


Utilisation d'un RNA pour l'identification des sols à partir de données pénétrométriques



Nicolas ROMANOWSKI

DÉPARTEMENT GÉNIE CIVIL

Mémoire présenté en vue de l'obtention du diplôme d'Ingénieur
Option Génie Civil

Février – Juin 2016

 : **Polytech'Clermont Ferrand - Département Génie Civil**
Rue des Meuniers – BP 206 - F 63174 Aubière Cedex (France)

 : Secrétariat : +33 (0)4 73 40 76 87 / 76 90 Scolarité : +33 (0)4 73 40 75 05
 : +33 (0)4 73 40 75 10 - Mèl : dept.gc@polytech.univ-bpclermont.fr

REMERCIEMENTS

Je tiens à remercier tout particulièrement Carlos SASTRE JURADO pour m'avoir conseillé, encouragé et soutenu durant ces cinq mois de stage. Ainsi que pour son aide à la rédaction de ce mémoire.

Je tiens également à remercier tout le personnel de l'entreprise SolSolution pour leurs sympathies et leur accueil, en particulier les personnes du service R&D.

Merci également à la direction de l'entreprise qui ont su se montrer disponibles durant ce stage, je pense particulièrement à Mr Roland GOURVES, directeur de SolSolution, ainsi que Mr Miguel BENZ NAVARRETE.

Enfin, je remercie Mr Claude BACCONNET et Mr Mohammed SAHLI, enseignants à Polytech' Clermont-Ferrand, qui ont suivi avec intérêt mon travail et pour leurs conseils qui m'ont aidé à la rédaction de ce mémoire.

Utilisation d'un RNA pour l'identification des sols à partir de données pénétrométriques

Nicolas ROMANOWSKI

Polytech Clermont-Ferrand (CUST) - Mémoire présenté en vue de l'obtention du diplôme d'ingénieur -
- Département Génie Civil - Juin 2016 -

- Résumé -

La caractérisation géotechnique des sols consiste à identifier leurs natures et leurs paramètres mécaniques. Il existe plusieurs moyens pour caractériser un sol, soit à partir d'essais in-situ, soit au laboratoire. Parmi les essais in-situ, les pénétromètres sont les plus utilisés. Ils consistent à enfoncer dans le sol un pieu à échelle réduite muni d'une pointe conique. Le fonçage peut être statique ou dynamique. Par ailleurs, dans de nombreux cas, il est difficile de réaliser des essais in-situ pour des raisons d'accessibilité, de coût ou de durée d'intervention.

C'est pour répondre à ces besoins qu'il a été développé le pénétromètre Panda®. Il s'agit d'un appareil léger, peu cher et rapide à mettre en oeuvre avec une capacité de pénétration suffisante pour ausculter la plupart des sols présents en surface. La résistance du sol se mesure grâce à la formule des Hollandais. Le principal atout du Panda® est qu'il mesure quasiment en continu la résistance du sol. Cependant, son principal inconvénient, comme pour la plupart des pénétrogrammes, est qu'il est dit « aveugle » car on ne prélève pas d'échantillons. Cependant, un utilisateur expérimenté peut détecter à partir de la signature du signal des indices lui permettant d'estimer la nature du sol sondé, sans qu'il soit possible de formuler cette estimation par un algorithme mathématique traditionnel.

C'est dans ce contexte que l'utilisation des réseaux de neurones artificiels (RNA) semble opportune. En effet, les RNA sont des systèmes d'intelligence artificielle qui permettent de recréer le fonctionnement d'un cerveau humain. Ils ont déjà été employés avec succès dans divers domaines de la géotechnique (fondations, liquéfaction, ...). L'objectif de ce travail est de mettre au point une méthodologie basée les RNA sur qui permettrait d'identifier la nature du sol à partir du signal pénétrométrique.

Pour cela, il faut un ensemble conséquent de données destiné à l'apprentissage du RNA. Le meilleur outil pour stocker et exploiter autant de données est un système de base de données. Les détails de la création de la base de données ainsi que son mode de fonctionnement sont détaillés.

Dans un second temps, les performances du RNA sont évaluées dans différents cas. En examinant l'influence de la pointe, de la régularisation du signal et son avantage sur les résultats obtenus. Enfin, une étude comparative avec d'autres systèmes de classification automatique est développée afin de justifier l'emploi des RNA. Une application du RNA sur des reconnaissances réelles montre que les résultats obtenus permettent de classer le sol avec un taux de réussite très satisfaisant.

Les perspectives d'avenir sur les RNA ainsi que les modifications possibles du modèle de RNA testé durant tout ce mémoire sont discutés ainsi que le choix du modèle de classification des sols retenu (GTR, USCS, LCPC,...).

Mots-clés

caractérisation des sols - pénétromètre dynamique - Panda 2® - base de données - réseaux de neurones - classification

Sol Solution

Service Innovation & Diagnostic d'Ouvrages

ZA des Portes de Riom Nord

63204 RIOM cedex

**Travaux dirigés par Carlos SASTRE JURADO et Miguel BENZ NAVARRETE - Service Recherche
et Développement**

Using an ANN for the identification of soil from penetrometer data

Nicolas ROMANOWSKI

Polytech Clermont-Ferrand (CUST) - Project report for the award of Diploma in Civil Engineering -
-Department of civil engineering - Jun 2016 -

- Abstract -

Geotechnical characterization of soils is to identify their natures and mechanical parameters. It exists several ways to characterize a soil either from tests either laboratory. Among in-situ tests, penetrometers are the most used. They consist to sink into the ground a reduced scale pole with a cone tip. Sinking can be static or dynamic. Furthermore, in many cases, it is difficult to achieve in-situ tests for reasons of accessibility, cost or duration of intervention.

It was to satisfy these needs that it was developed the Panda® penetrometer. This is a lightweight, low-priced and easy to implement with a sufficient capacity of penetration to examine most of surface soil. Soil resistance is measured with the Dutch formula. The main asset of Panda®, that it measures soil resistance almost continuous. However, its main disadvantage, as most penetrometers, is that it is called “blind” because we don’t collect samples. But, an experienced user can detect from pulse signature clues for estimating the nature of the soil. Nevertheless, this analysis called “expert” is very complex to model with traditional mathematical algorithms.

It is in this context that the use of artificial neural networks (ANN) appears appropriate. Indeed, ANN are artificial intelligence systems that enable recreate a human nervous system. They have been applied successfully to several problems in geotechnical engineering (foundations, liquefaction ...). The purpose of this work is to develop a methodology which would identify soil nature from penetrometer signal.

For this, you need a data set for ANN training. The best tool to store and use much data is a database system. The details of the creation of the database and its operating mode are described.

Second, ANN performances are evaluated in different cases. In examining tip impact, signal regularization and its advantage over results. Finally, a comparative study with other automatic classification systems is develop in order to justify ANN utilization. An ANN application on real cases shows that results have a satisfactory success rate.

Future prospects on ANN and possible changes of ANN model are discussed as well as choice of the successful soil classification model (GTR, USCS, LCPC ...).

Key-words

soil characterization - dynamic penetrometer - Panda 2® - database - neural networks - classification

Sol Solution

Service Innovation & Diagnostic d’Ouvrages

ZA des Portes de Riom Nord

63204 RIOM cedex

**Work led by Carlos SASTRE JURADO and Miguel BENZ NAVARRETE - Research and
Development Department**

TABLES DES MATIÈRES

REMERCIEMENTS	2
TABLES DES FIGURES.....	8
TABLES DES TABLEAUX	10
INTRODUCTION.....	11
PARTIE I : ÉTUDE BIBLIOGRAPHIQUE	13
INTRODUCTION.....	14
CHAPITRE 1 : CARACTÉRISATION D'UN SOL	14
I. DESCRIPTION QUALITATIVE DES SOLS.....	14
II. DESCRIPTION QUANTITATIVE ET IDENTIFICATION DES SOLS	14
II.1. Description de l'état du sol	14
II.2. Propriétés des particules du sol	15
III. CLASSIFICATION DES SOLS	16
IV. COMPORTEMENT MÉCANIQUE DES SOLS	18
V. LES ESSAIS PÉNÉTROMÉTRIQUES.....	19
V.1. Pénétromètre statique C.P.T.	19
V.2. Pénétromètre dynamique	21
VI. PÉNÉTROMÈTRE PANDA	23
VI.1. Généralités	23
VI.2. Principe de l'essai	23
VI.3. Les applications des pénétromètres dynamiques.....	24
VI.4. Fiabilité des mesures	25
CHAPITRE 2 : LES RÉSEAUX DE NEURONES ARTIFICIELS	26
I. PRÉSENTATION DES RNA	26
I.1. Historique	26
I.2. Domaines d'utilisation	27
I.3. Les types d'apprentissage.....	27
I.4. Le Perceptron	28
II. FACTEURS DE PERFORMANCES	30
II.1. Détermination des entrées du modèle.....	30
II.2. La division de données.....	31
II.3. Les données de pré-traitement.....	32
II.4. Détermination du modèle de l'architecture	32
II.5. Optimisation du modèle.....	33
II.6. Critères d'arrêt.....	33

II.7. La validation du modèle.....	33
III. LES RNA EN GÉOTECHNIQUE	34
CHAPITRE 3 : SYSTÈMES DE CLASSIFICATION AUTOMATIQUES.....	35
I. LES ARBRES DE DÉCISION	35
II. LES ANALYSES DISCRIMINANTES.....	36
III. LES MACHINES À VECTEURS DE SUPPORT.....	36
IV. LA MÉTHODE DES K PLUS PROCHES VOISINS	37
BILAN.....	38
PARTIE II : ÉTUDE.....	39
CHAPITRE 1 : LA BASE DE DONNÉES	40
I. INTRODUCTION	40
II. CONCEPTION D'UNE BASE DE DONNÉES	41
II.1. Objectifs de la base de données	42
II.2. Modélisation conceptuelle	42
II.3. Choix d'un SGBD	44
II.4. Modélisation physique	45
III. ANALYSE DES DONNÉES DISPONIBLES	46
III.1. Base de données laboratoires	46
III.2. Base de données in-situ	47
CHAPITRE 2 : PARAMÈTRES D'ENTRÉE ET DE SORTIE DU RNA.....	48
I. PRÉSENTATION DES PARAMÈTRES D'ENTRÉE DU RNA.....	48
I.1. Analyse statistique	49
I.2. Analyse non linéaire.....	50
I.3. Analyse morphologique du signal.....	50
I.4. Analyse spectrale	51
II. PRÉSENTATION DES PARAMÈTRES DE SORTIE DU RNA.....	52
CHAPITRE 3 : DÉTERMINATION DE L'INFLUENCE DE LA SECTION DE LA POINTE....	53
I. PRÉSENTATION DE LA RÉGULARISATION	53
I.1. Principe de fonctionnement.....	53
I.2. Exemples d'application	54
II. CALCUL DES ERREURS	54
III. VÉRIFICATION DE L'HOMOGENÉITÉ D'UN SONDAGE.....	55
III.1. Choix de la taille de la fenêtre de régularisation.....	56
III.2. Sable de Fontainebleau.....	57
III.3. Sable de Lyon	58
III.4. Grave de chantier	59
III.5. Conclusion	59
IV. INFLUENCE DE LA SECTION TRANSVERSALE DE LA POINTE	60

IV.1. Caractéristiques des matériaux et régularisation	60
IV.2. Influence sur les paramètres de l'analyse statistique	63
IV.3. Influence sur les paramètres de l'analyse morphologique	64
IV.4. Valeurs des erreurs	66
IV.5. Analyse 3D des paramètres statistique	67
IV.6. Résultats de classification du RNA	69
V. BILAN.....	71
CHAPITRE 4 : COMPARAISON PMC AVEC DIFFÉRENTS SYSTÈMES DE	
CLASSIFICATION AUTOMATIQUES	72
I. PRÉSENTATION DES DONNÉES	72
II. RÉSULTATS	73
II.1. Arbres de décision et analyse du discriminant	73
II.2. Les machines à vecteur de support	73
II.3. La méthode des K plus proches voisins	74
III. CONCLUSION	74
CHAPITRE 5 : COMPARAISON DES RÉSULTATS AVEC DIFFÉRENTS ÉTATS HYDRIQUES.	
76	
I. TEST SUR LA PROFONDEUR CRITIQUE	77
I.1. Test sur les trois états hydriques.....	77
I.2. Test sur l'ensemble des données disponibles.....	78
II. TEST SUR LA RÉGULARISATION.....	78
II.1. Test sur les trois états hydriques.....	78
II.2. Test sur l'ensemble des données disponibles.....	79
III. VALIDATION DES MODÈLES.....	79
BILAN.....	81
CONCLUSION ET PERSPECTIVES	82
CONCLUSION	83
PERSPECTIVES DE TRAVAIL	86
BIBLIOGRAPHIE	87
ANNEXES	90

TABLES DES FIGURES

Figure 1 : Courbe granulométrique et définition des paramètres D10, D30 et D60	15
Figure 2 : Schéma de principe de l'essai CPT	19
Figure 3 : Pénétrogramme type obtenu en chambre de calibration pour un matériau homogène ...	24
Figure 4 : Schéma d'un neurone formel.....	26
Figure 5 : Schéma des fonctions de transfert	26
Figure 6 : Différence entre la généralisation et la mémorisation.....	28
Figure 7 : Schéma de principe d'un Perceptron simple	29
Figure 8 : Schéma de principe d'un Perceptron Multi-Couche.....	29
Figure 9 : Structure arbres de décision	35
Figure 10 : Schéma des frontières des SVM.....	36
Figure 11 : Schéma de principe de développement d'une BDD.....	41
Figure 12 : Structure de la base de données in-situ	43
Figure 13 : Structure de la base de données laboratoire	43
Figure 14 : Répartition des données suivant la classification GTR.....	46
Figure 15 : Répartition des données suivant la section de la pointe	47
Figure 16 : Répartition des données disponibles suivant leurs classes GTR.....	47
Figure 17 : Analyse de la sensibilité des caractéristiques	48
Figure 18 : Schéma distribution du coefficient de Skewness	50
Figure 19 : Analyse spectrale	51
Figure 20 : Classification GTR des sols	52
Figure 21 : Répartition des données selon les classes proposés	52
Figure 22 : Principe de la régularisation et du lissage.....	54
Figure 23 : Pénétrogramme brut (à gauche) et régularisé (à droite)	54
Figure 24 : Nombre de points par pénétrogrammes avant et après régularisation	57
Figure 25 : Comparatif de l'analyse statistique du sable de Fontainebleau brut (à gauche) et régularisé (à droite).....	57
Figure 26 : Comparatif de l'analyse morphologique du sable de Fontainebleau brut (à gauche) et régularisé (à droite).....	58
Figure 27 : Comparatif statistique du Sable de Lyon brut (à gauche) et régularisé (à droite).....	58
Figure 28 : Comparatif morphologique du Sable de Lyon brut (à gauche) et régularisé (à droite)....	58
Figure 29 : Comparatif statistique de la Grave brut (à gauche) et régularisé (à droite).....	59
Figure 30 : Comparatif morphologique de la Grave brute (à gauche) et régularisée (à droite)	59
Figure 31 : Nombre de points des moules homogènes avant et après régularisation.....	62
Figure 32 : Nombre de points moules hétérogènes avant et après régularisation	62
Figure 33 : Statistique - Sable de Dunkerque brut (à gauche) et régularisé (à droite)	63
Figure 34 : Statistique - Sable de Dunkerque brut (à gauche) et régularisé (à droite)	63
Figure 35 : Statistique - Sable grossier brut (à gauche) et régularisé (à droite).....	64
Figure 36 : Morphologie - Sable de Dunkerque brut (à gauche) et régularisé (à droite)	64
Figure 37 : Morphologie - Sable de Dunkerque brut (à gauche) et régularisé (à droite)	65
Figure 38 : Morphologie - Sable grossier brut (à gauche) et régularisé (à droite).....	65
Figure 39 : Coefficient de détermination des moules homogènes entre pointe de 2 et 4 cm ²	66
Figure 40 : Résultats de l'analyse statistique en 3D du sable de Dunkerque brut	67
Figure 41 : Résultats de l'analyse statistique en 2D du sable de Dunkerque brut	67
Figure 42 : Résultats de l'analyse statistique en 3D du sable de Dunkerque régularisé.....	68

Figure 43 : Résultats de l'analyse statistique en 2D du sable de Dunkerque régularisé.....	68
Figure 44 : Matrice de confusion données brutes	69
Figure 45 : Matrice de confusion données régularisées.....	69
Figure 46 : Matrice de confusion données brutes	70
Figure 47 : Matrice de confusion données régularisées.....	70
Figure 48 : Graphique des résultats obtenus pour les différents machines à vecteurs de support..	73
Figure 49 : Graphique des résultats obtenus avec la méthode des K plus proches voisins	74
Figure 50 : Pénétrogramme du sable de Sermentizon pour trois états hydriques différents	76
Figure 51 : Répartition des données avant et après pré-traitement.....	77
Figure 52 : Résultats des ensembles d'entraînement et de validation	77
Figure 53 : Résultats entraînement et validation pour l'ensemble des données	78
Figure 54 : Résultats entraînement et validation pour les données brutes ou régularisées	78
Figure 55 : Résultats entraînement et validation pour les données bruts et régularisés.....	79
Figure 56 : Schématisation du phénomène d'apprentissage d'un RNA.....	79

TABLES DES TABLEAUX

Tableau 1 : Essais d'identification des sols	16
Tableau 2 : Caractéristiques des différents modèles d'arbres de décision.....	35
Tableau 3 : Caractéristiques des différentes analyses discriminantes	36
Tableau 4 : Caractéristiques des différents modèles de SVM.....	37
Tableau 5 : Caractéristiques des différents modèles de KNN.....	37
Tableau 6 : Synthèse des différents logiciels de SGDB.....	45
Tableau 7 : Caractéristiques des classes de sols testées	52
Tableau 8 : Caractéristiques granulométriques des 3 matériaux	56
Tableau 9 : Caractéristiques physiques des 3 matériaux	56
Tableau 10 : Caractéristiques physiques des moules homogènes	60
Tableau 11 : Agencement des matériaux des moules hétérogènes.....	61
Tableau 12 : Caractéristiques physiques et mécaniques des moules hétérogènes	61
Tableau 13 : Erreurs pour les moules homogènes bruts.....	66
Tableau 14 : Erreurs pour les moules homogènes régularisés	66
Tableau 15 : Comparaison des taux de réussite des différents modèles	75
Tableau 16 : Caractéristiques du sable de Sermentizon	76

INTRODUCTION

Les sols représentent tous les matériaux de l'écorce terrestre. Cependant, il faut distinguer deux types de sol : les sols meubles et les sols rocheux. La caractérisation des sols meubles est l'une des tâches essentielle de la géotechnique. Quant aux sols rocheux, leur caractérisation dépend de la mécanique des roches.

La nature discontinue du sol est le facteur prépondérant qui gouverne le comportement mécanique du matériau. L'approche micromécanique permet de prendre en compte cette spécificité et de déduire les lois de comportement du sol à partir de l'analyse des propriétés des grains et de leurs assemblages.

C'est dans ce contexte, que depuis plusieurs années les hommes mettent au point des instruments de caractérisation des sols afin d'en déterminer leurs paramètres mécaniques. Idéalement pour caractériser correctement un site, plusieurs essais de différentes natures sont nécessaires car ils n'apportent pas tous les mêmes informations. Néanmoins, dans la réalité le nombre d'essais est limité. La raison principale est financière. En effet, généralement les essais sont coûteux en moyens techniques et humains et pour la plupart, ils sont longs à effectuer. De plus, une mesure ponctuelle ne peut suffire. Par sa nature et son histoire, le sol présente une grande variabilité spatiale en terme de caractéristiques physiques et mécaniques.

L'essai de pénétration nous apporte des connaissances sur la résistance du sol en place. On distingue deux familles de pénétromètres, les statiques et les dynamiques. En 1991, Roland Gourvès développe le pénétromètre dynamique à énergie variable : le Panda. L'originalité de cet appareil repose sur la réduction de taille du pénétromètre par rapport aux autres, mais surtout sur l'idée d'adapter l'énergie de battage à la résistance du sol. De part sa rapidité d'emploi, sa répétitivité et sa fiabilité, il est devenu très populaire. À la fin d'un essai, on obtient un pénétrogramme qui représente la résistance de pointe du sol en fonction de la profondeur.

De plus, un des attraits de l'appareil est sa capacité à enregistrer presque en continu la résistance du sol. On obtient donc un pénétrogramme assez détaillé car il contient un nombre de points important. Le pénétrogramme contient donc des informations sur la stratigraphie du sol et sur ses différentes propriétés mécaniques. Cependant, on ne prélève aucun échantillon durant l'essai et on obtient aucune information sur la nature du sol sondé. Il s'agit donc d'un sondage « aveugle ».

On observe que la forme du pénétrogramme est différente selon le type de sol. En effet, un utilisateur expert va pouvoir détecter, en observant le pénétrogramme, des indices qui vont lui permettre d'identifier la nature du sol. Il existe donc une relation entre la signature pénétrométrique d'un essai et la nature des matériaux traversés. Cependant, cette relation n'est pas modélisable avec des algorithmes mathématiques traditionnels.

C'est de ce constat que l'entreprise Sol Solution, par le biais des travaux de Carlos SASTRE JURADO, tente de mettre au point un système de réseau de neurones artificiels (RNA) qui aura pour but de recréer le fonctionnement du cerveau d'un utilisateur aguerri.

La première partie de ce mémoire consiste en une synthèse bibliographique faisant un résumé des différents paramètres régissant les sols. Elle est suivie d'une présentation succincte des différents pénétromètres statiques et dynamiques et plus particulièrement du pénétromètre PANDA. Ensuite, un

état de l'art sur les réseaux de neurones artificiels est présenté et plus particulièrement sur les Perceptrons Multi-Couche ainsi que sur des exemples d'applications des RNA dans le domaine de la géotechnique.

Dans une seconde partie, nous présentons le développement de la base de données qui va servir au calibrage du réseau de neurones. Le développement d'un RNA passe dans un premier temps par la définition des entrées du modèle, il s'agit du second point développé dans cette partie. Ensuite, nous étudierons l'influence de la section de la pointe afin de savoir si le RNA est capable de prédire la classe du sol quelque soit la pointe utilisée pour effectuer le sondage. Mais aussi l'influence de la régularisation sur les performances du RNA et l'influence de la profondeur critique. Puis nous comparerons les résultats obtenus avec un PMC avec d'autres systèmes de classifieurs semi-automatiques afin de justifier l'emploi des RNA et plus particulièrement d'un PMC.

Dans une troisième partie, les conclusions générales et les perspectives de travail viennent clore se mémoire. Elle est suivie d'une synthèse de la bibliographie citée.

PARTIE I : ÉTUDE BIBLIOGRAPHIQUE

INTRODUCTION

La première étape de toute étude géotechnique a pour objectif de caractériser les sols du site étudié. Parmi les méthodes de reconnaissance géologique et géotechnique, les forages et les sondages tiennent une place importante du fait des renseignements qu'ils fournissent par eux-mêmes ou grâce à des systèmes complémentaires. Les sols prélevés dans les forages peuvent être décrits de plusieurs manières complémentaires : d'après la nature, les proportions et d'après les propriétés physiques de leurs constituants. Une fois un sol identifié, un système de classification permet d'identifier de manière globale un sol et ce qui facilite ainsi les échanges entre spécialistes.

CHAPITRE 1 : CARACTÉRISATION D'UN SOL

I. DESCRIPTION QUALITATIVE DES SOLS

[FAG 11] [MAG 97] [MAG 01]

Les sols sont des agrégats de particules (minérales ou organiques) de tailles et de formes variables. Les particules sont liées entre elles mais peuvent être aisément séparables. Les sols ont deux origines principales : désagrégation des roches par altération ou décomposition d'organismes vivants.

Le sol est un matériau à trois phases, on distingue :

- la phase solide : elle est constituée par des particules minérales ou organiques de l'agrégat ;
- la phase liquide : elle est constituée par de l'eau qui comble les vides de l'agrégat. On distingue deux états, si tous les vides sont remplis par de l'eau alors le sol est dit saturé sinon il est non saturé ;
- la phase gazeuse : pour les sols non saturés, une partie des vides est remplie par de l'air.

La coexistence de ces trois phases différentes d'un point de vue physique est à l'origine de la variabilité et de la complexité du comportement mécanique d'un sol.

II. DESCRIPTION QUANTITATIVE ET IDENTIFICATION DES SOLS

[FAG 14] [MAG 97] [MAG 01]

Identifier un sol correspond à déterminer un ensemble de propriétés physiques ou mécaniques qui permettent de caractériser au mieux le sol. Ces propriétés sont déterminées à partir d'essais d'identification. Ils conduisent à une description précise et chiffrée du sol. Une définition chiffrée est nécessaire car des sols voisins d'aspects peuvent présenter des comportements mécaniques très différents.

II.1. Description de l'état du sol

II.1.1. Paramètres d'état

L'arrangement des particules, de l'eau et du gaz est trop complexe pour être décrit de façon détaillée. Donc, on se satisfait d'une estimation globale des proportions de chacune des phases. On peut définir :

- masse volumique de l'eau,
- masse volumique du sol,
- masse volumique du sol sec,
- masse volumique de l'air,
- masse volumique des particules,
- masse volumique du sol saturé,
- masse volumique déjaugée du sol,
- la porosité,
- l'indice des vides,
- la teneur en eau,
- la teneur en eau de saturation,
- le degré de saturation.

Les masses volumiques permettent de calculer les contraintes dans les sols, elles sont pour cela multipliées par la pesanteur.

II.1.2. Relations entre les paramètres d'état

Les paramètres définis précédemment ne sont pas indépendants. Il existe entre eux un certain nombre de relations mathématiques permettant de retrouver des paramètres manquants. Cependant, certaines relations varient du fait que le sol soit saturé ou non, d'autres sont complètement indépendantes.

L'état d'un sol peut donc être défini par :

- trois paramètres : indice des vides, masse volumique des particules et par son degré de saturation ou sa teneur en eau, si le sol n'est pas saturé ;
- deux paramètres : masse volumique des particules et la teneur en eau ou l'indice des vides, si le sol est saturé ;
- et par un ou deux paramètres si on admet que la masse volumique des particules est constante.

II.2. Propriétés des particules du sol

Les essais caractérisant la nature du sol peuvent être effectués sur des échantillons intacts ou remaniés.

II.2.1. Granulométrie

La granulométrie d'un sol est la répartition de la dimension moyenne des particules de ce sol, exprimée en pourcentage de la masse totale du matériau testé. La granulométrie s'exprime par la courbe granulométrique, voir figure 1.

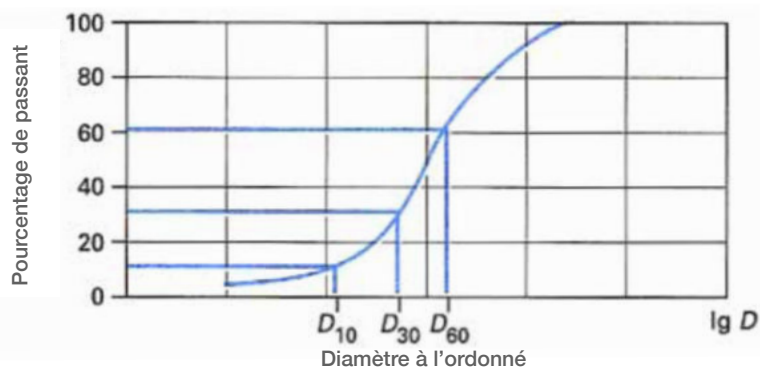


Figure 1 : Courbe granulométrique et définition des paramètres D_{10} , D_{30} et D_{60}

Si l'on note D_p le diamètre correspondant à l'ordonnée $p(\%)$, on appelle :

- D_{10} : diamètre efficace ;
- C_U : coefficient d'uniformité ou de Hazen, c'est le rapport de D_{60} sur D_{10} ;
- C_C : coefficient de courbure, il se calcule de la façon suivante : $C_C = (D_{30})^2 / (D_{10} D_{60})$

Suivant la forme de la courbe granulométrique, on parle de granulométrie uniforme ou étalée et discontinue ou continue.

II.2.2. Les essais d'identification

Les essais d'identification sont répartis en deux catégories : essais pour les sols pulvérulents (grenus) et les essais pour les sols cohérents (fins). Ils sont résumés dans le tableau 1.

Sol	Nature	État
Pulvérulent (grenu)	Granulométrie (analyse granulométrique)	Indice des vides
	Équivalent de sable	Teneur en eau
	Teneur en carbonates	Indice de densité
	Masse volumique des particules	
Cohérent (fin)	Granulométrie (sédimentométrie)	Indice des vides
	Limites d'Atterberg	Teneur en eau
	Teneur en carbonates	Degré de saturation
	Teneur en matières organiques	Indice de consistance
	Masse volumique des particules	

Tableau 1 : Essais d'identification des sols

Les essais d'identification sont conventionnés et doivent être réalisés en respectant un protocole opératoire précis. Ces protocoles sont définis par des normes.

III. CLASSIFICATION DES SOLS

[MAG 97]

Pour pouvoir calculer les ouvrages, il faut connaître les paramètres physico-chimiques du sol. En effet, la classification permet de simplifier la micromécanique en regroupant la description minéralogique, la description des grains (formes et granulométrie) et la sensibilité à l'eau, en plusieurs catégories.

Cependant, suivant l'objectif d'utilisation de la classification, on va distinguer deux grands types de classification.

III.1.1. Classification dans le cadre de la reconnaissance géotechnique

Il s'agit des classifications les plus complètes car elles utilisent de façon importante la granulométrie du sol avec leurs paramètres et la courbe granulométrique. Par exemple, la teneur en matières organiques est prise en compte. Il existe plusieurs types de classification : la classification du LCPC ¹ et la classification U.S.C.S ² ; pour le système U.S.C.S il s'agit d'un système de classification fiable et peu coûteux qui prend en compte :

- les pourcentages de classes granulométriques particulières,
- la forme de la courbe granulométrique,
- les caractéristiques de plasticité,
- les caractéristiques de compressibilité.

Le système U.S.C.S distingue trois principaux types de sol :

- les sols à gros grains (SGG), qui contiennent 50 pour cent ou moins de fins,
- les sols à grains fins (SGF), qui contiennent plus de 50 pour cent de fins,
- les sols fortement organiques, tels les sols tourbeux, les sols fangeux, les sols à humus ou les sols marécageux.

III.1.2. Classification dans le cadre de l'utilisation des matériaux en remblais

Il existe pour ce cas aussi plusieurs systèmes de classification qui sont utilisés suivant les pays. Il ne sera présenté ici que le système français du guide des terrassements routiers du SETRA³. Le GTR⁴ nous indique donc les cadres de l'utilisation de matériaux extraits en déterminant les caractéristiques de ces matériaux. Il définit les conditions de mise en oeuvre en prenant en compte les caractéristiques physiques, mécaniques et hydriques du matériau et les conditions météorologiques de mise en oeuvre de ces matériaux.

Le GTR distingue trois grandes familles :

- les sols de classe A, B, C et D :
 - classe A : sols fins (passant à 80 μ m supérieur à 35%),
 - classe B : sols sableux et graveleux avec fines,
 - classe C : sols comportant des fines et de gros éléments,
 - classe D : sols insensibles à l'eau.
- les matériaux rocheux de classe R ;
- les sols organiques et sous produits industriels de classe F.

Les paramètres d'identification des sols sont :

- les paramètres de nature, dont les caractéristiques sont la granulométrie et l'argilosité,
- les paramètres de comportement mécanique, dont les caractéristiques sont : la résistance à la fragmentation, la résistance à l'usure et la friabilité,
- les paramètres d'état hydrique.

¹ LCPC : Laboratoire Central des Ponts et Chaussées

² U.S.C.S : Unified Soil Classification System

³ SETRA : Service d'Etudes sur les Transports, les Routes et leurs Aménagements

⁴ GTR : Guide des Terrassements Routiers

IV. COMPORTEMENT MÉCANIQUE DES SOLS

[BNA 09]

Le comportement mécanique d'un sol peut être évalué théoriquement de deux façons différentes : en supposant le sol comme un milieu continu équivalent, soit en prenant en compte sa nature discontinue.

IV.1.1. Considération en milieu continu (MMC)

La MMC (Mécanique des Milieux Continus) s'appuie sur une hypothèse forte d'homogénéité du milieu étudié. Elle utilise l'approche thermodynamique pour caractériser un milieu, elle est basée sur la continuité des champs de contraintes et de déformations à l'intérieur du milieu. Cependant, on comprend que par la nature hétérogène d'un sol, l'hypothèse de continuité n'est pas réaliste. Néanmoins, si le rapport entre le volume élémentaire et la dimension des grains est grand, l'hypothèse d'un milieu continu peut être justifiée.

Mais ce type d'approche doit faire face à deux difficultés :

- les phénomènes tels que la dilatance ou la non-existence d'un domaine élastique ne pouvant pas être décrits par cette approche,
- la prise en compte de la réalité et des variables internes caractérisant l'état microstructural du matériau (variables d'état, variables d'écrouissage...).

Malgré cela, cette approche est toujours utilisée. La MMC régit le comportement des matériaux, en prenant en compte la loi unissant tenseur des contraintes et tenseur des déformations, selon trois classes de comportement : élasticité linéaire, plasticité et viscosité. Actuellement, les lois élasto-plastiques respectant le critère de Mohr-Coulomb et le critère de Tresca sont le plus utilisées conventionnellement.

IV.1.2. Considération en milieu discret (MMG)

Le sol est un milieu triphasique qui peut être interprété à l'aide de la micromécanique. La micromécanique a pour objectif de comprendre les interactions globales d'un sol mais aussi les phénomènes discrets ayant lieu à l'échelle du grain. Le changement d'échelle du problème physique est donc d'exprimer le comportement macroscopique à partir du comportement local, tel est l'objectif de la micromécanique. Si l'on considère un cube élémentaire d'un sol dont les dimensions sont suffisamment grandes devant la taille des grains, alors il est possible de définir un vecteur contrainte sur chacune de ses facettes. La somme des forces appliquées par l'intermédiaire des contacts inter-granulaires et des interactions avec les fluides peut être simplifiée en une résultante unique située au centre de la facette. Ce cube, qui s'appelle alors Volume Élémentaire Représentatif (V.E.R.), subira un état de déformation intégrant les déformations de ses phases.

On constate qu'en MMC, la solution d'un problème est unique, donc la réponse d'un VER à une sollicitation donnée l'est aussi. Les différents paramètres rhéologiques du VER sont intimement liés aux propriétés physiques des grains, de leurs assemblages ainsi que des interactions avec les fluides. De part les hypothèses de la micromécanique, il existe une réponse biunivoque entre une sollicitation donnée et un assemblage granulaire donné. Donc, en supposant qu'un sol est réversible (grains incassables, insensibles à l'attrition, réversibilité des phénomènes de succion), alors il serait possible qu'un sol soit reconstitué de manière industrielle. Lorsque les paramètres d'assemblages sont constants sur un échantillon fabriqué industriellement alors la réponse mécanique à la sollicitation est inchangée par rapport à un échantillon naturel subissant la même sollicitation.

V. LES ESSAIS PÉNÉTROMÉTRIQUES

[SAG 05] [ROB 11] [KOF 00] [MAG 93] [WAS 83]

L'essai pénétrométrique consiste à foncer dans le sol un pieu à échelle réduite muni d'une pointe conique. Le fonçage peut être statique par vérinage à vitesse contrôlée. Sinon il peut être dynamique, par battage sous énergie contrôlée.

Même si l'effort d'enfoncement concerne les tiges et la pointe, la mesure obtenue doit seulement être l'effort de pointe sans le possible frottement. La courbe obtenue lors de l'essai s'appelle un pénétrogramme, il s'agit de la résistance de pointe mesurée en fonction de la profondeur. Cette courbe peut être continue ou par faibles paliers, elle décrit le comportement du sol de façon locale.

Il s'agit d'essais dit aveugles, car lors de leurs réalisations on ne prélève pas d'échantillon. Il est souvent complété par des mesures supplémentaires telles que la pression interstitielle ou le frottement latéral.

La gamme des pénétromètres est diversifiée, il en existe de plusieurs tailles et de puissances de pénétration différentes. Leurs applications sont variées : reconnaissance de sol pour le calcul de capacité portante ou le contrôle de compactage.

V.1. Pénétromètre statique C.P.T.

Le C.P.T (Cone Penetration Test) est un essai de pénétration statique mis au point en 1932 par Buisman et Barentsen. Les pénétromètres statiques ont connu une forte évolution à partir de 1948.

V.1.1. Principe de l'essai C.P.T.

L'essai C.P.T. est réalisé dans tous les sols fins et les sols grenus dont la dimension moyenne des éléments ne dépasse pas 20 mm. Il consiste à enfoncer dans le sol, à vitesse constante et à l'aide d'un vérin hydraulique, une pointe terminée par un cône. Un dispositif particulier permet de mesurer la résistance à la pénétration du cône, ainsi qu'éventuellement, le frottement latéral mobilisé sur une longueur donnée. Cet essai peut être réalisé à terre avec des camions mais aussi en milieux marins grâce à des plateformes élévatoires avec des systèmes modulaires. Cependant, pour des profondeurs supérieures à 1500 m la technique diffère.

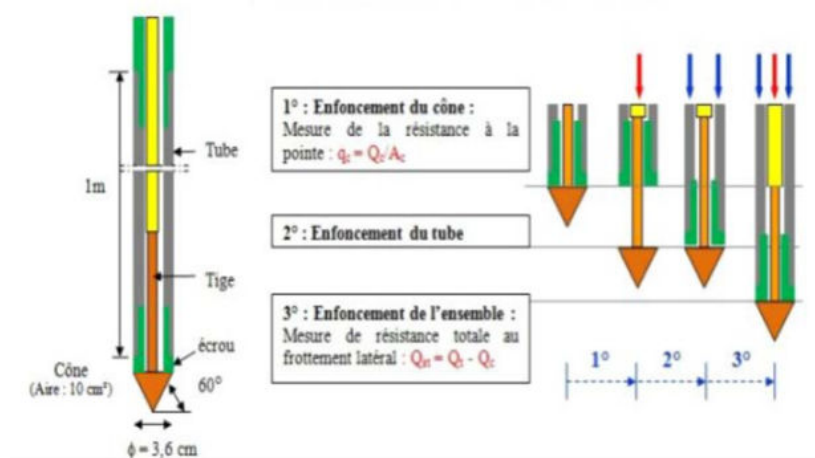


Figure 2 : Schéma de principe de l'essai CPT

La norme NF P 94-113 oblige à respecter plusieurs conditions dont par exemple la vitesse de pénétration qui doit être de 2 cm/s. Les pénétromètres normalisés peuvent avoir des modalités de fonçage différents ou avoir une technique de mesure de la résistance de pointe différente. Il existe plusieurs types de sondes qui sont choisies suivant le type de sol et la nécessité des études, soit une pointe mécanique soit une pointe électrique.

Lors d'un essai au pénétromètre statique, on obtient en plus de la résistance de pointe, le frottement latéral. Cette valeur permet de donner des informations importantes sur les natures du sol. Par exemple, pour les argiles, le frottement latéral va être élevé alors que dans les sols sableux le frottement sera faible.

V.1.2. Résultats

Voici les différents résultats obtenus avec un essai C.P.T. :

- effort total de pénétration Q_t : il s'agit de la force totale nécessaire pour enfoncer la pointe et les tiges dans le sol, sans choc ou vibration ou rotation des tiges ;
- effort apparent sur le cône Q_c : il s'agit de la force nécessaire pour enfoncer le cône de la pointe dans le sol ;
- résistance apparente à la pénétration du cône q_c : cette valeur est obtenue en divisant l'effort total apparent Q_c par la surface de la base du cône.
- effort total de pénétration Q_{st} : il s'agit de la force obtenue en faisant la différence de l'effort total de pénétration du train de tige Q_t et l'effort total apparent sur le cône seul Q_c , ces valeurs dépendent de la profondeur atteinte lors de la réalisation de l'essai.
- effort de frottement latéral Q_s et frottement latéral unitaire local f_s : le frottement latéral unitaire local f_s est obtenu conventionnellement en divisant la force Q_s nécessaire à l'enfoncement du manchon de frottement par sa surface latérale A_s . Cette valeur attribuée à la profondeur correspond au milieu du manchon de frottement. Le diamètre du manchon doit être supérieur ou égal au diamètre de la partie cylindrique du cône.

V.1.3. Critiques

L'essai C.P.T possède un inconvénient majeur qui peut être la non représentativité de l'essai, ces erreurs proviennent principalement de la nature dimensionnelle ainsi que de la rhéologie du sol. En effet, si l'essai est réalisé sur un point dur inconnu alors il amènera à étudier de mauvaises valeurs. Cependant, contrairement aux autres essais de pénétration il est relativement rapide à effectuer.

D'autres facteurs, autres que ceux dépendant de l'opérateur (processus de réalisation de l'essai), peuvent influencer de manière négative les résultats obtenus lors de la réalisation de l'essai. Il s'agit par exemple des variations de température, de l'inclinaison, des erreurs de calibrage ou d'usure du cône.

La forme du cône influe sur les résultats obtenus. Par exemple, suivant le type de cône utilisé, il peut augmenter ou diminuer les valeurs de résistance à la pénétration, le cône GOUDA, en général, surestime les valeurs. La vitesse de pénétration peut aussi influencer sur les valeurs de q_c :

- q_c augmente avec la vitesse de fonçage sur les argiles molles,
- sur les argiles raides et les sables secs, la vitesse n'a quasiment pas d'influence,
- sur les sables et limons aquifères, q_c diminue lorsque la vitesse de pénétration augmente.

V.2. Pénétromètre dynamique

Le principe repose sur la pénétration d'un train de tiges avec une section conique située à l'extrémité. L'enfoncement se réalise par battage. Ils existent plusieurs appareils permettant de réaliser un sondage de cette manière. Ils ont été inventés pour pouvoir mesurer la résistance de pénétration d'un sol ainsi que le niveau de compactage de ce sol. Le contrôle du compactage sur un échantillon est très difficile car l'extraction de l'échantillon, sans modifier ses caractéristiques mécaniques, est extrêmement compliquée d'où l'intérêt du pénétromètre dynamique. Les pénétromètres dynamiques sont également utilisés dans le cadre de la justification des fondations profondes pour la détermination de la capacité portante du sol et de la profondeur d'encastrement.

Voici une liste non exhaustive des différents appareils de sondages de types pénétromètres dynamiques :

- Pénétromètre dynamique léger DPL,
- Pénétromètre dynamique mi-lourd DPM,
- Pénétromètre dynamique lourd DPH,
- Essai de pénétration au carottier – Standard Pénétration Test S.P.T (NF P 94-116),
- Pénétromètre dynamique super-lourd A DPSH-A (NF P 94-114),
- Pénétromètre dynamique super-lourd B DPSH-B (NF P 94-115) ,
- Pénétromètre dynamique à énergie variable PANDA (NF XP P 94-105).

Comme on peut le constater, ils existent plusieurs pénétromètres dynamiques. Leurs différences sont par exemple la hauteur de chute du mouton ou le poids du mouton. Certains ont été normalisés pour pouvoir comparer les valeurs obtenues lors de l'essai.

Par ailleurs sous certaines simplifications, il est possible d'établir une relation entre la résistance de pointe et l'énergie de battage :

- l'énergie du choc du mouton doit être connue,
- la déformation élastique du système doit être négligeable par rapport à la déformation plastique du sol,
- le frottement latéral doit être négligeable,
- pour les sols fins la vitesse de pénétration peut créer des surpressions qui peuvent fausser les valeurs.

V.2.1. Le pénétromètre dynamique type S.P.T.

V.2.1.1. Principe de l'essai

L'essai de pénétration au carottier est utilisé dans tous les sols fins ou grenus dont la dimension moyenne des grains n'excède pas 20 mm. L'essai S.P.T. a pour objectif de battre dans le sol un carottier aux caractéristiques connues, il est constitué des appareils suivant :

- un équipement de forage et de tenue de la paroi,
- un dispositif de battage,
- un train de time et son guidage,
- un carottier,
- un système de mesure.

Une fois le forage réalisé et maintenu par une boue d'injection ou par un tubage, le carottier est installé dans le fond du trou réalisé. Le déroulement de l'essai est en 3 étapes :

- N_0 : amorce, enfoncement de 0 à 15 cm ;
- N_1 : premier essai, enfoncement de 15 à 30 cm ;
- N_2 : deuxième essai, enfoncement de 30 à 45 cm.

Le nombre de coup N_i permettant d'enfoncer la tige de 15 cm est mesuré. La résistance à la pénétration N est alors déterminé en sommant N_1 et N_2 . Cependant, lorsque les 15 cm d'enfoncement ne sont pas atteints après 50 coups, alors l'essai est stoppé et on mesure l'enfoncement obtenu. En fin de battage, la carotte est récupérée afin de pouvoir observer visuellement la nature du sol testé.

V.2.1.2. Analyse des résultats

Il a été établi, par de multiples essais réalisés dans des laboratoires, des corrélations entre le nombre de coups de mouton et les caractéristiques suivantes :

- la compacité des sables et leur angle de frottement interne,
- la résistance des sols à la compression simple,
- la capacité portante pour les fondations,
- le risque de liquéfaction des sables.

Par ailleurs, il est préférable d'utiliser avec prudence dans les sols cohérents.

V.2.2. Critiques

La technique de battage est la cause de plusieurs problèmes qui doivent être soit appréhendés, soit maîtrisés, afin que le pénétrogramme soit utilisable.

V.2.2.1. Le frottement latéral

Le frottement latéral représente le problème majeur rencontré par tous les pénétromètres. En effet, lors d'un essai au pénétromètre la seule valeur intéressante et recherchée est la résistance de pointe. Il existe plusieurs méthodes pour diminuer l'influence du frottement latéral :

- la pointe doit avoir un diamètre supérieur par rapport à celui des tiges,
- l'injection de boue bentonitique,
- la mesure du frottement latéral et la mesure de la résistance de pointe doivent être séparées en utilisant un dispositif de type C.P.T.

Afin d'avoir des résultats corrects, il est préférable que le diamètre de la pointe soit 1,3 fois supérieur à celui du train de tiges. Mais l'injection de boue bentonitique est la meilleure solution pour limiter l'influence du frottement. Cependant, cette méthode est la plus onéreuse.

Sinon, on peut déterminer l'influence du frottement latéral à l'aide d'une clé dynamométrique afin de savoir si le frottement latéral a une forte influence sur la résistance de pointe qui devra donc être sous estimée ou savoir si il est négligeable.

V.2.2.2. La pression interstitielle

Lors de l'enfoncement de la pointe dans un sol saturé, durant la réalisation de l'essai, il est malheureusement obligatoire de créer une surpression interstitielle autour de la pointe. De ce fait, la résistance de pointe sera influencée par la pression interstitielle.

V.2.2.3. Energie de battage et transfert des ondes de choc

Lors de l'essai, les ondes engendrées par le battage et transmises jusqu'au sol sont estimées ou calculées selon certaines hypothèses qui sont spécifiques à chaque appareil. Pour les pénétromètres à énergie constante, il est donc nécessaire de garantir plusieurs aspects :

- la chute libre du mouton,
- la hauteur de chute constante,
- un enfoncement régulier qui soit ni trop important ni trop faible.

VI. PÉNÉTROMÈTRE PANDA

[SAG 05] [ROB 11] [KOF 00] [MAG 93] [CHA 01]

VI.1. Généralités

Les essais in situ sont généralement peu nombreux pour des raisons économiques, d'accessibilité ou de durée d'exécution de l'essai. Dans la plupart des cas, il est nécessaire de caractériser les sols sur une faible profondeur pour pouvoir dimensionner les fondations des structures et avoir des valeurs correctes du comportement du sol et de sa variabilité spatiale.

C'est dans cette optique que le Panda 1 a été inventé en 1992. Il s'agit d'un pénétromètre dynamique à énergie variable qui peut tester tous les types de sols sur quelques mètres de profondeur. En 2002, le Panda 2 a été mis au point par l'entreprise Sol Solution.

Il s'agit d'un appareil léger, peu cher et facile d'utilisation pour une durée d'exécution rapide. En effet, la valise pèse une vingtaine de kilos et les essais au Panda sont réalisables par un seul opérateur.

Son domaine d'application est varié, on retrouve les domaines classiques des pénétromètres à savoir la reconnaissance de sol et le contrôle de compactage de remblais routiers. Mais aussi pour le contrôle de digues ou de barrages en terre et le diagnostic de remblais.

VI.2. Principe de l'essai

Le Panda est basé sur l'utilisation d'une énergie variable par le biais d'une accélération de la masse frappante et ainsi de diminuer son importance. Cela permet de remplacer l'utilisation du mouton à chute libre où l'énergie correspond à l'énergie potentielle.

L'essai consiste à enfoncer manuellement un train de tiges avec une pointe conique à son extrémité dans le sol pour mesurer la résistance du sol jusqu'à la profondeur souhaitée. Le battage s'effectue à l'aide d'un marteau normalisé.

Pour chaque coup donné, des capteurs mesurent la vitesse du marteau à l'instant du contact avec la tête de frappe. Grâce à ces capteurs, il est possible de déterminer l'énergie qui va être transmise à la pointe. Au même instant, des capteurs mesurent l'enfoncement de la pointe dans le sol.

À l'aide de ces deux informations, un boîtier calculateur enregistre ces informations et calcule la résistance de pointe q_d en utilisant la formule des Hollandais. Donc pour chaque coup de marteau, le boîtier enregistre le couple profondeur—résistance dynamique correspondant. La formule des Hollandais a cependant dû être adaptée car l'énergie est variable pour le pénétromètre Panda :

$$q_d = \frac{I}{A} \cdot \frac{E}{e} \quad \text{avec} \quad E = \frac{I}{2} \cdot M \cdot V^2 \cdot \frac{M}{M+P} \quad [1]$$

Avec :

- A : section droite de la pointe (m^2) ;
- E : énergie de battage ;
- e : enfoncement moyen sous un coup ;
- M : masse frappante (kg) ;
- V : vitesse d'impact (m/s^2) ;
- P : masse morte entraînée.

Sur un pénétrogramme réalisé sur un matériau homogène dans un moule, on va distinguer deux parties dans la courbe : la résistance de pointe à la surface q_{d0} et la résistance de pointe en profondeur q_{d1} . La zone de transition correspond à la profondeur critique Z_c . D'après la figure 3, on observe la zone d'augmentation proportionnelle à la profondeur et la partie après Z_c qui correspond à la stabilisation de la résistance.

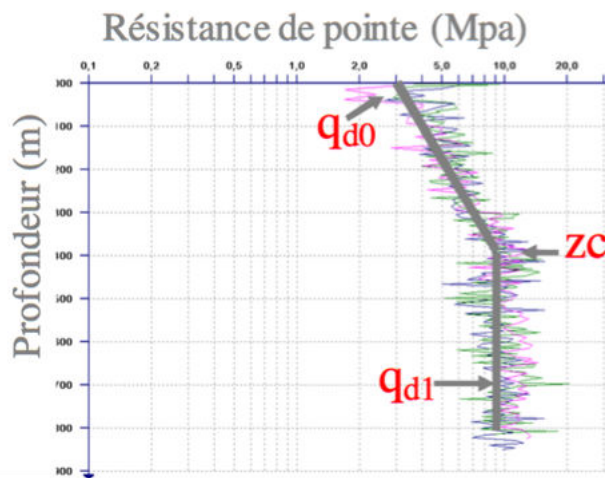


Figure 3 : Pénétrogramme type obtenu en chambre de calibration pour un matériau homogène

VI.3. Les applications des pénétromètres dynamiques

Les pénétromètres dynamiques comme le pénétromètre Panda sont utilisés principalement dans deux domaines d'applications : le contrôle de compactage et la reconnaissance de sol.

VI.3.1. Le contrôle de compactage

Le compactage d'un sol dépend de la valeur de la densité sèche du matériau à un état hydrique connu. Les sols de remblais ont un comportement réversible d'un point de vue de la micromécanique, donc il existe une relation entre la densité sèche d'un sol γ_d et la résistance de pointe q_d .

La qualité d'un compactage dépend des matériaux utilisés et du processus de réalisation des travaux. Contrairement au pénétromètre dynamique léger LRC, le pénétromètre Panda permet d'obtenir des valeurs mesurées plus fines mais aussi plus facilement reproductibles tout cela avec un coût raisonnable.

Le pénétrogramme nous permet aussi d'observer les hauteurs de couches de compactages. Actuellement, le contrôle de compactage est normalisé en France par la norme NF P 94-105.

VI.3.2. La reconnaissance des sols

L'utilisation des pénétromètres dynamiques pour la reconnaissance des sols permet de définir plusieurs caractéristiques essentielles dans le dimensionnement d'une structure de génie civil. En effet, il est possible de définir :

- l'implantation des constructions,
- le choix du type de fondations,
- le niveau des fondations,
- les dimensions de la fondation,
- le mode de réalisation des fondations.

Le pénétromètre Panda a l'avantage, de part sa taille, de réaliser des sondages dans des zones difficiles d'accès. La durée d'exécution de l'essai permet d'augmenter le nombre de sondages et donc d'assurer la fiabilité des mesures. En effet, le sol étant hétérogène, on comprend l'intérêt d'effectuer plusieurs sondages pour obtenir une bonne cartographie du sous-sol. Cependant, la seule utilisation du pénétromètre dynamique Panda ne permet pas de caractériser totalement le sol. Les sondages sont généralement couplés avec d'autres méthodes qui permettent d'identifier clairement les sols, même si le Panda le permet, cela reste de manière aveugle.

VI.4. Fiabilité des mesures

Les travaux sur la mesure de la fiabilité du Panda [CHA 01] ont mis évidence que les mesures de résistance de pointe q_d avaient une précision de l'ordre de 10%. La comparaison des essais Panda avec d'autres essais classiques de pénétration et des essais d'auto-corrélation ont mis en évidence la fiabilité et la reproductibilité du signal pénétrométrique Panda in-situ. Il faut néanmoins rappeler que ces travaux ont été réalisés sur le Panda 1.

CHAPITRE 2 : LES RÉSEAUX DE NEURONES ARTIFICIELS

I. PRÉSENTATION DES RNA

I.1. Historique

[BIS 95] [PER 03]

Dès 1943, Mac Culloch et Pitts proposent des neurones formels mimant les neurones biologiques et qui étaient capables de mémoriser des fonctions booléennes simples. Un neurone est une cellule qui concentre des signaux, les traite et qui les traduit en une réponse, voici un schéma d'un neurone formel :

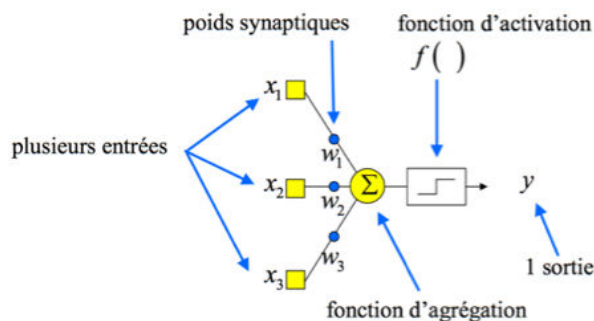


Figure 4 : Schéma d'un neurone formel

Un neurone formel est composé de neurones d'entrées x_i et d'une sortie y . La sortie correspond à la transformation par une fonction d'activation de la fonction d'agrégation. La fonction d'agrégation n'est rien d'autre que la somme des poids synaptiques multipliée par leurs entrées et l'addition du biais.

Les fonctions de transfert ou d'activation peuvent être de plusieurs types : Heaviside, linéaire ou sigmoïde. Cela va dépendre de la nature des entrées et du type de sortie désirée. La figure 5 représente les fonctions d'activation les plus utilisées.

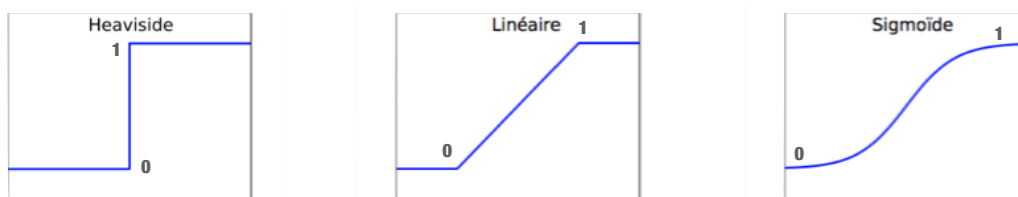


Figure 5 : Schéma des fonctions de transfert

Les réseaux de neurones artificiels (RNA) réalisés à partir de ce type de neurones sont inspirés du système nerveux. Ils sont conçus pour reproduire certaines caractéristiques des mémoires biologiques parce qu'ils sont :

- capables d'apprendre ;
- capables de traiter des informations incomplètes ;
- capables de mémoriser l'information dans les connexions entre les neurones ;
- tolérance au bruit ;
- massivement parallèles.

Le premier modèle opérationnel de RNA est le Perceptron, il est inspiré du modèle visuel et il est capable d'apprendre. Ce modèle a été inventé en 1958 par Rosenblatt. Les limites du Perceptron monocouche ont été mis en évidence par les mathématiciens Minsky et Papert en 1969. Le Perceptron simple ne pouvait que résoudre des problèmes linéairement séparables.

En 1982, les recherches de Hopfield mettent en avant le fait que des RNA pouvaient résoudre des problèmes d'optimisation. Dans le même temps, Kohonen montre que les RNA sont capables de solutionner des problèmes de classification et de reconnaissance. Les RNA sont devenus populaires grâce à l'algorithme de rétropropagation.

I.2. Domaines d'utilisation

De nos jours, les réseaux de neurones artificiels sont utilisés dans de nombreux domaines d'application et dans des secteurs très variés :

- traitement d'images : reconnaissance de caractères et de signatures, compression d'images, reconnaissance de forme, cryptage, classification,
- traitement du signal : filtrage, classification, identification de sources, traitement de la parole,
- contrôle : commande de processus, diagnostic, contrôle qualité, asservissement de robots,
- optimisation : planification, allocation de ressources, gestion et finance,
- simulation : simulation de boîte noire, prévision météorologique, prévision géotechnique, recopie de modèle,
- robotique.

I.3. Les types d'apprentissage

[RAK 01] [WIR 09]

I.3.1. Notion d'apprentissage

L'apprentissage dépend de l'architecture du réseau et de l'environnement du problème. Les règles d'apprentissage, celles de Hebb et de Widrow, ne concernent qu'un seul neurone. Chaque architecture étant unique par ses spécificités, elle nécessite une règle d'adaptation des poids spécifiques aussi.

L'apprentissage, dans le cadre des réseaux de neurones artificiels, est l'inverse de la logique déductive. En effet, pour les RNA, à partir d'observations limitées, on va tirer des généralisations plausibles. Alors que dans le cas de la logique déductive, c'est à partir de connaissances déjà établies que l'on tire des connaissances dérivées.

La notion d'apprentissage rassemble deux phénomènes :

- la mémorisation : il s'agit du fait d'assimiler sous une forme dense des informations éventuellement nombreuses.
- la généralisation : il s'agit de la capacité, à partir d'exemples appris, de traiter des exemples distincts et inédits mais semblables.

Le phénomène de mémorisation est donc un état de sur-apprentissage du modèle. Dans cet état, il ne sait pas prédire correctement, ce phénomène est illustré en figure 6.

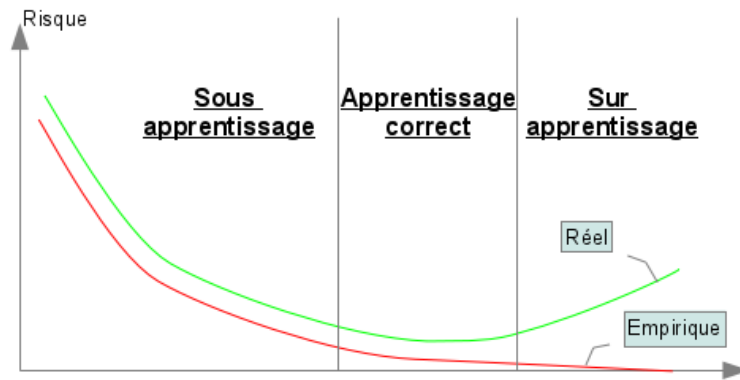


Figure 6 : Différence entre la généralisation et la mémorisation

Le choix des données d'apprentissage est aussi important. En effet, si les données ne sont pas représentatives du processus alors on n'apprendra à notre modèle qu'une partie des données fournies. Il faut donc bien choisir :

- le type de données,
- leur nature,
- leur pertinence.
- la quantité,
- l'ordre d'apprentissage,

I.3.2. Les 3 types d'apprentissage

On distingue trois types d'apprentissage : supervisé, non supervisé et par renforcement.

I.3.2.1. L'apprentissage supervisé

Ce type d'apprentissage doit disposer d'un comportement de référence précis pour pouvoir l'inculquer au réseau de neurones. On doit connaître précisément les données d'entrées et de sorties. L'apprentissage supervisé s'effectue sous le contrôle d'un expert.

I.3.2.2. L'apprentissage non supervisé

Contrairement à l'apprentissage supervisé, ce type d'apprentissage est autonome. Il s'agit d'un système d'apprentissage autodidacte. Les paramètres internes du réseau ne sont modifiés qu'avec les stimuli. L'apprentissage non supervisé construit donc une représentation interne de la connaissance issue de l'environnement. L'apprentissage non supervisé nécessite seulement de connaître des données d'entrées.

I.3.2.3. L'apprentissage par renforcement

Il s'agit d'un type d'apprentissage qui ne nécessite pas de comportement de référence explicite mais seulement d'informations grossières. Il peut également être appelé apprentissage par renforcement.

I.4. Le Perceptron

[ROU 01] [BOU 01] [KOR 01]

I.4.1. Présentation

Le Perceptron est l'un des réseaux de neurones les plus utilisés pour des problèmes de classification ou de régression. Il s'agit du modèle proposé par Rosenblatt et instrumenté avec la règle de Hebb. Nous

allons distinguer trois types de modèles de Perceptron : le Perceptron simple, le Perceptron en parallèle et le Perceptron Multi-Couches.

I.4.1.1. Le Perceptron simple

Il est composé de valeurs d'entrées, d'un biais assimilable à une autre entrée mais variable, d'un système de poids reliant la fonction d'agrégation aux données d'entrées et d'une fonction de transfert qui va donner la valeur de sortie. Voici un schéma de principe :

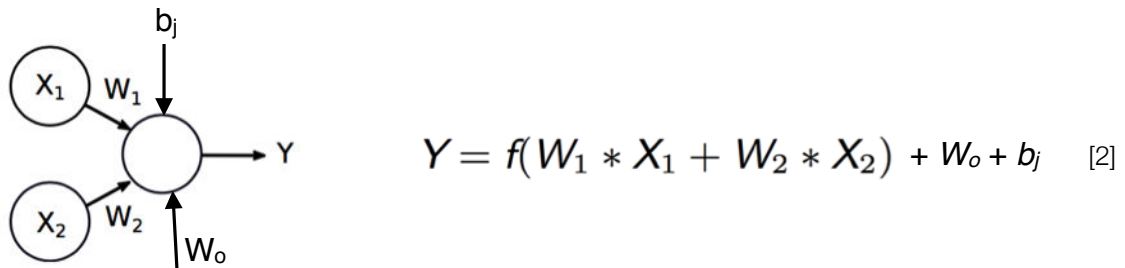


Figure 7 : Schéma de principe d'un Perceptron simple

Avec :

- X_i : entrée du Perceptron ;
- W_i : poids ;
- b_j : biais ;
- Y : sortie du Perceptron.

Ce type de modèle permet de classifier des éléments compris dans des catégories différentes. Dans la mesure où les catégories sont linéairement séparables. Si c'est le cas, alors il est possible de prédire à quelle catégorie appartient un élément.

I.4.1.2. Le Perceptron Multi-Couches

Un Perceptron ne peut construire qu'une frontière linéaire. Mais la majorité des problèmes ne sont pas linéairement séparables. Pour obtenir des frontières plus complexes, il est possible de superposer plusieurs couches de neurones dans un même Perceptron : le Perceptron Multi-Couches (PMC). Le fait d'assembler plusieurs neurones permet de multiplier les capacités d'apprentissage. Cependant les Perceptron qui ont plus de trois couches de neurones sont des Deep Neural Network, il s'agit des modèles utilisés par Google entre autre [MOT 15]. Mais ces types de RNA ne sont pas développés ou utilisés dans la suite de ce mémoire. Voici un exemple d'un réseau multi-couche complètement connecté, c'est à dire que chaque noeud d'une couche est relié à tous les noeuds de la couche précédente :

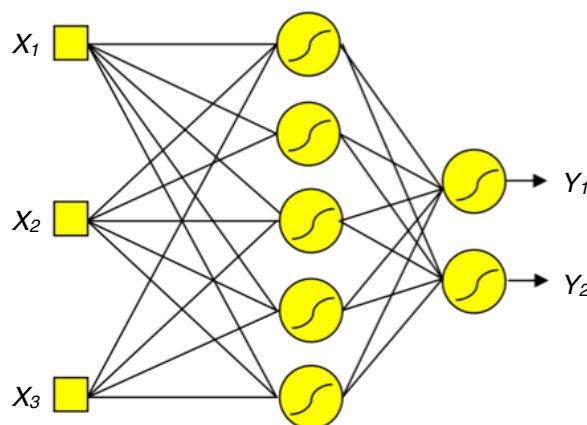


Figure 8 : Schéma de principe d'un Perceptron Multi-Couche

Le réseau ci-dessus est constitué de 5+2 neurones avec 3 entrées. Ce réseau est bien ordonné, on y distingue les couches, et les neurones sont entièrement connectés entre les couches.

Un réseau multi-couche est composé d'une succession de couches dont chacune prend ses entrées sur les sorties de la précédente. Chaque couche (i) est composée de N_i neurones, prenant leurs entrées sur les N_{i-1} neurones de la couche précédente. À chaque synapse est associée un poids synaptique, de sorte que les N_{i-1} sont multipliés par ces poids, puis additionnés par les neurones de niveau i , ce qui est équivalent à multiplier le vecteur d'entrée par une matrice de transformation.

Mettre différentes couches l'une derrière l'autre reviendrait à mettre en cascade plusieurs matrices de transformation et équivaldrait à une seule matrice, produit des autres, s'il n'y avait à chaque couche, la fonction d'activation qui introduit une non linéarité à chaque étape. Un réseau de neurones dont les sorties seraient linéaires n'aurait que peu d'intérêt.

Le PMC présentent divers avantages et inconvénients :

- Avantages :
 - ➡ capacité de généralisation avec des données bruitées,
 - ➡ apprentissage supervisé capable de prédire des phénomènes très complexes,
 - ➡ incrémentalité,
 - ➡ scalabilité (capacité à être mise en oeuvre sur des bases de grande capacité).
- Inconvénients :
 - ➡ modèle boîte noire (fonctionnement interne inconnu),
 - ➡ difficulté de paramétrage (nombre de neurones dans la couche cachée),
 - ➡ problème de convergence (optimum local),
 - ➡ apprentissage statistique (extrapolation,),
 - ➡ danger de sur-apprentissage.

II. FACTEURS DE PERFORMANCES

[SJM 09] [SJM 04]

Le Perceptron Multi-Couche (PMC) est l'un des types de réseaux de neurones le plus utilisé en géotechnique. L'utilisation de RNA doit tenir compte de facteurs importants tels que la détermination des entrées du modèle adéquat, la division des données et leur pré-traitement, le choix de l'architecture de réseau approprié, une sélection rigoureuse de certains paramètres internes qui contrôlent la méthode d'optimisation, les critères d'arrêt et la validation du modèle. Ces facteurs seront expliqués ci-dessous.

II.1. Détermination des entrées du modèle

La première étape importante est de sélectionner les variables d'entrées de notre modèle, c'est ce qui a le plus d'impact sur la performance de notre modèle. En effet, un bon sous-ensemble de variables d'entrées peut considérablement améliorer les performances de réseaux de neurones. Cependant, il ne suffit pas de rentrer un très grand nombre de variables possibles dans notre modèle. Cela va augmenter la taille du réseau mais surtout augmenter la durée de traitement des informations et réduire l'efficacité du réseau.

Habituellement, en géotechnique les variables d'entrées sont choisies à l'avance sur la base sûre des connaissances préalables. Une autre approche utilisée par plusieurs chercheurs est de créer de nombreux réseaux de neurones avec différentes combinaisons de variables d'entrées et au final, sélectionner le réseau qui a le meilleur rendement. Sinon, une technique par étapes peut également être utilisée dans des réseaux distincts qui ont une seule des variables disponibles comme entrée du modèle. Le réseau le plus performant par rapport aux différents critères est alors retenu. Certains auteurs utilisent, le critère d'Akaike pour déterminer si le réseau est performant pour ce problème. Ce processus est répété pour un nombre croissant de variables d'entrées, jusqu'à ce que l'addition de variables supplémentaires ne se traduise par aucune amélioration de la performance du modèle.

Les approches énoncées ci-dessus ont un inconvénient potentiel. En effet, la détermination quant à savoir si une entrée de paramètre est significative ou non, dépend de l'erreur d'un modèle formé, qui est non seulement une fonction des entrées, mais aussi de la structure de modèle et d'étalonnage. Cela peut potentiellement occulter l'impact des entrées de modèles différents. Ce problème peut être surmonté en utilisant des approches sans modèle, ils utilisent des mesures linéaires de dépendance, telle que la corrélation pour obtenir les entrées du modèle significatif avant d'élaborer le modèle du réseau de neurones définitif.

II.2. La division de données

Comme énoncés ci-dessus, les réseaux de neurones artificiels de types supervisés sont semblables à des modèles statistiques traditionnels. En effet, les paramètres du modèle sont réglés pendant la phase d'étalonnage du modèle afin de minimiser l'erreur entre la valeur obtenue par le réseau et celle normalement obtenue.

Les réseaux de neurones artificiels fonctionnent mieux quand le réseau n'est pas obligé d'extrapoler au-delà des valeurs de la plage de données utilisées pour l'apprentissage. En effet, seul les réseaux probabilistes en sont capables. En d'autres termes, si le nombre de degrés de liberté du modèle est trop grand par rapport au nombre de points des données utilisées pour l'étalonnage, le modèle pourrait ne plus correspondre à la tendance générale, comme souhaitée, mais pourrait apprendre les particularités des points de données utilisées pour l'étalonnage conduisant à une "mémorisation", plutôt qu'une «généralisation». Cela s'appelle le phénomène de sur-apprentissage. C'est pour cela qu'il est nécessaire de garder un ensemble de données de validation pour s'assurer que le modèle a appris correctement à partir des données d'apprentissage. On va donc distinguer les données disponibles en deux sous-ensembles :

- un ensemble de formation pour construire le modèle du réseau de neurones,
- un ensemble de validation pour estimer la performance du modèle dans un environnement connu.

Classiquement, l'ensemble de formation représente deux tiers des données disponibles et donc l'ensemble de validation représente un tiers des données disponibles.

Une autre méthode de répartition appelée validation croisée, inventée par Stone en 1974, consiste à séparer les données disponibles en trois sous-ensembles :

- un ensemble de formation, qui représente 70 % de données totales disponibles, pour ajuster les poids du modèle du réseau de neurones,

- un ensemble de test, qui représente 15 % de données totales disponibles, pour vérifier la performance du modèle à différents stades de la formation et de déterminer quand arrêter la formation pour éviter de trop ajuster le modèle,
- un ensemble de validation, qui représente 15 % des données totales disponibles, pour estimer la performance du modèle dans l'environnement de déploiement.

Dans le cadre de l'application pour la géotechnique, les données des sous-ensemble sont choisies généralement de façon arbitraire. Mais cela peut avoir un impact significatif sur les résultats fournis par le réseau. Par exemple, si dans les données disponibles il y a des cas extrêmes qui ne sont pas utilisés pour la phase d'étalonnage alors le modèle ne pourra pas bien performer. Les données de validation permettront de tester la capacité d'extrapolation du modèle, et non sa capacité d'interpolation.

II.3. Les données de pré-traitement

Une fois que toutes les données disponibles sont intégrées à leurs sous-ensembles, il faut pré-traiter ces données pour qu'elles soient adaptées à notre réseau de neurones artificiels. Les plages de pré-traitement sont variables et dépendent de notre fonction d'activation, les plages sont : $[-1;1]$, $[0,1]$.

Le pré-traitement est nécessaire pour que toutes les données reçoivent une attention égale lors du processus. De plus, le pré-traitement permet d'accélérer le processus d'étalonnage. Le pré-traitement des données n'est pas obligatoire mais c'est fortement recommandé. Cependant, certains chercheurs ont mis en avant le fait que l'ajustement des poids du modèle était identique si les données étaient pré-traitées ou non.

II.4. Détermination du modèle de l'architecture

La détermination du modèle de l'architecture du réseau de neurones est l'une des tâches les plus importantes et des plus difficiles dans l'élaboration du modèle de réseau de neurones artificiels. Elle correspond à la sélection du nombre idéal de couches et le nombre optimum de noeuds dans chacune des couches. Malheureusement, il n'y a aucune approche standardisée pour déterminer l'architecture optimale du réseau.

Dans les PMC, on retrouve toujours deux couches qui représentent les variables d'entrées et de sorties. Pour ce qui est des couches cachées, il a été prouvé qu'il ne suffit que d'une couche cachée et de l'exploitation d'une fonction d'activation sigmoïde pour approximer une fonction continue.

Cependant, l'utilisation de plusieurs couches cachées est nécessaire pour modéliser des fonctions complexes dans de nombreuses situations. Lapedes et Farber en 1988, ont proposé un modèle avec seulement deux couches cachées. Selon eux, cela serait suffisant. La première couche cachée a pour rôle d'extraire les caractéristiques locales des données d'entrées et la seconde couche cachée est utilisée pour extraire les caractéristiques globales du modèle de formation. Mais l'utilisation de plus d'une couche cachée ralentit le processus de formation et augmente les chances de se trouver coincé dans des minima locaux du système. Globalement le nombre de couches cachées va dépendre de la nature du problème à résoudre.

Le nombre de noeuds dans les couches d'entrées et de sorties est limité respectivement par le nombre d'entrées et de sorties du modèle. Il n'y a pas de manière directe et précise pour déterminer le nombre idéal de noeuds dans chaque couche. Dans le cadre de notre application des réseaux de neurones artificiels en géotechnique, une procédure d'essai et d'erreur est généralement utilisée pour déterminer le nombre et la connectivité des noeuds de la couche cachée.

II.5. Optimisation du modèle

Le processus d'optimisation des poids de connexion est équivalent à la phase d'estimation des paramètres des modèles statistiques classiques. L'objectif est de trouver une solution globale à ce qui est généralement un problème d'optimisation fortement non-linéaire. La méthode la plus couramment utilisée pour trouver la combinaison optimale des poids d'un PMC est l'algorithme de rétro-propagation. Cette méthode a l'avantage de pouvoir échapper aux minima locaux et donc les solutions obtenues seront optimales ou proches mais le taux de convergence est lent. Elle consiste à ré-ajuster les poids à partir de la valeur prédite et celle désirée en commençant par la fin du PMC.

II.6. Critères d'arrêt

Les critères d'arrêt sont utilisés pour déterminer quand il faut arrêter le processus de formation. Plusieurs approches peuvent être utilisées, la formation peut s'arrêter :

- après la présentation d'un nombre fixe de données d'apprentissage,
- lorsque l'erreur de formation atteint une valeur suffisamment faible,
- après l'absence de modifications de l'erreur d'apprentissage.

Cependant, les critères présentés ci-dessus peuvent conduire le modèle à s'arrêter prématurément ou à sur-apprendre. La technique de validation croisée, énoncée précédemment, est une approche qui peut être aussi utilisée pour surmonter ces problèmes. La technique de validation croisée exige que les données soient divisées en trois ensembles : la formation, les tests et la validation. L'ensemble de la formation est utilisé pour ajuster les poids de connexion. L'ensemble de test mesure la capacité du modèle à généraliser. La formation est arrêtée lorsque l'erreur de l'ensemble de test commence à augmenter. L'ensemble de tests est également utilisé pour déterminer le nombre optimal de nœuds de couche cachée et les valeurs optimales des paramètres internes (taux, durée d'impulsion et poids initiaux de l'apprentissage). L'ensemble de validation est utilisé pour évaluer la performance du modèle, une fois la formation accomplie. Il existe d'autres types de critères d'arrêt.

II.7. La validation du modèle

Une fois la phase d'apprentissage du modèle terminée, la performance du modèle obtenu doit être validée. L'objectif de la phase de validation du modèle est de veiller à ce que le modèle ait la capacité de généraliser dans les limites fixées par les données d'apprentissage. Généralement, la performance du réseau est testée sur un ensemble de données indépendantes qui n'a pas été utilisé pour l'apprentissage du modèle. Si les résultats obtenus sont corrects alors le réseau est considéré comme robuste.

Le coefficient de corrélation r , l'erreur quadratique moyenne **RMSE** et l'erreur absolue moyenne **MAE** sont les principaux critères qui sont généralement utilisés pour estimer la performance de prédiction du réseau de neurones artificiels.

L'erreur quadratique moyenne est la mesure la plus connue de l'erreur et elle a l'avantage que les grandes erreurs reçoivent beaucoup plus d'attentions que les petites erreurs. Contrairement à l'erreur quadratique moyenne, l'erreur absolue moyenne élimine l'importance accordée à des erreurs importantes.

Shahin, en 2005 [SJM 05], a mené une étude pour prédire le tassement des fondations superficielles sur les sols granulaires. Il a constaté que la bonne performance des réseaux de neurones artificiels sur les

données de validation ne garantit pas que le modèle performe bien sur une plage de données similaires à celle utilisée pour l'étalonnage.

C'est pour cette raison que Shahin a proposé une méthode pour tester la robustesse de la capacité de prédiction des modèles de RNA en effectuant une analyse de sensibilité pour étudier la réponse des sorties du modèle à des changements dans ses entrées.

III. LES RNA EN GÉOTECHNIQUE

[SJM 08] [SUL 11]

Contrairement à l'apprentissage traditionnel qui est basé sur des connaissances préalables sur la nature des relations entre les données, les relations qui existent en géotechnique entre les paramètres de granulométrie d'un sol et son pénétrogramme sont inconnues. C'est dans ce cadre là que les RNA montrent tout leurs intérêts. Cependant, les matériaux rencontrés en géotechnique sont complexes à modéliser de part leur nature propre et leurs fortes variabilités. Les propriétés géotechniques des sols sont contrôlées par des facteurs tels que la minéralogie, la structure, la forme des grains et la pression interstitielle.

C'est dans ce cadre que l'application des réseaux de neurones artificiels est utile. En effet, leurs capacités à apprendre de l'expérience les rendent supérieures à d'autres modèles de modélisation car il n'y a pas besoin de faire des hypothèses sur des données sous-jacentes.

Depuis les années 1990, les réseaux de neurones artificiels ont été appliqués avec succès pour quasiment tous les problèmes de géotechnique. Voici une liste des principaux domaines d'utilisation des RNA en géotechnique :

- La prédiction du degré de sur-consolidation (OCR),
- La détermination du potentiel de liquéfaction d'un sol,
- La prédiction du tassement des fondations,
- L'évaluation de la capacité portante des pieux,
- La prédiction des paramètres de compactage pour un sol cohérent,
- Le contrôle de compactage des remblais réalisés en matériaux non-cohésifs.

III.1.1. Caractérisation des sols à partir d'un essai de pénétration CPT

[KGT 01] [NCH 07] [BBS 06]

Ils existent des travaux qui ont utilisé des systèmes de réseaux de neurones pour prédire la nature des sols à partir d'un essai CPT. Les données prédites sont le pourcentage de sable, limon et d'argile. Ces études s'appuient sur les travaux de Robertson. Certaines de ces études ont comparé les résultats obtenus avec d'autres système de classification (arbre de décision, machines à vecteurs de support, ...) et les meilleures performances ont été obtenu avec un RNA. Suivant les études les variables d'entrées varient mais la résistance de pointe et le frottement sont toujours pris en compte.

CHAPITRE 3 : SYSTÈMES DE CLASSIFICATION AUTOMATIQUES

Les RNA sont les outils de modélisation les plus utilisés mais il existe d'autres modèles de classifieur automatique. Ces méthodes, comme pour les RNA, ont pour objectif de rechercher à partir de données disponibles un modèle explicatif ou prédictif. Dans cette partie, les principaux modèles qui sont proposés par le logiciel Matlab via son application « *Classification Learner* » sont expliqués.

I. LES ARBRES DE DÉCISION

Les arbres de décision font partis de la famille des graphes d'induction, ceux sont les modèles les plus populaires. Les raisons de ce succès sont qu'ils sont faciles à mettre en oeuvre et les résultats qu'ils fournissent sont aisés à interpréter. Il s'agit de méthode statistique non-paramétrique, ils permettent de classer un ensemble de données décrits par des variables qualitatives et quantitatives. La figure 9 représente le fonctionnement d'un arbre de décision de type booléen qui permet de classer les trois valeurs A, B et C.

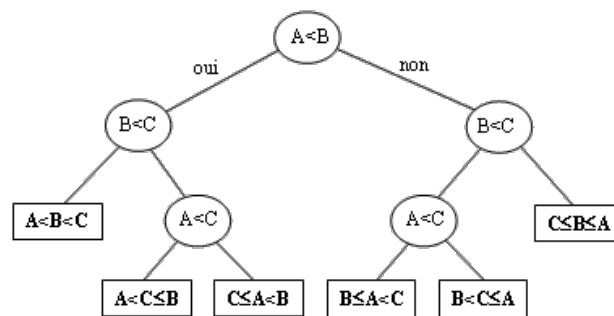


Figure 9 : Structure arbres de décision

Dans l'application « *Classification Learner* », il est proposé de tester trois modèles d'arbres de décision qui ont chacun des spécificités :

Type de classifieur	Interprétabilité	Flexibilité du modèle
Simple Tree	Facile	Faible (4 classes de sorties au maximum)
Medium Tree	Facile	Moyenne (20 classes de sorties au maximum)
Complex Tree	Facile	Forte (100 classes de sorties au maximum)

Tableau 2 : Caractéristiques des différents modèles d'arbres de décision

II. LES ANALYSES DISCRIMINANTES

L'analyse discriminante est l'une des plus anciennes techniques de discrimination. Ce modèle a pour objectif de caractériser de manière multidimensionnelle l'appartenance d'individus à des groupes prédéfinis. Il s'agit d'un algorithme très populaire car il est facilement utilisable, interprétable et performant. L'analyse du discriminant est adaptée aux larges ensembles de données. Pour entraîner ce type de classifieur, la fonction d'estimation optimale estime les paramètres de distribution de la fonction gaussienne de chaque classe. Il existe deux types d'analyse du discriminant : linéaire et quadratique.

Type de classifieur	Interprétabilité	Flexibilité du modèle
Linear Discriminant	Facile	Faible (création de frontières linéaires entre chaque classe)
Quadratic Discriminant	Facile	Faible (création de frontières non-linéaires entre chaque classe)

Tableau 3 : Caractéristiques des différentes analyses discriminantes

III. LES MACHINES À VECTEURS DE SUPPORT

Les machines à vecteurs de support (Support Vectoriel Machine SVM) sont une méthode de classification relativement récente. Leur objectif est de déterminer le meilleur hyperplan qui sépare les valeurs des données en fonction de leurs classes d'appartenance. L'hyperplan le plus performant est celui qui a la marge, par rapport à l'hyperplan séparant les différentes classes, la plus grande. Aucun point ne doit être contenu dans la marge. L'hyperplan va donc devenir la frontière entre les différentes classes. Ils vont transformer de façon non-linéaire les données d'entrées et c'est par la suite qu'ils vont créer la frontière optimale.

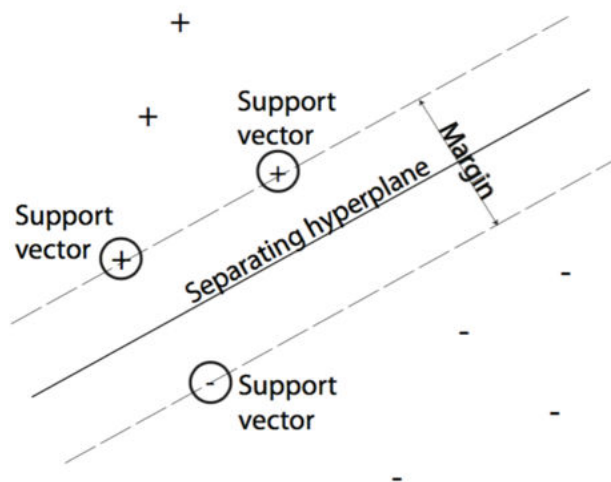


Figure 10 : Schéma des frontières des SVM

Voici les caractéristiques des six modèles de machines à vecteurs de support proposés par le « Classification Learner » :

Type de classifieur	Interprétabilité	Flexibilité du modèle
Linear SVM	Facile	Faible (création de frontières linéaires entre chaque classe)
Quadratic SVM	Difficile	Moyenne
Cubic SVM	Difficile	Moyenne
Fine Gaussian SVM	Difficile	Difficile - Diminue l'arrangement des échantillons
Medium Gaussian SVM	Difficile	Moyenne
Coarse Gaussian SVM	Difficile	Faible

Tableau 4 : Caractéristiques des différents modèles de SVM

IV. LA MÉTHODE DES K PLUS PROCHES VOISINS

Il s'agit d'une méthode qui a pour objectif de classer des données cibles (classe méconnue) en fonction de leurs distances par rapport à des points constituant un échantillon d'apprentissage (c'est-à-dire dont la classe est connue). Elle s'appelle Nearest Neighbor Classifiers (KNN) en anglais. Cependant cette méthode perd de sa fiabilité lorsque le nombre de dimensions est élevé. Les modèles sont difficiles à interpréter.

Type de classifieur	Interprétabilité	Flexibilité du modèle
Fine KNN	Difficile	Distinctions fines entre classes. Le nombre de voisin est fixé à 1.
Medium KNN	Difficile	Distinctions moyennement détaillées entre classes. Le nombre de voisin est fixé à 10.
Coarse KNN	Difficile	Distinctions grossières entre classes. Le nombre de voisin est fixé à 100.
Cosine KNN	Difficile	Distinctions moyennement détaillées entre classes, utilisant une distance au cosinus. Le nombre de voisin est fixé à 10.
Cubic KNN	Difficile	Distinctions moyennement détaillées entre classes, utilisant une distance cubique. Le nombre de voisin est fixé à 10.
Weighed KNN	Difficile	Distinctions moyennement détaillées entre classes, utilisant un poids de distance. Le nombre de voisin est fixé à 10.

Tableau 5 : Caractéristiques des différents modèles de KNN

BILAN

L'objectif de ce travail est de déterminer la nature d'un sol à partir d'un essai de pénétration. Aujourd'hui, le seul moyen d'obtenir la nature des sols présents est de réaliser des carottages lorsqu'on effectue un essai Panda. Certains essais de pénétration (CPT et SPT) fournissent en plus de la résistance du sol une carotte permettant d'identifier les matériaux sondés. Mais contrairement à l'essai Panda, ils sont plus cher, plus long à réaliser et demande des conditions d'accès plus contraignantes. Cependant, malgré tous les avantages du Panda, on ne prélève pas d'échantillon de sol lors du sondage et donc on ne dispose d'aucune information sur la nature du sol.

Cependant, on a constaté empiriquement qu'il existe des liens entre un type de sol et la forme signal pénétrométrique. Toutefois, ces liens ne sont pas connus, même si un utilisateur averti va pouvoir déduire du signal pénétrométrique des indices lui permettant d'estimer la nature du sol sondé.

L'objectif de la méthodologie proposée s'appuie sur l'idée de simuler le fonctionnement d'un cerveau d'un utilisateur expérimenté. Pour cela, nous allons nous appuyer sur la théorie des RNA. L'utilisation des RNA dans le domaine de la géotechnique a déjà été employé avec succès. Le choix d'utiliser des RNA est à priori opportun. Ils devraient nous permettre de prédire la nature des matériaux sondés.

Cependant, on sait que la phase d'entraînement est primordiale pour un bon fonctionnement du modèle. La quantité et la qualité du set d'apprentissage doivent être déterminés judicieusement parce que les PMC sont basés sur un apprentissage statistique.

Dans la partie suivante, nous allons dans un premier temps détailler la conception et la réalisation de la base de données. Puis nous expliquerons les paramètres d'entrée du RNA, les aspects du signal qu'ils représentent et la manière de les calculer. Et les sorties du modèle, leurs significations d'un point de vue de la GTR. Puis nous allons étudier, les variations obtenus sur le paramétrage et les performances du PMC pour deux section de pointes différentes et si il faut prendre en compte ou non la profondeur critique pour un essai réalisé dans une chambre de calibration sur un matériau homogène. Enfin nous évaluerons les performances du PMC par rapport à d'autres systèmes de classifieurs automatiques.

PARTIE II : ÉTUDE

CHAPITRE 1 : LA BASE DE DONNÉES

On peut se demander pourquoi préférer une base de données plutôt qu'un tableur ou qu'un document. Les bases de données sont bien plus compétentes pour retrouver ou trier des informations. Les informations sont stockées de la même façon qu'un tableur, c'est à dire en liste, mais les bases de données ont comme principale capacité de pouvoir afficher les informations sous formes de formulaires. Les formulaires peuvent contenir seulement les informations souhaitées et facilitent leurs saisies. De plus, ils permettent d'accélérer la saisie des données tout en maintenant leurs cohérences.

De nos jours, nous sommes en contact permanent avec des bases de données : un carnet d'adresses, un calendrier, ... Cependant pour ces exemples, leurs fonctions sont très limitées et ne peuvent stocker qu'un nombre limité d'informations.

A contrario, les bases de données peuvent stocker une grande quantité d'informations et leurs présentations sont spécifiques à chaque cas. De plus, on peut relier diverses informations entre elles, aspect qui est quasiment impossible avec un simple tableur [PRA 16].

I. INTRODUCTION

[SLA 00]

L'informatique a toujours connue une évolution rapide et impressionnante. Bien sûr le domaine des bases de données a suivi ce mouvement.

En 1964, le terme base de données, « database » en anglais, est apparu dans notre langage. Il désigne une collection d'informations partagées par différents utilisateurs d'un système d'informations.

Il existe quatre types de bases de données :

1. Les bases de données hiérarchiques : il s'agit du système le plus ancien, les données sont liées dans une structure arborescente où chaque enregistrement n'a qu'un seul possesseur. Par exemple, le *canard* appartient à la famille des *anatidés* qui elle-même appartient à l'ordre des *ansériformes* qui lui-même appartient à la classe des *oiseaux* qui elle-même appartient au sous-embouchement des *vertébrés* qui lui-même appartient au règne *animal*. Ces bases de données sont cependant inadaptés pour décrire des structures complexes.
2. Les bases de données relationnelles : les données sont stockées sous forme de tableaux à deux dimensions. Ces modèles ont été introduits par Edgar F. Codd en 1970. Les logiciels qui permettent de créer et d'utiliser ces bases de données sont appelés système de gestion de bases de données relationnelles (S.G.B.D.R). Ces systèmes sont créées par un langage SQL, ce langage permet de créer ou d'interroger les tables. Contrairement aux bases de données hiérarchiques, une donnée peut avoir plusieurs possesseurs.
3. Les bases de données déductives : il s'agit d'une base d'information permettant d'extraire ou de déduire des informations de données stockés précédemment dans la base. Ce système utilise deux types de spécification : les faits et les règles.

4. Les bases de données objets : les données sont stockées sous forme de collections d'objets. Il s'agit d'un conteneur logique qui englobe des informations et des traitements relatifs à une chose du monde réel. Les logiciels et les langages sont spécifiques aux bases de données orientés objet.

Dans la suite de cet état de l'art, nous allons nous intéresser seulement aux bases de données relationnelles en décrivant leurs conceptions et leurs utilisations. En effet, parmi les quatre structures proposées il s'agit du modèle le plus adéquat à l'utilisation que nous souhaitons en faire. De plus, le langage SQL est le code le plus utilisé dans le monde de nos jours.

II. CONCEPTION D'UNE BASE DE DONNÉES

[DLS 16]

La base de données constitue le coeur de notre système d'information. Sa conception est la tâche la plus complexe du développement du système d'information.

Il est préconisé de concevoir notre base de données en plusieurs étapes. Chaque étape correspond à un niveau d'abstraction différent :

1. Concept : cette étape représente le contenu de la base en termes conceptuels, aucune notion informatiques n'est prise en considération lors de cette étapes.
2. Choix : il s'agit d'effectuer un choix de logiciel de gestion de base de données
3. Ecriture : cette étape est utilisée pour décrire les méthodes d'organisation et d'accès aux données de la base.

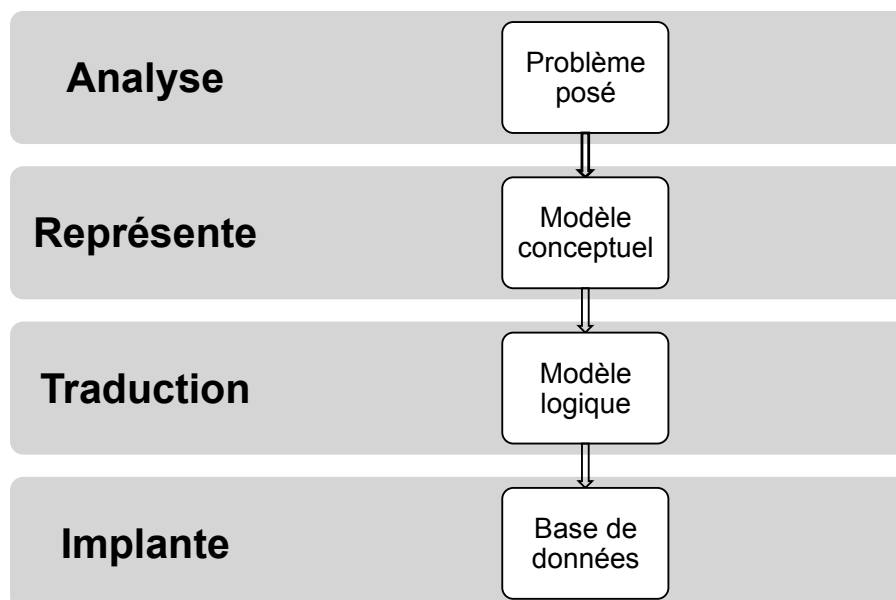


Figure 11 : Schéma de principe de développement d'une BDD

II.1. Objectifs de la base de données

Le système de RNA mis en place nécessite une base de données permettant l'apprentissage, le test et la validation du modèle. Actuellement, les informations étaient stockées de différentes façons dans divers dossiers et sous divers formats.

Dans ce contexte, on comprend la nécessité de la création d'une base de données regroupant toutes les informations de façon claire et précise permettant ainsi leurs accès plus faciles et leurs tris tout en étant sûr de la fiabilité des données et qu'il s'agisse de l'ensemble des données disponibles. La mise en place d'une BDD est précieuse pour une entreprise. Elle devra contenir diverses informations sur les résultats des études menées par Sol-Solution. Son utilisation servira dans un premier temps à alimenter le RNA mais elle pourra servir également à d'autres projets.

Il y aura à la fin deux bases de données pour notre étude, une base de données nommée « essais in situ » et une autre « essais laboratoire ». La première base de données correspond aux essais réalisés in situ par l'entreprise Sol-Solution, plus particulièrement par le pôle étude géotechnique, dans le cadre d'études spécifiques, exemple : création d'un lotissement. Pour la seconde, il s'agit de la base de données des essais réalisés en laboratoire par Polytech' Clermont-Ferrand pour les travaux de thèse de Chaigneau (2001) principalement. Les informations pour des essais réalisés in-situ ou en laboratoire sont différentes, c'est pour cela qu'il a été préféré de séparer ces deux cas et donc de créer deux bases de données.

II.2. Modélisation conceptuelle

La modélisation est une étape capitale de la conception de la base de données. En effet, lors de cette étape, le contenu de la base ainsi que les relations entre les données sont déterminées. Voici les définitions des éléments de base du modèle :

- les entités : il s'agit d'un objet pouvant être identifié. On distingue les entités régulières et faibles, la différence entre ces deux entités est que pour une entité faible son existence dépend de l'existence d'une autre entité.
- les attributs : il s'agit des caractéristiques des entités. Un attribut peut être obligatoire ou facultatif.
- les relations : il s'agit des liens entre les entités.
- l'identifiant : il s'agit d'un attribut d'une entité permettant d'identifier une seule entité à l'intérieur de l'ensemble.

II.2.1. Modélisation de la base de données *in-situ*

La structure de la base de données des essais in-situ est représentée en figure 12. On distingue quatre entités qui sont : Sites, Sondages, mesures_Panda et mesures_labo. Elles sont reliées entre elles suivant différents attributs.

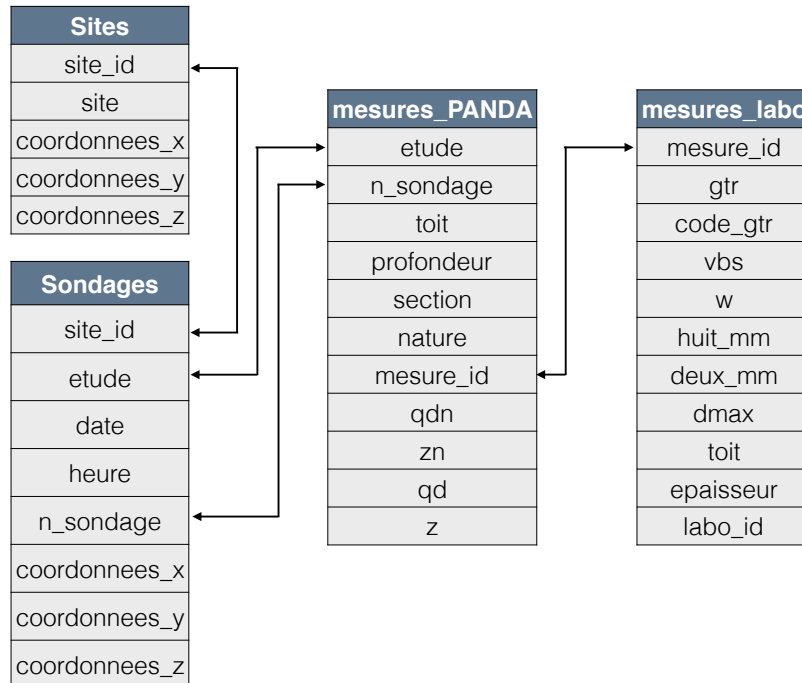


Figure 12 : Structure de la base de données in-situ

II.2.2. Modélisation de la base de données de laboratoire

La structure de la base de données des essais de laboratoire est représentée en figure 13. On distingue trois entités qui sont : mesures_granulo, mesures_Panda et mesures_labo. Elles sont reliées entre elles suivant différents attributs.

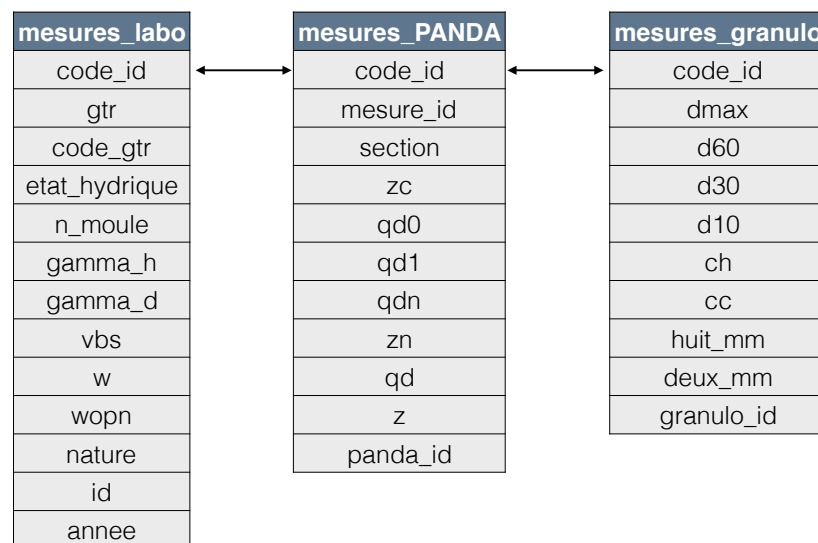


Figure 13 : Structure de la base de données laboratoire



II.3.Choix d'un SGBD

Maintenant que le système conceptuel est déterminé et fixé, l'étape suivante est de choisir parmi tous les logiciels de SGBD⁵ celui qui correspond le mieux à nos besoins. Il faut donc déterminer quels sont les besoins d'un utilisateur lambda et les contraintes de réalisation.

Dans notre cas, voici les différentes contraintes initiales :

- un système de gestion simple et intuitif,
- un système permettant de gérer de grandes quantités,
- un système pouvant être installé sur tous les ordinateurs,
- un système pouvant être mis en réseau,
- un logiciel gratuit.

Une fois nos contraintes déterminées, il faut choisir le logiciel le plus adéquat. Voici un tableau comparatif non exhaustif :

Logiciel	Avantages	Inconvénients
SQL Server 	<ul style="list-style-type: none"> • Administration aisée • Indépendance entre les diverses bases • Une des bases les plus performantes sous Windows • Niveau de SQL très près de la norme SQL • Services Web • Compression des données et des sauvegardes 	<ul style="list-style-type: none"> • Distributions fortement liées au système d'exploitation • Jungle des versions, mais fonctionnalités cantonnées dans les éditions Enterprise, Developer et Standard • Mono-plateforme (MS Windows) • Pas de prise directe sur les tables système • Pas certifié SQLJ, pas d'intégration Java, orientation C# • Pas de contraintes d'unicité multi null
Access 	<ul style="list-style-type: none"> • Très puissant et très ludique • Il contient une grande série d'outils de conversion de données, pour récupérer ou exporter depuis presque n'importe quel format vers pratiquement n'importe quel format. • Les macros permettent à des néophytes de se lancer dans une forme 'allégée' de l'automatisation. • Quantité d'assistants dirigeant l'utilisateur vers une première solution. • Forte intégration à la Suite Microsoft Office/VBA, déjà fortement répandue en entreprise 	<ul style="list-style-type: none"> • Le moteur JET étant un moteur "Fichier", il est gourmand en ressources réseau et ne convient pas pour les applications distantes. • Le système de verrouillage des enregistrements peut induire des messages d'avertissements si deux utilisateurs cherchent à écrire simultanément dans une même page de données. La consultation est par contre multi-utilisateur. • Mono-plateforme (MS Windows) • N'implémente pas complètement les normes SQL.

⁵ SGDB : Système de Gestion de Base de Données


Logiciel	Avantages	Inconvénients
MySQL 	<ul style="list-style-type: none"> • Solution très courante en hébergement public • OpenSource, bien que les critères de licence soient de plus en plus difficiles à supporter • Facilité de déploiement et de prise en main. • Plusieurs moteurs de stockage adaptés aux différentes problématiques, configurable au niveau table. 	<ul style="list-style-type: none"> • Ne supporte qu'une faible partie des standards SQL-92 • Support incomplet des triggers et procédures stockées • Assez peu de richesse fonctionnelle • Manque de robustesse avec de fortes volumétries • Pas d'héritage de table • Pas de vue matérialisée • Pas de sauvegarde consistante à chaud
PostgreSQL 	<ul style="list-style-type: none"> • OpenSource et gratuit • Fiable et relativement performant, tout en restant simple d'utilisation • Supporte la majorité du standard SQL-92 et possède en plus un certain nombre d'extensions (Java, Ruby, PL-SQL). • Très riche fonctionnellement, notions d'héritage de tables, multitude de modules • Simple d'utilisation et d'administration 	<ul style="list-style-type: none"> • La modification du fichier de sécurité pg_hba.conf nécessite un reboot pour être prise en compte • Sauvegardes peu évoluées • Supporte les bases de moyenne importance • Pas d'ordonnanceur intégré • Pas de requêtes récursives
SQLite 	<ul style="list-style-type: none"> • OpenSource et gratuit • Le plus petit moteur SGBDR du marché • Simple d'utilisation et d'administration • Aisément installable 	<ul style="list-style-type: none"> • Fonctionnalités minimales • Pas d'intégrité référentielle • Ne supporte pas les jointures externes • Pas d'ordonnanceur intégré • Volumétrie • Pas de vue matérialisée • Pas de partitionnement • Pas de notion de rôles, pas de hiérarchisation de groupes, gestion de la sécurité minimaliste

Tableau 6 : Synthèse des différents logiciels de SGDB

Au vue des informations du tableau ci-dessus, le choix du logiciel c'est porté sur le logiciel PostgreSQL. En effet, il s'agit d'un des logiciels répondant le mieux aux besoins de l'entreprise et son principal atout est d'être en opensource et gratuit.

II.4.Modélisation physique

Lors de cette phase, le schéma conceptuel va être traduit de façon à être compris par notre logiciel de SGBD : PostgreSQL. La création d'une base de données peut s'effectuer de deux manières sous PostgreSQL : soit de manière manuelle en saisissant le script de création, soit de manière graphique grâce à l'interface graphique.

Il a été choisi d'écrire le script, car cela nous permettaient de mieux gérer la base de données et le langage SQL nous étaient déjà connu.

Le SQL est un langage proche du discours humain, son acronyme signifie : Structured Query Language. Il est reconnu comme étant la norme officielle de langage de requête relationnelle.

Les scripts de création des bases de données se trouvent en annexe 1 et 2.

III. ANALYSE DES DONNÉES DISPONIBLES

Une fois les différentes structures de base de données créées et remplies, il s'agit de faire un constat du nombre de données disponibles.

III.1. Base de données laboratoires

Il s'agit d'une base de données d'essais pénétrométriques Panda réalisés dans des moules de 80 cm de profondeur et de 40 cm de diamètre. Les essais ont été réalisés sur divers matériaux : sable, argile, limon, grave, ... avec des états hydriques différents.

Pour chacun des matériaux à un état hydrique spécifique, cinq moules de ce matériau ont été réalisés. Mais chaque moule a une densité humide et sèche différente. Pour chaque moule, trois essais Panda ont été réalisés. Les trois essais forment un triangle dans le moule, leur espacement permet d'éviter les efforts de bord. Au total, nous avons 249 matériaux testés et donc au total 744 pénétrogrammes. La répartition des pénétrogrammes selon les classes GTR est présentée en figure 14.

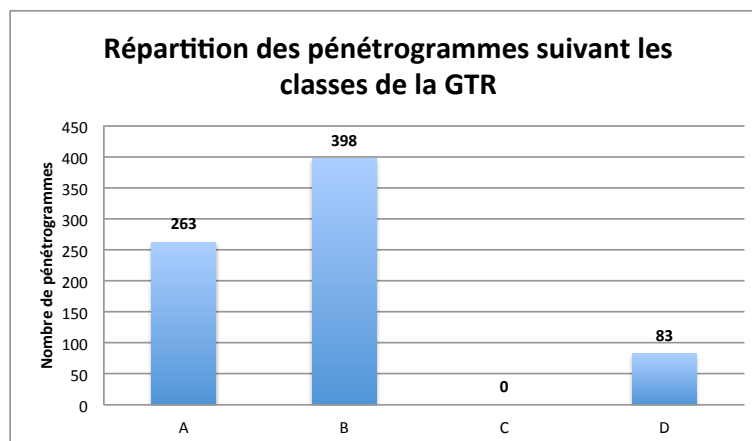


Figure 14 : Répartition des données suivant la classification GTR

Etant donnée la faible profondeur des essais, ils sont réalisés avec des pointes de 2 cm². Mais pour certains matériaux trop lâches, il est possible d'utiliser des pointes plus grosses : 4 cm² ou 10 cm². Les essais réalisés avec une pointe de 2 cm² représentent 97,80% des données disponibles, soit 729 pénétrogrammes. La répartition des pénétrogrammes, selon la section de la pointe, est illustrée en figure 15.

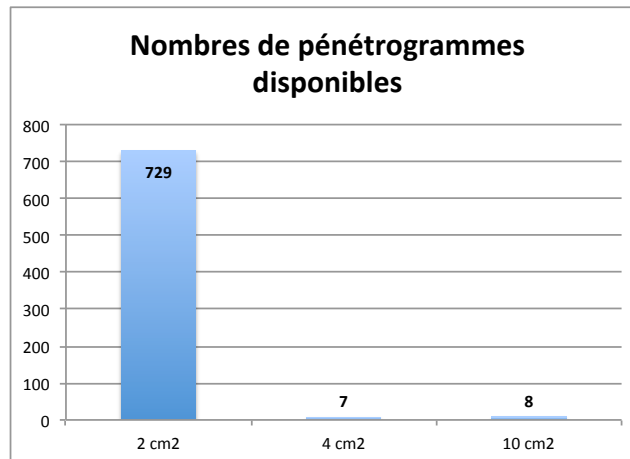


Figure 15 : Répartition des données suivant la section de la pointe

III.2. Base de données in-situ

Cette base de données s'appuie sur la totalité des études réalisées par Sol-Solution soit environ 3000 études, mais avec des conditions spécifiques. Ces conditions sont les suivantes :

- le sondage Panda doit être réalisé à côté d'un carottage,
- au moins un des matériaux de la carotte doit être caractérisé suivant la classification GTR.

Si une de ces deux conditions n'est pas respectée alors le sondage n'est pas conservé dans notre étude et n'est pas saisi dans la base de données.

Au total, 140 sondages ont été conservés, répartis sur 122 sites. Les matériaux rencontrés sont présentés dans la figure 16.

Répartition des données suivant les classes GTR

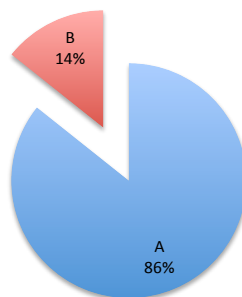


Figure 16 : Répartition des données disponibles suivant leurs classes GTR

On peut observer que principalement les sols rencontrés et classifiés selon la GTR sont des sols de type A. Il s'agit de sols fins présents dans la région d'Auvergne.

CHAPITRE 2 : PARAMÈTRES D'ENTRÉE ET DE SORTIE DU RNA

I. PRÉSENTATION DES PARAMÈTRES D'ENTRÉE DU RNA

Le signal pénétrométrique ne peut pas être transmis directement au modèle de RNA. Le signal pénétrométrique doit être caractérisé en fonction de plusieurs paramètres qui vont créer une carte d'identité unique à chaque pénétrogramme et représentant au mieux celui-ci. Ces paramètres sont les données d'entrée du RNA.

Initialement, 26 paramètres permettaient de créer une empreinte du signal. Les paramètres proviennent de différents domaines comme la reconnaissance de forme ou vocale. Une analyse de sensibilité des 26 paramètres par la méthode One-Factor-At-A-Time (OAT) a été réalisée en vue d'exclure des paramètres qui ne sont pas influents sur les performances du RNA. En effet, lors de la création d'un RNA il faut optimiser au maximum le nombre de paramètres d'entrées. Les objectifs de l'analyse de sensibilité sont :

1. La réduction de la complexité du modèle,
2. L'amélioration de la compression de l'information car la performance des RNA baissent de façon exponentielle avec l'augmentation du nombre de paramètres d'entrées,
3. L'augmentation de la vitesse d'exécution,
4. L'augmentation de la performance de classification.

La méthode consiste à modifier chaque paramètre d'entrée d'une valeur de 10, 50 et 100% de son écart-type et la variance. Les paramètres internes du RNA sont fixes durant cette l'analyse. L'effet de chaque modification est analysé en fonction de la variation de l'erreur finale du modèle, afin de déterminer les paramètres les plus pertinents pour le RNA. Les résultats obtenus sont présentés en figure 17.

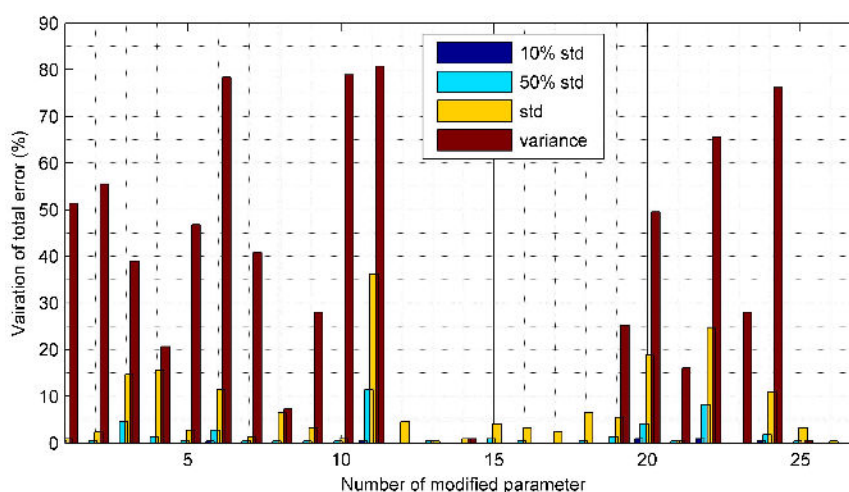


Figure 17 : Analyse de la sensibilité des caractéristiques

On peut observer que certains paramètres ne sont pas influencés, par la modification de leur valeur. Ces paramètres peuvent donc être supprimés. Mais l'analyse de sensibilité permet de mettre en évidence que certains paramètres sont très sensibles aux variations donc qu'ils sont influents sur les résultats de

prédiction du RNA. Au final, 17 paramètres ont été retenus pour modéliser le signal. Ils sont divisés en 4 catégories qui analysent séparément un aspect du signal. Ils vont être détaillés par la suite.

I.1. Analyse statistique

L'analyse statistique du signal pénétrométrique de q_d comprend 9 paramètres ayant comme objectif de décrire un ensemble de valeurs. Les deux premiers paramètres sont des caractéristiques de position. Ensuite, les paramètres 3 à 7 sont des caractéristiques de dispersion des valeurs. Puis les deux derniers paramètres statistiques (8 et 9) sont des caractéristiques de forme.

1. La valeur moyenne de q_d [MPa],
2. La valeur médiane de q_d [MPa],
3. L'écart-type de q_d [MPa],
4. Le coefficient de variation de q_d [%],
5. La variance de q_d [MPa],
6. L'étendue de q_d [MPa],
7. L'écart interquartile de q_d [MPa],
8. L'asymétrie de q_d [-],
9. Le coefficient d'aplatissement de q_d [-].

Le paramètre 4, correspondant au coefficient de variation de q_d , peut être également appelé écart-type relatif, il s'agit d'une mesure de dispersion relative. Le coefficient de variation n'a pas d'unité, il se calcule comme le ratio de l'écart-type rapporté à la moyenne. Il permet de prendre en compte l'ordre de grandeur de la moyenne.

Le paramètre 5, correspondant à la variance de q_d , est une mesure qui indique de quelle manière la série de données se disperse autour de sa moyenne. Sa formule consiste en la moyenne de la somme du carré des déviations de la série, elle s'écrit de la manière suivante :

$$V = \frac{1}{P-1} \sum_{i=1}^P (q_{di} - \overline{q_d})^2 \quad [3]$$

P : nombre de données disponibles ($P = 1, \dots, i$)

q_{di} : valeur q_d du point i

$\overline{q_d}$: valeur moyenne de l'ensemble $\{q_{di}\}$

Le paramètre 6, correspondant à l'étendue de q_d , correspond à la plage de valeurs des données, ce paramètre donne une information sur l'étendue des valeurs. Il se calcule en effectuant la différence de la valeur maximale de q_d et la valeur minimale de q_d .

Le paramètre 7, correspondant l'écart interquartile de q_d , correspond à la taille de l'intervalle situé au centre de la série et incluant 50% des observations. Plus l'écart interquartile est grand, plus la dispersion des observations est forte. Il se calcule de la manière suivante :

$$E_{interquartile} = Q_3 - Q_1 \quad [4]$$

Q_1 : quart de valeurs inférieures au premier quartile

Q_3 : quart de valeurs supérieures au troisième quartile

Le paramètre 8, correspondant à l'asymétrie de q_d , permet d'analyser la symétrie de la série, il est également appelé coefficient de Skewness. Lorsque la distribution est symétrique (voir figure 18) alors le coefficient est nul, cas (b). Lorsque la distribution possède une forte queue vers la droite alors le coefficient est positif, c'est-à-dire que la distribution est décalée à gauche de la médiane, cas (c). Si le coefficient est négatif alors il s'agit du cas inverse, cas (a).

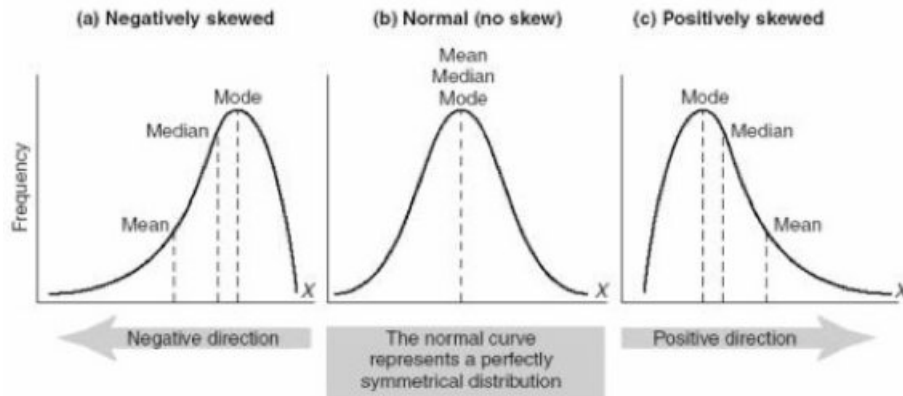


Figure 18 : Schéma distribution du coefficient de Skewness

Le paramètre 9, correspondant au coefficient d'aplatissement de q_d , donne une information sur les queues de distribution, il est également appelé kurtosis. Une valeur de coefficient positive indique que les queues comptent plus d'observations que dans une distribution normale. Une valeur négative indique que les queues comptent moins d'observations qu'une distribution normale.

I.2. Analyse non linéaire

L'objectif de l'analyse non linéaire est de quantifier le contenu moyen en information du signal, à partir de la théorie probabiliste de l'entropie de Shannon. L'entropie de Shannon permet de mesurer l'état d'organisation d'une structure, elle est utilisée dans le domaine de la reconnaissance de forme (RdF) pour optimiser les systèmes. De plus, une entropie des paramètres statistiques est calculée. Les formules utilisées sont les suivantes :

$$E_{Shannon}(q_d) = - \sum_i q_{di}^2 \log(q_{di}^2) \quad [5] \quad E_{log}(q_d) = \sum_i \log(x_i^2) \quad [6]$$

Chaque x_i représente un paramètre statistique et q_{di} les mesures de résistance à une profondeur z donnée. On obtient donc les paramètres 10 et 11 nommés de la façon suivante :

10. L'entropie de Shannon [bits],
11. L'entropie des paramètres statistiques [bits].

I.3. Analyse morphologique du signal

L'analyse morphologique a comme objectif de fournir des paramètres descriptifs de la forme du signal. Pour effectuer cette analyse, le signal pénétrométrique va être divisé en portions de 10 cm consécutives pour évaluer les paramètres proposés pour l'analyse morphologique. Ensuite, une moyenne de toutes les valeurs de chaque morceau de signal est calculée. Les paramètres retenus sont les suivants :

12. Changement du signe de la pente du pénétrogramme [-],
13. Nombre de croisements du signal pénétrométrique avec la ligne de tendance [-],

14. Longitude de la forme de l'onde [-],
15. Coefficient de la droite de régression du pénétrogramme [-],
16. Coefficient indépendant de la droite de régression [-].

Le paramètre 12, correspondant au changement du signe de la pente du pénétrogramme, signifie le nombre de fois que le pénétrogramme va passer d'une pente positive à une pente négative alternativement. Cette valeur est divisée par le nombre d'intervalles de 10 cm que contient le pénétrogramme.

Pour le paramètre 13, correspondant au nombre de croisements du signal pénétrométrique avec la ligne de tendance, il s'agit de compter le nombre de fois que le signal brut va croiser la ligne de tendance sur l'intervalle choisie précédemment. Cette valeur est moyennée par le nombre d'intervalles total. La ligne de tendance peut être calculée de différentes manières, par exemple par la méthode des moindres carrés.

Le paramètre 14 peut être estimé à une forme d'entropie, il se calcule de la façon suivante :

$$\Delta q_d = \sum_i q d_{i+1} - q d_i \quad [7]$$

Les paramètres 15 et 16 sont basés sur l'hypothèse que le signal pénétrométrique suit une tendance linéaire. Cette hypothèse est réaliste pour les essais réalisés en laboratoire dans des moules de calibration avec un matériau homogène et seulement pour la partie supérieure à la profondeur critique Z_c . Cependant, on comprend vite que pour des essais in-situ l'hypothèse est difficilement valide. Ces paramètres sont donc modélisés à partir d'une tendance non linéaire automatique.

I.4. Analyse spectrale

Avant de pouvoir calculer le paramètre 17, le signal doit être modifié car le pas d'échantillonnage doit être constant pour tous les essais (figure 19). L'enfoncement minimal vaut 1 mm. Donc, les pénétrogrammes sont modifiés avec une fréquence d'échantillonnage de 1 mm. Les valeurs manquantes sont calculées par interpolation linéaire. Après l'application de la Transformée Rapide de Fourier (FFT), la puissance spectrale maximale du bruit du signal est calculée pour obtenir la valeur du paramètre 17 [MPa²/Hz].

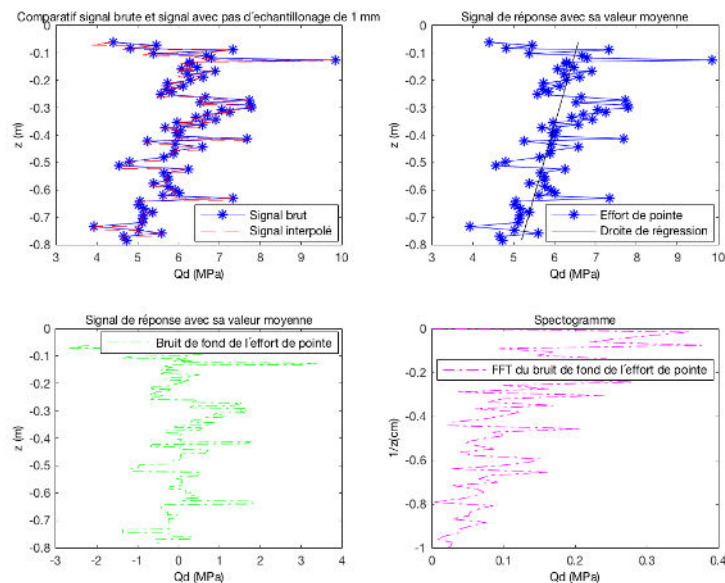


Figure 19 : Analyse spectrale

II. PRÉSENTATION DES PARAMÈTRES DE SORTIE DU RNA

Les sorties désirées sont la nature du sol. Pour l'ensemble des études, et donc des données disponibles, les sols ont été classifiés selon la nomenclature GTR (NF-P 11-300). Il a été proposé 4 classes de sorties différentes sur la base des paramètres granulométriques des matériaux. Nos références sont le passant à 80 μm et à 2 mm (voir figure 20).

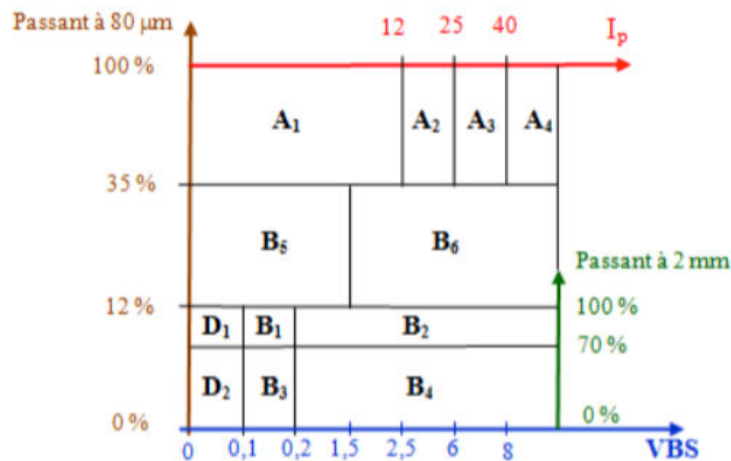


Figure 20 : Classification GTR des sols

Les caractéristiques des 4 classes de sols testées sont résumées dans le tableau 7.

Classe proposée	Classe GTR	Nature	Codification
Classe 1	A1 A2 A3 A4	Sols fins	1 0 0 0
Classe 2	B5 B6	Sables et graves	0 1 0 0
Classe 3	D1 B1 B2	Sables	0 0 1 0
Classe 4	D2 B3 B4	Graves	0 0 0 1

Tableau 7 : Caractéristiques des classes de sols testées

En figure 21, on peut observer la répartition des données selon les classes proposées pour les ensembles in-situ et laboratoire.

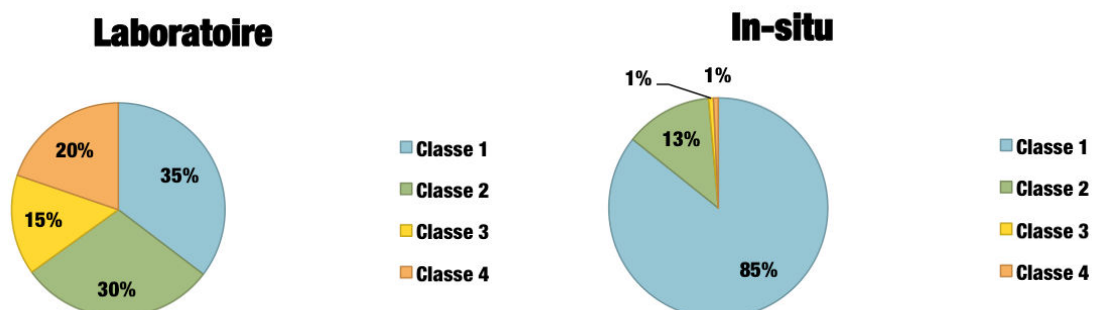


Figure 21 : Répartition des données selon les classes proposés

On observe que les données laboratoire sont réparties de manière plus homogène contrairement aux données in-situ où principalement nous avons des sols de classe 1 (120 essais sur 140 disponibles).

CHAPITRE 3 : DÉTERMINATION DE L'INFLUENCE DE LA SECTION DE LA POINTE

Le pénétromètre Panda peut être utilisé avec trois pointes différentes : 2 cm², 4 cm² et 10 cm². Les pointes de 2 cm² sont utilisées lorsque la profondeur maximum atteinte est inférieure à 1,5 m. Au delà, il faut utiliser les pointes de 4 cm². C'est pour cela que généralement les pointes de 2 cm² sont utilisées dans des moules de calibration pour réaliser des essais en laboratoire. Les pointes de 4 cm² sont utilisées pour les essais in situ. Les pointes de 10 cm² sont rarement utilisées, elles sont utiles pour les sols très mous dans lesquelles les pointes de section inférieure s'enfoncent trop rapidement.

Le Panda est un pénétromètre à énergie variable, le nombre de points va dépendre de la résistance du sol. En effet pour les faibles résistances, qui correspondent à des grands enfoncements, le pénétrogramme est lissé et cela indique que les espacements entre chaque points est important. Dans le cas contraire de fortes résistances, correspondant à de faibles enfoncements, le pénétrogramme est plus bruité donc l'espacement entre chaque points est plus petit.

Dans un premier temps, on peut donc se demander quelle est l'influence de cette variabilité sur les valeurs des 17 paramètres, qui servent de données d'entrées au RNA. Pour tenter de répondre à cette question, nous allons étudier les résultats obtenus pour un signal brut et un signal régularisé pour différents types de sol. Les données utilisées ont été obtenues à partir d'essais réalisés dans des moules de calibration contenant un seul matériau.

Dans l'objectif de pouvoir utiliser les deux bases de données créées, il faut vérifier si le paramétrage est identique pour un essai réalisé avec une pointe de 2 cm² et une pointe de 4 cm². Cela nous permettrait donc de mesurer l'influence ou non de la taille de la pointe sur les capacités de prédiction du PMC. Pour répondre à cette question, nous allons comparer les résultats obtenus avec une pointe de 2 cm² et 4 cm² pour différents types de sols. Les matériaux sont des matériaux de catégorie GTR D1 ou D2, ils sont dit insensibles à l'eau. Ces matériaux sont présents pour les trois cas étudiés. Au total, nous disposons de 25 comparaisons possibles entre les résultats pour une pointe de 2 cm² et de 4 cm².

I. PRÉSENTATION DE LA RÉGULARISATION

[DEP 91]

I.1. Principe de fonctionnement

Les mesures brutes sont des résistances de pointes réparties de manière non uniforme. La formule permettant la régularisation est la suivante :

$$qd_{moy} = \frac{\sum_i^n e_i \cdot qd_i}{\sum_i^n e_i} \quad [8]$$

Une fenêtre glissante de hauteur ξ avance suivant un pas μ . Pour chaque position, on réalise une moyenne additive des valeurs se trouvant dans la fenêtre. Il faut faire la distinction entre lissage et

régularisation (figure 22). Le lissage correspond à faire une moyenne lorsque le pas d'avancement est plus petit que la largeur de la fenêtre. Alors que la régularisation s'effectue lorsque le pas d'avancement et la largeur de la fenêtre sont de tailles identiques.

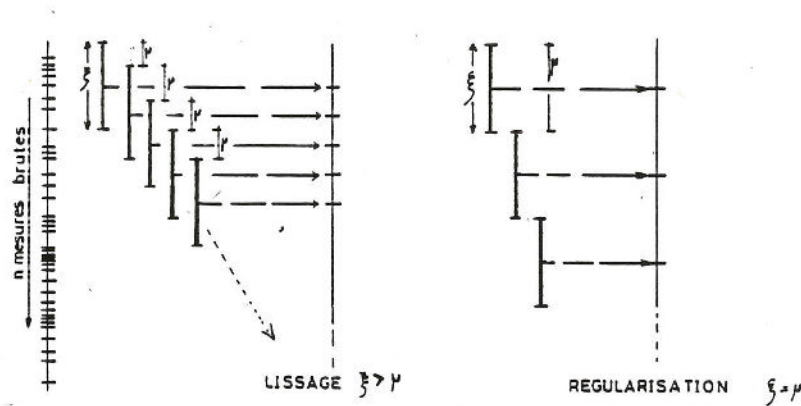


Figure 22 : Principe de la régularisation et du lissage

I.2. Exemples d'application

Pour pouvoir mesurer l'impact de la régularisation, trois types de matériaux sont pris comme cas d'étude. En figure 23, on peut observer qu'il n'y a aucune différence entre un pénétrogramme brut ou régularisé. Donc d'un point de vue visuelle, la régularisation n'a aucun impact. Son influence doit donc être analysée en comparant les valeurs pour les paramètres d'entrées dans les deux configurations possibles. D'autres pénétrogrammes sont illustrés en annexe 4.

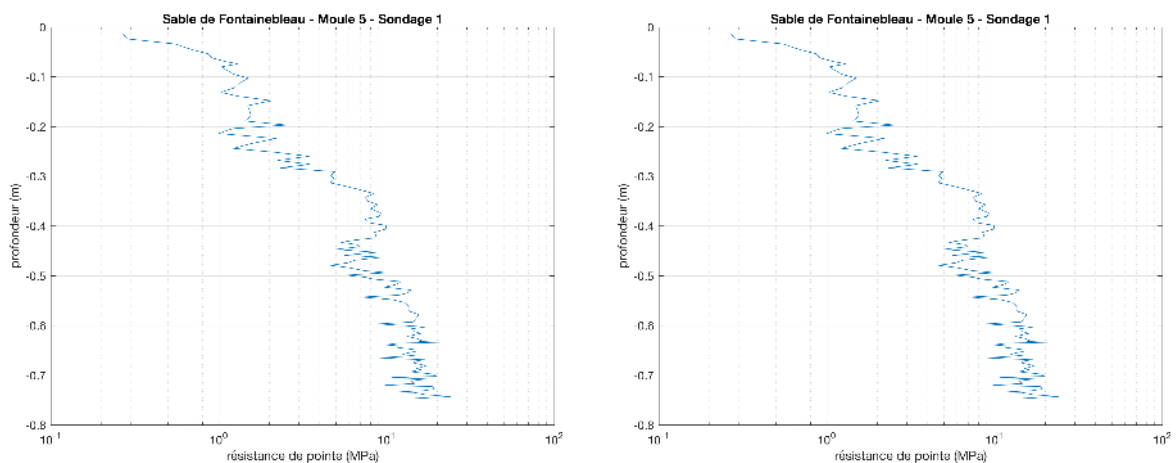


Figure 23 : Pénétrogramme brut (à gauche) et régularisé (à droite)

II. CALCUL DES ERREURS

Afin de mesurer les différences entre les valeurs des 17 paramètres obtenus dans chacun des cas, nous allons calculer différentes valeurs d'erreur. Pour cela, nous utiliserons quatre types d'erreurs : $RMSE$, MAE , R et R^2 , il s'agit des calculs d'erreurs les plus utilisés dans la littérature.

L'erreur quadratique moyenne :

$$RMSE = \sqrt{\frac{1}{P} \sum_{i=p}^P (t_i^{(p)} - y_i^{(p)})^2} \quad [9]$$

P : nombre de données disponibles ($p = 1, \dots, P$)

t_i : valeur désirée pour la donnée p

y_i : valeur obtenue pour la donnée p

L'erreur absolue moyenne :

$$MAE = \sum_{p=1}^P \frac{|t_i^{(p)} - y_i^{(p)}|}{P} \quad [10]$$

P : nombre de données disponibles ($p = 1, \dots, P$)

t_i : valeur désirée pour la donnée p

y_i : valeur obtenue pour la donnée p

Le coefficient de corrélation :

$$R = \frac{\sum_{p=1}^P (t_i^{(p)} - \bar{t}_i) - (y_i^{(p)} - \bar{y}_i)}{\sqrt{\sum_{p=1}^P (t_i^{(p)} - \bar{t}_i)^2 - \sum_{p=1}^P (y_i^{(p)} - \bar{y}_i)^2}} \quad [11]$$

P : nombre de données disponibles ($p = 1, \dots, P$)

t_i : valeur désirée pour la donnée p

\bar{t}_i : valeur moyenne de l'ensemble $\{t_i^{(p)}\}$

y_i : valeur obtenue pour la donnée p

\bar{y}_i : valeur moyenne de l'ensemble $\{y_i^{(p)}\}$

Le coefficient de détermination :

$$R^2 = \left(\frac{\sum_{p=1}^P (t_i^{(p)} - \bar{t}_i) - (y_i^{(p)} - \bar{y}_i)}{\sqrt{\sum_{p=1}^P (t_i^{(p)} - \bar{t}_i)^2 - \sum_{p=1}^P (y_i^{(p)} - \bar{y}_i)^2}} \right)^2 \quad [12]$$

P : nombre de données disponibles ($p = 1, \dots, P$)

t_i : valeur désirée pour la donnée p

\bar{t}_i : valeur moyenne de l'ensemble $\{t_i^{(p)}\}$

y_i : valeur obtenue pour la donnée p

\bar{y}_i : valeur moyenne de l'ensemble $\{y_i^{(p)}\}$

III. VÉRIFICATION DE L'HOMOGENÉITÉ D'UN SONDAGE

Afin de pouvoir constater les différences de valeurs des 17 paramètres, l'analyse s'appuiera sur trois types de matériau :

- Sable de Fontainebleau (D1-1) : 5 moules avec 3 sondages par moule ;
- Sable de Lyon (D1-2) : 4 moules avec 3 sondages par moule ;

- Grave de chantier (D2-1) : 4 moules avec 3 sondages par moule.

Pour chacun des moules, les valeurs de densité sèche sont différentes. En effet, elles augmentent progressivement avec les moules. Chaque sondage a été réalisé avec une pointe de 2 cm². Cependant les moules 1 des deux types de sables (Fontainebleau et Lyon) ne sont pas pris en compte pour l'étude car leur nombre de points sont trop faibles (inférieur à 20).

Notre ensemble de données, qui comprend au total 33 pénétrogrammes, a dans un premier temps subi un pré-traitement ayant pour objectif de supprimer l'influence du fond du moule sur les mesures. Donc nous avons gardé seulement les valeurs qui sont incluses dans les 75 premiers centimètres.

Les caractéristiques de chacune des éprouvettes conçues et analysées se trouvent en tableau 8 et 9. Les courbes granulométriques des matériaux sont présentés en annexe 5.

Caractéristiques	Sable de Fontainebleau	Sable de Lyon	Grave de chantier
Dmax	0,31	4,00	18,89
D60	0,23	0,70	12,64
D30	0,17	0,30	5,77
D10	0,14	0,15	1,43
CH	1,66	4,67	8,85
CC	0,93	0,86	1,85
Passant à 2 mm (%)	100	84,10	12,36
Passant à 80 µm (%)	0,12	1,80	0,07
VBS	0,08	0,09	0,06

Tableau 8 : Caractéristiques granulométriques des 3 matériaux

Caractéristiques	Sable de Fontainebleau				Sable de Lyon			Grave de chantier			
	N2	N3	N4	N5	N2	N3	N4	N1	N2	N3	N4
Teneur en eau (%)	4,66	4,95	6,40	6,56	12,05	11,53	10,96	3,07	3,02	3,01	3,53
Densité sèche (T/m³)	13,60	14,75	15,72	16,22	15,43	16,82	17,72	15,57	17,11	18,54	20,08
Densité humide (T/m³)	14,24	15,48	16,72	17,29	17,29	18,76	19,96	16,05	17,63	19,10	20,79

Tableau 9 : Caractéristiques physiques des 3 matériaux

III.1. Choix de la taille de la fenêtre de régularisation

Le choix de la taille de la fenêtre de régularisation peut avoir une influence importante sur le signal. En effet, si le pas est trop grand alors le pénétrogramme va être lissé de façon importante et de l'information sera perdue. Et dans le cas contraire si l'on choisit un pas trop petit alors on va créer de l'information qui ne sera pas forcément exacte. Les travaux de Deplagne [DEP 91] ont mis en évidence le choix optimum de la taille de la fenêtre, la dimension idéale est de 5 mm. En effet plusieurs raisons sont à l'origine de cette décision. C'est une valeur qui ne rajoute presque pas de points par rapport au nombre de mesures initiales contrairement au pas de 3 et 4 mm. C'est un pas qui est dans la moyenne des données brutes.

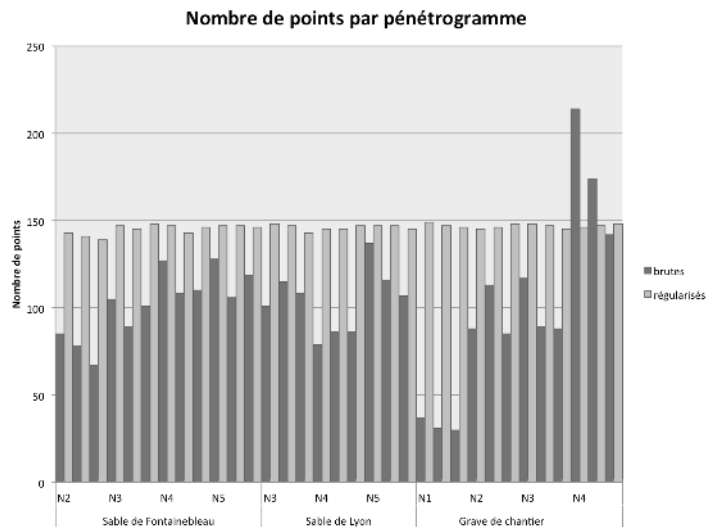


Figure 24 : Nombre de points par pénétrogrammes avant et après régularisation

On constate, dans la figure 24, que le nombre de points par pénétrogramme lorsqu'ils sont régularisés est globalement constant pour tous les cas. En moyenne, la régularisation modifie le pénétrogramme à un nombre de 146 points.

Les pénétrogrammes des trois types sols, avant et après régularisation, se trouvent en annexe 4. On peut observer que la régularisation ne modifie pas le signal et même après un nombre de points cinq fois supérieurs après régularisation.

Les valeurs des 17 paramètres d'entrée obtenues pour les 3 matériaux dans les deux cas (brut et régularisé) se trouvent en annexe 6.

III.2. Sable de Fontainebleau

Dans la figure 25, on présente la différence obtenue entre le sondage le 1 et 2 et entre sondage 1 et 3 pour les paramètres statistiques du Sable de Fontainebleau. Le sondage 1 est choisi comme sondage de référence étant donné que pour ce sondage le matériau n'a pas été perturbé par la réalisation de sondages ultérieurs.

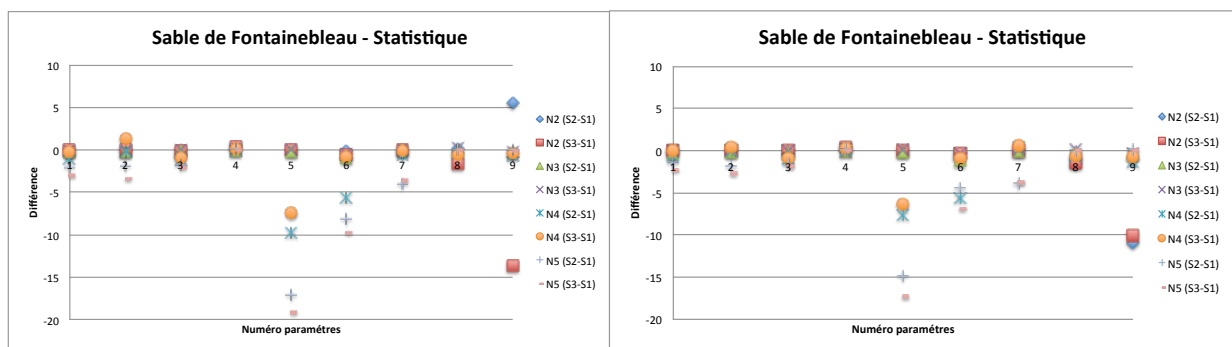


Figure 25 : Comparatif de l'analyse statistique du sable de Fontainebleau brut (à gauche) et régularisé (à droite)

On observe que la régularisation a permis d'homogénéiser les résultats obtenus entre les différents sondages. Les paramètres 5, 6, 7 et 9 dans les deux cas présentent des grandes différences entre les résultats des différents moules.

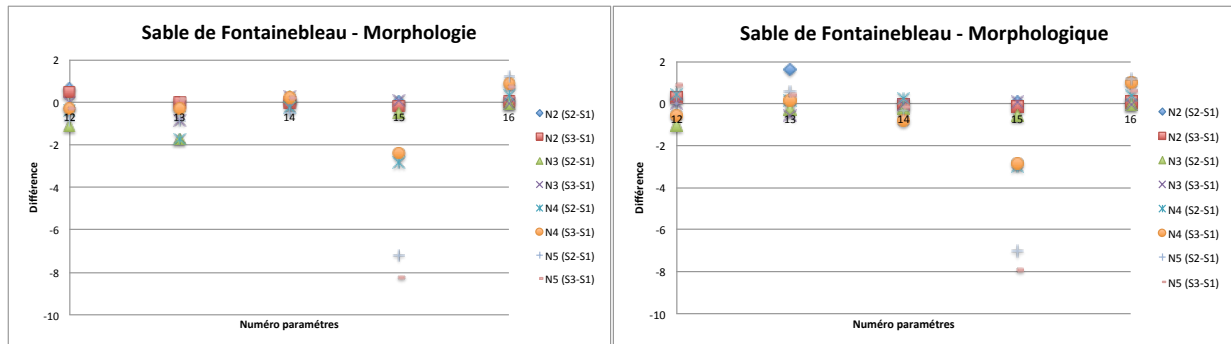


Figure 26 : Comparatif de l'analyse morphologique du sable de Fontainebleau brut (à gauche) et régularisé (à droite)

D'après la figure 26, on peut observer que la régularisation n'a pas non plus un impact considérable sur les paramètres morphologiques. Ce qui est logique étant donné que la régularisation ne modifie pas la morphologie du signal.

III.3. Sable de Lyon

En figure 27, il s'agit des de la différence obtenue entre les sondages 1 et 2 et entre les sondages 1 et 3 pour les paramètres statistiques du Sable de Lyon. Le sondage 1 est choisi comme sondage de référence étant donné que pour ce sondage le matériau n'a pas été perturbé par le fongage d'autres tiges.

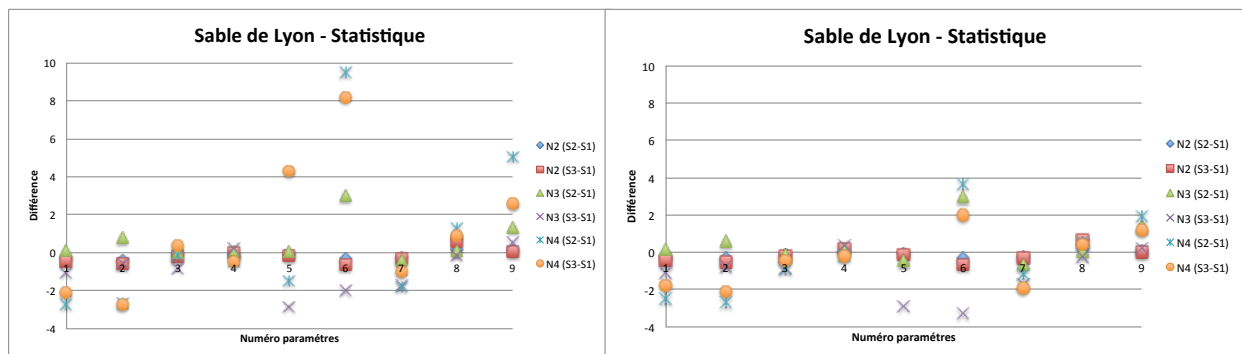


Figure 27 : Comparatif statistique du Sable de Lyon brut (à gauche) et régularisé (à droite)

La régularisation a eu pour conséquence de diminuer la plage de variation des valeurs de paramètre. Ce qui veut dire qu'après régularisation, les sondages dans les moules sont semblables d'un point de vue de l'analyse statistique. Il s'agit d'un impact positif, car cela veut dire que pour un même matériau dans plusieurs états de compacité, ils sont homogènes d'un point de vue des paramètres statistiques.

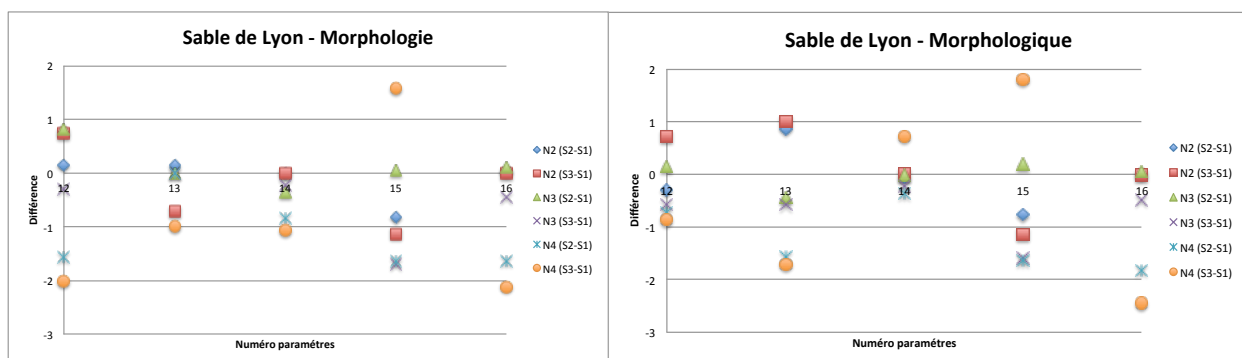


Figure 28 : Comparatif morphologique du Sable de Lyon brut (à gauche) et régularisé (à droite)

D'après la figure 28, on peut observer que la régularisation n'a pas un impact considérable sur les paramètres morphologiques 15 et 16. Cependant, on observe une variation inverse des résultats pour les paramètres 12 et 13 après régularisation.

III.4. Grave de chantier

En figure 29, il s'agit de la différence obtenue entre les sondages 1 et 2 et entre les sondages 1 et 3 pour les paramètres statistiques de la Grave de chantier. Le sondage 1 est choisi comme sondage de référence étant donné que pour ce sondage le matériau n'a pas été perturbé par le fonçage d'autres tiges.

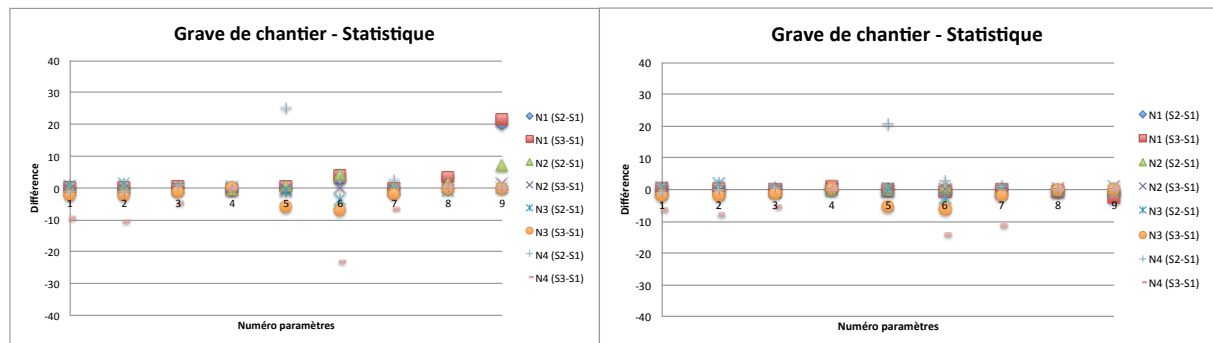


Figure 29 : Comparatif statistique de la Grave brut (à gauche) et régularisé (à droite)

La valeur du comparatif «N4 (S3-S1)» pour le paramètre 5 est en dehors de l'échelle du graphique pour les deux cas, ces valeurs sont aberrantes et donc afin de ne pas fausser l'interprétation, elles ont été exclues. Les paramètres 5, 6 et 9, qui sont les plus variables comme pour les exemples précédents, ont une plage de variabilité diminuée après régularisation.

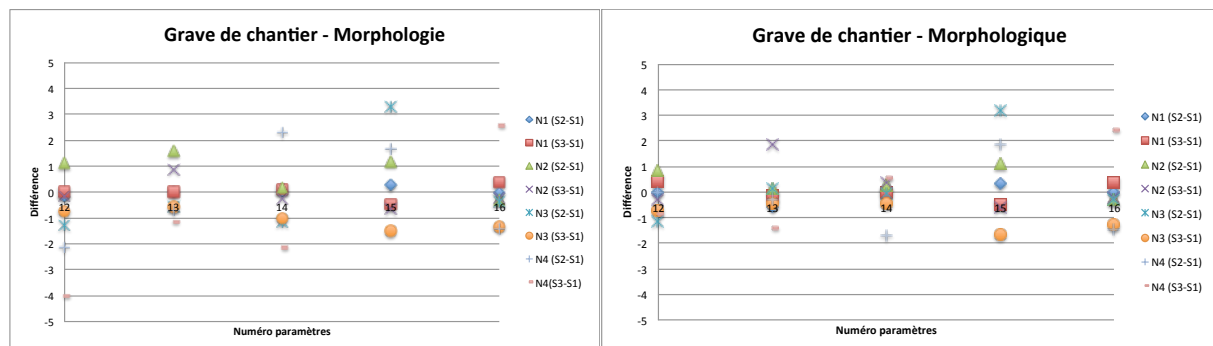


Figure 30 : Comparatif morphologique de la Grave brute (à gauche) et régularisée (à droite)

D'après la figure 30, on peut observer que la régularisation n'a pas un impact considérable sur les paramètres morphologiques sauf pour le paramètre 14 qui a vu ses valeurs pour le moule N4 se rapprocher des autres moules.

III.5. Conclusion

On peut donc conclure que la régularisation a un impact positif d'un point de vue de l'homogénéité des sondages. Ce qui veut dire que pour un matériau, si son pénétrogramme est régularisé, alors il y aura une meilleure ressemblance entre différents sondages effectués sur un même matériau. Donc, pour notre système de classification automatique que nous souhaitons mettre au point, la régularisation devrait améliorer la performance du modèle.

IV. INFLUENCE DE LA SECTION TRANSVERSALE DE LA POINTE

Nous avons constaté que la régularisation a un impact positif sur l'homogénéité des sondages pour un même matériau. On sait que l'on peut réaliser les essais au Panda avec des pointes de section différentes. Nous nous intéresserons seulement aux deux types de pointes les plus utilisées dans la pratique. Il s'agit des pointes de 2 cm² qui sont généralement utilisées pour les moules et les pointes de 4 cm² pour les essais in-situ.

La comparaison entre les 2 types de pointes ne se fera que sur les analyses statistiques et morphologiques. En effet, les analyses non linéaires et spectrales ne sont que respectivement 2 et 1 paramètre et il n'est pas intéressant de les tracer. De plus, les paramètres non linéaires ont une plage de valeurs très importante.

IV.1. Caractéristiques des matériaux et régularisation

IV.1.1. Moules homogènes

Le matériau des moules homogène est le sable coquillé de Dunkerque. Au total 17 moules ont été réalisés, mais après vérification des résultats 4 moules (M6, M11, M12 et M17) sont exclus de la base de données car ils présentaient des résultats erronés. Les caractéristiques de chacun des moules exploitables sont résumées dans le tableau 10.

	Poids total (kg)	Hauteur total (m)	Teneur en eau (%)	Nombre de couches	Densité sèche (T/m ³)	Densité humide(T/m ³)
Moule 1	120	0.69	6	4	14.6	14.5
Moule 2	150	0.84	6	5	14.8	14.75
Moule 3	180	0.99	6	6	15.1	15.04
Moule 4	180	0.94	6	6	16	15.9
Moule 5	180	0.99	6	6	15.2	15.08
Moule 7	120	0.69	4.5	4	14.6	14.52
Moule 8	150	0.84	4.5	5	15	14.89
Moule 9	180	0.99	4.5	6	15.2	15.12
Moule 10	180	0.99	4.5	6	15.3	15.19
Moule 13	180	0.95	4.5	6	15.8	15.76
Moule 14	180	0.91	4.5	6	16.5	16.41
Moule 15	180	0.96	4.5	6	15.6	15.53
Moule 16	180	1.01	4.5	6	14.9	14.84

Tableau 10 : Caractéristiques physiques des moules homogènes

Pour chacun des moules les points ayant une profondeur supérieur à 0,75 m n'ont pas été conservés pour l'étude afin de s'affranchir de l'effet du fond du moule sur les mesures de résistance de pointe. Cependant la partie correspondant à la résistance de pointe de surface a été conservée.

IV.1.2. Moules hétérogènes

Les moules contiennent 4 types de matériaux : sable grossier, ballast, microballast et sable de Dunkerque. Au total 4 moules ont été réalisés et l'ordre des matériaux dans les moules sont variés pour chacun des moules. Le compactage a été réalisé de manière dynamique avec un masse de 20 kg tombant d'une hauteur de chute de 60 cm, chaque couche a subi 6 coups de masse. Voici les caractéristiques des moules avec l'ordre des matériaux dans les quatre moules :

	Moule 1	Moule 2	Moule 3	Moule 4
Couche 1	Sable grossier	Microballast	Ballast	Sable grossier
Couche 2	Microballast	Sable grossier	Sable Dunkerque	Ballast
Couche 3	Sable Dunkerque	Ballast	Sable grossier	Microballast
Couche 4	Ballast	Sable Dunkerque	Microballast	Sable Dunkerque

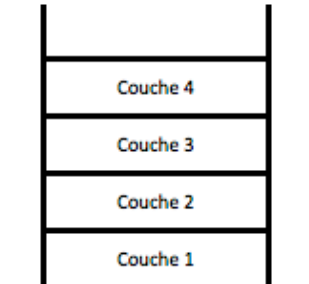


Tableau 11 : Agencement des matériaux des moules hétérogènes

Dans chaque moule, trois essais pénétrométriques sont effectués, le premier avec une pointe de 2 cm² fixe, le deuxième avec une pointe de 2 cm² perdue et le dernier avec une pointe de 4 cm² perdue. Nous nous intéresserons seulement aux cas 1 et 3, car la pointe de 2 perdue est utilisée seulement à titre de recherche. Les caractéristiques mécaniques et physiques de chacun des moules sont référencés dans le tableau 12. Lors des essais dans un moule, les tiges des essais précédents ont été conservés afin de ne pas perturber le milieu et l'agencement granulaire qui s'est formé autour de la tige.

	Moule 1	Moule 2	Moule 3	Moule 4
Hauteur (m)	1,1	1,12	1,07	1,08
Poids (kg)	197,7	197,1	196,9	196,7
W (%)	1	1	1	1
Densité humide (T/m³)	15,01	14,73	15,31	15,26
Densité sèche (T/m³)	15	14,7	15,3	15,3

Tableau 12 : Caractéristiques physiques et mécaniques des moules hétérogènes

Dans le cadre de l'étude sur la prédiction de la nature du sol, nous nous intéresserons seulement aux matériaux sableux : sable grossier 0/2 et sable de Dunkerque. Les matériaux recomposés : ballast et microballast sortent du cadre de la base de données qui a servi pour l'entraînement du RNA.

Pour tous les pénétrogrammes, on ne s'intéresse qu'aux sables donc afin d'obtenir un signal plus « propre » à fournir au réseau. Les zones de transition ont été enlevées de la plage de valeur. Cependant, il faut garder une plage de données suffisamment importante pour avoir un signal représentatif du sol en place.

IV.1.3. Régularisation

Pour les moules homogènes et hétérogènes, les valeurs des paramètres seront exploitées pour des données brutes et régularisées. On sait que la généralisation ne modifie pas l'aspect du signal mais permet seulement de réguler le pas d'enfoncement. Pour nos données, le pas d'avancement choisi est de 5 mm.

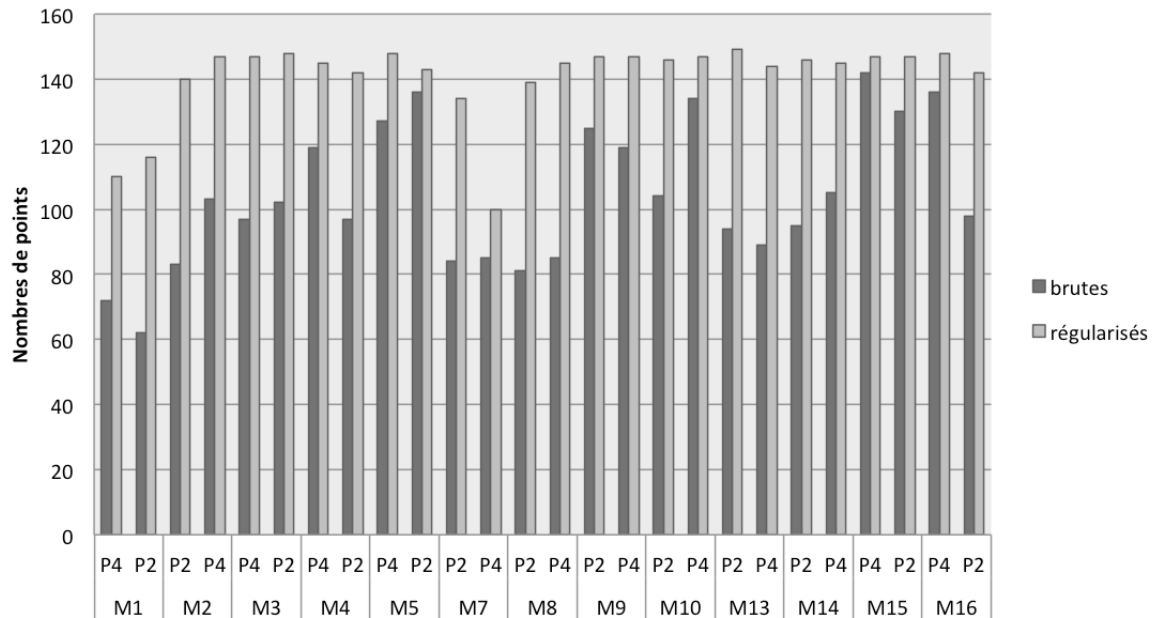


Figure 31 : Nombre de points des moules homogènes avant et après régularisation

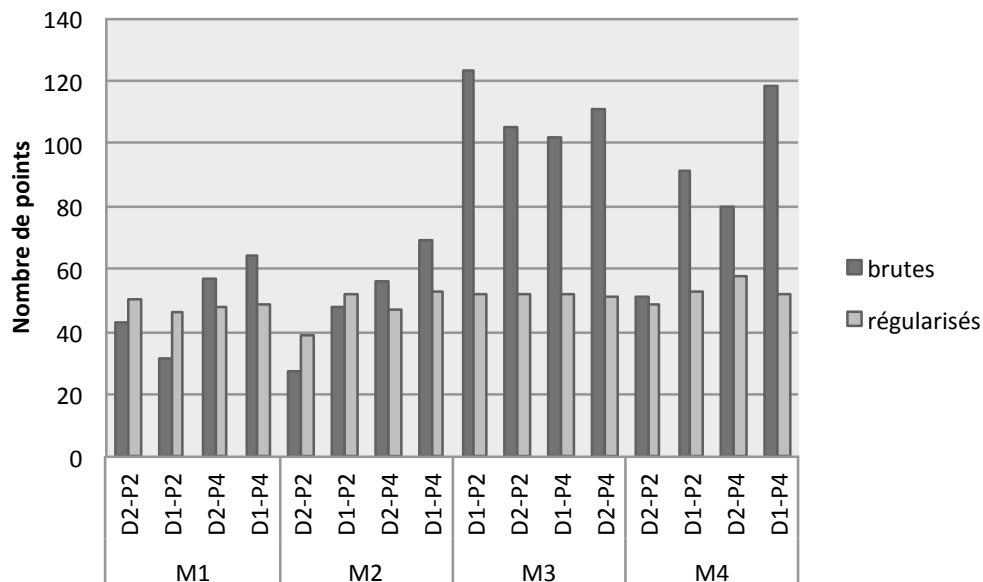


Figure 32 : Nombre de points moules hétérogènes avant et après régularisation

IV.2. Influence sur les paramètres de l'analyse statistique

Dans cette partie, nous allons nous intéresser aux valeurs obtenues pour les paramètres statistiques (1 à 9). Les informations sont recueillies à partir d'essais réalisés dans des moules, or les essais en moules de calibration sont normalement réalisés avec une pointe de 2 cm² étant donné la faible profondeur de l'essai. C'est pourquoi c'est cette section qui est choisie comme valeur de référence. Les valeurs des 17 paramètres d'entrée pour les moules homogènes et hétérogènes sont en annexe 8.

IV.2.1. Moules homogènes

Afin de mieux observer la différence de valeurs entre pointe de 2 et 4 cm², les moules 1, 5 et 7 sont exclus du graphique. Ils présentent des valeurs extrêmes et aberrantes, cela peut s'expliquer par l'état lâche des matériaux pour les moules 1 et 7 mais pour le moule 5 ce n'est pas le cas. Il peut s'agir d'une erreur provenant soit de l'opérateur soit du calcul de cette valeur.

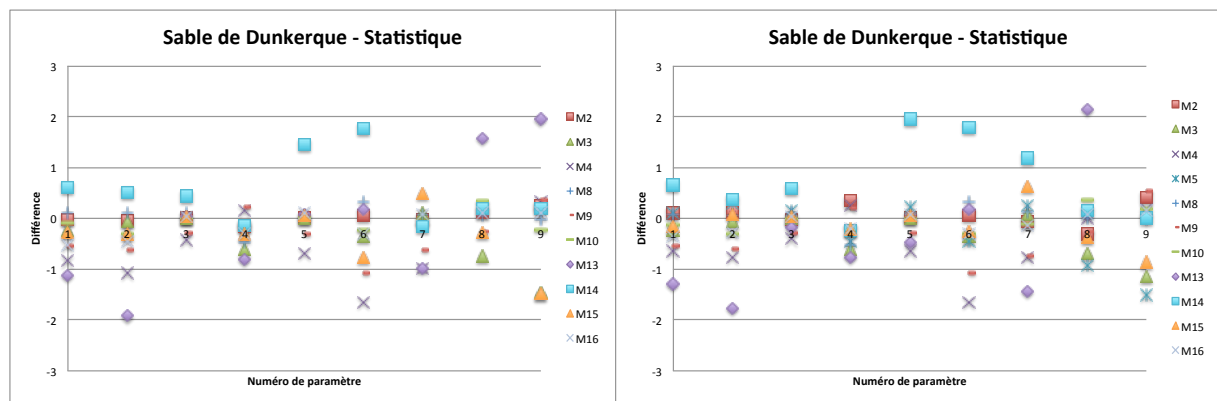


Figure 33 : Statistique - Sable de Dunkerque brut (à gauche) et régularisé (à droite)

Pour l'analyse statistique des moules homogènes, on n'observe que très peu de différence avant ou après régularisation. Les valeurs extrêmes correspondent toujours aux mêmes moules dans les deux cas.

IV.2.2. Moules hétérogènes

Afin de mieux observer la différence de valeurs entre pointe de 2 et 4 cm², les valeurs du paramètre 5 et 9 du moule 4 sont exclues du graphique pour le sable de Dunkerque. Elles présentent des valeurs extrêmes qui créent un problème d'échelle du graphique et ne permettent pas d'analyser correctement les résultats.

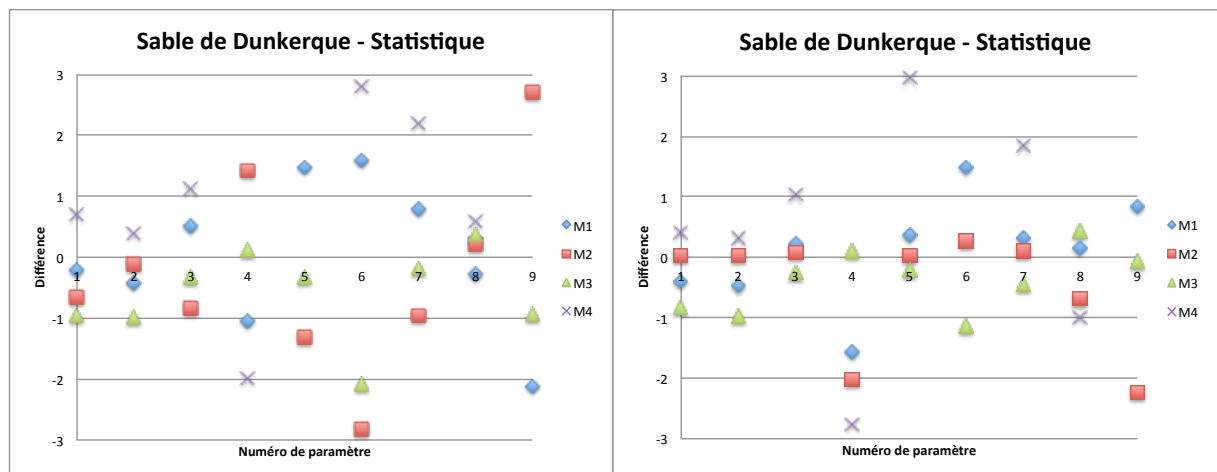


Figure 34 : Statistique - Sable de Dunkerque brut (à gauche) et régularisé (à droite)

Sur la couche du sable de Dunkerque, on observe que la régularisation a un impact positif sur certains paramètres, surtout sur le paramètre 6.

Pour le sable grossier, le paramètre 9 du moule 1 n'est pas pris en compte pour les mêmes raisons que pour le sable de Dunkerque des moules hétérogènes.

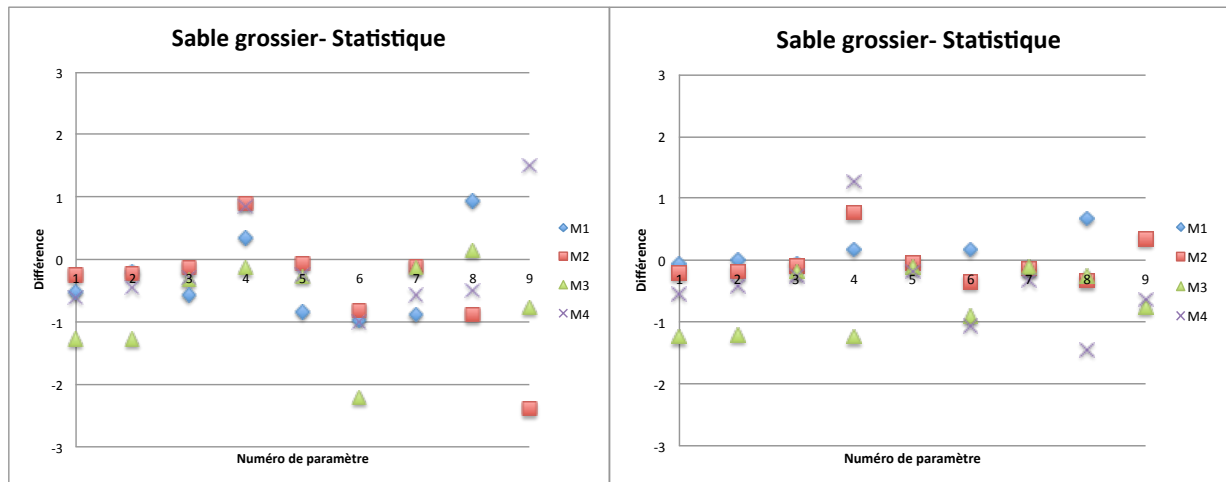


Figure 35 : Statistique - Sable grossier brut (à gauche) et régularisé (à droite)

L'impact de la régularisation sur la couche du sable grossier est plus importante. Les résultats pour les paramètres 5, 6, 7 et 9 sont moins variables.

IV.2.3. Bilan analyse statistique

Globalement, la régularisation a un impact positif sur la variabilité des valeurs de chacun des paramètres statistiques proposés pour le paramétrage des pénétrographes. Les résultats obtenus sur les moules hétérogènes doivent être nuancés car l'analyse statistique n'a eu lieu que sur des pénétrographes n'excédant pas 20 cm. Cette longueur n'est pas suffisante pour juger de la fiabilité des résultats obtenus.

IV.3. Influence sur les paramètres de l'analyse morphologique

Dans cette partie, nous allons nous intéresser aux valeurs obtenues de l'analyse morphologique avec une pointe de 2 cm² et une pointe de 4 cm². Les informations sont recueillies à partir d'essais réalisés dans des moules, normalement les essais en moules de calibration sont normalement réalisés avec une pointe de 2 cm² étant donné la faible profondeur de l'essai. C'est pourquoi c'est cette section qui est choisie comme valeur de référence.

IV.3.1. Moules homogènes

Aucun moule n'a été exclu de l'analyse ici présentée.

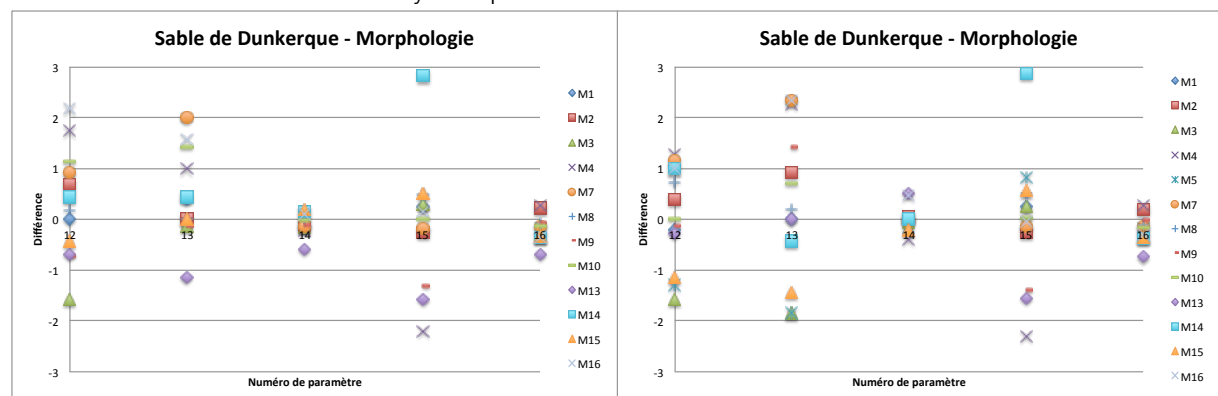


Figure 36 : Morphologie - Sable de Dunkerque brut (à gauche) et régularisé (à droite)

Pour l'analyse morphologique des moules homogènes, on peut observer que les paramètres ne sont pas impactés par la régularisation.

IV.3.2. Moules hétérogènes

L'analyse morphologique du sable de Dunkerque se trouve en figure 37 et celle du sable grossier en figure 38.

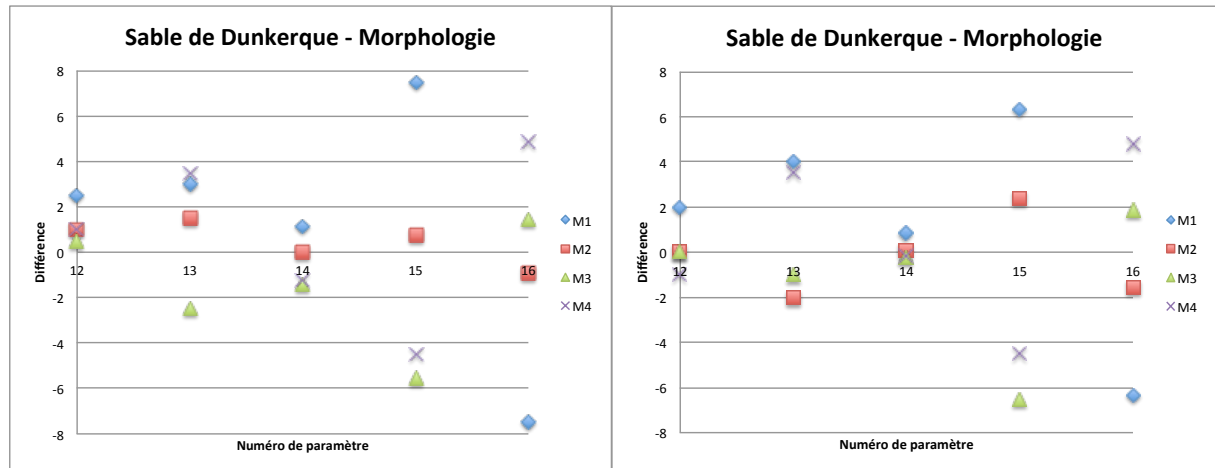


Figure 37 : Morphologie - Sable de Dunkerque brut (à gauche) et régularisé (à droite)

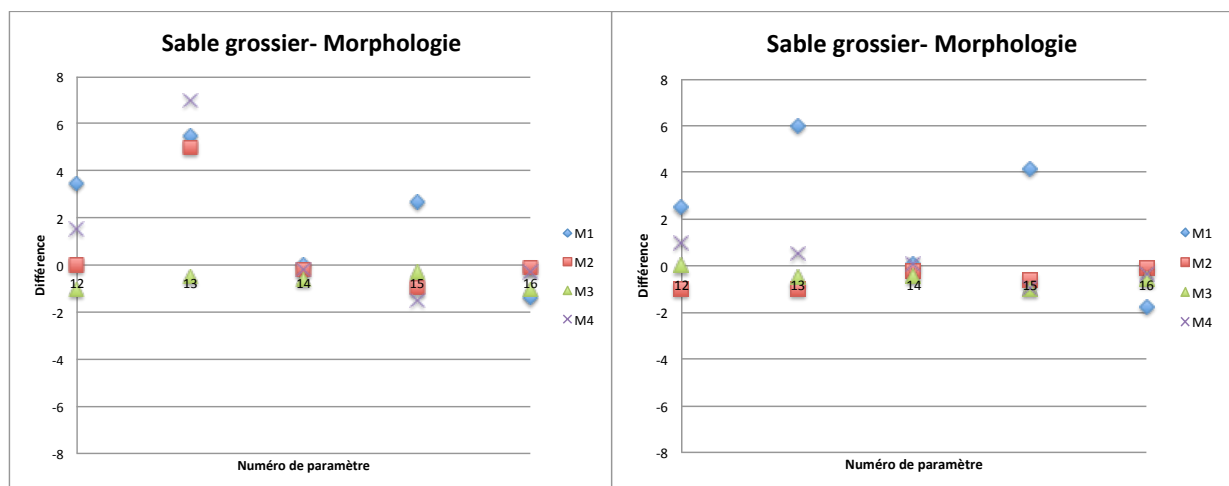


Figure 38 : Morphologie - Sable grossier brut (à gauche) et régularisé (à droite)

Pour les deux matériaux des moules hétérogènes, la régularisation n'a pas influencé sur les valeurs de l'analyse morphologique.

IV.3.3. Bilan analyse morphologique

Globalement, la régularisation ne modifie pas les résultats obtenus. Les différences tendent vers 0, nous n'avons pas de tendance forte. Il existe donc peu de différence morphologique entre un essai réalisé avec une pointe de 4 ou 2 cm².

IV.4. Valeurs des erreurs

On peut analyser les erreurs pour chacun des paramètres en calculant les erreurs types, celle qui sont cités précédemment (voir V. Chapitre 3). Nous étudierons seulement les résultats des moules homogènes étant donné ceux obtenus pour les moules hétérogènes qui sont mauvais. Les erreurs obtenues sont résumées dans les tableaux 13 et 14.

	1 [MPa]	2 [MPa]	3 [MPa]	4 [%]	5 [MPa]	6 [MPa]	7 [MPa]	8 [-]	9 [-]	10 [bits]	11 [bits]	12 [-]	13 [-]	14 [-]	15 [-]	16 [-]	17 [MPa ² /Hz]
MSE	0,24	0,46	0,05	0,17	0,25	7,80	0,21	4,57	338,71	727987,75	167,16	1,15	0,94	0,04	1,44	0,09	0,01
MAE	0,37	0,46	0,17	0,35	0,29	1,53	0,31	1,14	7,96	528,77	9,95	0,89	0,71	0,16	0,85	0,26	0,05
RMS E	0,49	0,68	0,23	0,41	0,50	2,79	0,46	2,14	18,40	853,22	12,93	1,07	0,97	0,21	1,20	0,31	0,12
R	0,893	0,789	0,922	0,557	0,920	0,529	0,868	0,549	0,513	0,885	0,961	0,476	-0,294	0,726	0,890	0,492	0,475
R²	0,798	0,623	0,851	0,310	0,847	0,279	0,754	0,301	0,263	0,783	0,924	0,226	0,087	0,528	0,791	0,242	0,226

Tableau 13 : Erreurs pour les moules homogènes bruts

	1 [MPa]	2 [MPa]	3 [MPa]	4 [%]	5 [MPa]	6 [MPa]	7 [MPa]	8 [-]	9 [-]	10 [bits]	11 [bits]	12 [-]	13 [-]	14 [-]	15 [-]	16 [-]	17 [MPa ² /Hz]
MSE	0,24	0,35	0,05	0,25	0,36	0,72	0,40	0,71	16,67	1429852,44	99,48	0,87	2,20	0,06	1,48	0,10	0,02
MAE	0,36	0,40	0,15	0,42	0,30	0,63	0,43	0,55	2,03	677,03	8,14	0,78	1,21	0,16	0,84	0,26	0,06
RMS E	0,49	0,59	0,23	0,50	0,60	0,85	0,63	0,84	4,08	1195,76	9,97	0,93	1,48	0,24	1,22	0,31	0,13
R	0,876	0,820	0,920	0,586	0,881	0,935	0,727	0,295	0,344	0,834	0,927	0,373	0,314	0,705	0,887	0,434	0,431
R²	0,767	0,672	0,847	0,343	0,777	0,875	0,529	0,087	0,119	0,696	0,860	0,139	0,099	0,496	0,787	0,188	0,185

Tableau 14 : Erreurs pour les moules homogènes régularisés

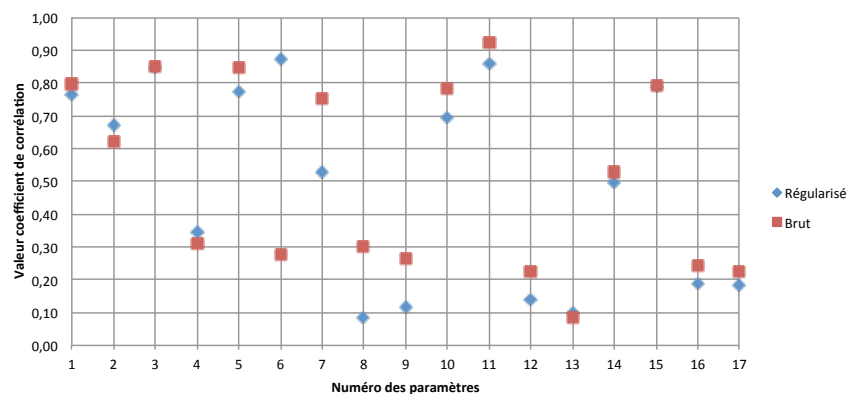


Figure 39 : Coefficient de détermination des moules homogènes entre pointe de 2 et 4 cm²

On peut observer que le régularisation fait chuter la valeurs du coefficient pour quasiment tous les paramètres. Sauf pour le 6, son coefficient est bien meilleur lorsque les données sont régularisés. On peut aussi observer que les paramètres morphologiques et spectraux ont un coefficient de détermination très faible. Il existe donc une variation importante pour cela entre un essai avec une pointe de 2 cm² et un essai réalisé avec une pointe de 4 cm².

IV.5. Analyse 3D des paramètres statistique

L'objectif de cette partie est de tracer des graphiques de types nuage de points en 3D afin d'observer si des tendances apparaissent lors de la visualisation. L'analyse se portera sur la valeur de l'erreur relative pour chaque moule, donc pour une densité spécifique, et pour les paramètres statistiques. L'axe X correspond au numéro de paramètre de l'analyse statistique (1 à 9). L'axe Y correspond à la valeur de la densité sèche de chaque moule. L'axe Z correspond à l'erreur relative.

Etant donné les résultats obtenus pour les moules hétérogènes, l'étude 3D de ses paramètres statistiques ne sera pas faite.

IV.5.1. Moules homogènes avec des données brutes

Dans cette partie, nous nous intéressons au sable de Dunkerque lorsqu'il n'est pas régularisé. Chaque couleur de point correspond une valeur de densité sèche. Voici le graphique obtenu :

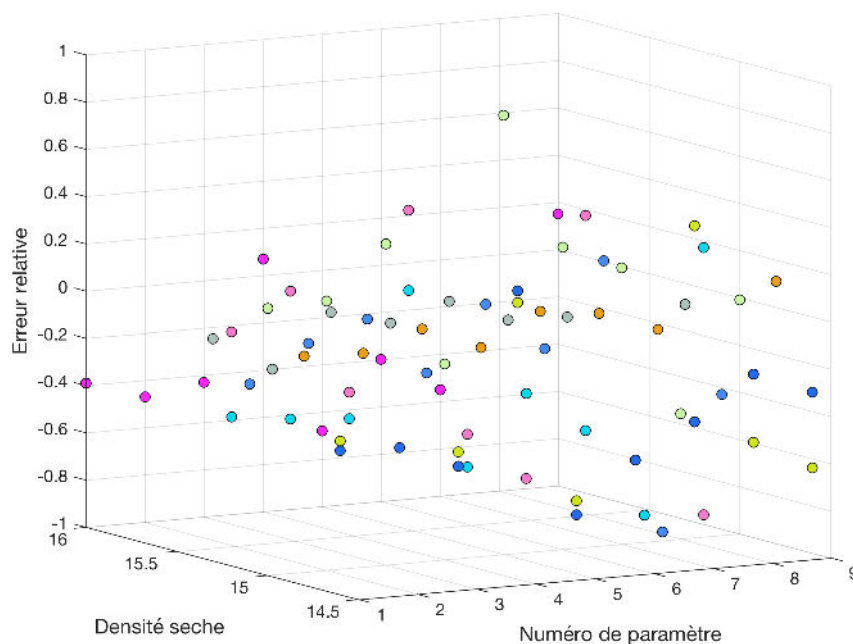


Figure 40 : Résultats de l'analyse statistique en 3D du sable de Dunkerque brut

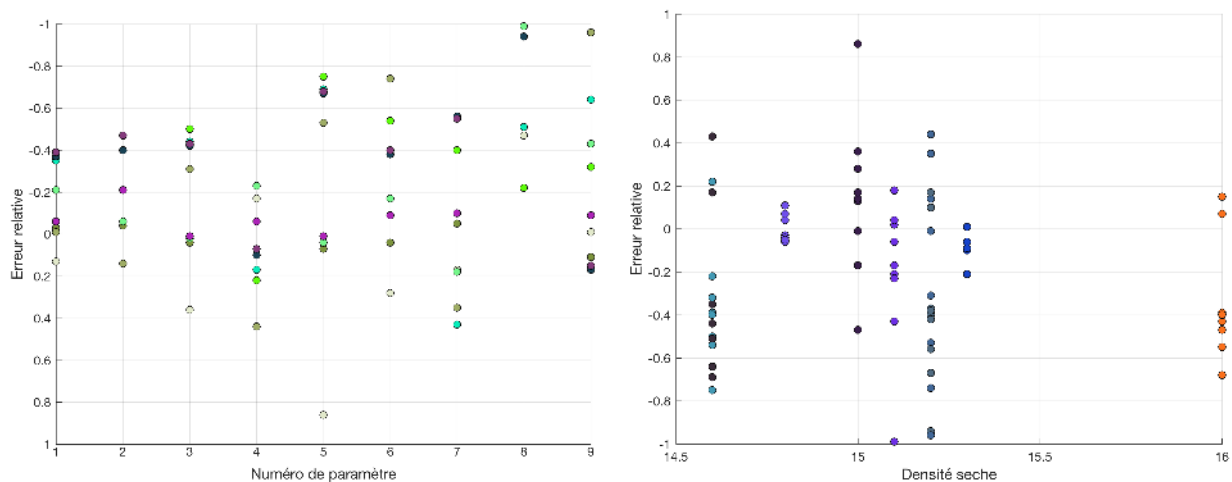


Figure 41 : Résultats de l'analyse statistique en 2D du sable de Dunkerque brut

IV.5.2. Moules homogènes avec des données régularisées

Dans cette partie, nous nous intéressons au sable de Dunkerque lorsqu'il est régularisé. Chaque couleur de point correspond une valeur de densité sèche. Voici le graphique obtenu :

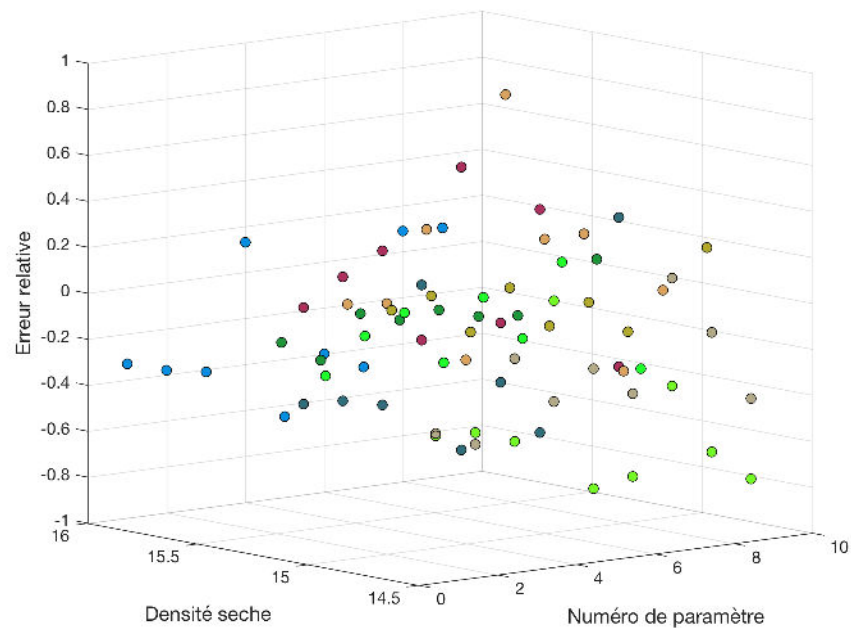


Figure 42 : Résultats de l'analyse statistique en 3D du sable de Dunkerque régularisé

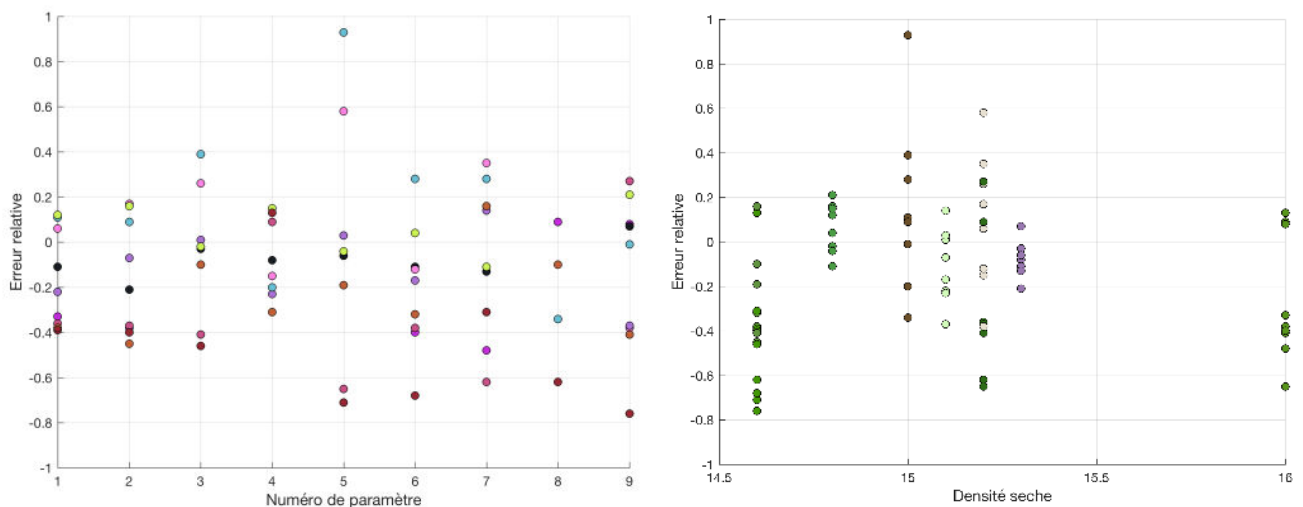


Figure 43 : Résultats de l'analyse statistique en 2D du sable de Dunkerque régularisé

IV.5.3. Conclusion de l'analyse 3D des paramètres statistiques

On peut observer que les données régularisées ont une plage d'erreur relative qui diminue et se rapproche de 0. Cela veut donc dire que la régularisation diminue les dispersions entre les valeurs des deux types de pointes. On peut observer que les erreurs sont plus faibles pour un état de densité moyen, les états lâches et très compacts présentent les erreurs les plus élevés.

IV.6. Résultats de classification du RNA

Maintenant, il s'agit de tester la performance de classification de notre RNA. Pour cela, nous allons lui fournir les 17 paramètres d'entrée pour les moules homogènes et hétérogènes. Il faut savoir au préalable que le sable de Dunkerque (ou sable coquillé de Dunkerque) est un matériau inconnu du modèle. En effet, aucun échantillon de ce matériau n'a servi à l'apprentissage du RNA. Les résultats sont présentés sous forme de matrice de confusion, elles nous permettent d'observer rapidement si le système parvient à classer correctement. Chaque colonne de la matrice représente le nombre d'occurrences d'une classe estimée, tandis que chaque ligne représente le nombre d'occurrences d'une classe réelle.

IV.6.1. Moules homogènes

Voici les résultats de la matrice de confusion pour le sable de Dunkerque avec des données brutes et régularisées :

Matrice de confusion

	1	2	3	4	
1	0 0.0%	0 0.0%	0 0.0%	3 11.5%	0.0% 100%
2	0 0.0%	0 0.0%	0 0.0%	2 7.7%	0.0% 100%
3	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
4	0 0.0%	0 0.0%	0 0.0%	21 80.8%	100% 0.0%
	NaN% NaN%	NaN% NaN%	NaN% NaN%	80.8% 19.2%	80.8% 19.2%
	1	2	3	4	

Classe estimée

Classe référence

Figure 44 : Matrice de confusion données brutes

Matrice de confusion

	1	2	3	4	
1	0 0.0%	0 0.0%	0 0.0%	2 7.7%	0.0% 100%
2	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
3	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
4	0 0.0%	0 0.0%	0 0.0%	24 92.3%	100% 0.0%
	NaN% NaN%	NaN% NaN%	NaN% NaN%	92.3% 7.7%	92.3% 7.7%
	1	2	3	4	

Classe estimée

Classe référence

Figure 45 : Matrice de confusion données régularisées

Pour les données brutes, les 5 échantillons mal classés sont les suivants :

- Moule 1 et pointe de 4 cm²
- Moule 1 et pointe de 2 cm²
- Moule 5 et pointe de 2 cm²
- Moule 7 et pointe de 2 cm²
- Moule 7 et pointe de 4 cm²

Pour les données régularisées, les 2 échantillons mal classés sont les suivants :

- Moule 1 et pointe de 2 cm²
- Moule 7 et pointe de 2 cm²

Les résultats sont très bons et surtout après régularisation. En effet, seulement deux moules ont été mal classés et il s'agit des moules correspondant à l'état lâche de ce matériau. L'état foisonné est un cas particulier des matériaux qui n'a pas été pris dans le set d'entraînement du RNA.

On peut également constater qu'il a majoritairement bien classé, que les sondages soient réalisés avec une pointe de 2 cm² ou 4 cm². On peut donc conclure que le RNA est performant et il l'est encore plus avec des données d'entrée régularisées.

IV.6.2. Moules hétérogènes

Comme précisé précédemment, nous nous intéressons aux matériaux sableux : sable grossier (GTR : D1) et sable de Dunkerque (GTR : D2). Voici les résultats de la matrice de confusion pour les données brutes et régularisées :

Matrice de confusion					
Classe estimée	1	2	3	4	
	0 0.0%	0 0.0%	0 0.0%	3 18.8%	0.0% 100%
	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
	0 0.0%	0 0.0%	8 50.0%	5 31.2%	38.5% 61.5%
	NaN% NaN%	NaN% NaN%	0.0% 100%	62.5% 37.5%	31.2% 68.8%
					Classe référence
	1	2	3	4	

Figure 46 : Matrice de confusion données brutes

Matrice de confusion					
Classe estimée	1	2	3	4	
	0 0.0%	0 0.0%	3 18.8%	5 31.2%	0.0% 100%
	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
	0 0.0%	0 0.0%	5 31.2%	3 18.8%	37.5% 62.5%
	NaN% NaN%	NaN% NaN%	0.0% 100%	37.5% 62.5%	18.8% 81.2%
Classe référence					
	1	2	3	4	

Figure 47 : Matrice de confusion données régularisées

Pour les données brutes, les 11 échantillons mal classés sont les suivants :

- Moule 1 : sol D1 et pointe de 4 cm²
- Moule 1 : sol D2 et pointe de 2 cm²
- Moule 1 : sol D1 et pointe de 2 cm²
- Moule 2 : sol D1 et pointe de 2 cm²
- Moule 2 : sol D1 et pointe de 4 cm²
- Moule 2 : sol D2 et pointe de 4 cm²
- Moule 3 : sol D1 et pointe de 2 cm²
- Moule 3 : sol D1 et pointe de 4 cm²
- Moule 4 : sol D1 et pointe de 4 cm²
- Moule 4 : sol D2 et pointe de 2 cm²
- Moule 4 : sol D1 et pointe de 2 cm²

Pour les données régularisées, les 13 échantillons mal classés sont les suivants :

- Moule 1 : sol D1 et pointe de 4 cm²
- Moule 1 : sol D2 et pointe de 2 cm²
- Moule 1 : sol D1 et pointe de 2 cm²
- Moule 2 : sol D2 et pointe de 2 cm²
- Moule 2 : sol D1 et pointe de 2 cm²
- Moule 2 : sol D1 et pointe de 4 cm²
- Moule 2 : sol D2 et pointe de 4 cm²
- Moule 3 : sol D1 et pointe de 2 cm²
- Moule 3 : sol D1 et pointe de 4 cm²
- Moule 3 : sol D2 et pointe de 4 cm²
- Moule 4 : sol D1 et pointe de 4 cm²
- Moule 4 : sol D2 et pointe de 2 cm²
- Moule 4 : sol D1 et pointe de 2 cm²

Les pénétrographes des matériaux des moules hétérogènes, une fois les zones de transition enlevées, ils sont de taille trop faible pour être clairement identifiés par le RNA. Un échantillon d'une longueur de 20 cm est trop faible pour être classé. Les analyses statistiques, non-linéaires, morphologiques et spectrales ne peuvent fournir des valeurs correctes avec si peu de points.

V. BILAN

Il faut tout d'abord rappeler que les analyses des paramètres d'entrée ont été réalisés sur des matériaux insensibles à l'eau donc un cas particulier. L'objectif de cette analyse était de comparer un essai réalisé avec une pointe de 2 cm² et un essai réalisé avec une pointe de 4 cm². En outre nous avons étudié l'impact de la régularisation sur ces essais. Pour ce faire, les paramètres statistiques et morphologiques ont été analysés avant et après régularisation pour la différence entre les deux pointes. Les paramètres non-linéaires et spectraux n'ont pas été analysés car ils sont difficilement interprétables aux vues de leurs valeurs et qu'ils ne sont que respectivement deux paramètres et un paramètre.

On peut conclure de manière générale que la régularisation a un impact positif sur les paramètres statistiques. En effet, la régularisation fait tendre la différence, entre les deux pointes, vers 0. Pour les paramètres morphologiques, la régularisation n'a pas d'impact visible sur les résultats. Il faut cependant rappeler que les valeurs des différences obtenues et analysées ne sont pas normalisées, et que le PMC les normalise avant de prédire la classe du sol.

En résumé, on peut dire que notre modèle RNA est plus performant sur les signaux testés régularisés. En effet, un signal brut va dépendre des conditions d'essais. La régularisation gomme ces imperfections en changeant la distribution initiale des points. Puis pour la question de la pointe, on peut conclure que le RNA est performant pour des pénétrogrammes obtenus avec une pointe de 2 ou 4 cm². Les données en pointe de 2 cm² sont présentes en plus grande quantité et sont plus représentatives des 4 classes possibles. Ces données peuvent donc servir de base d'apprentissage pour un modèle de RNA afin de prédire des natures de sol provenant d'essais in-situ réalisés avec des pointes de 4 cm² généralement. Cependant, une limite apparaît, il s'agit de la taille des signaux à classer. Si ils sont trop court, inférieurs à 20 cm, le RNA ne peut calculer correctement certaines valeurs de paramètres comme les différentes valeurs d'entropies. De plus, le RNA a été entraîné avec des signaux qui ont une taille d'au moins 40 cm. D'un point de vue géologique, une couche de sol de 20 cm est peu probable. C'est une limite de la méthode d'identification de la nature des couches.

CHAPITRE 4 : COMPARAISON PMC AVEC DIFFÉRENTS SYSTÈMES DE CLASSIFICATION AUTOMATIQUES

Les RNA ne sont pas l'unique outil pour classifier ou prédire des données. Depuis le début, nous nous concentrons sur les RNA et plus particulièrement sur les PMC, mais il faut vérifier que ce choix est le plus opportun. Le logiciel MATLAB propose de tester plusieurs systèmes de classifications de manière semi-automatique et rapide. L'application s'appelle « *Classification Learner* ». Il est dit semi-automatique car il nous offre plusieurs possibilités de choix initiales sur différents points :

- Choix des données : il est possible de prendre en compte ou non certains paramètres de l'étude. Par exemple, dans le cas des 17 paramètres d'entrées utilisés, si on le souhaite on peut lui indiquer de ne pas prendre le paramètre n°9.
- Méthode de validation : il est possible de choisir deux méthodes de validation. Ces deux méthodes sont la validation croisée « *cross-validation* » et la « *holdout validation* ». Il est possible aussi de ne pas choisir une méthode de validation et donc de simplement entraîner avec le modèle avec l'ensemble de données. Cependant, dans ce cas il est possible de sur-entraîner notre modèle.

Ces différentes options permettent de tester le modèle dans différents cas. Etant donné que dans notre cas, nous souhaitons comparer les autres modèles par rapport au PMC, nous n'allons choisir aucune méthode de validation afin de pouvoir sélectionner l'ensemble de validation utilisée pour la validation du PMC. Les modèles seront entraînés avec le même ensemble d'entraînement que celui utilisé pour entraîner le PMC.

I. PRÉSENTATION DES DONNÉES

L'ensemble de données disponibles est divisé en deux sous-ensembles : un ensemble d'entraînement qui permet de construire le modèle de RNA et un ensemble indépendant de validation qui permet d'estimer la performance du modèle. Avec la méthode de validation croisée, les données sont divisées en trois sous-ensembles : entraînement, test et validation. L'ensemble d'entraînement permet d'ajuster les poids des connexions du RNA alors que l'ensemble de test est utilisé pour vérifier les performances du modèle à différentes étapes de l'entraînement. De plus, il permet de déterminer quand arrêter la phase d'entraînement afin d'éviter le problème de sur-apprentissage. Dans notre cas, nous avons choisi de garder seulement deux sous-ensembles. L'indépendance de l'ensemble de validation permet de vérifier la capacité de généralisation du modèle et de vérifier que notre modèle n'est pas en train de mémoriser. Lorsque le modèle mémorise, il associe une entrée à une sortie sans créer de relation mathématique entre elles.

Il est essentiel que ces deux sous-ensembles représentent la même population, c'est à dire que d'un point de vue statistique, les deux sous-ensembles doivent être similaires. Les propriétés statistiques, qui permettent d'estimer si nos deux sous-ensembles sont similaires, sont la moyenne, l'écart-type, la valeur minimum, la valeur maximum et la plage de valeurs. La qualité de ces ensembles ne doit pas être négligée. En effet, il est connu que les RNA sont plus performants lorsqu'ils ne doivent pas extrapoler au delà de la

plage des données d'entraînement. C'est pour cela que les ensembles de validation doivent être contenus dans l'ensemble d'entraînement afin d'obtenir le meilleur modèle possible.

Dans cette étude comparative qui a pour objectif d'évaluer les performances d'autres systèmes de classification, il est donc important de vérifier que les ensembles de validation qui vont permettre de comparer les systèmes soient similaires. Les paramètres statistiques considérés sont les mêmes que ceux énoncés précédemment et il s'agit des paramètres proposés par Shahin [SJM 04].

II. RÉSULTATS

II.1. Arbres de décision et analyse du discriminant

Les résultats obtenus pour les arbres de décision et les analyses du discriminant se trouvent en annexe 3. Globalement, les résultats ne sont pas bons pour ces deux outils. Étant donné la complexité de notre modèle on comprend que les performances ne soient pas bonnes. La possibilité d'utiliser ces outils par la suite est écartée.

II.2. Les machines à vecteur de support

Les résultats obtenus pour les machines à vecteur de support sont présentés sur la figure 48. Ils sont de meilleure qualité par rapport aux outils présentés précédemment.

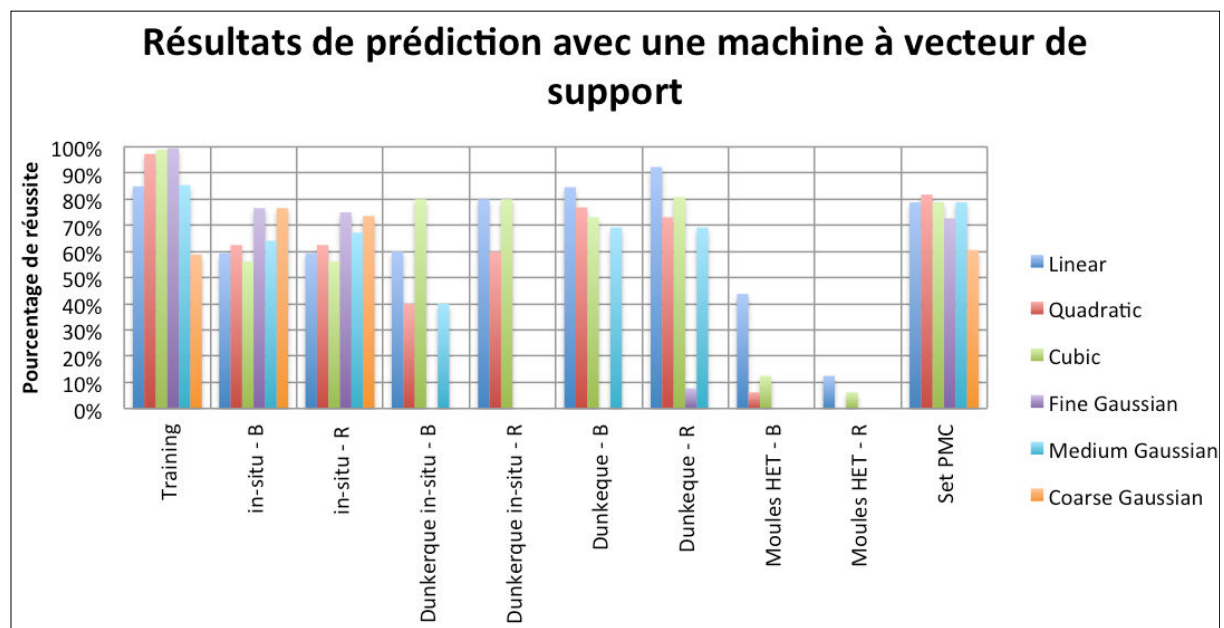


Figure 48 : Graphique des résultats obtenus pour les différentes machines à vecteurs de support

Les résultats pour les moules hétérogènes ne sont pas bons mais cela était prévisible étant donné la faible hauteur de signal disponible pour extraire les paramètres d'entrée. Les résultats pour l'ensemble d'entraînement et de validation utilisés pour le PMC sont corrects pour la plupart des machines sauf pour la machine utilisant une fonction gaussienne grossière. Pour les données restantes, le pourcentage de réussite est moins élevé. On peut conclure que les machines à vecteur de support les plus performantes pour le problème posé sont les suivantes : linéaire, quadratique et cubique.

II.3. La méthode des K plus proches voisins

Les résultats obtenus pour la méthode de K plus proches voisins sont présentés sur la figure 49. Ils sont globalement similaires par rapport aux machines à vecteur de support.

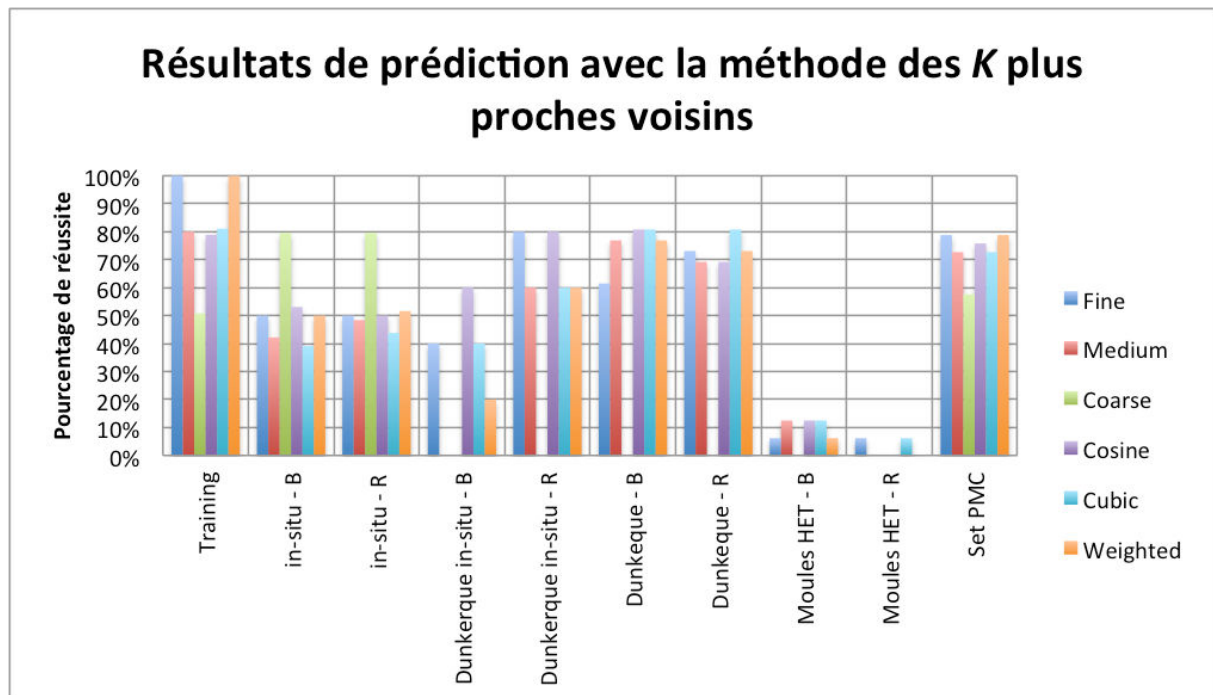


Figure 49 : Graphique des résultats obtenus avec la méthode des K plus proches voisins

Comme précédemment, les résultats des moules hétérogènes sont écartés pour les mêmes raisons que pour les machines à vecteurs de support. Les résultats pour l'ensemble d'entraînement et de validation utilisés pour le PMC sont similaires avec ceux obtenus avec la méthode des K plus proches voisins pour les modèles cosinus et pondéré.

III. CONCLUSION

Au terme des comparaisons réalisés, les modèles de types SVM et KNN présentent des résultats corrects pour les essais sur les Sables de Dunkerque pour le cas in-situ et laboratoire. Les résultats ne sont pas aussi bon que pour le PMC (voir tableau 15) néanmoins ils restent corrects sachant que les modèles SVM et KNN ne sont pas optimisés pour notre problème contrairement au PMC. Les modèles les plus performants sont les suivants :

- SVM linéaire,
- SVM cubique,
- SVM quadratique,
- KNN cosinus,
- KNN pondéré.

Si l'on compare les résultats avec le PMC sur différents cas en effectuant un somme pondéré de nos taux de réussite obtenus avec le nombre de données associés qui sont respectivement de 26 pour les essais Dunkerque laboratoire, 5 pour les essais Dunkerque in-situ, 185 pour l'ensemble d'entraînement et 33 pour l'ensemble de validation. Nous avons les résultats suivants :

Données	SVM linéaire	SVM cubique	SVM quadratique	KNN cosinus	KNN pondéré	RNA
Dunkerque labo – Brut	84,60 %	76,90 %	73,10 %	80,80 %	76,90 %	80,80 %
Dunkerque labo – Régularisé	92,30 %	73,10 %	80,80 %	80,00 %	60,00 %	92,30 %
Dunkerque in-situ – Brut	60,00 %	40,00 %	80,00 %	60,00 %	20,00 %	80,00 %
Dunkerque in-situ – Régularisé	80,00 %	84,00 %	80,00 %	69,20 %	73,10 %	80,00 %
Ensemble entraînement	84,90 %	98,90 %	97,30 %	78,90 %	81,10 %	98,90 %
Ensemble validation	78,80 %	78,80 %	81,80 %	75,80 %	72,70 %	93,20 %
Total pondéré	84,30 %	90,80 %	91,10 %	78,30 %	76,50 %	95,20 %

Tableau 15 : Comparaison des taux de réussite des différents modèles

D'après le tableau 15, le modèle le plus performant autre que le PMC est le SVM quadratique. Il s'agit d'un résultat logique car les machines à vecteur de support ont été créées pour des problèmes de classification et sont donc bien adaptés à notre problème. Néanmoins, le PMC obtient le meilleur de taux de réussite pondéré, le choix d'utiliser ce modèle est donc opportun. De plus, il a obtenu des bons taux de prédiction sur des matériaux inconnus (sable de Dunkerque) avec différents degrés de compactage. Cela nous indique également que notre problème n'est pas aberrant à résoudre et que le PMC a bien généralisé. Les performances des KNN sont moins bonnes, mais cela s'explique par le fait qu'ils ne sont pas adaptés pour travailler avec un grand nombre de paramètres d'entrées.

On peut également observer, que globalement la régularisation des données a permis d'augmenter le pourcentage de prédiction correcte.

CHAPITRE 5 : COMPARAISON DES RÉSULTATS AVEC DIFFÉRENTS ÉTATS HYDRIQUES

Les sols peuvent avoir une certaine teneur en eau, pour catégoriser ces différentes possibilités, la GTR les répertorie dans 5 états hydriques possibles : très sec, sec, moyen, humide et très humide. Un matériau n'aura pas la même résistance suivant son état hydrique et donc les pénétrographes vont être différents. Pour schématiser ces propos, il suffit d'observer la figure 50 qui correspond aux pénétrographes du Sable de Sermentizon (B5) dans trois états hydriques : très sec (ts), sec (s) et moyen (m). On observe une grande différence entre les trois pénétrographes. Donc il paraît obligatoire dans l'objectif de prédire la nature des sols de prendre en compte cette variabilité induite par l'état hydrique du matériau.

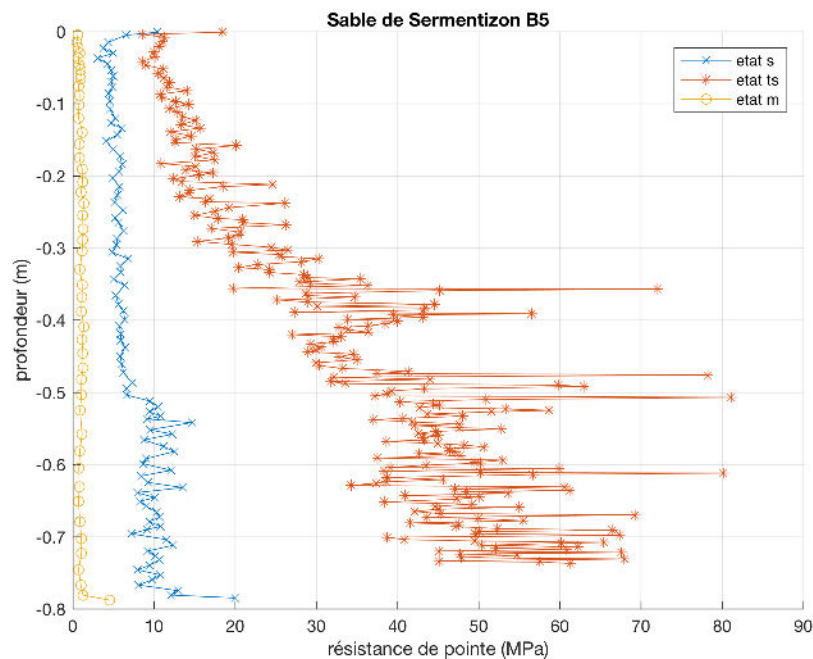


Figure 50 : Pénétrogramme du sable de Sermentizon pour trois états hydriques différents

Etat hydrique	Densité totale (T/m ³)	Densité sèche (T/m ³)	Teneur en eau (%)
Très sec	23,50	21,82	7,70
Sec	22,60	20,64	9,84
Moyen	21,13	19,02	11,07

Tableau 16 : Caractéristiques du sable de Sermentizon

Dans un premier temps, nous allons nous intéresser aux résultats de prédiction du PMC sur trois classes différentes : sec, moyen et humide. Les classes extrêmes, très sec et très humide, ne sont pas prises en compte au vu du faible nombre de données disponibles pour ces cas. Pour les trois états hydriques

étudiés, les données disponibles vont subir un pré-traitement qui a comme objectif de supprimer les effets du fond de moule en conservant les 0,75 premiers mètres. Puis les pénétrographes qui ont moins de 20 points ne sont également pas pris en compte. La répartition des données avant et après pré-traitement est présentée en figure 51 :

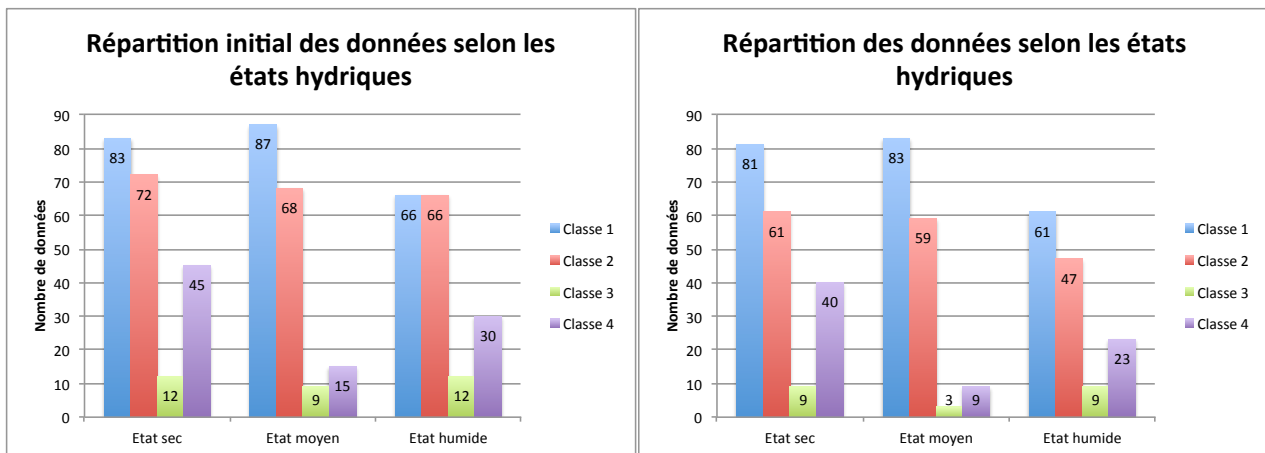


Figure 51 : Répartition des données avant et après pré-traitement

I. TEST SUR LA PROFONDEUR CRITIQUE

I.1. Test sur les trois états hydriques

L'objectif est de déterminer si un RNA est plus performant lorsque l'ensemble du signal lui est fourni, après le pré-traitement, ou si il faut en plus enlever la partie inférieure à la valeur de la profondeur critique. Pour cela, nous allons tester plusieurs PMC avec les 3 états hydriques et en compilant ces 3 états. Les tests vont faire varier le nombre de neurones de 9 à 14 avec 100 itérations pour chaque test. Les points en rouge correspondent aux pénétrographes dont le signal est homogène, c'est-à-dire que la partie inférieure à Zc a été supprimé. Les points en bleu correspondent aux signaux complets.

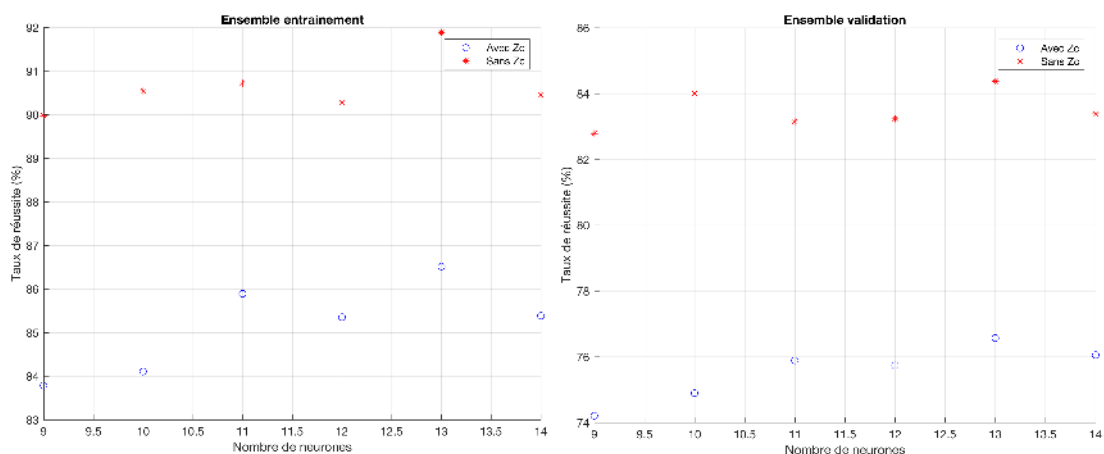


Figure 52 : Résultats des ensembles d'entraînement et de validation

D'après la figure 52, on peut constater les meilleures performances obtenues pour l'ensemble des états hydriques lorsque le signal est homogène. La partie inférieure à la profondeur critique n'est donc pas représentative du matériau testé, elle ne fournit pas d'informations utiles.

I.2. Test sur l'ensemble des données disponibles

On rassemble les trois matrices des états hydriques pour en obtenir une seule. Les deux cas sont testés : signal complet (avec Z_c) et signal homogène (sans Z_c). Les résultats pour l'ensemble d'entraînement et de validation sont présentés en figure 53.

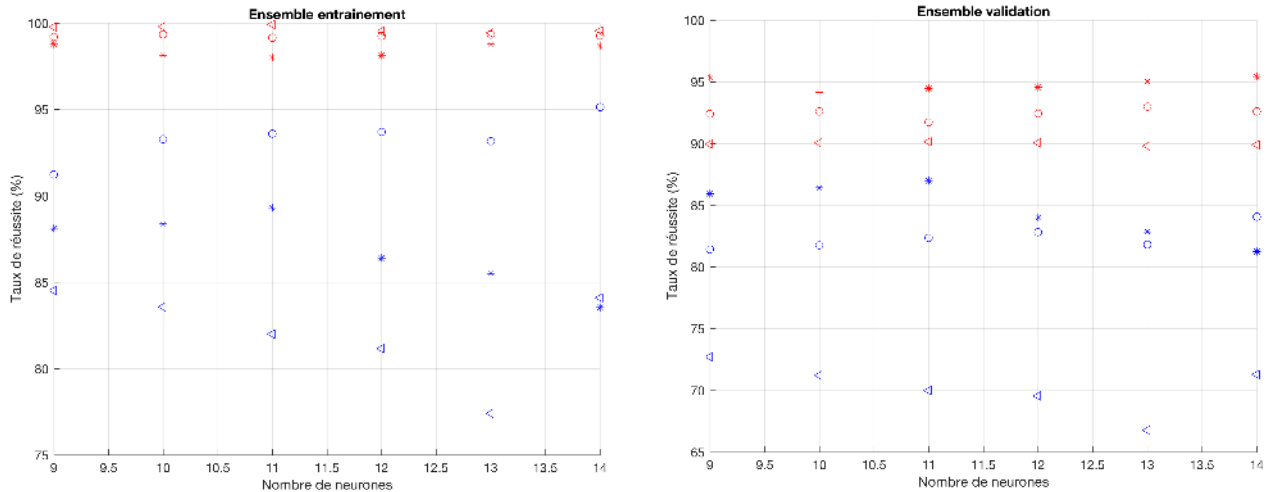


Figure 53 : Résultats entraînement et validation pour l'ensemble des données

Le résultat est clair, les performances sont accrues lorsque les pénétrogrammes fournis au réseau contiennent seulement la partie homogène du signal (sans Z_c).

II. TEST SUR LA RÉGULARISATION

Maintenant, on sait que la profondeur critique n'est pas intéressante pour notre problème de classification. Il faut savoir si l'on obtient de meilleures performances lorsque le signal est régularisé. Pour cela nous allons comparer les résultats obtenus pour les trois états hydriques avec les signaux homogènes bruts et régularisés. Le pas de régularisation est choisi à 5 mm, les signaux subissent le même pré-traitement que pour les tests précédents.

II.1. Test sur les trois états hydriques

Les résultats obtenus pour les différents PMC, variant de 9 à 14 neurones avec 100 itérations pour chaque cas, sont représentés en figure 54.

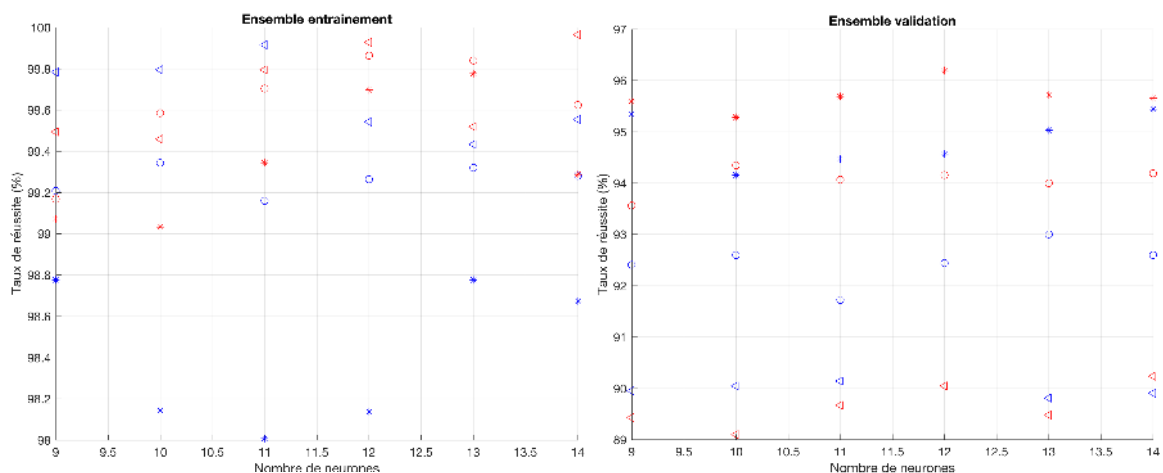


Figure 54 : Résultats entraînement et validation pour les données brutes ou régularisées

Les résultats pour l'ensemble d'entraînement sont très bon, supérieur à 98%, mais pour l'ensemble de validation on observe que les résultats sont regroupés par états hydriques. En effet, les étoiles correspondent à l'état moyen, les ronds à l'état sec et les triangles à l'état humide. Les résultats ne montrent pas clairement si la régularisation a un effet positif ou non.

II.2. Test sur l'ensemble des données disponibles

Afin de mesurer l'impact de la régularisation, on compare les résultats obtenus pour l'ensemble des données dans les deux cas étudiés : bruts ou régularisés. Ils sont représentés en figure 55.

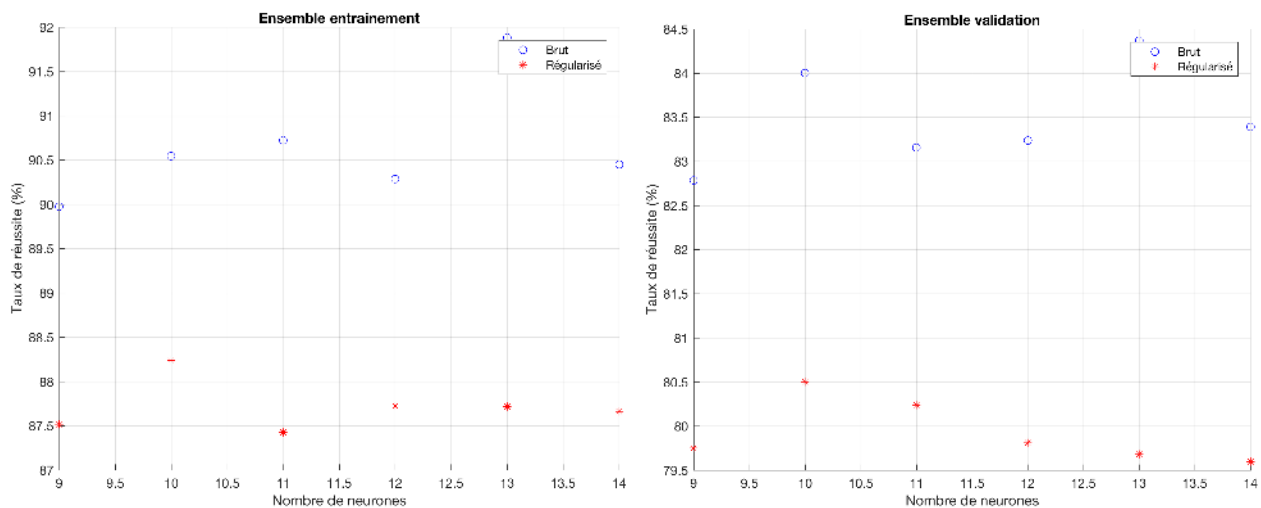


Figure 55 : Résultats entraînement et validation pour les données brutes et régularisées

Les résultats sont clairs, les performances des différents PMC sont accrues lorsque les données sont brutes.

III. VALIDATION DES MODÈLES

L'objectif est vérifier les résultats obtenus sur des données inconnues du réseau. Ces données sont des essais réalisés dans des moules de calibration sur du Sable de Dunkerque. Les résultats obtenus ne sont pas bons, dans la plupart des cas il n'arrive pas à classer correctement les 26 cas ou il classe correctement avec un taux de réussite inférieur à 20%. Les modèles testés ont établi un phénomène de sur-apprentissage, comme illustré en figure 56.

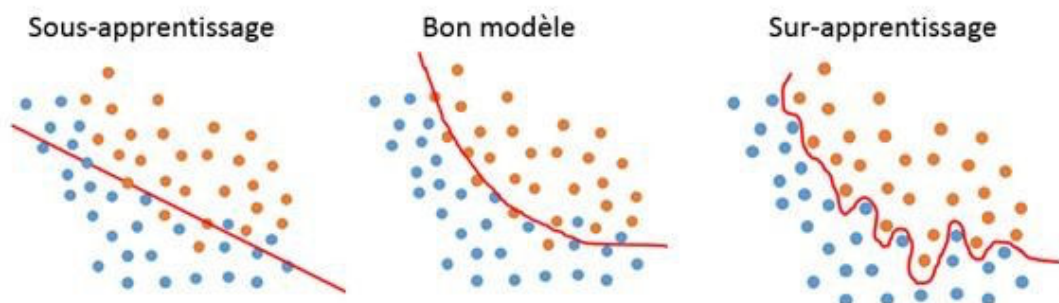


Figure 56 : Schématisation du phénomène d'apprentissage d'un RNA

Il semblerait que les réseaux ont mémorisé les cas d'apprentissage et non pas au contraire généralisé. Il existe plusieurs raisons possibles à cette mémorisation :

- la taille des signaux : après avoir supprimé la partie correspondante à la profondeur critique et les valeurs supérieures à 0,75 m, les signaux ont une taille trop petite pour calculer les paramètres d'entrées comme l'entropie ou la puissance spectrale du signal.
- les paramètres d'entrée : il est possible que les paramètres d'entrée ne soient pas adaptés au problème posé.
- la prise en compte de l'état lâche : on sait qu'il s'agit d'un état particulier et qui n'a pas de sens d'un point de vue géologique et géotechnique in-situ. De plus, les pénétrogrammes de cet état de consolidation sont particuliers.

BILAN

Les essais réalisés sur les moules de calibration nous donnent des résultats corrects et même pour des matériaux inconnus de la base d'entraînement. De plus, les résultats conservent un bon taux de réussite que ce soit pour des sondages réalisés avec une pointe de 2 cm² ou 4 cm².

Pour le cas de la régularisation, on constate que la régularisation des pénétrogrammes nous donne de meilleurs résultats par rapport aux pénétrogrammes initiaux pour le sable de Dunkerque. La régularisation est surtout efficace sur le paramètre n°6.

Cependant, un certain nombre d'hypothèses ont été choisies. Pour les essais réalisés dans des moules de calibration, nous avons gardé seulement les 75 premiers centimètres afin d'éviter de prendre en compte les effets du fond du moule. Pour la régularisation, nous avons choisi un pas de 5 mm comme dans la littérature.

Dans un second temps, l'étude sur l'influence de la profondeur critique nous montre que les RNA sont plus performants lorsque cette profondeur n'est pas prise en compte. Elle n'apporte pas d'informations permettant de caractériser les signaux car elle est trop influencée par les conditions d'essais et de réalisation.

Puis, nous avons justifié l'emploi des RNA par rapport à d'autres systèmes de classifieurs semi-automatique. Les résultats obtenus sont globalement meilleurs pour le PMC dans les différents cas étudiés. Mais il faut relativiser ces résultats car le PMC a été optimisé pour notre cas d'étude contrairement aux autres systèmes étudiés.

CONCLUSION ET PERSPECTIVES

CONCLUSION

Notre étude s'est focalisée sur le développement d'une méthodologie d'utilisation de réseaux de neurones artificiels (RNA) en vue de déterminer des classes de sol en termes de granulométrie. Pour cela nous nous appuyons sur la classification GTR et sur le signal pénétrométrique de l'essai Panda 2. Pour mettre au point cette méthodologie, nous avons utilisé un système de RNA et plus particulièrement un Perceptron multi-couche (PMC). Les RNA ont l'avantage d'être capable de travailler avec des données bruitées mais aussi de pouvoir résoudre et classer des problèmes très complexes avec une bonne fiabilité lorsqu'ils sont correctement programmés. Une des étapes cruciales du développement d'un système de RNA est la mise au point de l'ensemble d'apprentissage. En effet, les PMC fonctionnent avec un apprentissage supervisé, c'est-à-dire que durant sa phase d'apprentissage il va s'ajuster afin d'obtenir les meilleures prédictions possibles sur l'ensemble d'apprentissage. On comprend donc l'intérêt d'avoir un ensemble de données d'entraînement qui soit suffisamment représentatif du problème à résoudre. Cela va donc permettre au modèle créé de modéliser le plus efficacement possible notre problème.

La comparaison avec d'autres systèmes de classification automatiques :

Les systèmes de RNA ont été appliqués avec succès dans divers domaines de la géotechnique. Néanmoins, ce ne sont pas les seuls systèmes de classification automatiques. Il y a les arbres de décision, les analyses du discriminant, les machines à vecteur de support et la méthode des K plus proches voisins pour ne citer que ces méthodes les plus populaires. Nous avons pu tester ces quatre méthodes afin de justifier l'emploi des RNA.

Les résultats nous confirment que le RNA, et notre PMC particulièrement, obtient le taux de réussite le plus élevé pour les différents cas testés avec 95,20%. Néanmoins les machines à vecteur de support et tout particulièrement les cubiques et quadratiques obtiennent des taux qui sont plutôt bons 90,80% et 91,10% respectivement. Ils peuvent donc s'avérer des outils intéressants à explorer dans un futur.

Création de la base de données :

La première étape de ce travail a consisté à développer un système de base de données. Le but de cette base de données est de pouvoir stocker toutes les informations nécessaires à la mise au point de la méthodologie de prédiction via le système de RNA. C'est-à-dire les pénétrogrammes avec leurs classifications GTR et d'autres informations sur la granulométrie comme le passant à 2 mm ou la valeur au bleu. Mais les informations sur les matériaux sondés sont très différentes suivant qu'il s'agit d'un essai in-situ ou de laboratoire. La base de données peut également servir pour des études géotechniques car les essais in-situ sont référencés suivant leur position.

Pour les données laboratoire, une base de données pour le contrôle de compactage a déjà été réalisée par l'école Polytech et l'entreprise Sol-Solution, mais il s'agit d'une base de données Excel. Pour ces essais les matériaux testés sont bien étudiés et connus. Et pour les données in-situ, aucun système de gestion de base de données n'a été mis au point. Il a donc fallu mettre au point deux architectures pour la base de données : une base pour les essais laboratoire et une autre pour les essais in-situ. Ces systèmes de gestion de base de données vont être utilisés dans la suite de ce mémoire, ils doivent donc être conçus correctement et de façon adéquate à la future exploitation.

Lors d'une campagne de reconnaissance géotechnique, la caractérisation GTR des couches traversées n'est réalisée que sur les couches présentant la plus faible résistance. Il s'agit d'identifier la présence d'argiles sensibles à l'eau, car ces argiles ont la particularité de pouvoir gonfler ou se rétracter suivant sa teneur en eau. Il s'agit du phénomène de retrait gonflement. Au final sur un pénétrogramme complet, une seule partie de ce signal est caractérisée. Pour notre système de base de données, il a fallu associer la partie du pénétrogramme dont la couche a été caractérisée. Un premier problème se pose sur le découpage de ces pénétrogrammes car les zones d'influences ne doivent pas être prises en compte. En effet, le découpage est différent suivant l'utilisateur qui effectuera la tâche. Toutes les études réalisées par Sol-Solution ont été analysées afin de récupérer le plus d'informations possibles. Une partie de ce travail a été réalisé par Carlos SASTRE JURADO et l'autre partie par moi même.

Au final, on dispose d'une base de données in-situ qui contient 140 morceaux de pénétrogrammes dont on connaît la classe GTR. Et pour la base laboratoire, nous disposons de 749 pénétrogrammes réalisés dans des moules de calibration. Toutes ces informations nous ont servi pour entraîner et concevoir notre modèle de RNA.

Développement du RNA et paramétrage des données :

Comme il a été énoncé précédemment, on utilise généralement deux types de pointes avec le Panda. Dans le cadre du développement du système de RNA, il est nécessaire d'avoir un ensemble d'apprentissage représentatif d'un point de vue statistique et d'un point de vue des classes de sorties de notre modèle.

L'objectif de prédire la nature des sols à partir d'un essai pénétrométrique Panda est applicable aux essais réalisés in-situ pour lesquels on ne connaît justement pas les types de matériaux présents. Et au vu du faible nombre de données disponibles, provenant d'études in-situ, pour entraîner notre RNA il paraît donc obligatoire d'utiliser la base de données laboratoire. Cependant, un problème se pose : les essais ne sont pas réalisés avec la même pointe et on ne sait pas si l'on peut rassembler ces deux bases. Car la signature d'un pénétrogramme est très variable en fonction de la pointe, vis-à-vis de notre méthode.

Pour répondre à cette question, nous avons préalablement vérifié la répétabilité des valeurs d'entrées obtenus par le paramétrage pour un même matériau. Il s'agit de vérifier l'hypothèse que les 17 paramètres d'entrée sont semblables si l'on effectue plusieurs essais dans un moule contenant un unique. En effet, le signal brut n'est pas transmis directement au RNA mais on effectue un paramétrage du signal en calculant 17 valeurs qui représentent de la meilleure manière possible le signal.

Pour cela, nous avons comparé les valeurs obtenus pour trois matériaux : le sable de Fontainebleau, le sable de Lyon et la grave de chantier. Au vu des résultats obtenus, on peut dire qu'il existe une répétabilité entre les différents sondages réalisés dans un même matériau, et cette répétabilité est améliorée lorsque les pénétrogrammes sont régularisés. Il faut rappeler cependant que ces résultats ont été obtenus sur trois matériaux insensibles à l'eau.

Maintenant que l'on sait qu'il existe une reproductibilité entre différents essais sur les valeurs d'entrées calculées, il s'agit de vérifier si cette reproductibilité existe pour un même matériau sondé avec deux pointes de section différente. Pour cela nous nous sommes intéressés à des essais réalisés avec une pointe de 2 cm² et 4 cm² dans des moules homogènes contenant uniquement du sable coquillé de

Dunkerque et à des essais réalisés dans des moules hétérogènes contenant quatre types de matériau. Nous avons pour les deux types de moules réalisés une étude sur les valeurs des paramètres d'entrées du RNA et sur les résultats de prédiction du RNA.

Après avoir analysé les différences entre les valeurs obtenues avec la pointe de 2 cm² et celles obtenues avec la pointe de 4 cm², on constate que les valeurs du paramétrage des pénétrogrammes sont semblables pour les deux pointes. Si le signal est régularisé, on constate la même chose. Néanmoins, on observe à partir d'une étude des valeurs des erreurs de type RMSE ou MAE que le paramètre n°6, représentant l'intervalle de variation de q_d , était grandement influencé positivement par la régularisation contrairement aux autres paramètres d'entrée.

Evaluation des performances du RNA

Les résultats de prédiction des moules montrent que le sable coquillé de Dunkerque a été très bien classé avant régularisation 80,80% et encore mieux classé après régularisation des pénétrogrammes 92,30%. Il s'agissait en plus d'un matériau inconnu du PMC, on peut donc constater la bonne capacité de généralisation du modèle.

Les résultats des moules hétérogènes ne sont pas bons, 31,20% et 18,80%. Il apparaît alors une condition à notre problème de classification : la taille minimale du pénétrogramme utilisable. En effet, les pénétrogrammes des moules hétérogènes ne faisaient pas plus de 20 cm une fois les zones de transition enlevées. Or le PMC avec lequel nous avons obtenu ces résultats a été entraîné avec des pénétrogrammes d'une taille minimale de 40 cm. On comprend donc le faible taux de réussite pour les moules hétérogènes. De plus, le calcul des certains paramètres d'entrées, notamment les entropies, est faussé pour des signaux si petit. Il faut néanmoins rappeler que ces conclusions proviennent d'essais réalisés sur des matériaux insensibles à l'eau.

Prise en compte de la profondeur critique :

Sur un pénétrogramme réalisé dans un moule sur un matériau unique, on distingue deux parties dans le pénétrogramme. La partie homogène du signal correspond aux valeurs mesurées en-dessous de la profondeur critique c'est-à-dire à la résistance en profondeur. Il faut savoir si notre système de RNA doit être entraîné avec les pénétrogrammes au complet ou seulement avec la partie homogène des pénétrogrammes. Et surtout si cela améliore les performances du modèle. Pour cela nous avons récupéré tous les essais laboratoire, via la base de données, aux états hydriques sec à humide. Ensuite, nous les avons testés sur différents modèles de PMC. On a constaté que la méthode actuelle est plus performante lorsque seule la partie homogène a servi à l'entraînement.

De plus, dans le cadre de l'utilisation d'un PMC pour prédire la nature d'un sol provenant d'un essai in-situ il est préférable d'utiliser seulement les parties homogènes des pénétrogrammes pour l'apprentissage du modèle. En effet, la partie homogène d'un pénétrogramme est plus représentative du matériau testé et sera donc plus utile pour obtenir un ensemble d'entraînement représentatif des différents sols rencontrés.

PERSPECTIVES DE TRAVAIL

L'ensemble de ce travail a montré qu'il était possible de prédire la nature du sol à partir du signal pénétrométrique. Cependant à notre échelle, plusieurs aspects de la méthodologie peuvent être améliorés, en voici une liste non exhaustive :

- La base de données : le nombre de données stocké reste faible pour optimiser l'apprentissage du RNA et certains types de matériaux ne sont pas suffisamment présents pour avoir un ensemble d'entraînement représentatif. De plus, aucune méthodologie n'a été mise au point pour le découpage des pénétrogrammes et pour prendre en compte les zones d'influences.
- Les entrées du RNA : les 17 paramètres pourraient être optimisés afin de représenter au mieux la complexité du signal pénétrométrique. Cela peut passer par une diminution ou une augmentation du nombre de paramètres.
- Les sorties du RNA : au lieu d'utiliser la GTR en sortie, qui est inadaptée à la reconnaissance du sol, il faudrait changer de système soit en utilisant la classification U.S.C.S ou en pourcentage de passant par tamisât.
- La prise en compte de la nappe d'eau pour les essais in-situ : on sait que la signature d'un pénétrogramme change en fonction du degré de saturation.
- Le modèle de RNA : il existe d'autres types d'outils de Machine Learning qui sembleraient intéressante pour notre problématique.
- Le couplage de données : il serait possible également d'utiliser, en plus des informations fournis par le pénétrogramme, les résultats obtenus par d'autres essais comme par exemple les images fournies par l'endoscope.

Les RNA sont des outils puissants de modélisation. Malgré de bons résultats dans de nombreux cas d'application, ils possèdent un certain nombre d'inconvénients incluant la difficulté à extraire des connaissances du modèle entraîné, l'inaptitude à extrapoler en dehors de la plage des données d'entraînement et la gestion de l'incertitude.

BIBLIOGRAPHIE

- [ARA 12]** Abdelhadi Lotfi, Redouane Tlemsani et Abdelkader Benyettou. *Un Nouvel Algorithme d'Apprentissage pour les Réseaux Probabilistes, Vers une Réduction des PNN*. CARI 2012. 8 pages.
- [BBS 06]** B. Bhattacharya et D.P. Solomatine, *Machine learning in soil classification*. 2006 special issue. Neural networks, chapitre 19 p.186-195.
- [BIS 95]** BISHOP C. Robert, *Neural Networks and Pattern Classification*, Chapter 4: The Multi-Layer Perceptron. 1995.
- [BNA 09]** BENZ-NAVARRETE Miguel-Angel, *Mesures dynamiques lors du battage du pénétromètre PANDA 2* [Thèse]. Université Blaise Pascal - Clermont II, Ecole doctorale sciences pour l'ingénieur de Clermont-Ferrand. 12 juin 2009.
- [BOU 05]** BOUZY Bruno, *Réseaux de neurones*. 18 octobre 2005. Chapitre du cours d'apprentissage automatique, Master MISV[document en ligne].
- [BOU 14]** BOUZY Bruno, *Réseaux bayésiens*. 3 février 2014. Chapitre du cours d'apprentissage automatique, Master MISV[document en ligne]. 13 pages.
- [BRE 81]** BOREL Serge et REIFFSTECK Philippe, *Caractérisation de la déformabilité des sols au moyen d'essais en place*. Etudes et recherches des laboratoires des ponts et chaussées. Géotechnique et risques naturels, GT 81.
- [CHA 01]** CHAIGNEAU Laurent, *Caractérisation des milieux granulaires de surface à l'aide d'un pénétromètre* [Thèse]. Université Blaise Pascal - Clermont II, Ecole doctorale sciences pour l'ingénieur de Clermont-Ferrand. 26 septembre 2001.
- [DBE 10]** DERRAS Boumédiène et BEKKOUCHE Abdelmalek, *Estimation de la vitesse maximale du sol par la méthode des réseaux de neurones artificiels*. Journées Nationales de Géotechnique et de Géologie de l'Ingénieur, Grenoble 7-9 juillet 2010
- [DEP 91]** DEPLAGNE Franck, *Analyse structurale d'essais de pénétration*. Mémoire de Diplôme d'Études Approfondies. Université Blaise Pascal - Clermont II. Septembre 1991.
- [DLS 16]** Dana Torres, Luis Gonzalez et SaraTassini. *Les bases de données, cours de bases de données*. [document en ligne, consulté le 23 février 2016]
- [FAG 11]** J.C. FAUGERAS et R. GOURVES, *Caractéristiques physiques des sols identification et classification*. Centre universitaire des sciences et techniques, filière génie civil. Fascicule n°1.
- [FAG 12]** J.C. FAUGERAS et R. GOURVES, *L'eau dans les sols*. Centre universitaire des sciences et techniques, filière génie civil. Fascicule n°2.
- [FAG 14]** J.C. FAUGERAS, R. AZZOUZ et R. GOURVES, *Caractéristiques physiques des sols identification et classification*. Centre universitaire des sciences et techniques, filière génie civil. Fascicule n°4.
- [KGT 01]** P.U. Kurup, E.P. Griffin et M.T. Tumay, *Novel methodologies for soil characterization from CPT data*.
- [KOF 00]** D. KOLYMBAS & W. FELLIN, *Compaction of soils, granulates and powders. International workshop on compaction of soils, granulates and powders*. Innsbruck, 28-29 february 2000

- [KOR 01] KORCZACK Jerzy, *Réseaux neuronaux perceptron multi-couche*. Cours du LSIT, ULP.
- [NCH 07] N. CAGLAR et H. ARMAN, *The applicability of neural networks in the determination of soil profiles*. Bull Eng Geol Environ, 25 novembre 2006.
- [MAG 01] MAGNAN Jean-Pierre, *Cours de mécanique des sols et des roches, Volume 1*. Ecole Nationale des Ponts et Chaussées.
- [MAG 93] MAGNAN Jean-Pierre, *Corrélations entre les propriétés des sols*. Cours mécanique des sols et géotechnique. Techniques de l'ingénieur. 10 février 1993.
- [MAG 97] MAGNAN Jean-Pierre, *Description, identification et classification des sols*. Cours mécanique des sols et géotechnique. Techniques de l'ingénieur. 10 février 1997.
- [MOT 15] MORDVINSTEV A., OLAH C. et TYKA M., *Inceptionism: Going Deeper into Neural Networks*. Google Research Blog, 17 juin 2015
- [PER 03] PERNOD Erik, *Réseaux de neurones. Calculs scientifiques, 3ème année*. [document en ligne].
- [POC 13] P.-O. Chasset, *PNN : une nouvelle bibliothèque R pour la modélisation d'un réseau de neurones probabilistes de Specht*. Rencontres R, Lyon, le 27-28/06/2013. 2 pages.
- [PRA 16] Pragmacom. Article. *Les avantages d'une base de données*. [consulté le 22 février 2016]
- [RAK 01] RAKOTOMALALA Ricco, *Application du réseaux de neurones à l'apprentissage supervisé. Réseaux de neurones artificiels, Perceptron simple et multi-couches*. Tutoriels Tanagra. [document en ligne].
- [ROB 11] ROBERT Jean-Paul, *Forages et sondages*. Cours mécanique des sols et géotechnique. Techniques de l'ingénieur. 10 mai 2011.
- [ROU 01] ROUGIER P. Nicolas, *Perceptron simple Perceptron multi-couche*. Master 2 - Sciences cognitives Université de Bordeaux.
- [SAG 05] SHAHROUR Isam et GOURVES Roland, *Reconnaissance des terrains in situ*. Mécanique et Ingénierie des Matériaux, Traité MIM. Edition hermès sciences. Septembre 2005.
- [SJM 04] SHAHIN Mohammed A., JAKSA Mark et MAIER Holger, *Data Division for Developing Neural Networks Applied to Geotechnical Engineering*. Journal of computing in civil engineering Avril 2004. pages 105-114.
- [SJM 08] SHAHIN Mohammed A., JAKSA Mark et MAIER Holger, *State of the Art of Artificial Neural Networks in Geotechnical Engineering*. 2008, EJGE.
- [SJM 09] SHAHIN Mohammed A., JAKSA Mark et MAIER Holger, *Recent Advances and Future Challenges for Artificial Neural Systems in Geotechnical Engineering Applications*. Hindawi Publishing Corporation Advances in Artificial Neural Systems Volume 2009, Article ID 308239, 9 pages
- [SJM 10] SHAHIN Mohammed A., JAKSA Mark et MAIER Holger, *Investigation into the robustness of Artificial Neural Networks for case study in Civil Engineering*. 2010.
- [SLA 00] Stéphane Lambert. *Un peu d'histoire des SGBD sur le web et ailleurs*. Dossier "SAM l'Informaticien" du 13 novembre au 26 novembre 2000
- [SPS 15] Shreepad S. Sawant et Preeti S. Topannavar. *Introduction to probalistic neural networks - Used for image classifications*. International Journal of Advanced Research in Computer Science and Software Engineering. Volume 5, Issue 4, 2015. 5 pages.

- [SUL 11]** SULEWSKA Maria, *Applying Artificial Neural Networks for analysis of geotechnical problems*. Computer Assisted Mechanics and Engineering Sciences, 18: 231–241, 2011. Institute of Fundamental Technological Research, Polish Academy of Sciences.
- [TBA 04]** *Nature et classification des sols*. Techniques du bâtiment : préparer la construction. Techniques de l'ingénieur. 01 décembre 2004.
- [WAS 83]** WASCHKOWSKI Edwin, *le pénétromètre dynamique*. Bulletin de liaison Laboratoire des Ponts et Chaussées. Mai-juin 1983. Référence : 2805-2806.
- [WIR 09]** WIRA Patrick, *Réseaux de neurones artificiels : architectures et applications*. Cours de master recherche Sciences et Technologies de l'Université de Haute Alsace. Avril 2009.

ANNEXES

Annexe n°1 : Scripts de création SQL BDD in-situ

Annexe n°2 : Scripts de création SQL BDD laboratoire

Annexe n°3 : Graphiques des analyses arbres de décision et analyse de discriminant

Annexe n°4 : Pénétrogrammes bruts et régularisés du sable de Fontainebleau, sable de Lyon et de la grave de chantier

Annexe n°5 : Courbes granulométriques du sable de Fontainebleau, Lyon et de la grave de chantier

Annexe n°6 : Tableau résultats bruts et régularisés des 17 paramètres du sable de Fontainebleau, sable de Lyon et de la grave

Annexe n°7 : Pénétrogrammes des moules hétérogènes

Annexe n°8 : Tableau résultats des 17 paramètres du sable de Dunkerque et des moules hétérogènes

ANNEXE N°1

– Création de la table des sites

```
create TABLE donnees.sites (
    site_id SERIAL PRIMARY KEY,
    site VARCHAR(255),
    coordonnees_X FLOAT,
    UNIQUE (site,coordonnees_X),
    coordonnees_Y FLOAT,
    coordonnees_Z FLOAT
);
```

– Création de la table des sondages

```
CREATE TABLE donnees.sondages (
    site_id INTEGER,
    etude VARCHAR (8),
    date DATE,
    heure TIME DEFAULT NULL,
    n_sondage VARCHAR (10),
    UNIQUE (etude,n_sondage), – permet d'éviter les doublons
    coordonnees_X FLOAT,
    coordonnees_Y FLOAT,
    coordonnees_Z FLOAT,
    CONSTRAINT PK_ETUDE_SONDAGE PRIMARY KEY (etude,n_sondage),
    CONSTRAINT FK_SITE FOREIGN KEY (site_id) REFERENCES donnees.sites
);
```

– Création de la table des mesures du PANDA

```
CREATE TABLE donnees.mesures_PANDA (
    etude varchar (8),
    n_sondage varchar (10),
    toit float DEFAULT 0.00,
    profondeur float,
    section varchar(10) DEFAULT '4 cm2',
    nature varchar(255),
    mesure_id serial primary key,
    qdn varchar(8000),
    zn varchar(8000),
    qdc varchar(8000),
    zc varchar(8000),
    Constraint FK_etude_sondage foreign key (etude,n_sondage) references donnees.sondages
);
```

– Création de la table de valeurs des mesures de labo

```
CREATE TABLE donnees.mesures_labo (
    mesure_id integer,
    GTR varchar (6),
    Code_GTR varchar(8),
    VBS float,
    W float,
    huit_mm float,
```

```
deux_mm float,  
Dmax varchar(10),  
toit float,  
epaisseur float,  
labo_id serial NOT NULL PRIMARY KEY,  
Constraint FK_mesure_id foreign key (mesure_id) references donnees.mesures_PANDA  
];
```

ANNEXE N°2

```

CREATE TABLE donnees.mesures_labo
(
    code_id serial NOT NULL,
    gtr varchar(6),
    code_gtr varchar(8),
    etat_hydrique varchar(2),
    N_moule varchar(8),
    gamma_h double precision,
    gamma_d double precision,
    vbs double precision,
    w double precision,
    Wopn double precision,
    nature varchar(255),
    ip double precision,
    annee integer,
    CONSTRAINT mesures_labo_pkey PRIMARY KEY (code_id),
    CONSTRAINT mesures_laob_code_id UNIQUE (code_id)
);

CREATE TABLE donnees.mesures_granulo
(
    code_id serial NOT NULL,
    Dmax double precision,
    D60 double precision,
    D30 double precision,
    D10 double precision,
    Ch double precision,
    Cc double precision,
    huit_mm double precision,
    deux_mm double precision,
    granulo_id integer,
    CONSTRAINT mesures_granulo_pkey PRIMARY KEY (granulo_id),
    CONSTRAINT mesures_granulo_fk FOREIGN KEY (code_id)
    REFERENCES donnees.mesures_labo (code_id) MATCH SIMPLE
    ON UPDATE NO ACTION ON DELETE NO ACTION
);

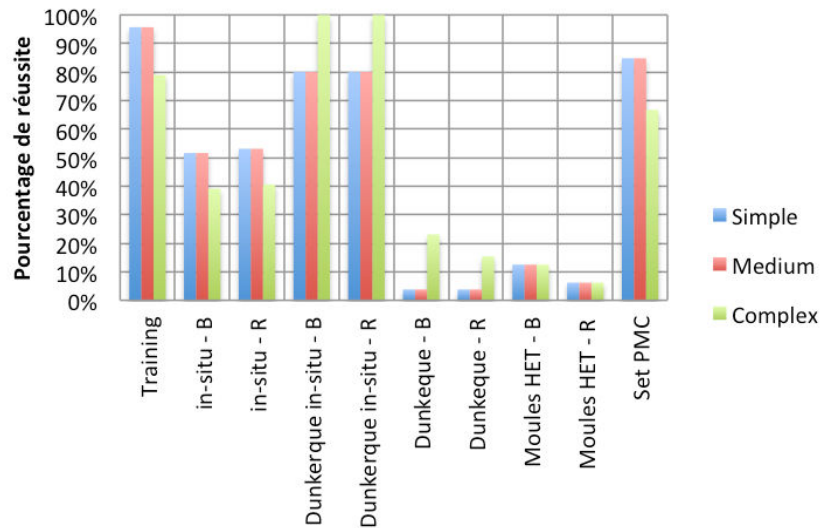
CREATE TABLE donnees.mesures_PANDA
(
    code_id serial NOT NULL,
    mesure_id integer,
    section varchar(10) DEFAULT '2 cm2'::character varying,
    zc double precision,
    qd0 double precision,
    qd1 double precision,
    qdn character varying(8000),
    zn character varying(8000),

```

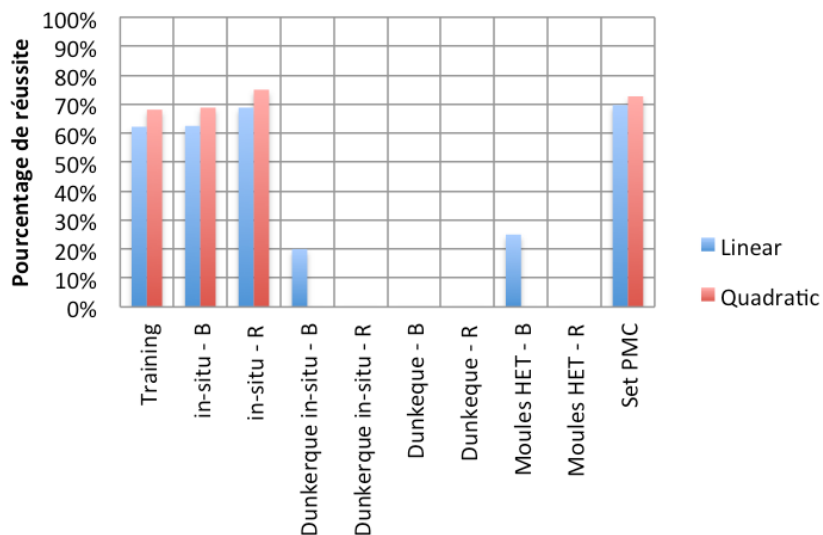
```
qd character varying(8000),  
z character varying(8000),  
panda_id integer,  
CONSTRAINT mesures_panda_pkey PRIMARY KEY (panda_id),  
CONSTRAINT mesures_panda_fk FOREIGN KEY (code_id)  
REFERENCES donnees.mesures_labo (code_id) MATCH SIMPLE  
ON UPDATE NO ACTION ON DELETE NO ACTION  
);
```

ANNEXE N°3 :

Résultats de prédiction pour les arbres de décision

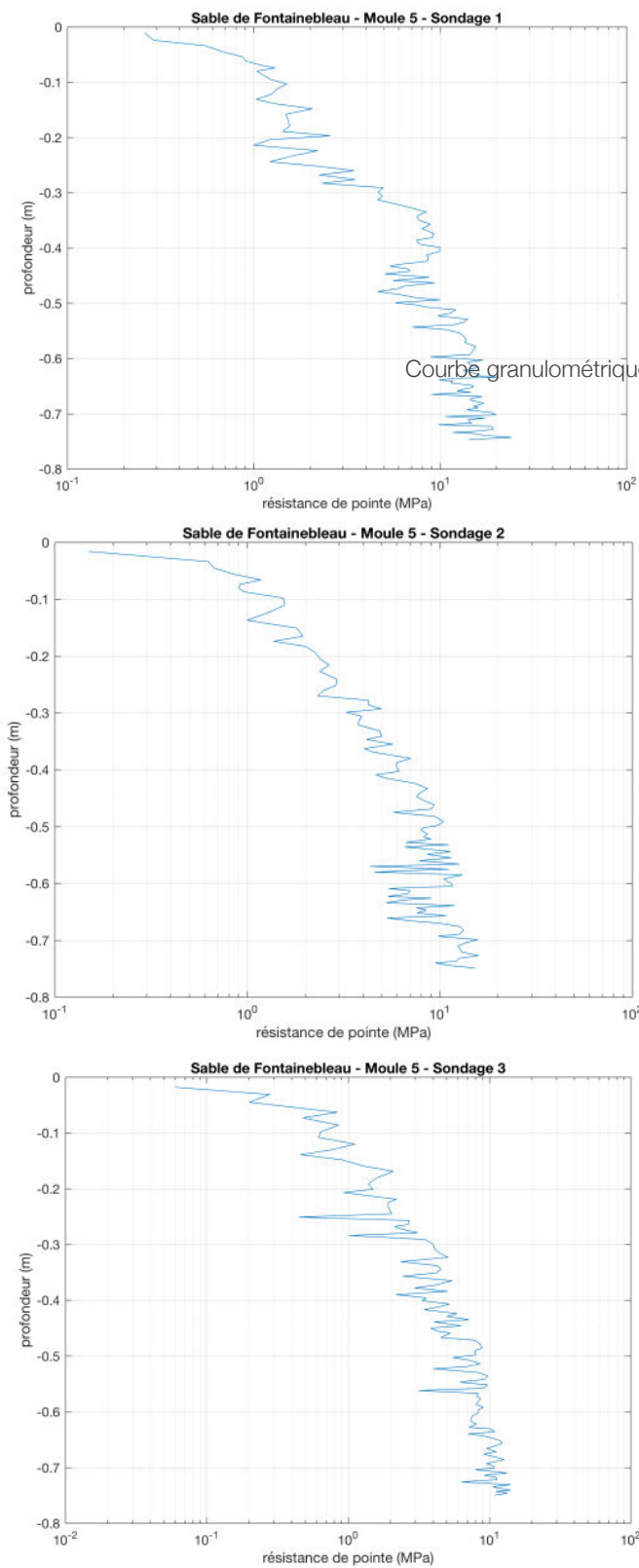


Résultats de prédiction avec une analyse de discriminant

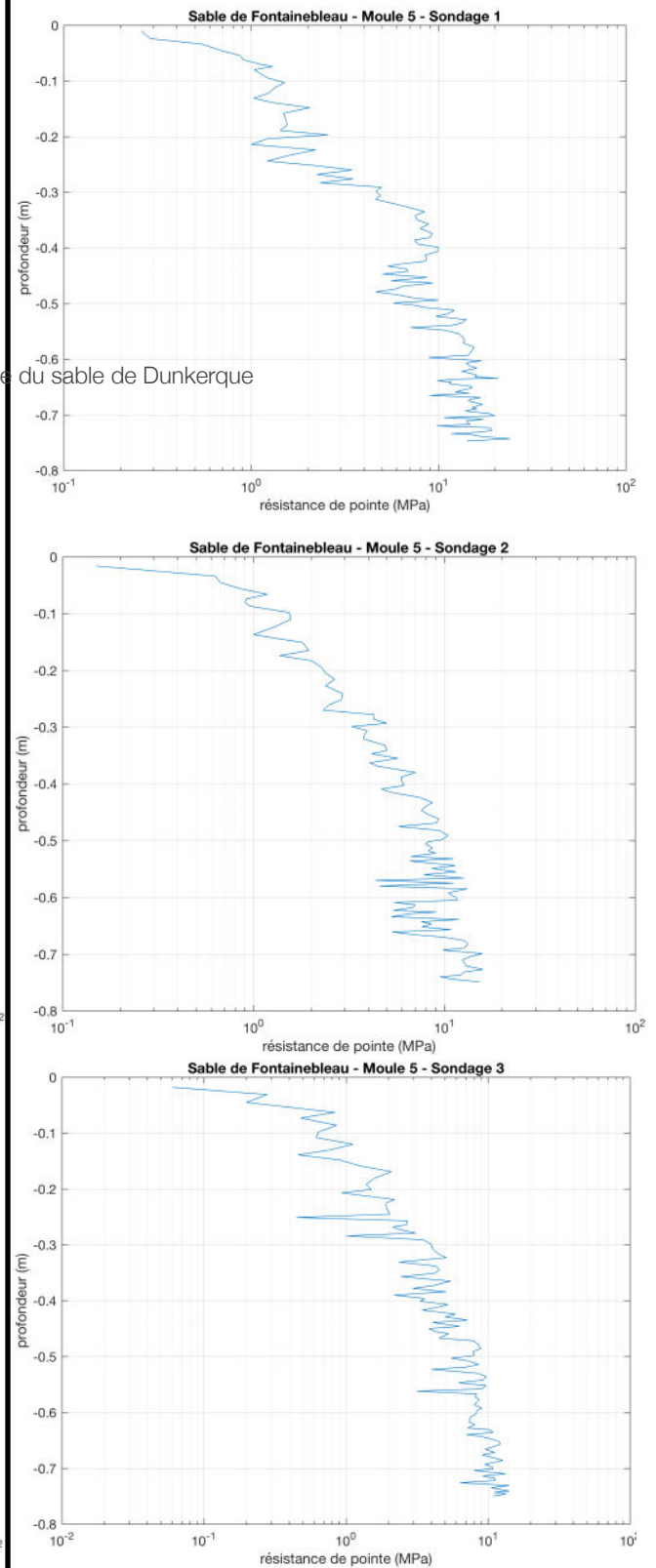


ANNEXE N°4 :

Bruts

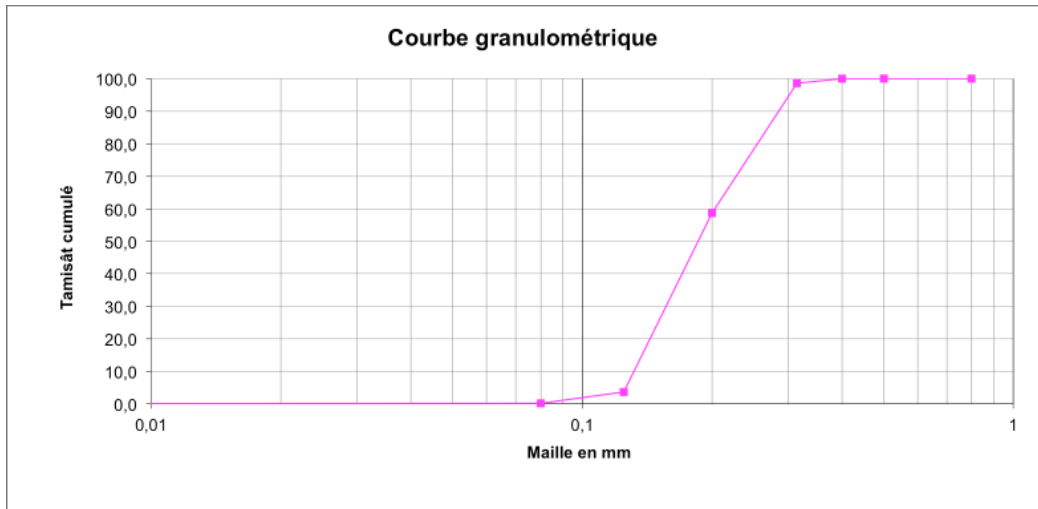


Régularisés

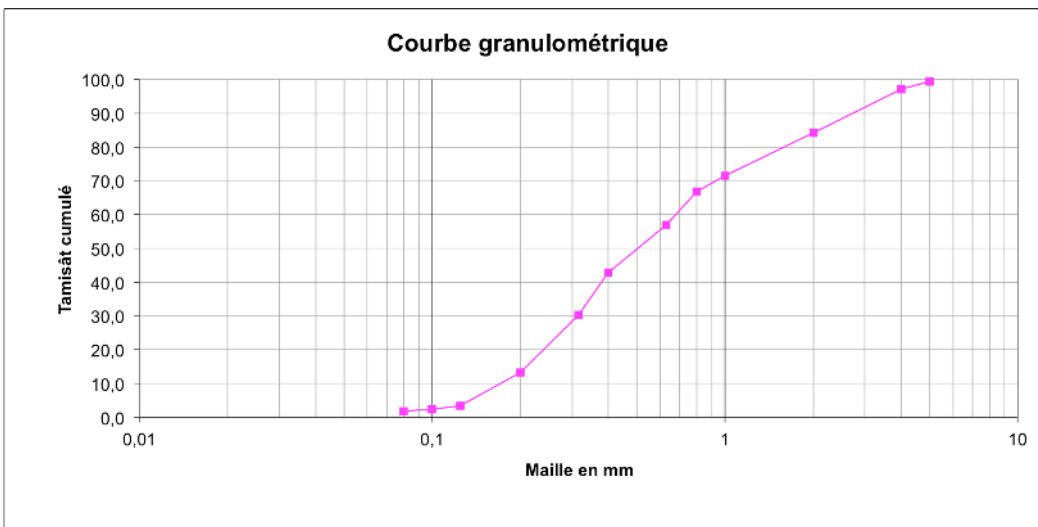


Courbe granulométrique du sable de Dunkerque

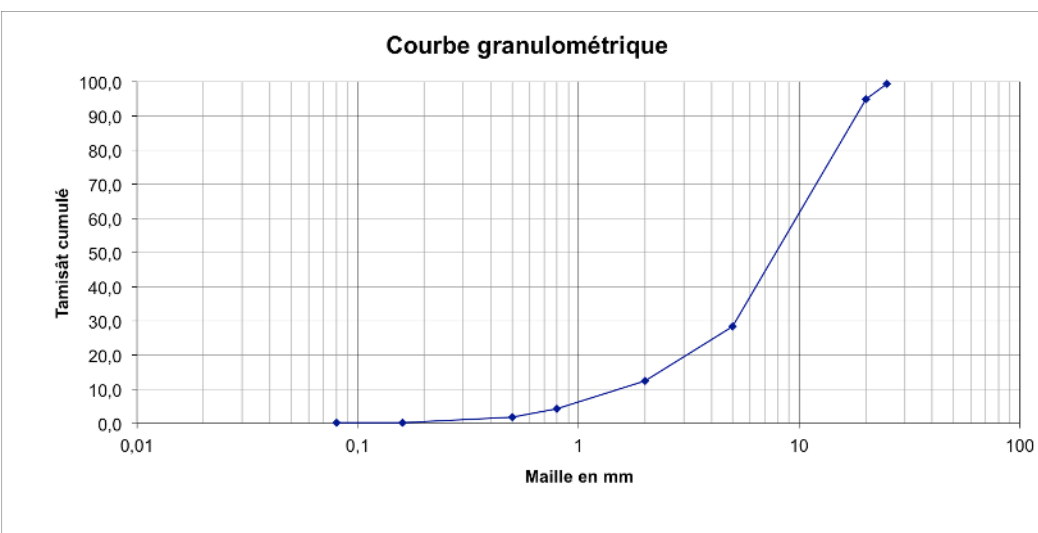
ANNEXE N°5 :



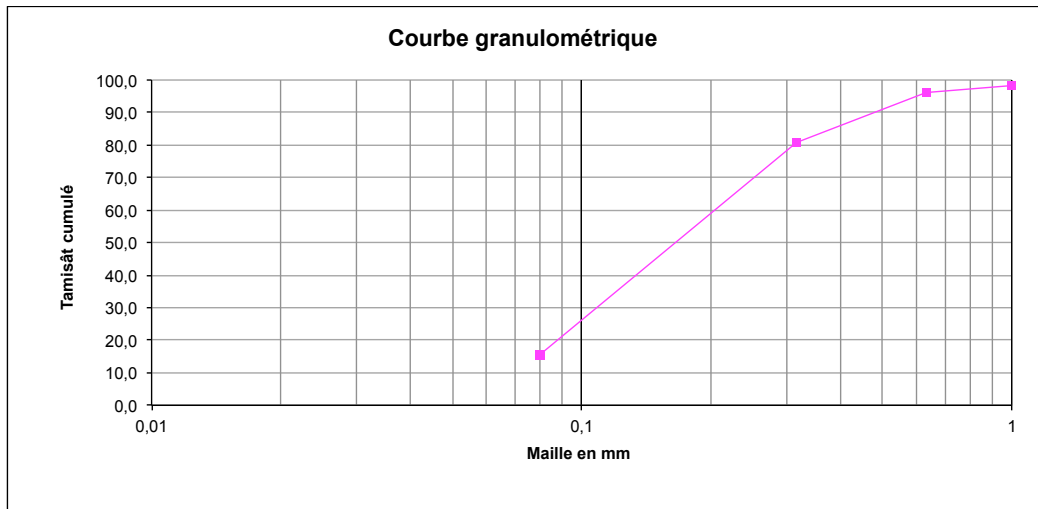
Courbe granulométrique du sable de Fontainebleau



Courbe granulométrique du sable de Lyon



Courbe granulométrique de la grave de chantier



Courbe granulométrique du sable de Dunkerque

ANNEXE N°6 :

		DONNÉES BRUTES																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
Sable de Fontainebleau (D1-1)	N2	S1	0,14	0,12	0,13	1,08	0,02	1,02	0,06	4,44	27,28	0,00	-16,08	1,67	0,00	0,02	0,18	0,04	0,00
		S2	0,15	0,13	0,10	1,46	0,01	0,85	0,06	4,66	32,86	0,00	-32,79	2,33	0,00	0,04	0,22	0,04	0,00
		S3	0,13	0,12	0,08	1,58	0,01	0,51	0,05	2,79	13,62	0,00	-55,29	2,17	0,00	0,00	0,00	0,10	0,00
	Différence	S2-S1	0,01	0,01	-0,03	0,37	-0,01	-0,17	0,00	0,22	5,58	0,00	-16,70	0,67	0,00	0,02	0,05	-0,01	0,00
		S3-S1	-0,01	0,00	-0,05	0,50	-0,01	-0,51	-0,01	-1,65	-13,66	0,00	-39,21	0,50	0,00	-0,01	-0,17	0,06	0,00
		(S2-S1)/S1	3,71%	8,33%	-22,76%	34,27%	-40,34%	-16,67%	0,00%	4,95%	20,44%	#DIV/0!	103,88%	40,00%	#DIV/0!	133,33%	26,37%	-16,28%	-47,89%
	Erreur relative	(S3-S1)/S1	-9,92%	0,00%	-38,35%	46,11%	-61,99%	-50,00%	-16,67%	-37,13%	-50,07%	#DIV/0!	243,80%	30,00%	#DIV/0!	-77,78%	-98,59%	123,16%	-17,89%
		S1	1,57	1,53	0,77	2,03	0,60	3,78	1,08	0,23	2,73	-475,01	-51,74	4,29	1,71	0,12	3,43	0,14	0,01
		S2	1,27	1,32	0,64	1,98	0,41	2,83	1,07	0,21	2,52	-225,68	-51,00	3,14	0,00	0,37	2,90	0,08	0,01
		S3	1,51	1,44	0,80	1,89	0,64	3,43	1,12	0,47	2,48	-459,11	-46,53	4,29	0,86	0,37	3,54	0,10	0,01
	Différence	S2-S1	-0,30	-0,21	-0,13	-0,05	-0,19	-0,95	-0,01	-0,02	-2,41	249,33	0,74	-1,14	-1,71	0,24	-0,54	-0,06	0,00
		S3-S1	-0,06	-0,09	0,02	-0,13	0,04	-0,35	0,04	0,23	-0,25	15,90	5,21	0,00	-0,86	0,25	0,11	-0,04	0,01
		(S2-S1)/S1	-1383,59%	-267,70%	329,21%	-166,72%	-666,00%	-743,50%	-157,80%	-602,58%	#DIV/0!	-2191,69%	947,62%	#DIV/0!	-146,93%	-393,83%	-93,17%	-103,80%	-107,38%
	Erreur relative	(S3-S1)/S1	-1624,39%	-308,03%	310,91%	-202,66%	-786,00%	-770,50%	-225,98%	-595,53%	#DIV/0!	-2008,62%	1328,57%	#DIV/0!	-147,94%	-458,93%	-91,58%	-107,38%	-107,38%
		S1	5,84	4,68	4,60	1,27	21,16	17,93	6,34	0,83	2,76	-33570,43	-19,17	4,71	1,71	2,01	17,72	-1,81	1,66
		S2	4,78	4,41	3,38	1,41	11,43	12,36	5,81	0,43	2,13	-15045,28	3,30	4,33	0,00	1,79	14,90	-1,48	0,40
		S3	5,55	5,95	3,71	1,50	13,75	17,13	6,28	0,36	2,44	-21045,48	30,44	4,43	1,43	2,20	15,33	-0,95	0,31
	Différence	S2-S1	-1,06	-0,27	-1,22	0,14	-9,72	-5,57	-0,53	-0,40	-0,63	18525,15	22,47	-0,38	-1,71	-0,22	-2,82	0,32	-1,26
		S3-S1	-0,28	1,27	-0,89	0,23	-7,41	-0,80	-0,05	-0,47	-0,32	12524,94	49,61	-0,29	-0,29	0,20	-2,39	0,85	-1,35
		(S2-S1)/S1	-129,40%	-209,78%	-54,57%	-664,25%	-257,25%	-175,34%	-119,09%	-135,69%	#DIV/0!	-116,40%	-67,38%	#DIV/0!	-220,77%	-424,74%	61,82%	-128,88%	-128,88%
	Erreur relative	(S3-S1)/S1	-134,20%	-220,36%	-51,81%	-778,30%	-317,94%	-181,51%	-115,89%	-140,89%	#DIV/0!	-251,52%	-66,67%	#DIV/0!	-249,00%	-434,00%	4,22%	-128,85%	-128,85%
		S1	9,05	8,87	5,79	1,56	33,54	23,92	10,09	0,12	2,01	-77508,74	65,58	4,71	2,00	1,70	24,61	-2,09	1,23
		S2	6,95	6,93	4,06	1,71	16,48	15,79	6,08	0,18	2,07	-31977,15	98,31	3,86	1,14	1,17	17,40	-0,91	0,50
		S3	6,03	5,51	3,80	1,58	14,47	14,14	6,54	0,18	1,92	-26338,86	130,31	4,43	2,14	1,93	16,36	-1,39	0,40
	Différence	S2-S1	-2,11	-1,95	-1,73	0,15	-17,07	-8,13	-4,01	0,06	0,05	45531,58	32,74	-0,86	-0,86	-0,53	-7,21	1,18	-0,73
		S3-S1	-3,03	-3,36	-1,99	0,02	-19,07	-9,78	-3,56	0,06	-0,09	51169,88	64,73	-0,29	0,14	0,23	-8,24	0,69	-0,83
		(S2-S1)/S1	-617,65%	-284,20%	-430,29%	-311,71%	-596,64%	-434,98%	-115,46%	-246,82%	#DIV/0!	-4008,75%	-678,57%	#DIV/0!	-147,10%	-500,81%	-2248,14%	-138,91%	-138,91%
	Erreur relative	(S3-S1)/S1	-549,09%	-272,62%	-405,78%	-285,91%	-544,74%	-460,04%	-115,12%	-236,36%	#DIV/0!	-5280,70%	-764,29%	#DIV/0!	-177,45%	-477,01%	-3399,95%	-130,89%	-130,89%

		DONNÉES REGULARISÉES																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
Sable de Fontainebleau (D1-1)	N2	S1	0,11	0,11	0,06	1,75	0,00	0,52	0,07	2,25	14,55	0,00	-23,81	2,00	4,71	0,03	0,14	0,06	0,00
		S2	0,13	0,12	0,06	2,24	0,00	0,29	0,06	0,61	3,62	0,00	-54,38	2,33	6,33	0,02	0,20	0,05	0,00
		S3	0,11	0,11	0,05	2,25	0,00	0,26	0,06	0,74	4,42	0,00	-85,91	2,33	6,83	0,00	-0,02	0,11	0,00
	Différence	S2-S1	0,01	0,01	-0,01	0,49	0,00	-0,23	-0,01	-1,64	-10,93	0,00	-30,57	0,33	1,62	-0,01	0,06	-0,01	0,00
		S3-S1	-0,01	0,00	-0,02	0,50	0,00	-0,26	-0,01	-1,50	-10,13	0,00	-62,10	0,33	2,12	-0,03	-0,16	0,05	0,00
		(S2-S1)/S1	11,77%	9,09%	-12,48%	27,71%	-23,41%	-44,23%	-13,25%	-72,83%	-75,12%	#DIV/0!	128,36%	16,67%	34,34%	-23,40%	45,49%	-16,31%	-22,21%
	Erreur relative	(S3-S1)/S1	-5,55%	0,00%	-26,37%	28,27%	-45,78%	-50,00%	-13,25%	-66,92%	-69,60%	#DIV/0!	260,77%	16,67%	44,95%	-92,08%	-112,85%	96,95%	23,61%
		S1	1,44	1,40	0,76	1,91	0,57	3,78	1,09	0,25	2,67	-552,64	-82,84	4,00	7,29	0,32	3,39	0,17	0,01
		S2	1,20	1,30	0,62	1,92	0,39	2,62	1,05	0,13	2,25	-303,31	-83,89	3,00	7,00	0,12	2,85	0,11	0,01
		S3	1,45	1,33	0,81	1,80	0,65	3,43	1,22	0,44	2,37	-638,75	-79,73	4,00	6,71	0,29	3,47	0,13	0,01
	Différence	S2-S1	-0,25	-0,10	-0,13	0,01	-0,18	-1,16	-0,04	-0,12	-0,42	249,33	-1,05	-1,00	-0,29	-0,19	-0,54	-0,06	0,00
		S3-S1	0,01	-0,06	0,05	-0,11	0,08	-0,35	0,13	0,19	-0,29	-86,11	3,12	0,00	-0,57	-0,02	0,08	-0,03	0,00
		(S2-S1)/S1	-17,00%	-7,17%	-17,46%	0,56%	-31,88%	-30,69%	-3,90%	-47,92%	-15,60%	-45,12%	1,27%	-25,00%	-3,92%	-60,63%	-15,93%	-35,90%	-21,73%
	Erreur relative	(S3-S1)/S1	0,59%	-4,66%	6,86%	-5,87%	14,20%	-9,26%	11,93%	73,76%	-11,04%	15,58%	-3,76%	0,00%	-7,84%	-6,79%	2,33%	-20,51%	38,04%
		S1	4,99	4,13	4,28	1,16	18,35	17,93	5,15	1,10	3,49	-29135,26	-52,91	4,57	8,00	1,74	17,99	-1,86	1,92
		S2	4,34	4,19	3,28	1,32	10,77	12,36	5,60	0,48	2,15	-16901,67	-30,89	5,00	8,14	1,96	15,01	-1,49	0,43
		S3	4,95	4,49	3,46	1,43	11,97	17,13	5,75	0,52	2,69	-22117,14	-6,68	4,00	8,14	0,95	15,10	-0,85	0,27
	Différence	S2-S1	-0,64	0,06	-1,00	0,16	-7,58	-5,57	0,45	-0,62	-1,33	12233,59	22,02	0,43	0,14	0,23	-2,98	0,37	-1,49
		S3-S1	-0,04	0,36	-0,82	0,27	-6,37	-0,80	0,60	-0,58	-0,79	7018,11	46,23	-0,57	0,14	-0,78	-2,89	1,01	-1,64
		(S2-S1)/S1	-12,93%	1,45%	-23,39%	13,66%	-41,31%	-31,07%	8,81%	-56,41%	-38,26%	-41,99%	-41,62%	9,38%	1,79%	12,97%	-16,57%	-19,93%	-77,63%
	Erreur relative	(S3-S1)/S1	-0,76%	8,72%	-19,21%	22,84%	-34,73%	-4,46%	11,65%	-52,79%	-22,73%	-24,09%	-87,37%	-12,50%	1,79%	-45,01%	-16,04%	-54,37%	-85,69%
		S1	7,26	6,92	5,59	1,30	31,29	20,33	10,57	0,42	1,97	-63024,00	22,36	3,71	6,43	1,71	24,43	-1,99	1,82
		S2	5,84	5,00	4,05	1,44	16,42	15,79	6,72	0,48	2,13	-33786,12	53,81	3,71	7,00	1,82	17,39	-0,82	0,63
		S3	4,93	4,20	3,75	1,32	14,03	13,36	6,84	0,47	2,00	-23776,85	85,20	4,57	6,86	2,01	16,48	-1,39	0,40
	Différence	S2-S1	-1,43	-1,92	-1,54	0,14	-14,87	-4,54	-3,85	0,06	0,17	29237,89	31,45	0,00	0,57	-0,36	-7,04	1,17	-1,19
		S3-S1	-2,34	-2,73	-1,85	0,02	-17,27	-6,97	-3,73	0,05	0,03	39247,16	62,85	0,86	0,43	-0,17	-7,95	0,61	-1,42
		(S2-S1)/S1	-19,65%	-27,75%	-27,56%	10,92%	-47,52%	-22,33%	-36,41%	13,97%	8,47%	-46,39%	140,68%	0,00%	8,89%	-16,43%	-28,81%	-58,80%	-65,27%
	Erreur relative	(S3-S1)/S1	-32,17%	-39,38%	-33,05%	1,31%	-55,17%	-34,30%	-35,30%	11,59%	1,67%	-62,27%	281,10%	23,08%	6,67%	-7,59%	-32,56%	-30,36%	-77,87%

		DONNÉES BRUTES																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
Sable de Lyon (D1-2)	N2	S1	1,05	1,17	0,40	2,64	0,16	1,66	0,64	-0,52	2,27	-55,45	138,64	3,43	0,71	0,17	1,48	0,46	0,10
		S2	0,77	0,81	0,30	2,53	0,09	1,35	0,42	-0,32	2,37	0,00	142,86	3,57	0,86	0,16	0,65	0,50	0,09
		S3	0,61	0,59	0,23	2,66	0,05	1,02	0,34	0,10	2,30	0,00	138,73	4,17	0,00	0,17	0,35	0,46	0,05
	Différence	S2-S1	-0,28	-0,36	-0,09	-0,11	-0,07	-0,31	-0,23	0,20	0,11	55,45	4,22	0,14	0,14	-0,01	-0,83	0,04	-0,01
		S3-S1	-0,45	-0,59	-0,17	0,02	-0,11	-0,64	-0,30	0,62	0,04	55,45	0,29	0,74	-0,71	0,00	-1,13	0,00	-0,05
		(S2-S1)/S1	-26,84%	-30,77%	-23,53%	-4,33%	-41,52%	-18,67%	-35,16%	-38,55%	4,80%	-100,00%	3,04%	4,17%	20,00%	-7,63%	-55,99%	9,29%	-9,64%
	Erreur relative	(S3-S1)/S1	-42,28%	-50,00%	-42,65%	0,64%	-67,11%	-38,55%	-46,88%	-118,98%	1,55%	-100,00%	0,06%	21,53%	-100,00%	-2,12%	-76,40%	-0,14%	-49,37%
		S1	4,00	3,33	2,20	1,82	4,83	8,44	3,77	0,24	1,93	-5599,92	165,09	3,00	0,00	0,48	4,29	2,35	6,99
		S2	4,11	4,13	2,21	1,86	4,88	11,44	3,41	0,39	3,27	-6529,82	194,75	3,83	0,00	0,13	4,36	2,44	5,72
	N3	S3	2,91	2,77	1,40	2,07	1,97	6,46	2,04	0,12	2,47	-2505,93	215,61	2,71	0,00	0,24	2,59	1,91	2,88
		S2-S1	0,12	0,80	0,01	0,05	0,04	3,00	-0,36	0,14	1,34	-929,90	29,66	0,83	0,00	-0,35	0,06	0,09	-1,27
		S3-S1	-1,08	-0,56	-0,79	0,26	-2,86	-1,98	-1,73	-0,13	0,54	3093,99	50,53	-0,29	0,00	-0,24	-1,70	-0,44	-4,11
Différence	(S2-S1)/S1	2,94%	23,87%	0,43%	2,49%	0,87%	35,55%	-9,55%	59,65%	69,41%	16,61%	17,97%	27,78%	HDIV/01	-73,17%	1,51%	4,02%	-18,11%	
	(S3-S1)/S1	-27,10%	-16,82%	-36,12%	14,12%	-59,19%	-23,46%	-45,89%	-52,27%	28,18%	-55,25%	30,61%	-9,52%	HDIV/01	-49,13%	-39,66%	-18,66%	-58,86%	
	S1	12,08	11,60	5,79	2,09	3,31	25,73	8,78	0,20	2,25	-135974,67	257,45	6,14	0,00	1,07	7,97	6,89	50,18	
N4	S2	9,35	8,94	5,66	1,65	32,04	35,24	6,96	1,48	7,31	-74274,22	304,58	4,57	1,00	0,22	6,33	5,27	34,44	
	S3	10,00	8,87	6,15	1,63	37,84	33,98	7,78	1,11	4,82	-80346,30	351,37	4,14	0,00	0,02	9,54	4,77	37,01	
	S2-S1	-2,73	-2,66	-0,13	-0,44	-1,47	9,51	-1,83	1,29	5,06	61700,45	47,13	-1,57	0,00	-0,85	-1,64	-1,62	-15,73	
Différence	S3-S1	-2,08	-2,73	0,36	-0,46	4,32	8,20	-1,00	0,92	2,57	55628,37	93,92	-2,00	-1,00	-1,05	1,57	-2,12	-13,17	
	(S2-S1)/S1	-22,62%	-22,93%	-2,22%	-20,86%	-4,38%	36,96%	-20,79%	65,949%	225,33%	-45,38%	18,31%	-25,58%	0,00%	-79,09%	-20,56%	-23,54%	-31,36%	
	(S3-S1)/S1	-17,21%	-23,53%	6,26%	-22,09%	12,90%	31,87%	-11,36%	46,95%	114,29%	-40,91%	36,48%	-32,56%	-100,00%	-98,29%	19,70%	-30,83%	-26,25%	

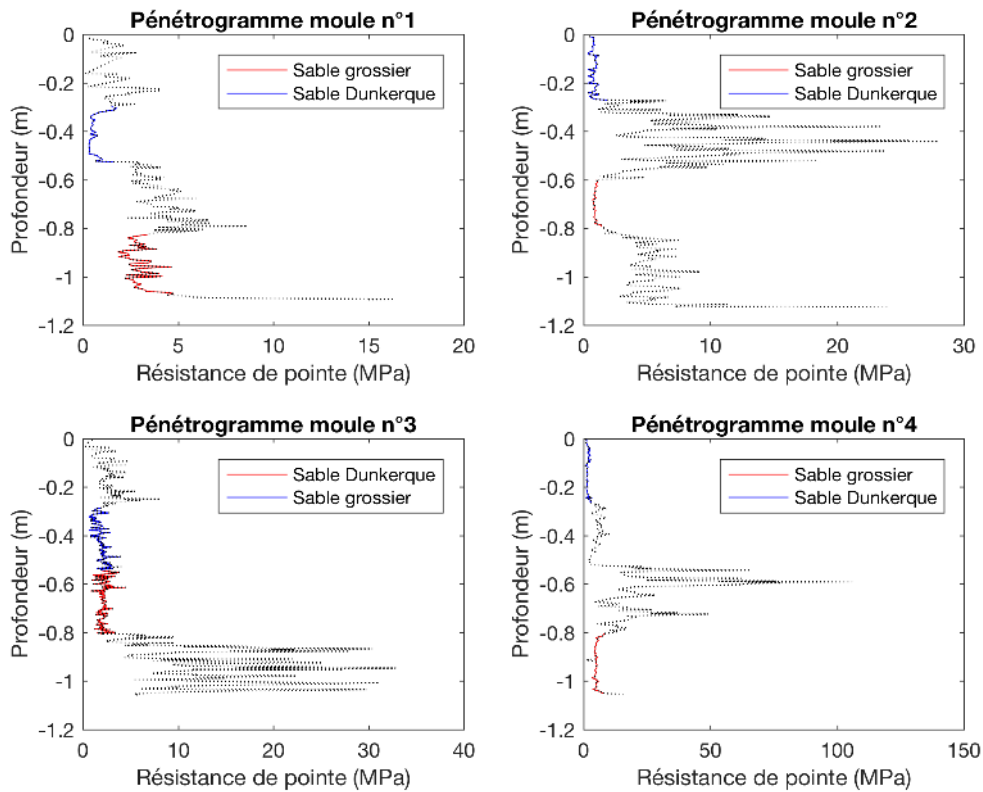
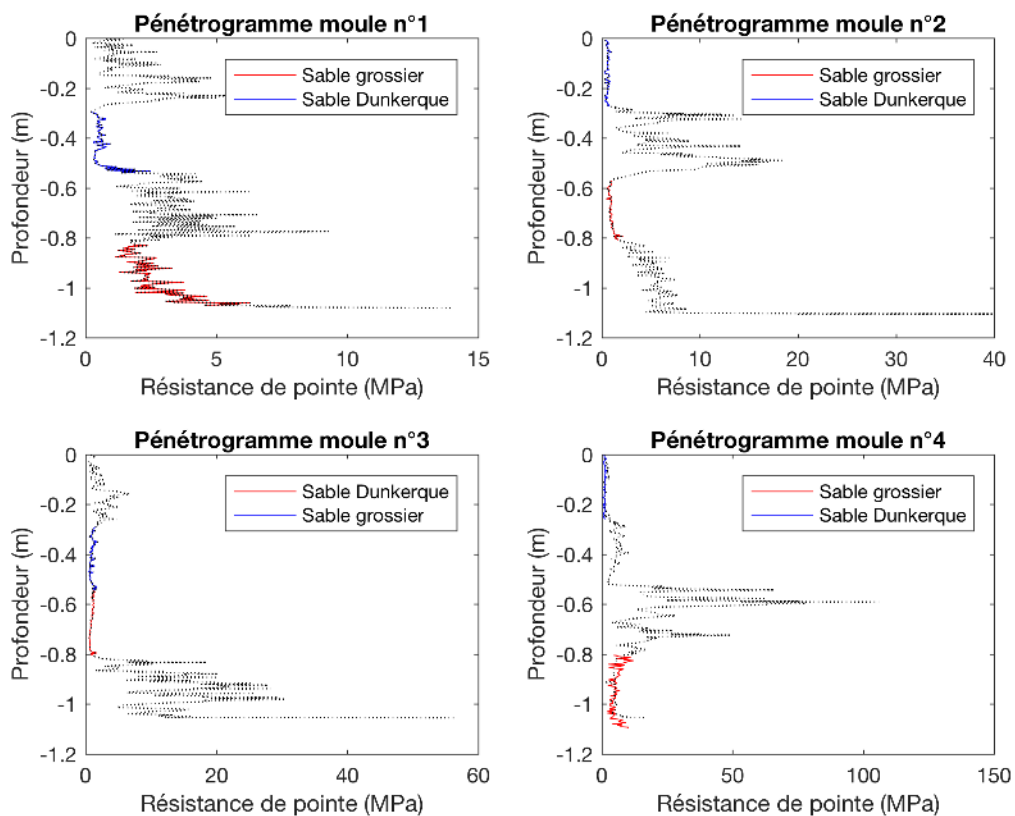
Projet de recherche et développement

		DONNÉES REGULARISÉES																		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17		
Sable de Lyon (D2-1)	N2	S1	1,03	1,14	0,42	2,48	0,17	1,66	0,66	-0,58	2,27	-77,63	92,79	3,86	5,86	0,11	1,47	0,48	0,09	
		S2	0,76	0,80	0,30	2,55	0,09	1,35	0,40	-0,31	2,37	0,00	95,95	3,57	6,71	0,05	0,70	0,50	0,09	
		S3	0,60	0,59	0,22	2,71	0,05	1,02	0,33	0,12	2,31	0,00	91,11	4,57	6,86	0,12	0,33	0,47	0,05	
	Différence	S2-S1	-0,27	-0,34	-0,12	0,07	-0,08	-0,31	-0,26	0,28	0,10	77,63	3,16	-0,29	0,86	-0,06	-0,78	0,02	-0,01	
		S3-S1	-0,43	-0,55	-0,20	0,23	-0,12	-0,64	-0,33	0,70	0,04	77,63	-1,68	0,71	1,00	0,01	-1,14	-0,01	-0,04	
		(S2-S1)/S1	-26,41%	-29,96%	-28,31%	2,64%	-8,60%	-18,67%	-39,39%	47,47%	4,32%	-100,00%	3,41%	-7,41%	14,63%	-56,00%	-52,64%	4,63%	-8,66%	
	Erreur relative	(S3-S1)/S1	-42,06%	-48,02%	-46,96%	9,23%	-71,87%	-38,55%	-50,00%	-119,74%	1,85%	-100,00%	-1,81%	18,52%	17,07%	12,00%	-77,44%	-1,59%	-46,03%	
		N3	S1	4,04	3,58	2,18	1,85	4,77	8,44	3,69	0,14	1,88	-10485,77	115,64	3,43	6,57	0,51	4,49	2,34	6,80
			S2	4,15	4,15	2,07	2,00	4,30	11,44	3,03	0,23	3,27	-10463,01	143,17	3,57	6,14	0,51	4,68	2,40	5,70
	S3		2,93	2,77	1,36	2,16	1,85	5,15	2,01	-0,11	2,05	-3977,15	166,78	2,86	6,00	0,28	2,90	1,85	2,73	
	Différence	S2-S1	0,11	0,57	-0,11	0,15	-0,47	3,00	-0,66	0,09	1,40	22,76	27,52	0,14	-0,43	0,00	0,19	0,06	-1,10	
		S3-S1	-1,10	-0,81	-0,83	0,31	-2,92	-3,29	-1,68	-0,26	0,18	6508,63	51,14	-0,57	-0,57	-0,23	-1,59	-0,49	-4,07	
		(S2-S1)/S1	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	
	Erreur relative	(S3-S1)/S1	-29,31%	-33,25%	-34,51%	7,94%	-57,11%	-54,98%	-33,72%	-147,62%	-37,18%	-61,99%	16,49%	-20,00%	-2,33%	-45,21%	-38,07%	-22,70%	-52,10%	
		N4	S1	10,35	9,41	5,66	1,83	32,06	22,77	8,26	0,37	2,26	-109826,57	207,89	5,00	9,14	0,42	8,20	7,22	50,84
			S2	7,87	6,73	4,72	1,67	22,31	26,38	7,06	0,89	4,16	-61486,21	249,91	4,29	7,57	0,05	6,55	5,39	31,87
S3	8,60		7,28	5,24	1,64	27,46	24,76	6,36	0,80	3,46	-75105,38	292,32	4,14	7,43	1,14	10,00	4,79	34,73		
Différence	S2-S1	-2,48	-2,69	-0,94	-0,16	-9,75	3,61	-1,20	0,52	1,91	48340,35	42,02	-0,71	-1,57	-0,37	-1,65	-1,83	-18,97		
	S3-S1	-1,75	-2,13	-0,42	-0,19	-4,59	1,99	-1,90	0,43	1,21	34721,19	84,42	-0,86	-1,71	0,72	1,80	-2,43	-16,11		
	(S2-S1)/S1	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%		
Erreur relative	(S3-S1)/S1	9,30%	8,25%	10,95%	-1,49%	23,10%	-6,14%	-9,92%	-9,65%	-16,81%	22,15%	16,97%	-3,33%	-1,89%	2159,89%	52,65%	-11,08%	8,97%		

		DONNÉES BRUTES																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
Grave de Chantier (D2-1)	N1	S1	0,43	0,34	0,36	1,18	0,13	1,31	0,53	1,05	3,22	0,00	702,88	0,83	0,00	0,06	0,18	0,18	0,04
		S2	0,52	0,37	0,73	0,71	0,53	4,11	0,27	4,51	23,47	-44,36	719,63	0,67	0,00	0,10	0,47	0,15	0,03
		S3	0,64	0,44	0,92	0,70	0,85	5,14	0,36	4,71	24,67	-80,47	739,77	0,83	0,00	0,17	-0,28	0,54	0,03
	Différence	S2-S1	0,09	0,03	0,37	-0,47	0,40	2,80	-0,26	3,45	20,24	-44,36	16,75	-0,17	0,00	0,04	0,29	-0,03	-0,01
		S3-S1	0,22	0,10	0,56	-0,48	0,71	3,83	-0,17	3,65	21,44	-80,47	36,89	0,00	0,00	0,10	-0,46	0,37	-0,01
	Erreur relative	(S2-S1)/S1	20,56%	8,82%	101,30%	-40,11%	305,20%	213,74%	-48,82%	327,40%	628,21%	#DIV/0!	2,38%	-20,00%	#DIV/0!	58,85%	156,19%	-15,38%	-35,67%
		(S3-S1)/S1	50,32%	27,94%	153,91%	-40,80%	544,71%	292,37%	-31,75%	346,47%	665,47%	#DIV/0!	5,25%	0,00%	#DIV/0!	160,42%	-251,61%	208,67%	-24,49%
	N2	S1	1,90	1,85	0,83	2,30	0,68	3,49	1,24	0,19	2,39	-706,76	757,40	3,17	0,00	0,49	2,32	0,89	0,60
		S2	2,07	2,06	1,15	1,80	1,32	7,88	1,41	1,67	9,43	-1414,37	785,90	4,29	1,57	0,65	3,52	0,60	0,10
		S3	1,43	1,47	0,60	2,37	0,36	3,35	0,93	0,48	3,67	-283,54	802,80	3,00	0,86	0,23	1,67	0,70	0,18
	Différence	S2-S1	0,16	0,21	0,32	-0,50	0,63	4,39	0,18	1,48	7,04	-707,61	28,50	1,12	1,57	0,15	1,20	-0,29	-0,50
		S3-S1	-0,48	-0,38	-0,23	0,07	-0,32	-0,14	-0,31	0,29	1,28	423,23	45,40	-0,17	0,86	-0,27	-0,66	-0,19	-0,42
		Erreur relative	(S2-S1)/S1	8,59%	11,35%	38,77%	-21,74%	92,57%	125,79%	14,37%	771,53%	294,45%	100,12%	3,76%	35,34%	#DIV/0!	31,18%	51,43%	-33,01%
	N3	(S3-S1)/S1	-24,95%	-20,54%	-27,21%	3,10%	-47,01%	-4,01%	-24,90%	150,01%	53,56%	-59,88%	5,99%	-5,26%	#DIV/0!	-54,25%	-28,21%	-21,45%	-69,48%
		S1	6,62	6,96	3,00	2,21	9,00	14,96	4,60	0,09	2,59	-26382,64	837,57	4,57	0,57	1,91	5,77	3,27	5,61
		S2	7,64	8,41	2,80	2,73	7,83	11,17	4,06	-0,66	2,55	-25495,39	877,04	3,29	0,00	0,76	10,87	2,89	3,89
	Différence	S3	4,43	4,61	1,73	2,56	2,99	8,01	2,72	-0,07	2,48	-6574,24	909,76	3,83	0,00	0,90	6,08	1,95	0,41
		S2-S1	1,03	1,45	-0,20	0,53	-1,17	-3,79	-0,54	-0,75	-0,05	887,25	39,47	-1,29	-0,57	-1,15	3,30	-0,38	-1,72
		S3-S1	-2,19	-2,36	-1,27	0,35	-6,01	-6,95	-1,89	-0,16	-0,11	19808,40	72,19	-0,74	-0,57	-1,02	-1,49	-1,32	-5,20
	Erreur relative	(S2-S1)/S1	15,49%	20,83%	-6,75%	23,85%	-13,04%	-25,33%	-11,68%	-830,55%	-1,79%	-3,36%	4,71%	-28,12%	-100,00%	-60,12%	43,57%	-11,53%	-30,66%
		(S3-S1)/S1	-33,12%	-33,84%	-42,34%	15,98%	-66,75%	-46,46%	-40,98%	-176,28%	-4,36%	-75,08%	8,62%	-16,15%	-100,00%	-53,14%	-19,64%	-40,28%	-92,65%
	N4	S1	21,46	23,23	10,73	2,00	115,03	50,94	12,94	-0,32	2,52	-823727,70	966,27	9,14	2,57	3,90	46,82	-1,75	12,12
		S2	20,32	21,91	11,85	1,72	140,31	53,17	15,31	-0,09	2,61	-647564,03	1023,45	7,00	1,86	6,21	48,49	-3,19	23,26
		S3	12,18	12,63	6,10	2,00	37,21	27,80	6,31	-0,04	2,61	-147526,93	1071,86	5,14	1,43	1,78	23,45	0,80	4,91
	Différence	S2-S1	-1,14	-1,32	1,12	-0,29	25,28	2,23	2,37	0,23	0,08	176163,68	57,18	-2,14	-0,71	2,31	1,67	-1,44	11,14
		S3-S1	-9,28	-10,60	-4,62	0,00	-77,81	-23,14	-6,63	0,28	0,08	676200,77	105,59	-4,00	-1,14	-2,12	-23,37	2,55	-7,20
		Erreur relative	(S2-S1)/S1	-5,31%	-5,68%	10,44%	-14,26%	21,98%	4,38%	18,32%	-71,47%	3,26%	21,39%	5,92%	-23,44%	-27,78%	59,16%	3,56%	82,36%
		(S3-S1)/S1	-43,23%	-45,63%	-43,12%	-0,19%	-67,65%	-45,43%	-51,24%	-86,94%	3,28%	-82,09%	10,93%	-43,75%	-44,44%	-54,32%	-49,92%	-146,02%	-59,45%

		DONNÉES REGULARISÉES																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
Grave de Chantier (D2-1)	N1	S1	0,25	0,13	0,26	0,96	0,07	1,31	0,27	2,09	7,88	0,00	615,73	0,71	3,43	0,10	0,11	0,20	0,03
		S2	0,34	0,33	0,18	1,87	0,03	1,10	0,25	1,51	6,76	0,00	617,38	0,71	2,86	0,08	0,42	0,18	0,03
		S3	0,44	0,39	0,20	2,21	0,04	1,00	0,18	1,61	5,58	0,00	619,54	1,14	3,29	0,08	-0,38	0,57	0,02
	Différence	S2-S1	0,09	0,21	-0,08	0,90	-0,03	-0,21	-0,02	-0,58	-1,12	0,00	1,65	0,00	-0,57	-0,03	0,30	-0,02	-0,01
		S3-S1	0,19	0,26	-0,06	1,25	-0,03	-0,31	-0,09	-0,48	-2,31	0,00	3,81	0,43	-0,14	-0,03	-0,50	0,37	-0,01
		Erreur relative	(S2-S1)/S1	37,15%	164,00%	-29,22%	93,77%	-49,90%	-16,03%	-5,98%	-27,72%	-14,21%	#DIV/0!	0,27%	0,00%	-16,67%	-27,17%	270,42%	-10,42%
	(S3-S1)/S1		74,87%	211,08%	-23,71%	129,22%	-41,80%	-23,66%	-31,64%	-23,03%	-29,27%	#DIV/0!	0,62%	60,00%	-4,17%	-26,04%	-440,28%	184,68%	-42,50%
	N2		S1	1,79	1,72	0,84	2,13	0,71	3,49	1,20	0,18	2,31	-1026,51	636,25	3,43	5,00	0,00	2,24	0,92
		S2	1,93	2,04	0,90	2,15	0,80	4,22	1,29	0,16	2,73	-1206,36	654,53	4,29	5,14	0,29	3,35	0,63	0,08
		S3	1,33	1,33	0,58	2,28	0,34	3,35	0,86	0,42	3,79	-383,35	669,86	3,14	6,86	0,36	1,63	0,71	0,17
	Différence	S2-S1	0,14	0,32	0,06	0,03	0,10	0,73	0,09	-0,02	0,42	-179,85	18,28	0,86	0,14	0,29	1,11	-0,30	-0,52
		S3-S1	-0,46	-0,39	-0,26	0,16	-0,37	-0,14	-0,34	0,23	1,08	643,16	33,60	-0,29	1,86	0,36	-0,61	-0,21	-0,43
		Erreur relative	(S2-S1)/S1	8,11%	18,75%	6,63%	1,38%	13,71%	20,82%	7,50%	-12,23%	18,30%	17,52%	2,87%	25,00%	2,86%	11560,38%	49,73%	-32,02%
	(S3-S1)/S1		-25,59%	-22,67%	-30,71%	7,38%	-51,99%	-4,01%	-28,13%	126,97%	46,87%	-62,65%	2,88%	-8,33%	37,14%	14111,32%	-27,29%	-22,77%	-71,27%
	N3		S1	6,22	6,31	2,85	2,18	8,13	13,96	4,41	0,06	2,38	-28999,27	702,46	4,43	6,00	1,35	7,81	3,26
		S2	7,13	6,18	2,88	2,48	8,30	11,17	4,43	-0,46	2,08	-36850,85	741,13	3,29	6,14	1,32	11,00	2,97	3,57
		S3	4,33	4,56	1,66	2,61	2,76	8,01	2,68	-0,08	2,37	-10208,29	773,12	3,71	5,57	0,93	6,15	1,96	0,43
	Différence	S2-S1	0,91	1,86	0,03	0,29	0,17	-2,79	0,02	-0,52	-0,30	-7851,58	38,67	-1,14	0,14	-0,03	3,19	-0,29	-2,15
S3-S1		-1,90	-1,76	-1,19	0,42	-5,37	-5,95	-1,73	-0,14	-0,01	18790,98	70,67	-0,71	-0,43	-0,42	-1,66	-1,29	-5,30	
Erreur relative		(S2-S1)/S1	14,59%	29,50%	1,05%	13,40%	2,11%	-19,96%	0,57%	-851,78%	-12,43%	27,08%	5,51%	-25,81%	2,38%	-2,07%	40,82%	-8,97%	-37,65%
	(S3-S1)/S1	-30,45%	-27,84%	-41,77%	19,44%	-66,09%	-42,60%	-39,16%	-222,82%	-0,60%	-64,80%	10,06%	-16,13%	-7,14%	-31,30%	-21,30%	-39,70%	-92,53%	
	N4	S1	16,39	18,69	11,27	1,45	127,09	36,77	21,72	0,04	1,53	-374462,60	820,95	4,71	7,57	2,60	47,18	-1,55	15,98
S2		15,54	18,48	12,15	1,28	147,67	39,48	22,98	0,05	1,50	-374239,38	869,56	4,71	7,29	0,89	49,04	-3,02	25,35	
S3		9,88	10,59	6,01	1,64	36,07	22,37	10,28	0,16	1,99	-106726,76	912,67	3,86	6,14	3,17	23,64	0,89	5,79	
Différence	S2-S1	-0,85	-0,20	0,88	-0,18	20,58	2,71	1,26	0,01	-0,03	223,23	48,61	0,00	-0,29	-1,72	1,86	-1,46	9,37	
	S3-S1	-6,51	-8,09	-5,27	0,19	-91,03	-14,40	-11,44	0,12	0,46	267735,84	91,72	-0,86	-1,43	0,57	-23,54	2,44	-10,18	
	Erreur relative	(S2-S1)/S1	-5,19%	-1,10%	7,79%	-12,04%	16,19%	7,36%	5,82%	21,53%	-2,18%	-0,06%	5,92%	0,00%	-3,77%	-65,91%	3,94%	94,20%	58,64%
(S3-S1)/S1		-39,74%	-43,31%	-46,73%	13,12%	-71,62%	-39,16%	-52,69%	286,90%	29,70%	-71,50%	11,17%	-18,18%	-18,87%	21,81%	-49,90%	-157,15%	-63,74%	

ANNEXE N°7 :

Pénétrogrammes des moules hétérogènes réalisé avec une pointe de 2 cm²Pénétrogrammes des moules hétérogènes réalisé avec une pointe de 4 cm²

ANNEXE N°8 :

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
M1	P4	0,35	0,29	0,22	1,59	0,05	1,45	0,20	2,93	15,27	-5,55	-7,53	3,00	0,80	0,05	0,50	0,14	0,00
M2	P4	0,81	0,82	0,33	2,44	0,11	1,41	0,51	0,05	2,34	-16,64	-21,08	3,86	0,00	0,07	1,15	0,43	0,02
M3	P4	0,84	0,93	0,41	2,04	0,17	1,69	0,68	0,01	1,94	-22,18	-33,27	2,86	0,57	0,10	1,69	0,21	0,02
M4	P4	1,30	1,23	0,58	2,26	0,33	2,44	0,79	0,24	2,40	-261,36	-35,04	4,57	1,00	0,48	2,41	0,39	0,04
M5	P4	1,84	2,01	0,78	2,36	0,61	3,25	1,11	-0,28	2,34	-801,03	-27,29	5,00	0,43	0,38	3,10	0,74	0,10
M7	P4	0,38	0,32	0,26	1,47	0,07	1,86	0,18	3,85	22,01	-5,55	-6,62	4,25	2,00	0,03	0,14	0,29	0,00
M8	P4	0,86	0,90	0,38	2,28	0,14	1,51	0,49	-0,07	2,30	-22,18	-16,74	3,17	0,00	0,19	1,59	0,17	0,02
M9	P4	0,96	0,92	0,39	2,46	0,15	1,74	0,49	0,02	2,56	-49,91	-25,29	4,43	1,00	0,12	1,51	0,40	0,02
M10	P4	1,30	1,17	0,62	2,11	0,38	2,49	0,86	0,34	2,34	-308,86	-20,83	5,57	1,43	0,31	2,62	0,11	0,02
M13	P4	1,75	1,19	1,44	1,21	2,08	5,86	1,26	1,44	4,01	-1148,74	5,13	3,17	0,00	0,44	4,32	-0,23	0,50
M14	P4	3,33	3,14	1,84	1,81	3,40	7,01	2,46	0,34	2,19	-4766,59	38,80	4,00	1,14	0,64	8,97	-0,24	0,11
M15	P4	1,78	1,60	1,09	1,64	1,18	5,12	1,66	0,63	2,81	-1210,57	55,12	5,57	1,00	0,67	4,63	0,26	0,07
M16	P4	3,03	3,09	1,89	1,60	3,58	6,98	2,94	0,30	2,24	-5544,94	97,26	6,00	1,57	0,85	8,06	0,05	0,22

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
M1	P2	0,54	0,48	0,40	1,36	0,16	3,14	0,14	5,96	42,76	-19,78	-5,73	3,00	0,40	0,18	0,24	0,39	0,01
M2	P2	0,83	0,85	0,32	2,60	0,10	1,35	0,54	-0,08	2,10	-5,55	-10,51	3,17	0,00	0,20	1,41	0,21	0,02
M3	P2	1,07	0,99	0,40	2,64	0,16	2,04	0,57	0,74	3,40	-77,63	-34,51	4,43	0,71	0,26	1,40	0,55	0,02
M4	P2	2,13	2,31	1,01	2,10	1,02	4,10	1,76	0,01	2,08	-1127,99	-34,40	2,83	0,00	0,69	4,63	0,12	0,04
M5	P2	1,86	1,77	1,13	1,64	1,29	12,51	0,83	6,46	61,80	-1608,13	-6,52	5,86	1,43	0,49	2,22	0,96	0,07
M7	P2	0,63	0,54	0,52	1,21	0,27	4,03	0,30	4,95	32,15	-69,68	-1,78	3,33	0,00	0,18	0,33	0,42	0,01
M8	P2	0,76	0,79	0,28	2,76	0,08	1,18	0,42	-0,14	2,32	0,00	-13,06	3,00	0,00	0,18	1,10	0,26	0,03
M9	P2	1,51	1,53	0,67	2,24	0,45	2,81	1,11	0,27	2,19	-474,28	-14,62	5,14	1,14	0,23	2,83	0,46	0,02
M10	P2	1,38	1,48	0,61	2,25	0,38	2,73	0,96	0,00	2,59	-328,63	-21,33	4,43	0,00	0,33	2,62	0,25	0,04
M13	P2	2,87	3,10	1,41	2,03	2,00	5,69	2,25	-0,14	2,07	-2486,06	-6,42	3,86	1,14	1,04	5,91	0,45	0,10
M14	P2	2,72	2,63	1,39	1,95	1,94	5,24	2,61	0,14	1,99	-2377,49	18,34	3,57	0,71	0,50	6,14	0,12	0,09
M15	P2	2,07	1,91	1,06	1,94	1,13	5,89	1,16	0,90	4,30	-1593,49	74,13	6,00	1,00	0,46	4,12	0,58	0,22
M16	P2	3,54	3,54	1,86	1,90	3,47	7,28	2,85	0,21	2,14	-5065,27	120,93	3,83	0,00	0,75	7,90	0,41	0,21

Valeurs de 17 paramètres pour le sable de Dunkerque avec une pointe de 2 et 4 pour des données brutes

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
M1	P4	0,30	0,26	0,12	2,50	0,01	0,62	0,16	1,20	4,62	0,00	-17,23	3,00	5,40	0,03	0,44	0,15	0,00
M2	P4	0,87	0,90	0,32	2,70	0,10	1,41	0,47	-0,30	2,47	-16,64	-50,90	3,71	6,43	0,17	1,15	0,43	0,02
M3	P4	0,86	0,96	0,40	2,14	0,16	1,69	0,68	-0,14	1,93	-33,27	-53,66	3,00	4,86	0,15	1,69	0,22	0,02
M4	P4	1,30	1,25	0,60	2,18	0,36	2,44	0,84	0,15	2,25	-316,81	-56,78	4,86	8,43	0,09	2,36	0,40	0,05
M5	P4	1,92	2,12	0,78	2,47	0,60	3,25	1,00	-0,44	2,49	-1028,18	-44,13	4,57	7,14	0,17	3,15	0,73	0,11
M7	P4	0,32	0,30	0,12	2,79	0,01	0,57	0,15	1,01	4,22	0,00	-60,90	4,50	7,00	0,02	0,11	0,29	0,00
M8	P4	0,79	0,80	0,39	2,00	0,15	1,51	0,54	-0,01	2,15	-27,73	-71,82	3,57	5,86	0,12	1,61	0,17	0,02
M9	P4	0,98	0,99	0,39	2,50	0,15	1,74	0,44	-0,12	2,50	-61,00	-73,54	4,71	9,14	0,36	1,50	0,41	0,02
M10	P4	1,13	1,08	0,61	1,84	0,37	2,42	0,82	0,43	2,55	-250,27	-75,93	4,43	7,00	0,22	2,62	0,12	0,02
M13	P4	1,45	1,10	1,20	1,20	1,44	5,86	0,79	2,07	6,80	-1145,90	-51,28	3,71	6,14	0,80	4,39	-0,24	0,55
M14	P4	3,14	2,89	1,95	1,61	3,80	7,01	3,16	0,35	2,06	-6248,42	-26,70	4,43	5,57	0,56	8,97	-0,22	0,12
M15	P4	1,96	2,04	1,06	1,84	1,13	5,12	1,62	0,33	2,63	-1491,50	-18,41	4,71	6,71	0,56	4,54	0,27	0,07
M16	P4	3,09	3,23	1,85	1,67	3,42	6,98	2,66	0,27	2,33	-6109,33	5,46	5,14	9,00	1,21	8,09	0,04	0,19

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
M1	P2	0,48	0,47	0,13	3,61	0,02	0,91	0,14	1,33	7,85	0,00	-29,25	3,20	5,40	0,08	0,19	0,40	0,00
M2	P2	0,78	0,77	0,33	2,35	0,11	1,35	0,53	0,00	2,04	-5,55	-46,59	3,33	5,50	0,12	1,41	0,23	0,02
M3	P2	1,10	1,03	0,40	2,76	0,16	2,04	0,60	0,55	3,07	-133,08	-55,48	4,57	6,71	0,21	1,44	0,55	0,02
M4	P2	1,94	2,02	1,01	1,93	1,01	4,10	1,61	0,14	2,08	-1358,28	-51,43	3,57	6,17	0,50	4,65	0,13	0,03
M5	P2	1,81	1,81	0,62	2,92	0,38	3,69	0,74	0,49	3,99	-821,95	-39,12	5,86	9,00	0,33	2,32	0,91	0,06
M7	P2	0,53	0,50	0,21	2,47	0,05	1,80	0,22	2,67	17,58	-5,55	-43,37	3,33	4,67	0,08	0,25	0,43	0,00
M8	P2	0,71	0,73	0,28	2,51	0,08	1,18	0,42	-0,02	2,18	0,00	-67,79	2,83	5,67	0,17	1,10	0,27	0,02
M9	P2	1,53	1,58	0,66	2,30	0,44	2,81	1,17	0,15	1,97	-557,45	-70,95	4,86	7,71	0,33	2,89	0,43	0,01
M10	P2	1,27	1,37	0,63	2,01	0,40	2,73	0,94	0,05	2,38	-395,18	-76,15	4,43	6,29	0,19	2,59	0,27	0,04
M13	P2	2,73	2,87	1,38	1,98	1,91	5,69	2,23	-0,07	2,00	-3548,99	-62,04	4,00	6,14	0,28	5,94	0,49	0,09
M14	P2	2,49	2,51	1,35	1,84	1,83	5,22	1,96	0,21	2,07	-2967,41	-41,64	3,43	6,00	0,55	6,10	0,15	0,08
M15	P2	2,13	1,95	1,03	2,06	1,07	5,37	0,99	0,71	3,48	-1867,44	-8,78	5,86	8,14	0,78	3,99	0,62	0,18
M16	P2	3,45	3,52	1,83	1,89	3,33	7,28	2,80	0,20	2,17	-6817,39	23,10	4,14	6,67	0,74	8,14	0,34	0,24

Valeurs de 17 paramètres pour le sable de Dunkerque avec une pointe de 2 et 4 pour des données régularisées

Projet de recherche et développement

			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
D2	M1	P4	3,01	2,43	1,65	1,82	2,73	7,52	1,74	1,61	5,23	-2276,16	80,90	7,50	10,50	1,71	12,71	-9,35	0,35
	M2	P4	1,05	0,97	0,37	2,85	0,14	2,12	0,32	1,77	7,40	-41,96	136,17	5,50	6,50	0,54	-0,05	1,08	0,12
	M3	P4	1,01	0,98	0,32	3,13	0,10	1,30	0,51	0,41	2,22	-72,09	184,58	11,50	13,00	0,20	-1,45	1,57	0,04
	M4	P4	5,71	5,29	2,22	2,57	4,91	9,89	2,75	0,64	3,02	-11690,11	317,32	7,50	10,50	1,07	-8,41	13,59	2,07
D1	M1	P4	0,70	0,55	0,45	1,56	0,20	2,29	0,34	1,92	6,61	-22,18	88,47	7,00	9,50	0,06	1,17	0,06	0,02
	M2	P4	0,59	0,60	0,16	3,75	0,02	0,76	0,13	-0,28	3,19	0,00	133,15	6,00	11,50	0,22	-0,25	0,61	0,01
	M3	P4	0,99	0,99	0,27	3,64	0,07	1,46	0,42	0,60	3,61	-22,18	179,74	11,50	18,00	0,06	-1,69	2,06	0,03
	M4	P4	1,14	1,15	0,30	3,79	0,09	1,85	0,28	0,20	4,52	-49,91	323,25	14,00	21,50	0,46	-0,43	1,13	0,01

			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
D2	M1	P2	3,22	2,86	1,13	2,86	1,27	5,92	0,95	1,89	7,35	-1508,77	15,98	5,00	7,50	0,58	5,20	-1,90	0,27
	M2	P2	1,71	1,08	1,20	1,43	1,44	4,94	1,28	1,55	4,70	-494,48	108,14	4,50	5,00	0,56	-0,79	2,01	1,10
	M3	P2	1,97	1,95	0,65	3,01	0,43	3,37	0,70	0,03	3,15	-740,97	158,12	11,00	15,50	1,66	4,08	0,15	0,05
	M4	P2	5,00	4,88	1,10	4,56	1,20	7,08	0,54	0,06	6,83	-4572,58	205,81	6,50	7,00	2,29	-3,96	8,70	0,73
D1	M1	P2	1,24	0,76	1,02	1,21	1,05	3,26	1,23	0,98	2,55	-260,42	26,51	3,50	4,00	0,04	-1,54	1,43	0,36
	M2	P2	0,83	0,84	0,29	2,86	0,08	1,58	0,24	0,62	5,58	-5,55	111,96	6,00	6,50	0,42	0,70	0,70	0,01
	M3	P2	2,25	2,27	0,60	3,78	0,36	3,67	0,56	0,45	4,37	-1130,40	148,57	12,50	18,50	0,64	-1,36	3,08	0,04
	M4	P2	1,74	1,60	0,59	2,94	0,35	2,86	0,86	0,69	3,00	-470,20	220,77	12,50	14,50	0,60	1,10	1,45	0,15

Valeurs de 17 paramètres pour le sable grossier et le sable de Dunkerque avec une pointe de 2 et 4 cm² pour des données brutes

			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
D2	M1	P4	2,51	2,29	0,90	2,79	0,81	4,43	1,14	1,20	4,69	-790,88	25,64	7,00	11,50	1,53	8,62	-5,68	0,07
	M2	P4	0,97	0,95	0,22	4,44	0,05	0,98	0,25	0,82	3,67	-11,09	24,38	5,00	7,00	0,29	2,27	-0,60	0,01
	M3	P4	0,98	0,95	0,28	3,57	0,08	1,04	0,36	0,45	2,36	-16,64	24,87	5,00	7,00	0,18	-1,44	1,57	0,03
	M4	P4	5,50	5,23	1,95	2,83	3,79	8,13	2,46	0,58	2,96	-7429,30	86,43	4,50	8,00	1,91	-9,00	14,00	1,61
D1	M1	P4	0,58	0,49	0,31	1,84	0,10	1,62	0,26	2,30	8,42	-11,09	28,03	5,50	10,00	0,12	1,41	-0,02	0,03
	M2	P4	0,57	0,61	0,14	4,18	0,02	0,57	0,14	-0,77	3,06	0,00	16,75	5,00	7,00	0,20	-0,13	0,58	0,01
	M3	P4	0,92	0,88	0,24	3,82	0,06	0,99	0,41	0,37	2,25	-5,55	25,56	6,50	8,50	0,02	-2,28	2,45	0,01
	M4	P4	1,08	1,10	0,23	4,63	0,05	1,17	0,34	-0,43	3,14	-5,55	90,46	6,50	10,00	0,43	-0,62	1,13	0,02

			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
D2	M1	P2	2,92	2,77	0,67	4,35	0,45	2,94	0,82	1,05	3,84	-1055,74	9,28	5,00	7,50	0,70	2,32	0,69	0,12
	M2	P2	0,93	0,91	0,14	6,47	0,02	0,70	0,15	1,52	5,92	0,00	26,35	5,00	9,00	0,19	-0,08	0,98	0,01
	M3	P2	1,82	1,93	0,53	3,46	0,28	2,18	0,80	0,02	2,42	-262,09	27,42	5,00	8,00	0,42	5,08	-0,28	0,03
	M4	P2	5,09	4,91	0,91	5,59	0,83	4,81	0,62	1,58	6,32	-4406,46	47,39	5,50	4,50	2,06	-4,49	9,21	0,36
D1	M1	P2	0,63	0,49	0,38	1,66	0,14	1,45	0,39	1,62	4,77	-22,18	9,52	3,00	4,00	0,01	-2,76	1,74	0,08
	M2	P2	0,78	0,81	0,23	3,41	0,05	0,92	0,28	-0,44	2,71	0,00	24,88	6,00	8,00	0,42	0,46	0,71	0,02
	M3	P2	2,15	2,09	0,43	5,05	0,18	1,88	0,53	0,63	3,00	-419,56	24,78	6,50	9,00	0,43	-1,28	3,00	0,03
	M4	P2	1,62	1,53	0,48	3,36	0,23	2,22	0,66	1,02	3,77	-197,95	58,42	5,50	9,50	0,35	0,33	1,52	0,10

Valeurs de 17 paramètres pour le sable grossier et le sable de Dunkerque avec une pointe de 2 et 4 cm² pour des données régularisées