étant en général le paramètre le plus préjudiciable pour l'étude des sols compressibles, nous allons nous intéresser à son évolution avec la profondeur.

L'objectif de cette étude des plans horizontaux est donc d'établir des cartes de courbes isométriques de la résistance de pointe à différentes profondeur. Nous étudierons des plans tous les 2,5 m à partir du niveau zéro.

Dans ce cas précis, l'intérêt de l'analyse en composante principale est qu'elle donne les représentations factorielles où tout les individus apparaissent dans un plan unique.

Les ACP sont réalisées sur les sondages de la zone du parking et en prenant comme variable Rp et Fl. C'est le plus petit nombre possible de variables et par conséquent nos plans factoriel auront une inertie de 100% puisque les deux variables sont expliquées par deux facteurs.

En indiquant dans un premier temps, la valeur de la résistance de pointe comme libellé, nous pouvons tracer sur les plans factoriels des droites divisant le nuage de points en classes d'appartenance d'étendue 10 bars.

3.6.5.2. Courbes isométrique de la résistance de pointe :

Les courbes isométriques de résistance de point sont faites par le logiciel SURFER.

Ce dernier effectue des modèles numériques de terrain en établissant des grilles contenant les coordonnées cartésiennes (X, Y, Z) qui définissent l'espace à étudier.

Dans le cas échéant les données sont issues des valeurs pénétrométriques (résistance de pointe) pour chaque sondage effectué dans le site de l'aéroport d'Alger.

Lors cette étude pour la réalisation des courbes, nous avons pour chaque profondeur (Z) fixe on introduit la résistance de pointe associée à chaque position de sondage définie par X et Y.

Donc on effectue une grille composée de trois colonnes, dont la première et la deuxième contiennent les coordonnées de chaque sondage et la troisième colonne contient la résistance de point affecter a chaque profondeur, puis à partir des données introduites on fait les courbe d'isométrie de résistance par l'option (contour map).

Les figures 3.12 à 3.18 représentent les cartes des courbes isométriques de la résistance de pointe aux profondeurs : 5 m ; 7,5 m ; 10 m ; 12,5 m ; 15 m ; 17,5 m ; 20m.

Au travers de ces courbes isométriques apparaît en aulques sorte un mouvement du terrain à une profondeur donnée. Si l'on considère la zone délimité par la courbe correspondant à $R_p = 20$ bars pour chaque profondeur étudiée, nous pourrons alors voir l'évolution des sols les plus compressible avec la profondeur.

Les figures suivantes montrent les cartes des courbes isométriques déjà évoquées ci-dessus.

Encore une autrefois, nous soulignons un mouvement du terrain à une profondeur donnée. Cela nécessite d'approfondir les investigations dans ce sens. Il sort de des objectifs de ce projet de fin d'études surtout avec la contrainte temps.

Sur ces cartes, les courbes isométriques prennent quelques fois, à certains endroits et pour certaines profondeurs des formes insolites. Se sont sans doute encore les inclusions de sables et de graviers qui sont à l'origine de ce phénomène. En effet, la reconnaissance géotechnique avait montré leur présence irrégulière tant horizontalement que verticalement.

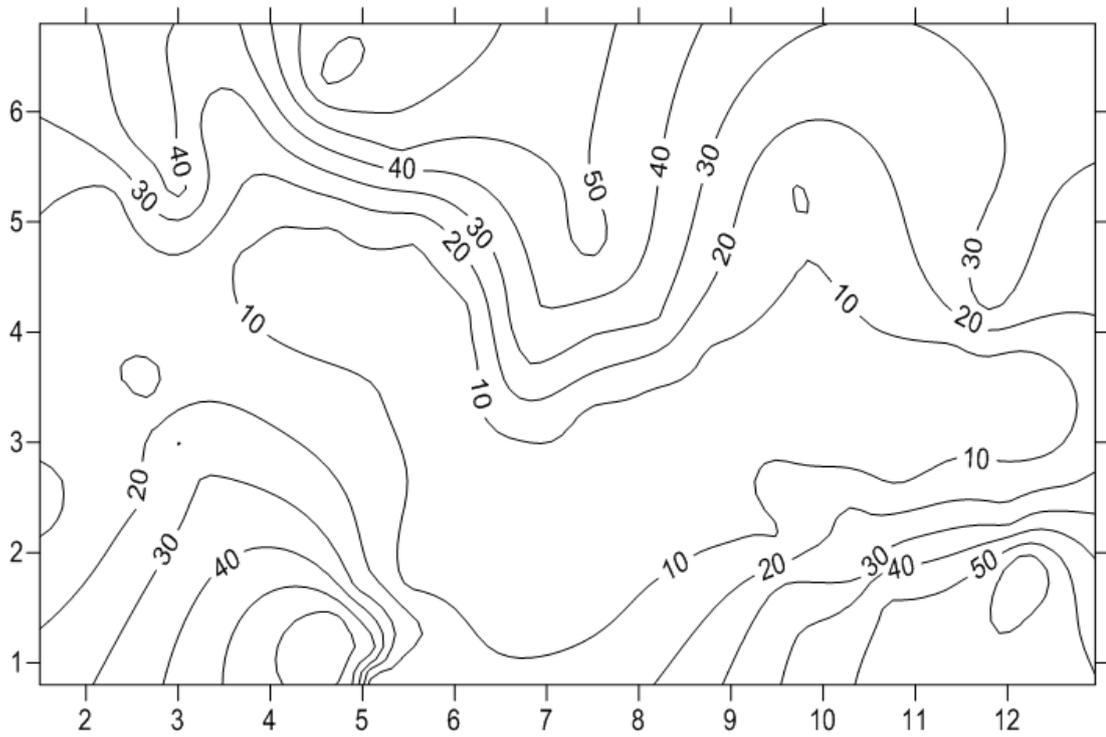


Figure 3 172: Carte de courbes isométriques de Rp à z=5 m

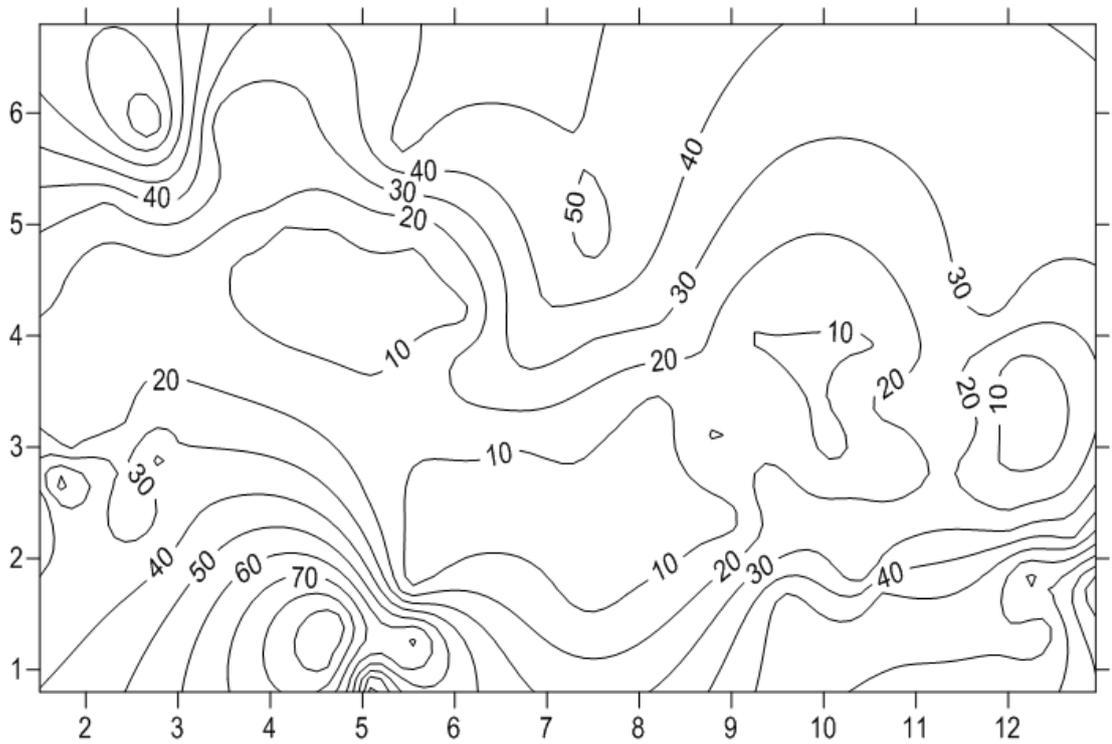


Figure 3 183: Carte de courbes isométriques de Rp à z=7,5 m

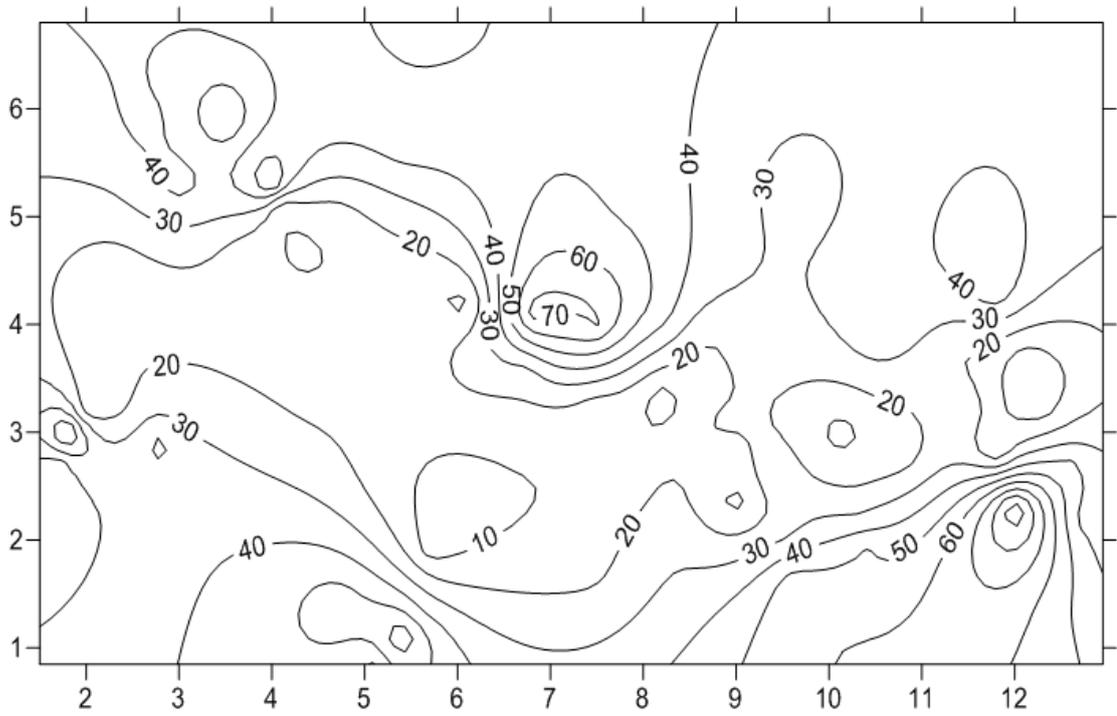


Figure 1.14: Carte de courbes isométriques de Rp à z=10 m.

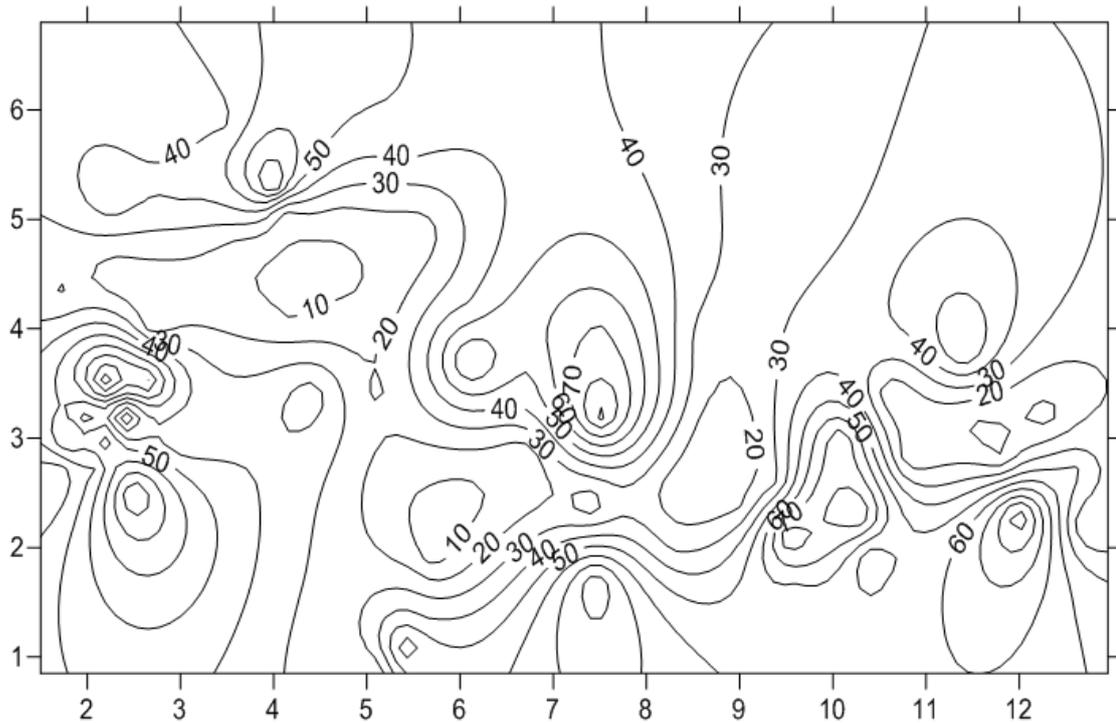


Figure 3 5: Carte de courbes isométriques de Rp à z=12,5 m.

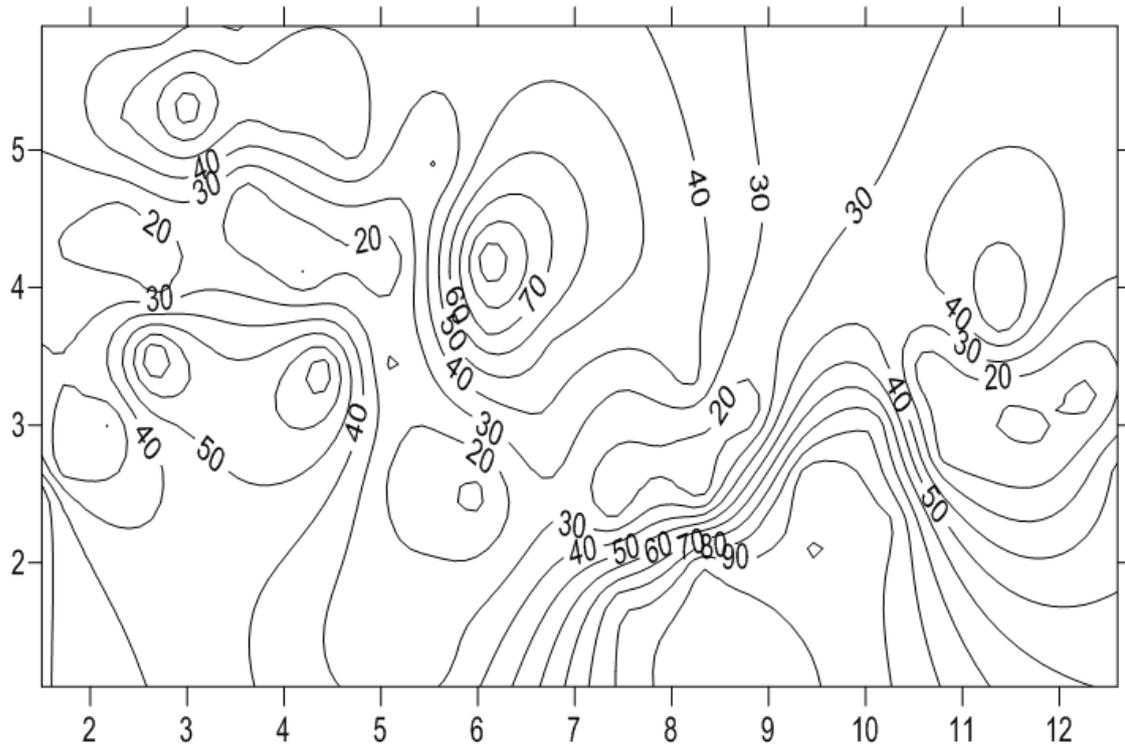


Figure 3 16 : Carte de courbes isométriques de Rp à z=15 m.

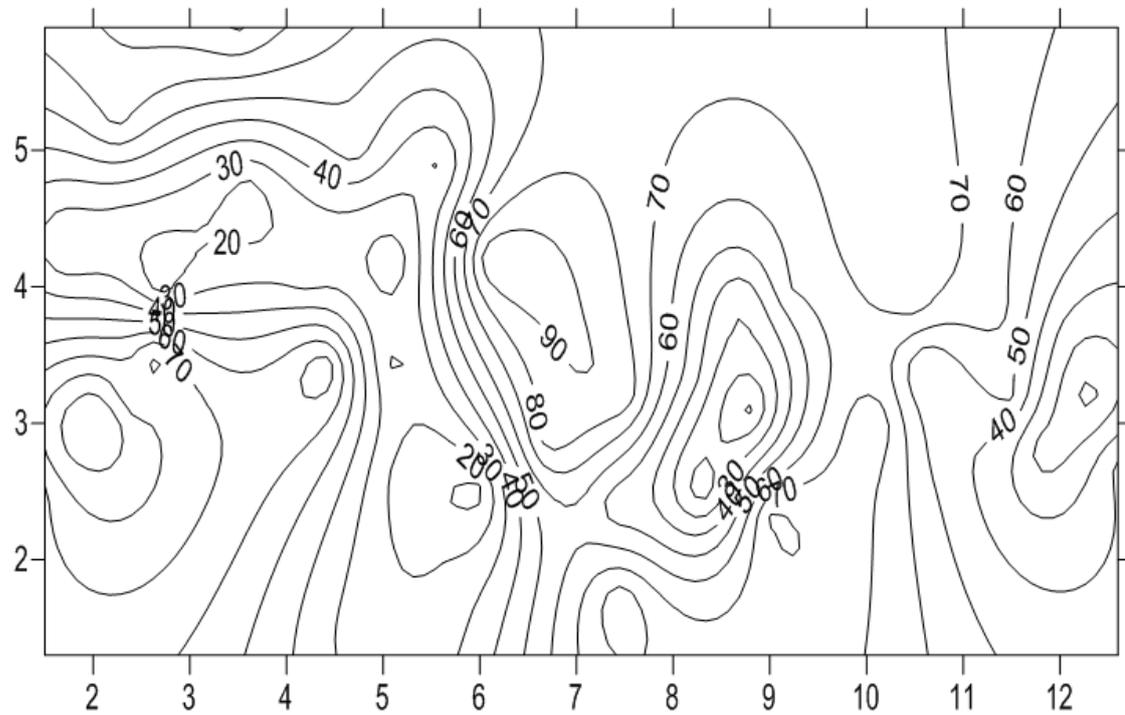


Figure 3 17: Carte de courbes isométriques de Rp à z=17,5 m.

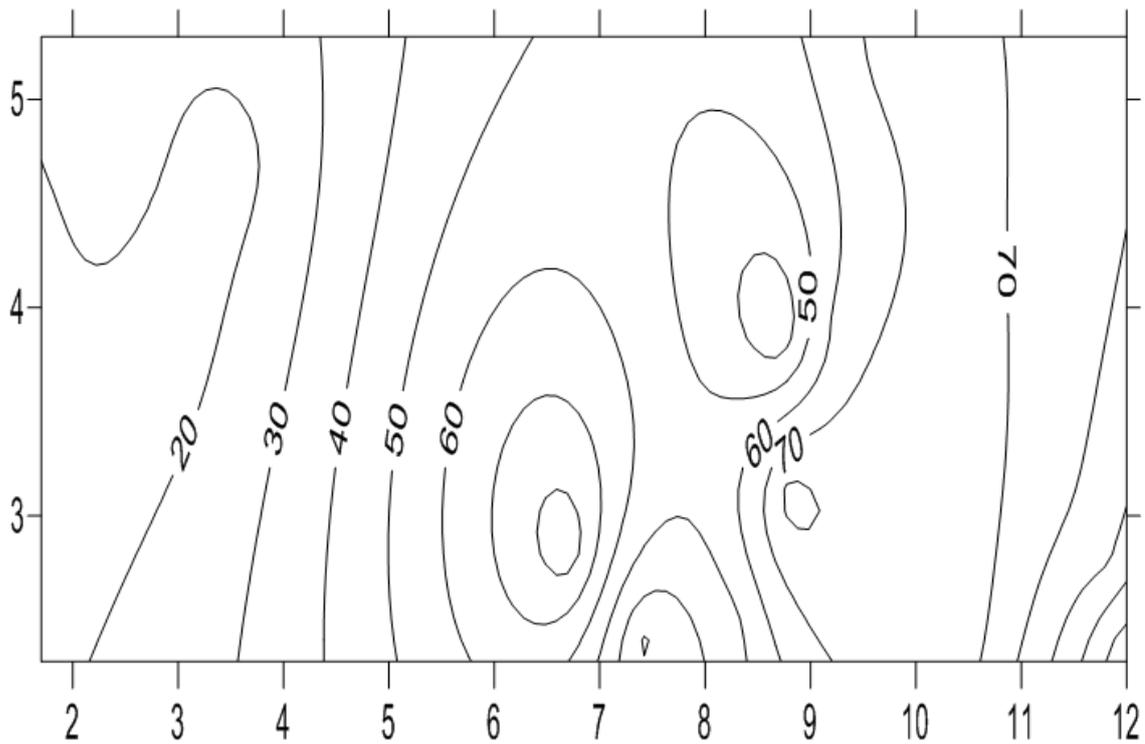


Figure 3 198: Carte de courbes isométriques de R_p à $z=20$ m.

La principale remarque à faire sur cette étude, est que l'analyse factorielle est ici un moyen et non une fin en soi. En effet, la représentation factorielle n'est à l'évidence pas le meilleur "outil" pour établir des cartes de courbes isométriques et il existe certainement des logiciels, tels que ceux de topographie par exemple, qui seraient beaucoup mieux adaptés à ce genre d'étude.

Toutefois, malgré que nous avons essayé d'exploiter en maximum les données que nous possédions, il reste un grand travail pour affiner l'étude des matériaux compressibles.

Enfin, nous désirons montrer l'avantage de l'ACP ou des méthodes dites de représentation factorielle qui permettent de mettre rapidement en évidence les classes d'individus lorsque l'on est en présence de peu de variable. Ce type d'exploitation des plans factoriels ne présente peut être pas, a priori, dans la pratique beaucoup d'avantage mais d'un point de vue statistique, elle pourrait sans doute être améliorée notamment en raisonnant non plus sur les valeurs mesurées mais sur ces dernières modifiées. (Prise en compte de la moyenne où de l'écart type).

Conclusion :

En conclusion l'étude par l'analyse en composante principale des données pénétrométriques de l'aéroport Houari Boumédiène d'Alger a atteint l'objectif fixé même partiellement malgré la difficulté de l'étude.

En effet, du fait de l'hétérogénéité des sols, nous avons pas pu valider la méthode de mise en évidence des couches puisque nous avons trouvé de résultats légèrement semblables à ceux établis par Magnan.

De plus, en exploitant la grande quantité de sondages, nous avons pu montrer un autre intérêt de l'ACP appliquée à des données géotechniques : la représentation issue de l'ACP met en évidence les classes d'individus et facilite ainsi l'établissement de cartes de courbes isométriques.

Toutefois, sauf erreur de notre part, nous pouvons reprocher à ces méthodes d'exploitation de l'ACP de laisser encore une bonne part à la subjectivité et de nécessiter un travail fastidieux.

CONCLUSION GENERALE

Outre l'analyse théorique et paramétrique menée dans les études de sites et d'ouvrages, les méthodes statistiques et probabilistes ont un apport considérable dans la compréhension de la géotechnique. Cependant et dans l'état actuel des connaissances ces approches « classiques » paraissent assez limitées dans la pratique et n'augmentent pas de manière significatives la fiabilité des études géotechniques.

Néanmoins, ces travaux ont permis de prendre conscience de l'importance de la variabilité des sols dans le calcul des ouvrages. Et, montrent la nécessité d'approfondir les études sur la répartition de la sécurité des ouvrages par l'introduction de méthodes modernes d'analyse de données afin de prendre les paramètres les plus pertinents sur la sécurité de l'ouvrage.

Dans la partie bibliographique l'analyse en composante principale permet de traduire un ensemble plus ou moins important de variables au moyen de quelques composantes (le plus souvent deux suffisent) qui sont des combinaisons des variables initiales et de représenter tous les individus relativement à ces composantes. Nous avons appliqué l'ACP à un cas réel de banque de données.

L'analyse factorielle dont fait partie l'analyse en composante principale « ACP », de la famille de la statistique multi-variée, est utilisée pour décrire la variabilité entre des variables observées au moyen de variables dissimulée. Elle permet ainsi de comprendre les interactions d'un grand nombre de paramètres et de déterminer les plus pertinents pour le dimensionnement de la structure.

Nous avons utilisé d'une manière générale l'ACP pour l'exploitation des données géotechniques. L'AFC n'étant pas adaptée à ce type de données, son application nécessite de transformer le tableau de données en un tableau de contingences.

Ensuite, nous avons exploité de données provenant de nombreux sondages sur le site de l'aéroport d'Alger. Ce site étant connu hétérogène, notre objectif était de valider la méthode de détermination des couches au moyen de l'ACP. Nous avons effectivement pu montrer l'existence des couches de sols compressibles que Magnan avait déterminée lors d'une expertise antérieure.

Nous avons aussi utilisé l'ACP pour établir des cartes de courbes isométriques de résistance de pointe afin d'élargir notre connaissance du terrain. Par ailleurs, cette dernière utilisation de l'ACP confirme que cette technique permet d'exploiter des données géotechniques de nombreuses façons.

Toutefois, si l'ACP paraît pouvoir servir dans les études de mécanique des sols, les démarches que nous avons observées pour traiter les données pénétrométriques demanderaient à être améliorées. En effet, non seulement ces études ont fait appel à une bonne part de subjectivité mais de plus, elles ont nécessité un travail fastidieux peu intéressant pour un géotechnicien.

Par ailleurs, nous pouvons nous demander si les techniques de classification hiérarchique de l'analyse de données ne conviendraient pas mieux que l'analyse factorielle pour ce type de recherche et plus spécialement pour la stratigraphie.

L'ACP utilisé dans cette étude a permis d'atténuer les incertitudes qui entachent le dimensionnement de l'ouvrage, ces causes sont dues aux spécificités des sols qui sont déterminées par une succession de couches à la fois irrégulières et hétérogènes à l'échelle spatiale.

L'application des ces deux outils de travail permettent de définir en tout points de l'espace la résistance de pointe du sol plus au moins exacte ce qui permet d'avoir une idée plus réelle de la variabilité du sol et détermine les coefficient de sécurité plus au moins proche de l'optimum, mais il reste à savoir si la prise en compte de cette variabilité influe réellement sur les coefficients de sécurité dans un sens plus représentatif de la réalité du comportement des massifs de sols naturels ?

Annexe A : ESSAIS EN PLACE

Les différentes couches de sol sont déterminées par des cartes d'isométrie qui définissent la même condensation de propriétés géotechniques. C'est propriétés sont définis par des essais in situ ou au laboratoire.

Les essais au laboratoire cumulent des erreurs lors de l'interprétation des résultats obtenus. Ils présentent, aussi, d'autres inconvénients tels que:

- ❖ Le remaniement des éprouvettes change les caractéristiques internes du sol,
- ❖ Nécessite la réalisation de plusieurs essais pour connaître une particularité du sol en un endroit donné.
- ❖ Le sol peut changer de caractéristiques chimique et physique au cours du temps.

C'est pour cette raison, les géotechniciens ont tendance à réaliser les essais sur place. Dans ce cas, les estimations sont plus réalistes et caractérisent efficacement le sol. Par ailleurs, les essais in-situ sont de plus en plus employés pour les avantages qu'ils offrent, en particulier:

- ❖ La plupart se déroulent de manière rapide et économique,
- ❖ Les échantillons subissent moins de remaniements,
- ❖ Ils donnent une évaluation réaliste des caractéristiques du massif de matériaux discontinus (comme les argiles raides fissurées, roches fissurées,...),
- ❖ Les caractéristiques internes du sol restent constantes,
- ❖ Ils sont faciles à mettre en œuvre et permettent de prendre en compte la variabilité spatiale du terrain.

De plus, les essais in-situ donnent des informations permettant la reconnaissance des sols ainsi que leurs caractéristiques. Ils se divisent en deux grandes familles. La première famille sollicite le sol au repos ou à l'état limite de service. Elle donne des courbes caractéristiques (contrainte en fonction de la déformation) où l'exploitation de ces courbes détermine les paramètres du sol. La deuxième famille donne les caractéristiques du sol à l'état limite ultime qui est la rupture.

A.1. LES ESSAIS EN PLACE :

A.1.1. Essai de tassement :

A.1.1.1. Essai de chargement à la plaque :

A.1.1.2. Essai tassométrique hydraulique :

A.1.1.3. Essai tassométrique magnétique :

A.1.2. Essais d'eau :

A.1.2.1. L'essai de pompage :

A.1.2.2. L'essai LEFRANC :

A.1.2.3. L'essai LUGEON :

A.1.2.4. Essai NASBERG :

A.1.2.5. Essai au Permafrost :

A.1.3. Les essais au scissomètre de chantier :

A.1.4. Les essais au phicomètre :

A.1.5. Essais de pressiomètre:

A.1.6. Essais pénétrométriques :

Ils se divisent en essais statiques ou dynamiques. Ils donnent une caractéristique du sol à l'état limite ultime ou de rupture donc il appartient à la deuxième famille d'essai.

Ce paragraphe donne un détail sur cet essai puisqu'il s'agit de l'essai à exploiter dans notre étude.

Les caractéristiques de déformation ne sont déterminées que par corrélation. L'essai consiste à enfoncer un train de tige métallique par une force constante (vérinage) qui donne une vitesse de pénétration ou par battage (choc répéter). La résistance de pointe est déduite de la contrainte de rupture du sol.

Il existe trois essais pénétrométriques qui sont:

- Essai par pénétromètre statique ou quasi statique,
- Essai par SPT (standard pénétration test) ou pénétration par carottier,
- Essai par pénétromètre dynamique.

A.1.6.1. Pénétromètre statique :

Il existe plusieurs types de pénétromètres statiques qui sont le Pénétromètre GOUDA, le pénétromètre ANDINA et le pénétromètre statique normalisé. Le pénétromètre statique normalisé s'effectue par procédé mécanique, électrique ou statique. Il a pour principe d'enfoncer dans le sol un train de tige qui contient à son extrémité inférieure une pointe en acier de 10 cm² de section. Le fonçage s'effectue par pas de profondeur et par une vitesse lente et constante de 2 cm/s. La figure suivante schématiser par (Samuel. Et al., 1980) présente l'appareillage d'un pénétromètre statique.

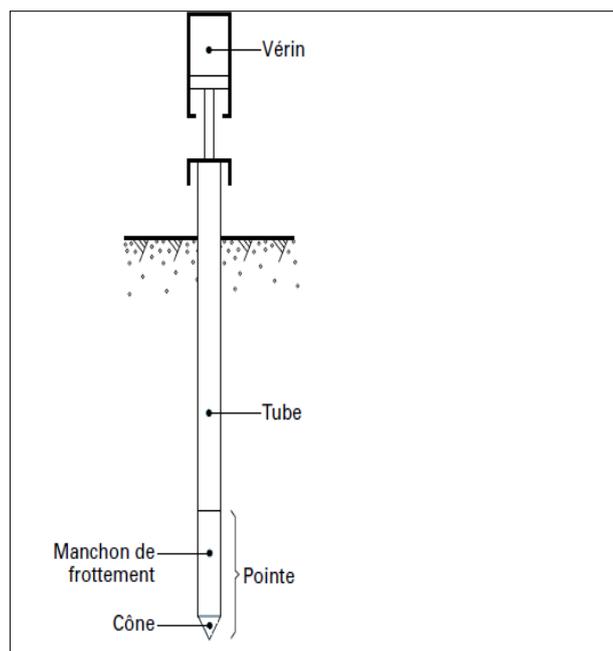


Figure A.1: Appareillage d'un pénétromètre statique.

La résistance opposée par le sol est mesurée soit par sondage continu, notée q_c en continu, ou par sondage discontinu notée q_t , avec :

$$q_t = q_c + q_s, \quad (A-1)$$

ou q_s est la résistance de frottement latéral.

La résistance de pointe statique est donnée par:

$$q_c = \frac{Q_c}{A_c} (Mpa) \quad (A-2)$$

$$Q_s = Q_t - Q_c (MN) \quad (A-3)$$

Q_c , Q_t et Q_s , sont respectivement l'effort de pointe, l'effort total et le frottement latéral en (MN), et A_c est la section de la pointe.

a) Interprétation des résultats:

A partir de Q_c et Q_t , on définit par des corrélations la nature des sols et leurs hétérogénéités (les coupes géologiques). Le niveau d'assise des fondations est déduit de la capacité portante du sol k_c , selon l'équation:

$$q_l = q_0 + k_c(q_c - q_0) \quad (A-4)$$

q_l : est la résistance limite du terrain sous la base de la fondation, semelle ou pieu,
 q_0 : pression totale verticale due au poids des terres.

Aussi on peut tirer la résistance au cisaillement du sol par la cohésion non drainée c_{uu} et l'angle de frottement φ par l'équation:

$$c_{uu} = \frac{q_c - q_0}{N_c} \quad (A-5)$$

N_c : Coefficient sans dimension dépend de φ

Pour la détermination de φ des sols pulvérulents, de nombreux auteurs ont utilisé la formule:

$$q_c = \gamma \cdot z \cdot N_q + q_0 \quad (A-6)$$

γ : poids volumique du sol,

N_q : le facteur de portance de profondeur,

z : Profondeur.

A.1.6.2. Pénétrromètre SPT:

Le pénétromètre au carottier, s'effectue par battage d'un carottier normalisé. Il détermine par corrélation la cohésion, l'angle de frottement, l'indice de densité ainsi que la nature du sol définie par N'_{70} . Ce dernier exprime le nombre de coups obtenus

avec une machine restituant 70% de son énergie théorique. Dans le cas de sols non cohérent, N_{70} peut être corrélé avec la densité du sol. Par contre, dans le cas de sols cohérent, N_{70} est corrélé avec la cohésion et l'angle de frottement.

Lors de l'essai on compte le nombre de coups pour enfoncer le carottier sur une certaine profondeur. Une fois remplis, on vide le dispositif et on recommence le travail. En suite, on trace un profil de pénétration qui se présente sous la courbe du nombre de coups en fonction de la profondeur.

Selon la norme NFP 94-116 l'essai ne s'applique qu'aux sols dont la granulométrie est inférieure à 20 mm et dont la valeur N est comprise entre 0 et 50. Le fond du sondage ne doit pas être trop remanié par la succession des opérations qui entraînent des affaissements ou gonflement dans les sols en général placés sous la nappe et pour des profondeurs de sondage ne dépassant pas 50 m.

La norme exige aussi, que le carottier placé au préalable est battu par l'intermédiaire d'un train à l'aide d'un marteau de 63,5 kg. Le marteau tombe en chute libre d'une hauteur de 76 cm sur la tête du train de tiges. Il représente une énergie potentielle de $E=474,5$ j. La profondeur de chaque passe de pénétration est de 45 cm. Le schéma ci dessous inspiré du document de (Samuel. Et al., 1994) explique l'appareillage du SPT. (Figure ci-dessous).

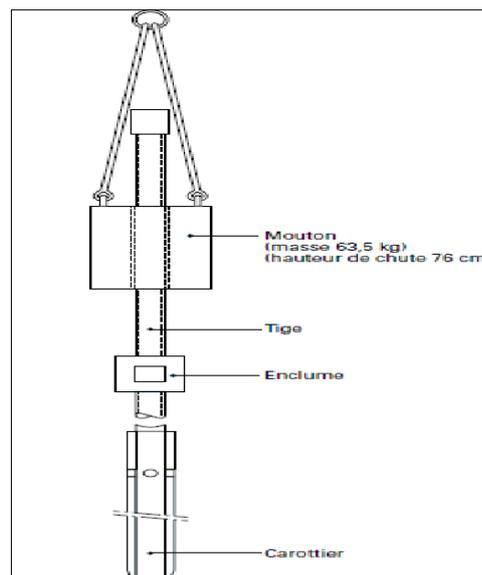


Figure A.2: Appareillage du SPT.

Les conditions de battage doivent être respectées pour éviter la dispersion des mesures. Pendant le battage, il-y a l'enfoncement de mise en place sous le poids propre de l'appareil (la pénétration est continuée aussitôt si le premier enfoncement dépasse 45 cm). L'enfoncement de départ fait pénétrer le carottier de 15 cm sous l'effet de N_0 coups de mouton. Enfin, l'enfoncement d'essai, où le carottier est enfoncé de 15 cm deux fois sous l'effet de N_1 Puis N_2 coups de mouton. Le nombre $N = N_1 + N_2$ est appelé résistance à la pénétration.

A.1.6.3. Pénétrromètre dynamique :

Le pénétrromètre dynamique consiste à enfoncer sous l'effet de chocs répétés un train de tige lisse dans le sol. Le battage s'effectue sur l'extrémité supérieure du dispositif par une masse (mouton de battage) qui tombe en chute libre d'une hauteur H constante.

Dans l'extrémité inférieure de l'appareil, est placée une pointe de forme conique et de diamètre supérieur à celui de la sonde. Elle crée un vide entre le train de tige et le sol. Le dispositif est enfoncé jusqu'à une profondeur donnée, en général limitée par l'aptitude de pénétration de l'appareil.

La figure suivante présente un schéma descriptif de l'appareillage du pénétrromètre dynamique (figure A.3):

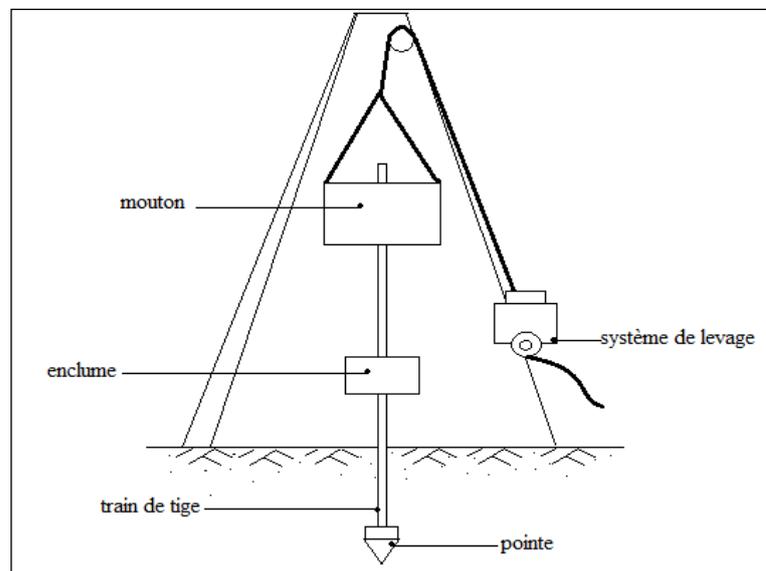


Figure A.3: Appareillage d'un pénétrromètre dynamique.

On trace le profil de résistance du sol correspondant en fonction de la profondeur atteinte par la pointe, avec une courbe dont l'axe des ordonnées peut être représenté par le nombre de coups N_d pour enfoncer l'appareil sur un pas de profondeur, ou par la résistance dynamique de la pointe notée q_d .

$$q_d = \frac{M}{e \cdot (M + M')} \cdot \frac{MgH}{A} \quad (\text{A-7})$$

Avec:

M : est le poids du mouton,

M' : est le poids des parties frappées (enclume, train de tige, pointe),

A : est la section droite de la pointe,

g : est l'accélération de gravité,

e : est l'enfoncement moyen par coup.

« e » est donné par la formule suivante:

$$e = \frac{H}{N_d} \quad (A-8)$$

Trois limitations doivent être averties:

- Limiter la profondeur d'étude par le pénétromètre à 30 m maximum. Le flambement des tiges sous l'effet des coups répétés provoque le frottement entre le sol et la tige et cause une surestimation de la résistance du sol,
- Ces essais ne s'effectuent pas dans des sols grenus dont la dimension des éléments est supérieure à 60 mm,
- Il n'existe pas de règle pour le dimensionnement des fondations à partir de la résistance dynamique q_d .

Il existe deux types de pénétromètre dynamique décrits dans les normes, NF P 94-114 et NF P 94-115 qui sont:

A) Pénétrromètres dynamiques de type A :

Ces appareils sont nommés dans les différentes littératures par les PDA, dont l'énergie de battage, la nature, la géométrie et les dimensions de la pointe sont normalisées. Ils ont en plus un dispositif qui sert à injecter petit à petit une boue bentonitique dans le vide qui se trouve entre le train de tige et le sol. La boue protège la tige du sol qui peut se gonfler ou s'ébouler. Elle assure, aussi, une transmission directe de l'énergie de battage à la pointe (contacte entre la tige et le sol est limité). Le Schéma suivant décrit le pénétromètre type A.4.

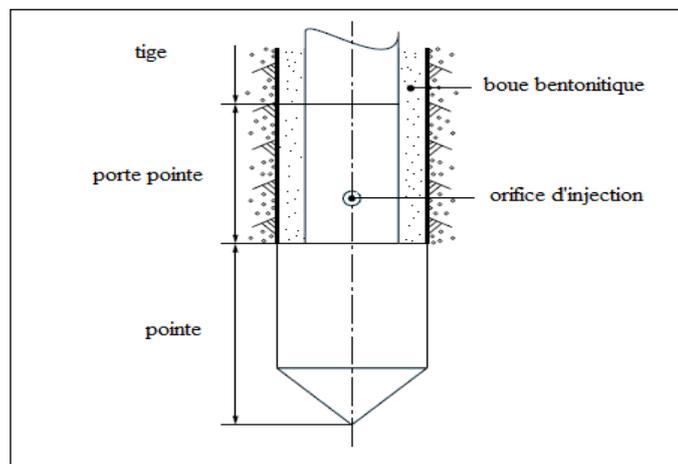


Figure A.4: Pénétrromètre type A.

B) Pénétrromètres dynamiques de type B :

Généralement nommés PDB, ces pénétromètres sont semblables à ceux du type A, à l'exception du fait qu'ils ne comportent pas de dispositif pour injecter la boue bentonitique. Ils présentent plusieurs inconvénients par rapport au PDA. Une partie de l'énergie de battage peut être mobilisée par le frottement latéral qui se produit entre le sol et le train de tiges. Dans certains cas de sol cohérent, l'appareil ne permet pas de

distinguer entre les couches de sol traversées. De plus, son aptitude de pénétration est plus limitée que celle de l'appareil PDA.

Le pénétromètre dynamique est utilisé pour donner une estimation qualitative du terrain étudié. Il réalise plusieurs tâches, qu'on cite si après:

- ❖ détermine l'épaisseur des différentes couches du terrain,
- ❖ localise les cavités et les discontinuités,
- ❖ reconnaît le niveau du toit du sol rocheux,
- ❖ donne des renseignements utiles pour provisionner les conditions de battage des pieux et des palplanches (Samuel. Et al., 1994).

Annexe B : DONNEES PENETROMETRIQUES DE L'AEROPORT D'ALGER.

	zone	n°	z(m)	Rp moy	Rp (bars) sigma ²	Fl moy	Fl (bars) sigma ²
sondage 1	1	1 à 3	1 à 2	43	175	3,37	0,01
	2	4 à 6	2,5 à 3,5	10,7	8,3	0,6	0
	3	7 à 9	6,5 à 7,5	35,3	25,3	2,27	0,97
	4	10 à 15	8 à 12	47	100,4	3,08	1,1
				36,6	26,5	2,48	1,62
sondage 2	1	16 à 19	0,5 à 2	15,3	0,9	1,15	0,02
	2	20 à 30	2,5 à 7,5	2,6	2,1	0,18	0,01
	3	31 à 34	8 à 9,5	15,5	1	0,8	0,01
	4	35 à 42	10 à 13,5	2,8	4,2	0,2	0,02
	5	43 à 53	14 à 19	23,4	660	1,2	1,4
				11,3	266	0,65	0,61
sondage 3	1	54 à 59	0,5 à 3	12,2	8,6	0,8	0,2
	2	60 à 68	3,5 à 7,5	2,1	0,9	0,14	0,01
	3	69 à 72	8 à 9,5	16	2,7	1,2	0,07
	4	73 à 91	10 à 19	5,8	12,3	0,33	0,08
				7	27,2	0,45	0,18
sondage 4	1	92 à 107	0,5 à 8	16,1	87,5	1,01	0,68
	2	108 à 131	8,5 à 20	42,6	665,9	2,16	1,74
				32	599,5	1,7	1,62
sondage 5	1	132 à 140	0,5 à 4,5	17,4	54,8	0,91	0,13
	2	141 à 152	5 à 10,5	37,3	37,2	2,25	0,24
	3	153 à 155	11 à 13	17,7	100,3	2,84	3,66
				27,4	148,2	1,83	1,04
sondage 6	1	156 à 158	0,5 à 1,5	15,3	4,3	1,07	0,06
	2	159 à 169	2 à 7	42,2	407,4	2,42	0,89
	3	170 à 172	7,5 à 8,5	100	0	4	1
	4	173 à 180	9 à 12,5	39,1	120,7	2,04	0,21
	5	181 à 183	13 à 14	92,3	134,3	4,07	0,9

				50	862,8	2,5	1,35
sondage 7	1	184 à 185	0,5 à 1,5	26,5	40,5	1,25	0,35
	2	186 à 211	1,5 à 14	54,1	404	2,65	0,76
				52,2	428,2	2,55	0,84
sondage 8	1	212 à 215	0,5 à 2	23,5	95	1,05	0,14
	2	216 à 217	2,5 à 3	39	0	3	0
	3	218 à 229	3,5 à 9	73,1	589,5	3,2	1,58
				58,3	881,3	2,72	1,9
sondage 9	1	230 à 251	0,5 à 11	24,9	135,3	1,74	0,43
	2	252 à 255	11,5 à 13	100	0	4,5	0,34
				36,5	877,1	0,16	1,44
sondage 10	1	256 à 260	0,5 à 2,5	20,4	92,3	1,16	0,08
	2	261 à 286	3 à 15,5	10,9	36,5	0,53	0,15
	3	287 à 289	16 à 17	39,3	137,3	2,17	1,33
	4	290 à 291	17,5 à 18	14,5	0,5	0,95	0,01
	5	292 à 295	18,5 à 20	53,8	110,9	2,45	0,23
				18,7	248,8	0,94	0,66
sondage 11	1	296 à 300	0,5 à 2,5	32	55,5	2,28	0,78
	2	301 à 309	3 à 7	7,1	25,1	0,27	0,03
	3	310 à 315	7,5 à 10	27,8	25,8	1,43	0,09
	4	316 à 326	10,5 à 15,5	10,1	39,1	0,37	0,07
	5	327 à 335	16 à 20	52,4	331,8	2,49	0,97
				24,4	409,7	1,22	1,25
sondage 12	1	336 à 339	0,5 à 2	19,5	17	2,35	0,16
	2	340 à 349	2,5 à 7	6,7	4,7	0,24	0,03
	3	350 à 355	7,5 à 10	20,7	61,1	1,07	0,07
	4	356 à 372	10,5 à 18,5	62,7	737,2	3,15	1,36
				36,1	983,4	1,94	2,28
sondage 16	1	373 à 377	0,5 à 2,5	20,8	41,2	1,82	0,45
	2	378 à 385	3 à 6,5	3,8	5,4	0,14	0,014

	3	386 à 403	7 à 15,5	14,3	33,5	0,76	0,27
	4	404 à 411	16 à 19,5	43,1	581,5	1,61	0,89
				20,9	466,6	1,04	1,07
sondage 17	1	413 à 418	0,5 à 3	22,5	66,3	1,1	0,22
	2	419 à 426	3,5 à 7	6,4	1,7	0,2	0,02
	3	427 à 432	7,5 à 10	23	132,8	1,28	0,29
	4	433 à 444	10,5 à 16	11,7	18,1	0,46	0,04
	5	445 à 448	16,5 à 19	37,3	52,9	1,8	0,23
				17	132,8	0,79	0,4
sondage 18	1	449 à 453	0,5 à 2,5	28,4	80,8	3,4	2,62
	2	454 à 463	3 à 7,5	9	36,4	0,5	0,26
	3	464 à 481	8 à 16,5	37,6	172,6	2,1	0,29
	4	482 à 486	17 à 19	76,6	99,8	3,7	0,29
				34	532,2	2,01	1,66
sondage 19	1	487 à 490	0,5 à 2	25,5	83,7	1,4	1,14
	2	491 à 497	2,5 à 5,5	48	497	2,24	2,24
	3	498 à 512	6 à 13	51,1	539,6	2,11	0,39
				46,3	515,2	2,04	0,49
sondage 20	1	513 à 514	0,5 à 1	24	50	1,4	0,08
	2	515 à 516	1,5 à 2	79	98	2,35	3,65
	3	517 à 527	2,5 à 8	19,3	83	0,98	0,33
	4	528 à 534	8,5 à 11,5	50,9	139,5	2,67	0,37
	5	535 à 538	12 à 14	40	26,7	1,83	0,1
				35,9	422,7	1,7	0,92
sondage 21	1	540 à 549	0,5 à 5	23,8	106,8	1,38	0,54
	2	550 à 556	5,5 à 8,5	27,1	155,1	1,44	0,55
	3	557 à 558	9 à 9,5	100	/	3,25	0,61
	4	559 à 561	10 à 11	26,7	154,3	1,17	0,22
	5	562 à 573	11,5 à 17	90,6	170,4	3,02	1,1
				52,8	1234,2	2,06	1,36
sondage	1	574 à 577	0,5 à 2	23,75	168,3	0,88	0,12

22	2	578 à 593	2,5 à 10	6,4	5,6	0,31	0,01
	3	594 à 597	10,5 à 12	27,3	38,3	1,25	0,04
	4	598 à 613	12,5 à 20	78,4	527,5	3,04	1
				39	1331,6	1,55	2,01
sondage 23	1	614 à 616	0,5 à 1,5	13,3	9,3	0,53	0,02
	2	617 à 625	2 à 6	36,4	107	1,63	0,39
	3	626 à 640	6,5 à 13,5	55,9	261,7	2,81	0,796
				44,7	383,6	2,16	1,17
sondage 24	1	641 à 645	0,5 à 2,5	18,4	18,3	1,74	0,23
	2	646 à 651	3 à 5,5	9,5	14,3	0,42	0,05
	3	652 à 663	6 à 12	25	38	1,37	0,15
	4	664 à 671	13 à 16,5	88	540,3	3,25	1,73
				37,2	1104,9	1,73	1,51
sondage 25	1	672 à 676	0,5 à 2,5	17	4	1,26	0,19
	2	677 à 687	3 à 8	5,6	2,25	0,15	0,01
	3	688 à 692	8,5 à 10,5	17,4	7,3	1,02	0,05
	4	693 à 709	11 à 19	88,3	342,7	3,48	1,71
				45,7	1680,3	1,9	2,98
sondage 26	1	710 à 711	0,5 à 1	71,5	1624,5	2,35	0,13
	2	712 à 719	1,5 à 5	12,4	7,98	0,85	0,11
	3	720 à 726	5,5 à 8,5	25,6	12,6	1,77	0,03
	4	727 à 729	9 à 10	27	75	1,07	0,01
	5	730 à 731	10,5 à 11	67,5	312,5	1,95	0,41
	6	732 à 737	11,5 à 14	28,7	73,1	1,32	0,18
	7	738 à 739	14,5 à 15	100	/	3,6	0,32
				33,6	720,9	1,54	0,62
sondage 27	1	741 à 744	0,5 à 1,5	33,8	74,9	1,43	0,04
	2	745 à 757	2 à 9	8,1	18,7	0,83	0,07
	3	758 à 767	9,5 à 14	27,9	452,3	1,42	0,27
	4	768 à 772	14,5 à 16,5	100	/	4,14	0,49
				31,8	105,9	1,43	1,82

sondage 28	1	773 à 778	0,5 à 3	20	229,2	1,25	31,1
	2	779 à 792	3,5 à 10	6,4	5,3	0,16	54,2
	3	793 à 797	10,5 à 12,5	19,8	28,2	1,22	29,8
	4	798 à 807	13 à 17,5	81,9	337	3,95	22
				32,2	1178,6	1,58	105,2
sondage 30	1	809 à 811	1 à 2	25,3	17,3	1,2	0,01
	2	812 à 823	2,5 à 8	10	7,5	0,37	0,03
	3	824 à 841	8,5 à 17	12,2	15,3	0,54	0,07
	4	842 à 846	17,5 à 19,5	43,4	744,8	2,2	2,48
				18,8	385,5	0,86	1,08
sondage 31	1	847 à 852	0,5 à 3	18,5	24,3	1,03	0,05
	2	853 à 861	3,5 à 7,5	6,2	4,9	0,19	0,01
	3	862 à 867	8 à 10,5	21,8	17,4	1,37	0,24
	4	868 à 881	11 à 17,5	13	7,2	0,49	0,02
	5	883 à 886	18 à 20	28,5	319	1,53	0,68
				17,4	259,9	0,85	0,78
sondage 32	1	887 à 890	0,5 à 2	18,8	20,9	1	0,09
	2	891 à 899	2,5 à 6,5	7,3	4,8	0,19	0,01
	3	900 à 906	7 à 10	21,1	121	1,43	0,16
	4	907 à 917	10,5 à 15,5	10,1	10,9	0,34	0,01
	5	918 à 920	16 à 17	43,3	1304,3	2,03	0,2
	6	921 à 923	17,5 à 18,5	7,7	0,3	0,3	0,01
	7	924 à 926	19 à 20	48,3	225,3	1,93	0,66
				17,5	259,5	0,82	0,54
sondage 33	1	928 à 932	1 à 3	15,4	6,8	1,22	0,45
	2	933 à 941	3,5 à 7,5	7	4,3	0,23	0,03
	3	942 à 953	8 à 13,5	12,2	9,1	0,58	0,09
	4	954 à 958	14 à 16	37,8	348,2	1,68	0,5
	5	959 à 966	16,5 à 20	84	318,3	3,68	0,43
				31,8	1086,8	1,45	1,99

sondage 34	1	967 à 970	0,5 à 2	15,5	1	0,8	0,33
	2	971 à 979	2,5 à 6,5	5,8	3,2	0,12	0,01
	3	980 à 986	7 à 10	12,4	9	0,5	0,05
	4	988 à 993	11 à 13,5	94,8	160,2	3,07	0,43
	5	995 à 998	14,5 à 16	27,8	74,9	1,28	0,11
	6	999 à 1006	16,5 à 20	87,5	265,7	4,25	0,19
				41,4	1489,9	1,78	2,18
sondage 35	1	1007 à 1010	0,5 à 2	20	50,7	1,5	0,79
	2	1011 à 1026	2,5 à 10	6,3	9,6	0,29	0,03
	3	1027 à 1045	10,5 à 19,5	44,8	668,4	2,34	1,69
				26,5	668,6	1,41	1,84
sondage 36	1	1047 à 1048	1 à 1,5	19	2	1,35	0,13
	2	1049 à 1070	2 à 12,5	5,6	12,3	0,29	0,02
	3	1071 à 1080	13 à 17,5	23,2	84,4	0,96	0,14
	4	1081 à 1085	18 à 20	55	99	2,64	1,45
				19,2	476,6	0,9	1,08
sondage 37	1	1086 à 1089	0,5 à 2	35	112,7	2,45	0,09
	2	1090 à 1093	2,5 à 4	16,3	0,25	0,9	0,13
	3	1094 à 1120	4,5 à 17,5	6,6	6,3	0,21	0,01
	4	1121 à 1124	18 à 19,5	12,8	20,9	0,48	0,04
				11,2	92,3	0,54	0,5
sondage 38	1	1125 à 1129	0,5 à 2,5	22	10	1,94	0,96
	2	1130 à 1137	3 à 6,5	4,4	12,6	0,14	0,01
	3	1138 à 1145	7 à 10,5	25,5	59,1	1,64	0,45
	4	1146 à 1157	11 à 16,5	73,2	808,2	2,93	1,68
				37,2	1119,5	1,79	1,98

sondage 39	1	1158 à 1162	0,5 à 2,5	18	13,5	1,24	0,73
	2	1163 à 1178	3 à 10,5	7,3	21,8	0,25	0,04
	3	1179 à 1186	11 à 14,5	41,1	119,3	1,74	0,21
	4	1187 à 1188	15 à 15,5	100	/	2,25	0,05
	5	1189 à 1192	16 à 17,5	99,5	1	4,2	0,35
				32,4	1180,4	1,3	1,78
sondage 40	1	1193 à 1198	0,5 à 3	39,5	169,9	1,87	0,38
	2	1199 à 1212	3,5 à 10	52,7	123,5	2,91	0,7
				48,8	167,8	2,6	0,82
sondage 41	1	1214 à 1217	0,5 à 2,5	42	482,7	0,58	0,15
	2	1218 à 1221	3 à 4,5	76,3	578,9	3,25	0,26
	3	1222 à 1229	5 à 8,5	38,9	23,8	2,26	0,11
	4	1230 à 1232	9 à 10	74,7	517,3	3,47	0,32
				55,4	638,5	2,33	1,18
sondage 42	1	1234 à 1241	1 à 4,5	17,4	67,7	0,83	0,64
	2	1242 à 1251	5 à 10	28,2	16,6	1,79	0,06
	3	1252 à 1263	10,5 à 16	49	416,2	2,38	0,24
				35,8	490,3	1,79	0,66
sondage 43	1	1264 à 1280	0,5 à 8,5	30,6	74,4	2,08	0,28
	2	1281 à 1288	9 à 12,5	71,1	224,4	3,63	0,3
				43,6	487,5	2,58	0,81
sondage 44	1	1289 à 1292	0,5 à 2	58,3	934,9	3,43	1,18
	2	1293 à 1303	2,5 à 7,5	30,3	718,6	1,03	0,34
	3	1304 à 1322	8 à 17	44,9	96,7	1,98	0,32
				41,7	438,1	1,84	0,92

sondage 45	1	1323 à 1326	0,5 à 2	27,5	171,7	2,13	0,48
	2	1327 à 1336	2,5 à 7	13,7	24,7	0,7	0,14
	3	1337 à 1342	7,5 à 10	16,3	14,7	0,48	0,01
	4	1343 à 1351	10,5 à 14,5	46,6	52,8	1,38	0,39
	5	1352 à 1360	15 à 19,5	79,3	310,8	3,22	0,88
				38,9	882,6	1,57	1,45
sondage 46	1	1361 à 1366	0,5 à 3,5	24,7	84,3	1,65	0,64
	2	1367 à 1376	4 à 8,5	10,9	5,7	0,38	0,02
	3	1377 à 1398	9 à 19,5	18	16,2	0,77	0,31
				17,2	42,1	0,81	0,43
sondage 47	1	1399 à 1403	1 à 3	21	12	1,66	0,12
	2	1404 à 1413	3,5 à 8	8	3,3	0,32	0,01
	3	1414 à 1417	8,5 à 10	13,8	0,3	0,68	0,02
	4	1418 à 1428	10,5 à 15,5	9,5	2,3	0,35	0,01
	5	1429 à 1434	16,5 à 19,5	22,3	101,5	1,23	0,49
				13,3	52,3	0,71	0,35
sondage 48	1	1435 à 1439	1 à 3	22	16	1,42	0,26
	2	1440 à 1448	3,5 à 7,5	7,8	7,7	0,31	0,01
	3	1449 à 1454	8 à 10,5	19,8	6,6	1,07	0,02
	4	1455 à 1469	11 à 19	12,9	8,6	0,57	0,02
				14,1	33,7	0,83	0,7
sondage 49	1	1471 à 1472	2,5 à 3	25,5	4,5	1,25	0,01
	2	1473 à 1482	3,5 à 8	11,4	9,6	0,41	0,03
	3	1483 à 1486	8,5 à 10	24,3	0,3	1,3	0,01
	4	1487 à	10,5 à 14	14,5	1,2	0,51	0,04

		1494					
	5	1495 à 1503	14,5 à 19	28,6	55,3	1,12	0,12
				19,2	71,8	0,79	0,19
sondage 50	1	1504 à 1510	0,5 à 3,5	19,6	10	0,36	0,08
	2	1511 à 1518	4 à 7,5	10,8	6,5	0,31	0,01
	3	1519 à 1529	8 à 13	14,2	7,6	0,71	0,08
	4	1530 à 1540	13,5 à 19,5	55,2	347	2,55	1,31
					26,6	463,6	1,29
sondage 51	1	1541 à 1548	0,5 à 4	31,9	28,7	1,6	0,66
	2	1549 à 1561	4,5 à 10,5	11,9	2	0,35	0,01
	3	1562 à 1566	11 à 13	19,6	5,3	1,04	0,12
	4	1567 à 1578	13,5 à 19	48,3	304,4	2,7	0,29
					28,6	334,25	1,44
sondage 52	1	1579 à 1588	0,5 à 5	16,5	40,1	0,56	0,07
	2	1589 à 1594	5,5 à 8	30,5	159,1	0,98	0,03
	3	1595 à 1599	8,5 à 10,5	61,6	100,3	3,16	0,35
	4	1600 à 1603	11 à 12,5	83,3	225,6	4,05	0,22
	5	1604 à 1608	13 à 15,5	60,6	15,3	3,1	0,03
					43,1	690,1	1,97
sondage 53	1	1609 à 1616	0,5 à 4	17,1	17,3	0,93	0,93
	2	1617 à 1625	4,5 à 8,5	10,1	4,4	0,28	0,07
	3	1626 à 1631	9 à 11,5	17,2	2,6	0,98	0,01
	4	1632 à 1636	12 à 14	13,6	2,3	0,48	0,01
	5	1637 à 1642	14,5 à 17	14,5	0,7	0,63	0,01
	6	1643 à 1648	17,5 à 20	17,2	4,6	0,83	0,04

				14,7	13,3	0,68	0,1
sondage 54	1	1650 à 1656	1 à 4	27,6	14,6	1,39	0,07
	2	1657 à 1661	4,5 à 8	43,8	173,2	2	0,2
	3	1662 à 1663	8,5 à 9	84,5	480,5	4,7	0,18
	4	1664 à 1673	9,5 à 14	43,7	79,1	2,3	0,26
	5	1674 à 1678	14,5 à 16,5	79,2	119,7	3,8	0,56
				48,5	456,4	2,46	1,19
sondage 55	1	1679 à 1683	0,5 à 2,5	23,8	38,2	1,44	0,33
	2	1684 à 1698	3 à 10	19,1	46,3	0,94	0,14
	3	1699 à 1710	10,5 à 16	38	160,2	1,52	0,45
	4	1711 à 1713	17 à 18	91,7	208,3	2,93	0,65
				32,5	497,1	1,38	0,59
sondage 56	1	17141716	0,5 à 1,5	24,3	82,3	1,43	0,22
	2	17171723	2 à 5	72,9	246,3	2,99	0,25
	3	17241726	7 à 8	32,3	108,3	1,23	0,17
	4	17271732	8,5 à 11	47,5	27,5	2,12	0,01
				50,8	477,3	2,19	0,63
sondage 57	1	1733 à 1735	0,5 à 1,5	20,3	20,3	1,03	0,02
	2	1736 à 1744	2 à 6	38,2	80,9	2,02	0,52
	3	1745 à 1752	6,5 à 10	43,1	28,1	2,08	0,27
	4	1753 à 1758	10,5 à 13	48,3	43,9	2,18	0,47
				40	111,5	1,96	0,46
sondage 58	1	1759 à 1767	0,5 à 4,5	27,7	62,5	1,17	0,05
	2	1768 à 1772	5 à 7	48	32,5	2,36	0,26
	3	1773 à 1774	7,5 à 8	90	200	4,55	0,01
	4	1775 à 1781	8,5 à 11,5	42,4	32	2,3	0,23

	5	1782 à 1789	12 à 15,5	30,5	9,7	1,2	0,05
	6	1790 à 1792	16 à 17	51,7	277,3	1,57	0,81
				40,1	288	1,82	0,9
sondage 59	1	1794 à 1799	1 à 3,5	25	26,4	2,13	0,88
	2	1800 à 1827	4 à 17,5	14,4	12,4	0,54	0,04
	3	1828 à 1831	18 à 19,5	44,3	490,9	0,88	0,25
				21,1	279,4	0,93	0,88
sondage 60	1	1832 à 1833	0,5 à 1	17,5	0,5	1,05	0,01
	2	1834 à 1839	1,5 à 4	67,3	287,9	2,6	0,14
	3	1840 à 1846	4,5 à 7,5	52,9	84,8	3,16	0,52
	4	1847 à 1856	8 à 12,5	46,2	51,3	2,05	0,33
				50,8	271,2	2,41	0,67
sondage 61	1	1857 à 1869	0,5 à 6,5	15,2	15,5	0,72	0,14
	2	1870 à 1879	7 à 11,5	29,2	32,2	1,47	0,05
	3	1880 à 1888	12 à 16	49,3	96,3	2,61	0,44
	4	1889 à 1892	16,5 à 18	72,3	348,3	3,2	0,15
				34	430,5	1,68	1,01
sondage 62	1	1893 à 1894	0,5 à 1	89,5	220,5	3,2	0,18
	2	1895 à 1909	1,5 à 8,5	19,9	44,8	0,9	0,09
	3	1910 à 1927	9 à 17,5	19,2	33,1	0,86	0,08
	4	1928 à 1932	18 à 20	16,6	0,8	0,78	0,03
				22,7	278,5	0,98	0,34
sondage 63	1	1933 à 1939	0,5 à 3,5	17	3,3	0,94	0,06
	2	1940 à 1947	4 à 7,5	11,1	1,6	0,31	0,01
	3	1948 à 1950	8 à 9	19,7	0,3	1,03	0,01

	4	1951 à 1964	9,5 à 16	13,4	2,6	0,47	0,01
	5	1965 à 1972	16,5 à 20	25,9	24,1	1,05	0,11
				16,6	34,2	0,68	0,13
sondage 64	1	1974 à 1980	1 à 4	21,4	54	1,29	0,17
	2	1981 à 2012	4,5 à 20	13,6	17,9	0,47	0,04
				17,1	212,2	0,76	0,99
sondage 65	1	2013 à 2017	0,5 à 2,5	29	75,5	1,92	0,66
	2	2018 à 2025	3 à 6,5	9,3	5,4	0,34	0,01
	3	2026 à 2032	7 à 10	25	27,3	1,61	0,08
	4	2033 à 2044	10,5 à 16	20,8	68	1,02	0,16
	5	2045 à 2052	16,5 à 20	73,5	700,3	3,46	0,8
				30,8	664,6	1,59	1,43
sondage 67	1	2054 à 2056	1 à 2	25,7	65,3	1,83	0,3
	2	2057 à 2066	2,5 à 7	10,5	4,9	0,41	0,01
	3	2067 à 2077	7,5 à 12,5	20,4	39,5	0,93	0,17
	4	2078 à 2092	13 à 20	30,3	154,8	1,3	0,52
				22,1	135,6	1,01	0,44
sondage 68	1	2093 à 2099	0,5 à 3,5	23	63,3	1,39	0,78
	2	2100 à 2111	4 à 9,5	9,6	1,9	0,33	0,01
	3	2112 à 2122	10 à 15	27,2	22,6	1	0,11
	4	2123 à 2132	15,5 à 20	60,2	365,5	2,73	0,98
				29,4	473,2	1,3	1,22
sondage 980	1	2133 à 2153	0,5 à 3	31,7	117,1	2,45	1,24
	2	2139 à 2142	3,5 à 5,5	68	821,3	4,05	1,1
	3	2143 à 2162	6,5 à 8,5	91,2	153,2	4,08	0,34

				61,2	970	3,42	1,45
sondage 981	1	2148 à 2153	0,5 à 3	52	684,8	2,37	0,68
	2	2154 à 2160	3,5 à 6,5	64,6	560,6	3,31	1,11
	3	2161 à 2162	7 à 7,5	100	/	4,5	/
				64,3	731,8	3,11	1,33
sondage 982	1	2164 à 2170	1 à 4	15,4	40,6	0,87	0,19
	2	2171 à 2174	4,5 à 6	26,3	141,6	1,55	0,3
	3	2175 à 2176	6,5 à 7	96	32	4,5	0,5
				31,9	849,8	1,83	2,28
sondage 983	1	2177 à 2181	0,5 à 2,5	28	46	2,68	2,21
	2	2182 à 2192	3 à 8	6,7	5,62	0,41	0,02
	3	2193 à 2204	8,5 à 14	23,1	52,8	1,58	0,44
	4	2205 à 2211	14,5 à 17,5	28,1	410,8	1,2	0,7
	5	2212 à 2216	18 à 20	85,8	126,2	3,52	0,33
				27,9	660,6	1,64	1,56
sondage 984	1	2217 à 2225	0,5 à 5	15,9	54,4	1,23	0,33
	2	2226 à 2235	6,5 à 11,5	57	46,7	2,56	0,93
	3	2236 à 2248	12,5 à 20	25,2	101,6	1,45	2,49
				32,5	768,6	1,73	1,65
sondage 985	1	2249 à 2254	0,5 à 3	37,5	/	2,12	0,43
	2	2255 à 2257	3,5 à 4,5	81	/	4,63	0,3
	3	2258 à 2264	5 à 8	56,7	/	3,03	1,1
				54,1	643,7	2,99	1,47
sondage 986	1	2265 à 2270	0,5 à 3	27,8	98,6	1,97	0,14
	2	2271 à 2280	3,5 à 8	52,7	258,2	2,74	0,22

				43,4	342,4	2,45	0,67
sondage 991	1	2281 à 2284	0,5 à 2	38,5	696,3	2,28	0,05
	2	2285 à 2297	2,5 à 8,5	9,9	34,9	0,68	0,09
	3	2298 à 2302	9 à 11	38,6	23,3	2,78	0,14
	4	2303 à 2307	11,5 à 13,5	19,8	11,2	1,4	0,13
				24,1	481,2	1,65	1,25
sondage 992	1	2309 à 2313	0,5 à 2,5	49	76	2,64	0,3
	2	2314 à 2321	3 à 6,5	22	32	1,38	0,27
	3	2322 à 2333	7 à 12,5	69,3	378,4	3,29	0,58
	4	2334 à 2338	13 à 15	27,6	18,8	0,96	0,34
				48	659	2,44	1,66
sondage 993	1	2341 à 2352	0,5 à 6	22,3	32,1	2,02	0,32
	2	2353 à 2358	6,5 à 9	71,8	257,4	3,75	0,98
	3	2359 à 2365	9,5 à 12,5	43	142,3	2,33	0,41
	4	2366 à 2377	13 à 18,5	26,2	203,1	1,18	0,21
				37,2	546,4	2,14	1,18
sondage 1010	1	2380 à 2390	0,5 à 6	5,1	18,7	0,42	0,24
	2	2391 à 2407	6,5 à 14,5	21,4	104,1	1,19	0,2
	3	2408 à 2416	15 à 19	68,1	669,9	2,94	0,28
				29,8	899,4	1,46	1,28
sondage 1011	1	2417 à 2430	0,5 à 7	9,4	37	0,88	0,3
	2	2431 à 2435	7,5 à 9,5	58	600	1,92	0,15
	3	2436 à 2447	10 à 15,5	22,5	52,6	1,11	0,13
	4	2448 à 2452	16 à 18	86,8	74,7	4,32	0,29
				31,3	867,2	1,58	1,56

sondage 1012	1	2453 à 2470	0,5 à 9	17	117,4	1,11	0,79
	2	2471 à 2477	95,5 à 12,5	65,7	1116,6	2,6	0,7
	3	2478 à 2482	13 à 15	25,6	280,3	0,94	0,13
	4	2483 à 2488	15,5 à 18	97	41,2	4,23	0,14
				41	1276,3	1,9	2,02
sondage 1013	1	2489 à 2493	0,5 à 2,5	25,2	88,7	2,08	0,74
	2	2494 à 2510	3 à 11	14,5	69,5	0,66	0,13
	3	2511 à 2512	11,5 à 12	96,5	24,5	4,45	0,61
	4	2513 à 2518	12,5 à 15	30,5	552,7	0,83	0,27
	5	2519 à 2523	15,5 à 17,5	92	88	3,6	0,55
				34,5	1062,8	1,53	1,91
sondage 1014	1	2525 à 2530	1 à 3,5	27,8	45,2	2,13	0,92
	2	2531 à 2545	4 à 11	11,4	94,5	0,37	0,07
	3	2546 à 2551	11,5 à 14	100	/	3,65	0,98
				37,1	104,5	1,58	2,43
sondage 1018	1	2552 à 2556	0,5 à 2,5	22,4	134,8	2,54	0,88
	2	2557 à 2561	3 à 5	6	7,5	0,8	0,03
	3	2562 à 2574	5,5 à 11,5	22,3	73,1	1,19	0,42
	4	2575 à 2587	12 à 18	83,8	648,6	2,87	1,42
				42,3	1298,1	1,87	1,66
sondage 1019	1	2558 à 2591	0,5 à 2	34,3	421,6	2,53	1,21
	2	2592 à 2609	2,5 à 11	19,3	96	1,09	0,15
	3	2610 à 2612	11,5 à 12,5	4,7	1,3	0,43	0,01
	4	2613 à 2617	13 à 15	43,8	113,2	0,38	0,07
	5	2618 à	15,5 à 19	70,4	417,1	3,28	0,49

		2625					
				35,9	745	1,79	1,48
sondage 1020	1	2627 à 2635	15	12,9	29,1	0,84	0,11
	2	2636 à 2645	5,51	25,6	55,2	1,42	0,07
	3	2646 à 2650	10,512,5	70	1323,5	3,52	1,29
	4	2651 à 2657	1316	19,9	64,2	1,04	0,14
	5	2658 à 2663	16,519	95,5	29,1	4,12	0,3
					40,4	1207	2,01
sondage 1021	1	2665 à 2674	1 à 5,5	17,8	60,8	1,18	0,39
	2	2675 à 2688	6 à 12,5	38,9	88,9	2,19	0,34
	3	2689 à 2694	13 à 15,5	25,2	23,4	1,32	0,22
					31,3	219,8	1,77
sondage 1039	1	2696 à 2697	0,5 à 1	75	1250	5	/
	2	2698 à 2724	1,5 à 14,5	30,3	116,8	1,82	0,75
					33,4	285,8	2,04
sondage 1040	1	2726 à 2737	1 à 6,5	22,7	80,2	1,13	0,32
	2	2738 à 2749	8 à 15,5	43,5	220,9	1,32	0,4
					38	580,8	1,39
sondage 1041	1	2751 à 2755	1 à 3	24,8	77,2	1,9	0,28
	2	2756 à 2766	3,5 à 8,5	18	121,4	0,89	0,25
	3	2767 à 2772	9 à 11,5	32	73,2	0,93	0,04
	4	2773 à 2775	12 à 13	92,7	76,3	1,57	0,3
					34,2	783,9	1,3

REFERENCE BIBLIOGRAPHIQUE

- 1) *van Wambek, e A. ; Kindermans, J.M., 1978. « essai de chargement à la plaque pour le controle du comactage », Bruxelles : Centre de recherche routièrès.,*
 - 2) *Abdul Baki, A. ; Magnan, J.P., Pouget, P., 1993. « Analyse probabiliste de la stabilité de deux remblais sur versant instable », Laboratoir centrale des ponts et chaussées, Paris.*
 - 3) *Berthier, P., 2012. « Les Essais Lefranc Et Nasberg », article, Bureau d'études GEOTEC Douai.*
 - 4) *Carpentier, F., 2005. « Introduction aux analyses multidimensionnelles ».*
 - 5) *Chauvet, P., 1999. « Aide mémoir de géostatistique linéaire ». Les Presses de l'École des Mines, Paris*
 - 6) *Cibois, P., 2006. « Principe de l'analyse factorielle », note de cours : université de versaille, Paris.*
 - 7) *Houmadi, Y., 2004. « Modelisation stochastique de la variabilité spatiale des paramètres géotechniques », mémoire de magistère. Tlemcen: Faculté de technologie.*
 - 8) *Immediato, H., 2010. « Cours statistiques », note de cours : Lyon: Univercité Claude Bernard de Lyon.*
 - 9) *Laffly, D., 2009. « Introduction A L'analyse Factorielle Des Correspondances », note de cours , Paris: Laboratoire Société Environnement Territoire.*
 - 10) *Monnet, J., 1993/94. « Essais in situ », France: Universite Joseph Fourier.*
 - 11) *Pierre Legendre, 2007. « Les lois de distribution », note de cours, Mmontréal: Université de Montréal.*
 - 12) *Samuel, A., Jézéquel, J-F et., 1980 « Propriétés mécaniques des sols déterminées en place ». aide mémoire, Paris: laboratoire central des ponts et chaussées.*
 - 13) *Magnan, J-P., 1980. « Essai de mécanique des sols en laboratoir et en place ». article, Paris: Laboratoir central des ponts et chaussées.*
 - 14) *Sohrab, B ; Magnan, J-P 1983. « Analyse probabiliste de la stabilité et des tassements des remblais du site expérimental de CUBZAC-les ponts ». article, Paris: laboratoire central des ponts et chaussées.*
- Duby, C., 2006. « Analyse en Composantes principale », note de cours, paris: Institut National Agronomique Paris - Grignon.*