

Influences de la méthode d'échantillonnage et de la présence
d'une phase liquide sur l'estimation de la teneur en matière
résiduelle d'un sol par analyse visuelle et intelligence
artificielle

par

Clélia MARCHAND

MÉMOIRE PRÉSENTÉ À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
COMME EXIGENCE PARTIELLE À L'OBTENTION DE
LA MAÎTRISE EN GÉNIE DE L'ENVIRONNEMENT
M. Sc. A.

MONTRÉAL, LE 27 NOVEMBRE 2023

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC



Clélia Marchand, 2023



Cette licence [Creative Commons](https://creativecommons.org/licenses/by-nc-nd/4.0/) signifie qu'il est permis de diffuser, d'imprimer ou de sauvegarder sur un autre support une partie ou la totalité de cette œuvre à condition de mentionner l'auteur, que ces utilisations soient faites à des fins non commerciales et que le contenu de l'œuvre n'ait pas été modifié.

PRÉSENTATION DU JURY

MÉMOIRE A ÉTÉ ÉVALUÉ

PAR UN JURY COMPOSÉ DE :

M. Jean-Sébastien Dubé, directeur de mémoire
Département de génie de la construction à l'École de technologie supérieure

M. François Duhaime, codirecteur de mémoire
Département de génie de la construction à l'École de technologie supérieure

M. Jonathan D. Aubertin, président du jury
Département de génie de la construction à l'École de technologie supérieure

M. Adrien Dimech, membre du jury
Département de génie de la construction à l'École de technologie supérieure

IL A FAIT L'OBJET D'UNE SOUTENANCE DEVANT JURY ET PUBLIC

LE 6 NOVEMBRE 2023

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

REMERCIEMENTS

Au cours de cette dernière année en tant qu'étudiante en maîtrise, j'ai pu explorer le domaine de la recherche et j'ai eu la chance de croiser la route de nombreuses personnes qui ont enrichi mon expérience. Je tiens à les remercier sincèrement.

Je tiens avant tout à remercier mon directeur de recherche, Monsieur Jean-Sébastien Dubé, et mon codirecteur, Monsieur François Duhaime pour leur encadrement, leur implication et leur disponibilité. J'exprime ma sincère gratitude envers Jean-Sébastien pour son engagement et sa supervision tout au long de cette année, qui m'ont offert à la fois des périodes d'indépendance et un soutien précieux lorsque j'en avais besoin. De même, je tiens à remercier chaleureusement François Duhaime pour son accompagnement lors de la phase de mon travail consacrée au réseau neuronal.

Je tiens également à témoigner ma profonde reconnaissance à celles et ceux qui m'ont aidée à me familiariser avec les procédures de laboratoire nécessaires à ma recherche. Pour tout ce qui touchait au laboratoire de géotechnique, je tiens tout particulièrement à remercier Alexis Vadeboncoeur pour son précieux soutien tout au long de cette année, mais également Mathilde Dol qui a été ma collègue à mes débuts et est aujourd'hui une précieuse amie, ainsi que Mirela Sona. Je remercie Marielle Jacques pour son aide lors de mon bref passage à la STEPPE.

Je souhaite exprimer ma reconnaissance envers mes amis et mes colocataires pour leur compagnie et leur soutien tout au long de mon travail. Un remerciement tout particulier revient à Céline Bouin, avec qui j'ai partagé des sessions de rédaction sur nos projets respectifs. J'ai grandement apprécié ces moments qui ont été une source majeure de motivation pour moi.

Enfin, un grand merci à ma famille et mes proches pour leur soutien constant et indéfectible, ainsi que leur confiance en moi tout au long de mon parcours académique, y compris lors de la réalisation de cette recherche qui en marque l'aboutissement.

Influences de la méthode d'échantillonnage et de la présence d'une phase liquide sur l'estimation de la teneur en matière résiduelle d'un sol par analyse visuelle et intelligence artificielle

Clélia MARCHAND

RÉSUMÉ

La caractérisation des sols urbains est cruciale pour leur gestion environnementale et la planification de leur réhabilitation. Cette recherche s'intéresse au cas de remblais contenant des matières résiduelles et plus précisément, des briques de construction. Dans le cadre de ce travail, dix essais d'échantillonnage combinant dix configurations de méthode d'échantillonnage, présence de liquide et, de manière secondaire, de présence de particules fines ont été réalisés. À travers ces dix essais, l'objectif a été d'étudier la représentativité des configurations via des analyses granulométriques et de concentration de matière résiduelle pour une analyse d'images ultérieure par un réseau neuronal convolutif. Les lots utilisés pour les essais étaient composés de particules de sol graveleux inférieures à 2,5 mm, et sans particules inférieures à 160 µm pour les essais de sol sans particules fines. Ces lots étaient artificiellement contaminés avec de la brique particulaire à une concentration massique de 30%. Les méthodes d'échantillonnage employées ont été le grappillage, le pelletage fractionné en 20 incréments et la séparation rotative. Les phases liquides introduites au cours des essais humides étaient l'eau et l'huile de canola, substitut sécuritaire aux produits pétroliers.

Cette recherche a confirmé la hiérarchie de représentativité des trois méthodes d'échantillonnage étudiées avec dans l'ordre croissant de représentativité, le grappillage étant la moins représentative, suivie du pelletage fractionné et enfin de la séparation rotative (applicable seulement aux essais secs). La présence d'une phase liquide, a considérablement amélioré la représentativité, en particulier pour le grappillage. La présence de fines a principalement affecté l'erreur fondamentale, mais cet effet était minime par rapport aux erreurs de ségrégation et de groupement. Cependant, des pertes de sol, y compris des particules fines, et de briques ont pu légèrement biaiser les résultats granulométriques et de quantification de briques, en particulier pour les essais humides.

Les résultats en lien avec l'analyse d'images ont montré l'importance de rééquilibrer la base de données photographiques utilisée pour l'entraînement et la validation du réseau neuronal pour que ce dernier soit en mesure de détecter toutes les concentrations en briques, notamment les plus extrêmes. Une base de données diversifiée et conséquente a nécessité un réseau neuronal complexe pour converger. L'ajustement des hyperparamètres a eu des effets spécifiques sur les performances. Les meilleures configurations de réseaux généraient des pertes autour de 0,012 et des racines carrées d'erreurs quadratiques moyennes à 0,035. De plus, le réseau a identifié en général des faux positifs et négatifs similaires à l'échantillonnage, renforçant ainsi la confiance en ce dernier. Cependant, l'erreur intrinsèque du réseau n'a fait encore que s'ajouter à l'erreur d'échantillonnage. Un objectif futur serait de la corriger.

VIII

Mots-clés : Matière résiduelle, méthode d'échantillonnage, phase liquide, particules fines, granulométrie, représentativité, analyse d'images, réseau neuronal convolutif, régression

Influences of sampling method and the presence of a liquid phase on the estimation of residual material content in soil through visual analysis and artificial intelligence

Clélia MARCHAND

ABSTRACT

The characterization of urban soils is crucial for their environmental management and rehabilitation planning. This research focuses on fill materials containing residual materials, specifically construction bricks. In the course of this work, ten sampling tests were conducted, combining ten configurations of sampling method, the presence of liquid, and, secondarily, the presence of fine particles. Through these ten tests, the goal was to study the representativeness of the configurations through granulometric and residual material concentration analyses for subsequent analysis by a convolutional neural network. The lots used for the tests were composed of gravelly soil particles smaller than 2.5 mm, with no particles smaller than 160 μm for tests without fine particles. These lots were artificially contaminated with particulate bricks at a mass concentration of 30%. The sampling methods employed were grab sampling, fractionated shoveling in 20 increments, and rotary separation. The liquid phases introduced during wet tests were water and canola oil, a safe substitute for petroleum products.

This research confirmed the hierarchy of representativeness of the three studied sampling methods, with grab sampling being the least representative, followed by fractionated shoveling and rotary separation (only applicable to dry tests). The presence of a liquid phase significantly improved representativeness, especially for grab sampling. The presence of fines mainly affected the fundamental error, but this effect was minimal compared to segregation and grouping errors. However, soil and brick losses, including fine particles, slightly biased granulometric and brick quantification results, especially for wet tests.

The results related to the image analysis highlighted the importance of rebalancing the photographic database used for training and validating the neural network to enable it to detect all brick concentrations, especially the most extreme ones. A diverse and extensive database required a complex neural network to converge. The adjustment of hyperparameters had specific effects on performance. The best network configurations generated losses around 0.012 and root mean square errors of around 0.035. Furthermore, the network generally identified false positives and false negatives similar to the sampling, thereby reinforcing confidence in the latter. However, the network's intrinsic error still added to the sampling error. A future objective would be to correct this intrinsic error.

Keywords: Residual material, sampling method, liquid phase, fines, granulometry, representativeness, image analysis, convolutional neural network, regression

TABLE DES MATIÈRES

	Page
INTRODUCTION	1
CHAPITRE 1 REVUE DE LITTERATURE	3
1.1 Contexte de la contamination des sols	3
1.1.1 Les mélanges de sol et matière résiduelle	4
1.1.2 Les produits pétroliers	5
1.2 L'échantillonnage de sol à des fins d'analyse granulométrique et quantification de matières résiduelles	6
1.2.1 Définitions	6
1.2.2 Théorie de l'échantillonnage des matières particulières	8
1.2.3 Méthodes d'échantillonnage	10
1.2.3.1 Grappillage	11
1.2.3.2 Pelletage fractionné	11
1.2.3.3 Séparation rotative	12
1.2.4 Comparaison des différentes méthodes d'échantillonnage	14
1.3 Analyse granulométrique et quantification des matières résiduelles particulières	16
1.3.1 Analyse granulométrique	16
1.3.2 Teneur en matières résiduelles particulières	17
1.3.2.1 Quantification de la teneur en matières résiduelles particulières	17
1.3.2.2 Éléments d'évaluation de la représentativité d'un échantillonnage	18
1.3.3 Analyse granulométrique et particulaire par analyse d'images	19
1.3.3.1 Réseau de neurones artificiels (ANN)	19
1.3.3.2 Réseau de neurones convolutif (CNN)	20
1.3.3.3 Bases de données photographiques	24
CHAPITRE 2 DESCRIPTION MÉTHODOLOGIQUE	25
2.1 Description et préparation des matériaux utilisés	25
2.1.1 Présentation des essais effectués	25
2.1.2 Description et préparation du sol	29
2.1.3 Description et préparation des fragments de brique ajoutés au sol	30
2.2 Déroulement des essais	33
2.2.1 Méthodes d'échantillonnage	33
2.2.1.1 Grappillage	33
2.2.1.2 Pelletage fractionné	34
2.2.1.3 Séparation rotative	35
2.2.2 Photographie	37
2.2.2.1 Le montage photo et les réglages	37
2.2.2.2 Procédure de prise de photographie	39

2.2.3	Enlèvement et quantification de la phase liquide	40
2.2.3.1	Eau	40
2.2.3.2	Huile.....	40
2.2.4	Granulométrie et séparation	42
2.3	Analyse d'image	45
2.3.1	Base de données photographique	46
2.3.1.1	Base de données photographique initiale (2072 photographies)	46
2.3.1.2	Base de données photographique rééquilibrée (3195 photographies)	50
2.3.1.3	Redimensionnement préalable des photographies (512×810 pixels)	52
2.3.2	Description du CNN	53
2.3.2.1	Données d'entraînement, de validation et de test	53
2.3.2.2	Description du modèle	54
2.3.2.3	Compilation et entraînement du modèle	56
2.4	Interprétation des résultats	58
2.4.1	Résultats granulométriques	58
2.4.1.1	Courbes granulométriques	58
2.4.1.2	Courbes d'erreur	59
2.4.2	Résultats de pourcentage de briques.....	60
2.4.2.1	Étude des biais relatifs, variances relatives et degrés de représentativité	60
2.4.2.2	Comparaison des coefficients de variations empiriques et théoriques	61
2.4.3	Analyse statistique des résultats granulométriques et de pourcentages de brique.....	62
2.4.4	Analyse d'images	64
2.4.4.1	Performance de l'entraînement.....	65
2.4.4.2	Performance du test.....	65
	CHAPITRE 3 RÉSULTATS	67
3.1	Analyse granulométrique.....	67
3.1.1	Granulométrie générale	67
3.1.2	Analyse descriptive des variances	70
3.1.2.1	Effet de la présence de fines	74
3.1.2.2	Effet de la présence d'une phase liquide.....	78
3.1.2.3	Effet de la méthode d'échantillonnage employée.....	80
3.1.3	Analyse statistique de la variance.....	83
3.1.3.1	Essais avec fines	83
3.1.3.2	Essais sans fines.....	86
3.2	Analyse des résultats de pourcentage de briques.....	87
3.2.1	Étude descriptive et statistique des variances.....	87
3.2.2	Étude des biais relatifs, variances relatives et degrés de représentativité	90
3.2.3	Coefficient de variation	92

	3.2.3.1	Coefficient de variation théorique CVthéo	92
	3.2.3.2	Coefficient de variation empirique CVemp	98
3.3		Analyse d'image	102
3.4		Lien entre les résultats d'échantillonnage et d'analyse d'images	107
	3.4.1	Biais relatifs, variances relatives et degrés de représentativité liés à l'échantillonnage et au réseau	107
	3.4.2	Nombre de faux positifs et négatifs générés lors de l'échantillonnage et l'analyse d'images	108
CHAPITRE 4 DISCUSSION			111
4.1		Discussion sur l'analyse granulométrique	111
	4.1.1	Pertes de sol et briques au cours de l'échantillonnage	111
	4.1.2	Pertes de brique particulaire pendant l'échantillonnage ainsi que dans les fractions inférieures à 1,25 mm	114
4.2		Discussion sur l'analyse d'images	115
4.3		Lecture croisée des résultats d'analyses granulométriques et d'images	115
CONCLUSION			117
RECOMMANDATIONS			121
ANNEXE I		ANALYSE SÉDIMENTOMÉTRIQUE D'UN ÉCHANTILLON	123
ANNEXE II		CONFIGURATIONS DE RÉSEAU NEURONAL	127
ANNEXE III		TEST DE VARIABILITÉ DES CONFIGURATIONS DE RÉSEAUX NEURONAUX	131
ANNEXE IV		RÉSULTATS DES TESTS D'ÉGALITÉ DES VARIANCES SUR LES RÉSULTATS GRANULOMÉTRIQUES	143
ANNEXE V		RÉSULTATS DES TESTS DE NORMALITÉ SUR LES POURCENTAGES DE BRIQUES	145
ANNEXE VI		RÉSULTATS DES TESTS D'ÉGALITÉ DES VARIANCES SUR LES RÉSULTATS DE POURCENTAGE DE BRIQUES	151
LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES			153

LISTE DES TABLEAUX

		Page
Tableau 1.1	Classement des auteurs des méthodes d'échantillonnage (de la meilleure à la pire) avec N le nombre d'incrément, N.R. « non recommandé » et N.A. « non applicable ».....	15
Tableau 2.1	Caractéristiques des dix essais de la recherche.....	27
Tableau 2.2	Description de la base de données photographique originale de sol avec briques	47
Tableau 2.3	Extrait de la base de données photographique de sol avec briques	48
Tableau 2.4	Information sur les essais réalisés et les pourcentages de briques des échantillons numéro 257, 641, 897, 1153, 1805 et 1873 présentés en Figure 2.15	50
Tableau 2.5	Information granulométriques des échantillons numéro 257, 641, 897, 1153, 1805 et 1873 présentés en Figure 2.15	50
Tableau 2.6	Nombre de photographies de sol contenant 30%, 20% et 50% de briques dans la base de photos originale et rééquilibrée, facteur de rééquilibre	51
Tableau 2.7	Description de la base de données photographique de sol avec briques complémentaire constituante de la base de données rééquilibrée.....	52
Tableau 2.8	Hypothèses nulle et alternative et seuil de signification des tests d'égalité des variances.....	63
Tableau 3.1	Granulométrie du lot de sol avec et sans fines.....	67
Tableau 3.2	Granulométrie du lot de sol et briques	69
Tableau 3.3	Biais d'échantillonnage, variance relative et degré de représentativité des résultats de pourcentage de briques des dix essais	91
Tableau 3.4	Pour chaque essai, la masse des lots, la fraction massique de la brique dans les lots, la masse par fraction granulométrique α et la fraction massique de la brique par fraction granulométrique α	94
Tableau 3.5	Calculs intermédiaires de l'invariant d'hétérogénéité, diamètre moyen des particules et invariant d'hétérogénéité.....	95
Tableau 3.6	Récapitulatif des variances théoriques liées à l'erreur fondamentale, de groupement et ségrégation, et totale pour les différentes	

	configurations d'essai calculées selon les équations (1.4), (1.3) et (1.2) respectivement.....	96
Tableau 3.7	Récapitulatif des coefficients de variation théoriques pour les différentes configurations d'essai calculés selon l'équation (2.9).....	97
Tableau 3.8	Pourcentage de briques moyen, minimal et maximal, variance et coefficient de variation empirique associés à chaque configuration d'essai	99
Tableau 3.9	Comparaison des coefficients de variation théoriques et empiriques des différentes configurations d'essai.....	100
Tableau 3.10	Biais relatifs, variances relatives et degrés de représentativité associés aux erreurs entre les %Briqueslot et %Briqueséch puis entre les %Briqueséch et %Briquesréseau pour différentes simulations du réseau neuronal n°3	108
Tableau 3.11	Nombre de faux négatifs générés par l'échantillonnage et l'analyse d'images.....	109
Tableau 4.1	Perte totale de sol et briques au cours des essais	111
Tableau 4.2	Pourcentages de perte lors des différentes étapes (souligné quand le ratio est fait avec des masses humides)	113
Tableau 4.3	Pourcentages de perte de phase liquide au cours des essais (souligné quand le ratio est fait avec des masses humides).....	113
Tableau 4.4	Pertes et omissions de particules de briques au cours des essais en pourcentage massique (souligné quand le ratio est fait avec des masses humides)	114

LISTE DES FIGURES

	Page
Figure 1.1	Schéma du processus d'échantillonnage par pelletage fractionné12
Figure 1.2	Schéma du processus d'échantillonnage par séparation rotative13
Figure 1.3	Architecture du CNN LeNet.....21
Figure 1.4	Illustration des premières couches de convolution et de Pooling du CNN LeNet ainsi que leur effet sur l'image d'entrée22
Figure 1.5	Résumé visuel de la propagation d'une image d'entrée à travers le CNN LeNet.....23
Figure 2.1	Origine des biocarburants mis sur le marché français en 201926
Figure 2.2	Photographies des bidons contenant le sol brut.....29
Figure 2.3	Dispositif de lavage d'importantes quantités de sol30
Figure 2.4	Brique entière ($\approx 20,5$ cm x $8,5$ cm ; $\approx 1,92$ kg)31
Figure 2.5	Pré-concassage manuel.....32
Figure 2.6	Concasseur mécanique32
Figure 2.7	Grappillage d'un mélange humide34
Figure 2.8	Pelletage fractionné sur des lots a) sec et b) humide.....35
Figure 2.9	a) Diviseur d'échantillons à cône rotatif Laborette 27 de FRITSCH (1 – entonnoir ; 2 – support d'entonnoir ; 3 – couvercle de tête de division ; 4 – tête de division ; 5 – flacons d'échantillonnage ; 6 – interrupteur principal ; 7 – fusible) et b) trémie d'alimentation à vibrations Laborette 24 de FRITSCH (1 – entonnoir ; 2 – goulotte ; 3 – appareil de base ; 4 – unité d'entonnoir ; 5 – support d'entonnoir réglable en hauteur)36
Figure 2.10	Montage pour la prise des photographies38
Figure 2.11	Photographie du four à moufle41
Figure 2.12	Échantillon de sol, briques et huile a) avant combustion, b) après combustion partielle et c) après combustion complète41

XVIII

Figure 2.13	Retenus a) 2,5 mm, b) 1,25 mm, c) 630 μm , d) 315 μm , e) 180 μm , f) 80 μm et de g) la panne pour un échantillon sans fines.....	43
Figure 2.14	Retenus a) 2,5 mm, b) 1,25 mm, c) 630 μm , d) 315 μm , e) 180 μm , f) 80 μm et de g) la panne pour un échantillon avec fines	44
Figure 2.15	Photographies des échantillons a) 257, b) 641, c) 897, d) 1153, e) 1805 et f) 1873	49
Figure 3.1	Courbe granulométrique du lot de sol avec fines	68
Figure 3.2	Courbe granulométrique du lot de sol sans fines.....	68
Figure 3.3	Courbe granulométrique du lot de sol avec fines sans briques (gris) et 30% de briques (rouge).....	69
Figure 3.4	Courbe granulométrique du lot de sol avec fines sans briques (gris) et 30% de briques (rouge).....	70
Figure 3.5	Graphiques en boîte à moustaches des pourcentages passants des sept essais avec fines des tamis a) 2,5 mm, b) 1,25 mm, c) 0,63 mm, d) 0,315 mm, e) 0,16 mm et f) 0,08 mm	72
Figure 3.6	Graphiques en boîte à moustaches des pourcentages passants des trois essais avec fines des tamis a) 2,5 mm, b) 1,25 mm, c) 0,63 mm et d) 0,315 mm	73
Figure 3.7	Granulométrie moyenne et intervalles de trois écart-types de l'essai n°2 de sol avec fines échantillonnés par grappillage	74
Figure 3.8	Granulométrie moyenne et intervalles de trois écart-types de l'essai n°4 de sol sans fines échantillonnés par grappillage.....	75
Figure 3.9	Granulométrie moyenne et intervalles de trois écart-types de l'essai n°3 de sol avec fines échantillonnés par pelletage fractionné	76
Figure 3.10	Granulométrie moyenne et intervalles de trois écart-types de l'essai n°6 de sol sans fines échantillonnés par pelletage fractionné.....	76
Figure 3.11	Granulométrie moyenne et intervalles de trois écart-types de l'essai n°1 de sol avec fines échantillonnés par séparation rotative	77
Figure 3.12	Granulométrie moyenne et intervalles de trois écart-types de l'essai n°5 de sol sans fines échantillonnés par séparation rotative.....	77

Figure 3.13	Comparaison des intervalles à trois écarts types centrés sur la moyenne pour les essais de sol avec fines échantillonnés par grappillage afin d'évaluer l'effet de la présence d'une phase liquide79
Figure 3.14	Comparaison des intervalles à trois écarts types centrés sur la moyenne pour les essais de sol avec fines échantillonnés par pelletage fractionné afin d'évaluer l'effet de la présence d'une phase liquide79
Figure 3.15	Comparaison des intervalles à trois écarts types centrés sur la moyenne pour les essais de sol sec avec fines de grappillage, pelletage fractionné et séparation rotative afin de comparer l'effet de la méthode d'échantillonnage employée.....81
Figure 3.16	Comparaison des intervalles à trois écarts types centrés sur la moyenne pour les essais de sol avec fines et 10% d'eau de grappillage et pelletage fractionné afin de comparer l'effet de la méthode d'échantillonnage employée81
Figure 3.17	Comparaison des intervalles à trois écarts types centrés sur la moyenne pour les essais de sol avec fines et 10% d'huile de grappillage et pelletage fractionné afin de comparer l'effet de la méthode d'échantillonnage employée82
Figure 3.18	Graphique des intervalles de comparaisons multiples des écarts types pour les pourcentages passants des sept essais avec fines des tamis a) 2,5 mm, b) 1,25 mm, c) 0,63 mm, d) 0,315 mm, e) 0,16 mm et f) 0,08 mm84
Figure 3.19	Graphique des intervalles de comparaisons multiples des écarts types pour les pourcentages passants des trois essais sans fines des tamis a) 2,5 mm, b) 1,25 mm, c) 0,63 mm et d) 0,315 mm86
Figure 3.20	Graphique en boîtes à moustaches des pourcentages de briques des dix essais88
Figure 3.21	Graphique des intervalles de confiance de Bonferroni à 95 % pour les écarts types des pourcentages de briques des dix essais89
Figure 3.22	a) Graphique des pertes d'entraînement et de validation au fil des époques et b) graphique de régression lors du test et performances du test pour la simulation n°1 effectuée avec la configuration de réseau n°1 - les points rouges représentent les échantillons secs sans fines, en gris les secs avec fines, en bleu les humides, en jaune ceux avec huile103
Figure 3.23	a) Graphique des pertes d'entraînement et de validation au fil des époques et b) graphique de régression lors du test et performances du

test pour la simulation n°1 effectuée avec la configuration de réseau n°2 - les points rouges représentent les échantillons secs sans fines, en gris les secs avec fines, en bleu les humides, en jaune ceux avec huile104

Figure 3.24 a) Graphique des pertes d'entraînement et de validation au fil des et
b) graphique de régression lors du test et performances du test pour la simulation n°6 effectuée avec la configuration de réseau n°3 - les points jaunes sont associés au grappillage, les points verts au pelletage fractionné, les points violets à la séparation rotative et les points bleus à la séparation à chutes104

Figure 3.25 Zoom sur le graphique de régression présenté dans le Tableau-A III-2 associé à la simulation n°3 du réseau neuronal n°2106

INTRODUCTION

Au Québec, le guide d'intervention sur la protection des sols et réhabilitation des terrains contaminés stipule que la caractérisation d'un terrain soupçonné de contenir des matières résiduelles est cruciale pour en planifier la réhabilitation. Une quantification représentative (voir définition en 1.2.1) des matières résiduelles présentes dans un sol est cruciale pour en déterminer le devenir, et constitue ainsi l'enjeu de cette recherche. Effectivement, pour certains types de matières résiduelles, si leurs concentrations dépassent 50% de la composition d'un sol, ce dernier est classé comme matière résiduelle à part entière et doit par conséquent être traité comme telle en vertu de l'article 16, deuxième alinéa, du guide d'application du règlement concernant la valorisation de matières résiduelles établi par le Ministère de l'Environnement et de la Lutte contre les Changements Climatiques (MELCC, 2021).

La quantification des matières résiduelles d'un sol se fait généralement par analyse visuelle d'échantillon selon la méthode décrite en annexe II du Règlement concernant la valorisation de matières résiduelles (MELCC, 2021). Cette façon de procéder est bien établie, cependant, elle est inévitablement laborieuse et chronophage pour l'analyste. Par ailleurs, l'influence de la qualité de l'échantillonnage sur la quantification des matières résiduelles de cette manière n'est pas établie.

À l'heure actuelle dans l'industrie, des logiciels d'analyse d'images par segmentation comme VisioRock et WipFrag (Guyot et al., 2004; Maerz et al., 1996) sont utilisés et ce, pour l'analyse granulométrique principalement mais également pour la quantification de matières résiduelles. Ils permettent ainsi de faciliter les tâches de caractérisation par rapport à l'identification visuelle, mais le processus de délimitation reste tout de même chronophage et fastidieux. L'analyse d'images par intelligence artificielle quant à elle, n'en est encore qu'au stade de développement en géotechnique. À l'ÉTS, des recherches se sont intéressées tout particulièrement au développement de réseaux neuronaux artificiels et convolutifs pour

l'analyse granulométrique ou la quantification de matières résiduelles dans des échantillons de sol (Bouyim, 2022; Dol, 2023; Zeryouh, 2023).

La quantification de la teneur d'un sol en matières résiduelles, qu'elle soit par analyse visuelle ou d'images, nécessite d'abord un échantillonnage représentatif (1.2.1).

Le **principal objectif** de ce mémoire est d'estimer l'erreur de quantification de la matière résiduelle présente dans un sol, générée par l'analyse d'une photo de ce sol via un réseau neuronal convolutif. Cette erreur serait comparée à celle produite par un processus d'échantillonnage, qui serait notamment utilisé pour produire les images analysées par le réseau. Pour répondre à cet objectif principal, deux sous-objectifs sont à atteindre dans un premier temps. Le **premier sous-objectif** est d'évaluer la représentativité de l'échantillonnage d'un sol contenant une matière résiduelle et ce, en fonction de la présence de phases liquides ainsi que de la méthode d'échantillonnage employée (grappillage, pelletage fractionné ou séparation rotative). Cette évaluation a été réalisée via l'étude de la variabilité de deux paramètres, soient la granulométrie des échantillons et le pourcentage de matière résiduelle qu'ils contiennent. Le **deuxième sous-objectif** de ce mémoire est d'élaborer un réseau neuronal convolutif permettant de quantifier le pourcentage de matière résiduelle dans le sol, incarnée ici par des particules de briques rouges.

Pour ce faire, la présente recherche se structure en six parties avec une revue de la littérature en chapitre un, une description méthodologique en chapitre deux, les résultats obtenus au cours de la recherche en chapitre trois, la discussion de ces résultats en chapitre quatre, la conclusion de ce mémoire et finalement, les recommandations.

CHAPITRE 1

REVUE DE LITTÉRATURE

1.1 Contexte de la contamination des sols

La contamination des sols constitue un problème majeur de par ses conséquences environnementales, sociales et économiques (Hébert & Bernard, 2013). Au Québec, sont considérés comme contaminants, « une matière solide, liquide ou gazeuse, un micro-organisme, un son, une vibration, un rayonnement, une chaleur, une odeur, une radiation ou toute combinaison de l'un ou l'autre susceptible d'altérer de quelque manière la qualité de l'environnement » d'après le deuxième paragraphe du chapitre 1 de loi sur la qualité de l'environnement. Les annexes I et II du Règlement sur la protection et la réhabilitation des terrains (RPRT) récapitulent les contaminants concernés à travers 12 catégories et en précisent les valeurs limites de concentration (MTESS, 2023). Ces catégories incluent les métaux et métalloïdes, autres composés inorganiques (comme le cyanure), composés organiques volatils (hydrocarbures aromatiques monocycliques, hydrocarbures aliphatiques chlorés) composés phénoliques (non chlorés, chlorés), hydrocarbures aromatiques polycycliques, composés benzéniques non chlorés, chlorobenzènes, biphényles polychlorés (BPC), pesticides (juste tébuthiuron), autres substances organiques (comme acrylonitrile), paramètres intégrateurs (juste les hydrocarbures pétroliers) et dioxines et furanes. Le ministère de l'Environnement et de la Lutte contre les changements climatiques met à disposition un répertoire des terrains contaminés au Québec (MELCC, 2023b) précisant notamment pour chacun d'eux, la nature des contaminants et l'état de réhabilitation. Des bilans sur les sites inscrits sont établis, le dernier porte sur l'année 2010 et concerne un total de 8334 sites. Ce dernier bilan met en avant le fait que les produits pétroliers sont les contaminants les plus récurrents avec une présence dans 80% des sites répertoriés, suivis des métaux, hydrocarbures aromatiques monocycliques et polycycliques tous retrouvés environ sur 25% des sites (Hébert & Bernard, 2013).

1.1.1 Les mélanges de sol et matière résiduelle

Les terrains contaminés ont souvent été confrontés à la présence de matières résiduelles. Dans le passé, les matières résiduelles ont été éliminées dans des sites dédiés ou de manière non réglementée. Cela dit, elles ont également été utilisées pour remblayer ou niveler les terrains, sans prendre les mesures nécessaires pour éviter la dispersion des contaminants, ce qui est considéré comme une élimination non contrôlée. Ces matières résiduelles peuvent provenir de contaminations à la fois récentes et historiques. Elles peuvent former des couches distinctes ou être mélangées avec les sols. Sont retrouvés fréquemment dans les terrains contaminés, des résidus de fonderie tels que des sables, des scories métallurgiques et de bouilloires, des mâchefers, ainsi que des débris provenant de la construction, de la rénovation et de la démolition, des déchets domestiques, des résidus miniers, et d'autres résidus liés à divers processus industriels ou d'entreposage (Ministère de l'environnement et de lutte contre les changements climatiques, 2021, partie 7.7).

L'appellation de « matière résiduelle » se décline en plusieurs catégories dont les matières résiduelles dangereuses, les déchets biomédicaux et l'amiante, qui sont les types de matières résiduelles systématiquement quantifiés. Cette appellation inclut également d'autres matières résiduelles inorganiques et non dangereuses qui sont, elles, quantifiées que dans certains cas. Selon la situation, différentes gestions des sols contenant de la matière résiduelle sont applicables. Ils peuvent être maintenus ou valorisés (pour les cas précisés dans les Lignes directrices relatives à la gestion de béton, de brique et d'asphalte issus des travaux de construction et de démolition et des résidus du secteur de la pierre de taille ou le Règlement concernant la valorisation de matières résiduelles (RVMR) et le Guide de valorisation des matières résiduelles inorganiques non dangereuses de source industrielle comme matériau de construction, selon le cas) ou excavés, traités in-situ ou ex-situ, ou enfouis selon leur catégorie (selon le Règlement sur les matières dangereuses ou le Règlement sur les déchets biomédicaux par exemple, selon le cas). La figure 12 du guide d'application donne un exemple de procédure à suivre lorsque des sols mélangés à des matières résiduelles sont excavés. (Ministère de l'environnement et de lutte contre les changements climatiques, 2021, partie 7.7)

1.1.2 Les produits pétroliers

L'article 2 de la loi *P-30.01* (2023) inclut dans l'appellation de « produit pétrolier », l'essence, le carburant diesel ou biodiesel, l'éthanol-carburant, le mazout, ainsi que tout autre mélange liquide d'hydrocarbures déterminé par règlement du gouvernement.

La contamination du sol par des liquides organiques et tout particulièrement les produits pétroliers constitue un problème majeur de par leur récurrence puisqu'ils concernent 80% des sites contaminés (Hébert & Bernard, 2013), mais également de par leur dangerosité pour l'environnement et la santé humaine. Cette toxicité se ressent à travers la Figure 4 du bilan au 31 décembre 2010 des terrains contaminés au Québec. En effet, dans près de 30% des cas d'intervention, l'élément déclencheur était la constatation d'impact notable sur la santé humaine, la faune, la flore l'environnement ou les biens (Hébert & Bernard, 2013).

Les produits pétroliers comme le diesel sont des produits complexes, contenant plus de 200 substances chacun, issus du raffinage du pétrole brut et utilisés comme carburant ou lubrifiant. Les modèles d'évaluation des risques pour la santé et l'écosystème sont conçus pour des sols avec des contaminants à doses spécifiques (voir les annexes I et II du *Q-2, r. 37*, 2023) et des effets toxiques connus. Cependant, les mélanges complexes de produits pétroliers rendent difficile l'évaluation des risques, en particulier en ce qui concerne la toxicité individuelle des constituants et les effets synergiques (Ministère de l'environnement et de lutte contre les changements climatiques, 2021).

Le MDDELCC exige que les sols contaminés dépassant les limites réglementaires définies par l'annexe I du *Q-2, r. 37* (2023) pour les hydrocarbures pétroliers C10-C50 (HP C10-C50) soient excavés ou traités conformément à la réglementation. Il est possible de conserver des HP C10-C50 dans les sols si leur présence est attribuée à de l'asphalte plutôt qu'à un carburant ou un lubrifiant, sous certaines conditions mais ces sols font l'objet de mesures de mitigation et d'avis de restriction d'utilisation inscrit au Registre foncier du Québec (Ministère de l'environnement et de lutte contre les changements climatiques, 2021).

1.2 L'échantillonnage de sol à des fins d'analyse granulométrique et quantification de matières résiduelles

1.2.1 Définitions

L'**échantillonnage** est défini par la norme ISO 3534-4:2014 comme une opération consistant à prélever un échantillon (Centre d'expertise en analyse environnementale du Québec, 2023). Il est souvent réalisé par étapes avec d'abord le prélèvement d'un échantillon primaire extrait d'un lot, puis éventuellement le prélèvement d'un échantillon secondaire extrait de l'échantillon primaire et ainsi de suite (Gy, 1992).

Le Cahier 5 sur l'échantillonnage des sols élaboré par le Centre d'expertise en analyse environnementale (Centre d'expertise en analyse environnementale du Québec, 2008) récapitule les étapes clés à suivre et les précautions à prendre pour réaliser une **campagne d'échantillonnage** de sol dans le cadre de l'analyse environnementale d'un site. La première étape à effectuer est l'évaluation préliminaire, suivi de l'élaboration de la campagne d'échantillonnage, puis du prélèvement d'échantillons et de leur analyse. Selon le Cahier 5, la qualité des échantillons devrait être garantie tout au long du processus par le respect du programme d'assurance et de contrôle de la qualité appliqué à toutes les étapes de prélèvement, de préservation et d'identification sur site comme en laboratoire.

Un **échantillon** correspond à la réunion d'un ou plusieurs incréments extraits d'un lot et est destiné à représenter ce dernier pour des opérations futures (Gy, 1992). Ces opérations concernent généralement des analyses permettant d'obtenir les caractéristiques physiques ou chimiques du lot. C'est pour cette raison que « l'échantillonnage des sols constitue une partie importante d'une étude de caractérisation d'un terrain » (Centre d'expertise en analyse environnementale du Québec, 2008). En effet, un échantillonnage non-représentatif peut mener à des diagnostics erronés de contamination (Boudreault, 2016).

Un **échantillon représentatif**, dans le contexte de l'échantillonnage de sol pour une analyse environnementale, est un prélèvement qui capture de manière fidèle et précise les

caractéristiques du sol à un emplacement spécifique. Il doit être obtenu en minimisant au maximum les altérations ou les changements induits lors de sa collecte afin que l'intégrité des propriétés physiques et chimiques du sol soient maintenues, tout en étant suffisamment volumineux pour des analyses approfondies ultérieures (Centre d'expertise en analyse environnementale du Québec, 2008; ISO 3534-4:2014).

L'échantillonnage correct correspond à des méthodes et des techniques appropriées pour collecter des échantillons de manière à minimiser les erreurs, les biais et les altérations potentielles afin que le ou les échantillons produits soient représentatifs du lot d'où ils proviennent. Il s'agit de suivre des procédures adéquates tout au long du processus d'échantillonnage, de la sélection des outils à la gestion des échantillons collectés (d'après *Échantillonnage*, 2023, fondé sur le texte de la norme ISO/IEC 17025:2017 sur les exigences générales concernant la compétence des laboratoires d'étalonnages et d'essais). Dans le cas de l'échantillonnage de sol plus précisément, des conseils généraux pour mettre en place des stratégies d'investigation de site et d'échantillonnage, par exemple, sur ce qu'il faut prélever, à quel emplacement et à quelle profondeur, et les types d'échantillons à prélever sont décrits de manière générale dans la norme ISO 18400-104:2018 et au Québec plus particulièrement, dans le guide d'échantillonnage à des fins d'analyses environnementales (Centre d'expertise en analyse environnementale du Québec, 2008). Dans le cadre de ce projet, des procédures d'échantillonnage d'intérêt (par grappillage en 1.2.3.1, pelletage fractionné en 1.2.3.2 et séparation rotative en 1.2.3.3) sont décrites dans l'ouvrage de Gerlach & Nocerino (2003) fondé sur les principes de la théorie d'échantillonnage.

1.2.2 Théorie de l'échantillonnage des matières particulières

La théorie d'échantillonnage de matières particulières définit les erreurs d'échantillonnage comme suit (Gerlach & Nocerino, 2003; Gy, 1992; Pitard, 2019) :

$$ETE = EFE + ESG + EP + ED + EE + ES_2 + ES_3 \quad (1.1)$$

Avec :

- *ETE* : l'erreur totale d'échantillonnage ;
- *EFE* : l'erreur fondamentale d'échantillonnage qui est une erreur inhérente ne pouvant être éliminée même en appliquant des procédures d'échantillonnage correctes (1.2.1), car elle est causée par la nature hétérogène même du matériau échantillonné ;
- *ESG* : l'erreur de ségrégation et de groupement qui désigne l'erreur induite par l'hétérogénéité de la distribution des particules, résultant en une séparation ou un regroupement sous l'influence de la gravité ;
- *EP* : l'erreur de préparation qui se produit lors de la manipulation d'un échantillon, entraînant sa contamination ou la perte d'une partie de celui-ci ;
- *ED* : l'erreur de délimitation qui survient lorsque le processus d'échantillonnage ne garantit pas une probabilité de sélection égale pour toutes les particules de l'échantillon ;
- *EE* : l'erreur d'extraction qui est la différence entre ce qui a été effectivement échantillonné et ce qui aurait dû être échantillonné ;
- *ES₂* : l'erreur de fluctuation de l'hétérogénéité selon la position sur le site de prélèvement ;
- *ES₃* : l'erreur de fluctuation de l'hétérogénéité selon le temps.

Dans le cas d'échantillonnages se faisant uniquement en laboratoire à partir de lots homogènes provenant d'un même site de prélèvement, $ES_2 = ES_3 = 0$. À part l'erreur fondamentale d'échantillonnage, qui dépend de l'hétérogénéité même du matériau échantillonné, ainsi que l'erreur de ségrégation et groupement, très souvent irréductible pour la plupart des méthodes d'échantillonnage, toutes les autres erreurs constituantes de l'erreur totale d'échantillonnage peuvent être minimisées et donc considérées comme nulles si des procédures correctes d'échantillonnage (1.2.1) sont employées (selon les procédures décrites en 1.2.3). Ainsi, $EP = ED = EE = 0$ ce qui fait que $ETE = EFE + ESG$.

La variance générée par l'erreur totale correspond donc à la somme de la variance liée à l'erreur fondamentale et à l'erreur de ségrégation et de groupement :

$$s^2(ETE) = s^2(EFE) + s^2(ESG) \quad (1.2)$$

Avec :

- $s^2(EFE)$: la variance générée par l'erreur fondamentale ;
- $s^2(ESG)$: la variance générée par l'erreur de ségrégation et groupement qui s'exprime comme suit.

$$s^2(ESG) = \frac{s^2(EFE)}{N} \quad (1.3)$$

Tirée de Pitard (2019)

Avec :

- $s^2(EFE)$: la variance générée par l'erreur fondamentale qui s'exprime comme suit ;
- N : le nombre d'incrément constitutif d'un échantillon.

$$s^2(EFE) = \left(\frac{1}{m_s} - \frac{1}{m_L} \right) \times HI_L \quad (1.4)$$

Tirée de Dubé & Esbensen (2022)

Avec :

- $s^2(EFE)$: variance générée par l'erreur fondamentale ;
- m_s : masse théorique d'un échantillon (g) ;
- m_L : masse du lot (g) ;
- HI_L : invariant d'hétérogénéité qui est déterminé en fonction des proportions des différents matériaux qui composent les différentes fractions granulométriques (g).

Dans son œuvre originale, cette formule se trouve à l'équation 12 issue de l'équation générale 8 simplifiée dans le cas qui est celui de cette recherche où les deux fractions de densité (la brique et le sol en l'occurrence) sont de même densité quelle que soit la fraction granulométrique.

Une expression générale et une autre applicable aux minerais de l'invariant d'hétérogénéité ont été proposées par Gy dans son ouvrage sur la théorie. Bien que l'expression applicable aux minerais présente des limites notamment par sa difficulté d'application, son inadaptation à

certaines matériaux, et l'écart qu'elle génère entre la variance théorique et observée qu'elle induit (Dubé & Esbensen, 2022), l'expression générale constitue le fondement de la théorie d'échantillonnage et fait souvent l'objet de reformulations (Boudreault, 2016; Dubé & Esbensen, 2022). L'invariant d'hétérogénéité s'exprime comme suit :

$$HI_L = f\rho \sum_{\alpha} d_{\alpha}^3 \frac{m_{L\alpha}}{m_L} \frac{(a_{\alpha} - a_L)^2}{a_L^2} \quad (1.5)$$

Tirée de Dubé & Esbensen (2022, p. 8)

Avec :

- HI_L : invariant d'hétérogénéité (g) ;
- f : facteur de forme des particules du lot (sans unité) ;
- ρ : masse volumique des particules (g/cm^3) ;
- α : fraction granulométrique α ;
- d_{α} : diamètre moyen des particules dans α (cm) ;
- $m_{L\alpha}$: masse de α (g) ;
- m_L : masse du lot (g) ;
- a_{α} : fraction massique de la matière résiduelle dans α (sans unité) ;
- a_L : fraction massique de la matière résiduelle dans le lot (sans unité).

1.2.3 Méthodes d'échantillonnage

La théorie de l'échantillonnage des matières particulières présente plusieurs méthodes d'échantillonnage (Gerlach & Nocerino, 2003; Pitard, 2019). Pour les besoins de la présente recherche, seules les méthodes de grappillage, pelletage fractionné et de séparation rotatives réalisées sur des lots de même provenance et homogènes seront évoquées et ce, dans le cadre d'échantillonnage en laboratoire uniquement. L'échantillonnage de terrain n'est pas abordé dans la présente étude.

1.2.3.1 Grappillage

Le grappillage est la méthode d'échantillonnage la plus couramment utilisée en caractérisation des sites contaminés à l'heure actuelle. Il consiste à prélever un échantillon au jugé à partir de la surface accessible du lot d'intérêt.

Cette approche présente des limites importantes en termes de représentativité des échantillons (1.2.1). En effet, lors de l'échantillonnage par grappillage, chaque particule de l'échantillon n'a pas la même probabilité d'être sélectionnée, ce qui peut entraîner des biais significatifs dans les résultats. De plus, cette méthode est particulièrement influencée par les effets de la ségrégation des particules dans le lot et les biais de l'échantillonneur. Par ailleurs, l'outil d'échantillonnage utilisé, qui est également conçu de manière incorrecte selon les principes de la théorie d'échantillonnage, constitue également une source d'erreur. Le tout peut générer des erreurs de délimitation (ED) et d'extraction d'incrément (EE) et n'exerce aucun contrôle sur les erreurs fondamentales (EF) et de ségrégation et de groupement (ESG).

Cette subjectivité fait de cette méthode une méthode non-correcte et non-probabiliste (Gerlach & Nocerino, 2003; Petersen et al., 2004).

1.2.3.2 Pelletage fractionné

Le pelletage fractionné consiste à diviser un lot en plusieurs sous-échantillons composés de plusieurs incréments. Un minimum de dix incréments par sous-échantillon est recommandé par Gerlach & Nocerino (2003). Une fois que le lot initial est épuisé, l'un des sous-échantillons produits est sélectionné au hasard. Cette technique fonctionne autant pour l'échantillonnage de matériau sec qu'humide.

La Figure 1.1 donne une explication visuelle de la méthode pour une meilleure compréhension du processus :

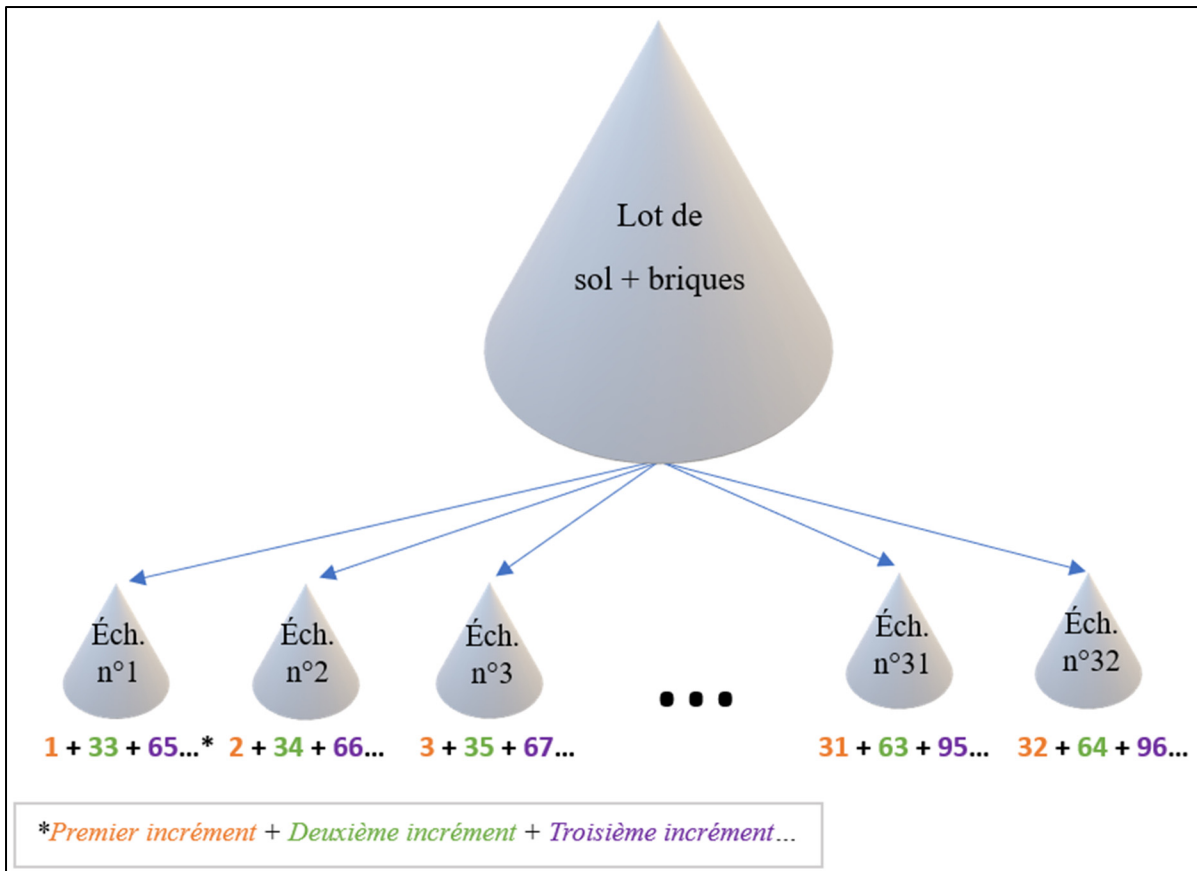


Figure 1.1 Schéma du processus d'échantillonnage par pelletage fractionné
Adaptée de Gerlach & Nocerino (2003, p.67)

Cette méthode est plus performante que le grappillage pour plusieurs raisons. Cette méthode permet de constituer les n échantillons en parallèle les uns des autres en une pluralité d'incrémentes à partir de la totalité du lot. Ce processus permet de réduire les erreurs de ségrégation et groupement de manière significative plus le nombre d'incrémentes est élevé. Cependant, le pelletage fractionné peut être relativement chronophage pour l'échantillonneur (Gerlach & Nocerino, 2003; Petersen et al., 2004).

1.2.3.3 Séparation rotative

Un séparateur sectoriel rotatif est un dispositif composé d'une tête rotative conique avec des crêtes et des vallées. Le lot est acheminé vers la tête rotative via une trémie d'alimentation vibrante, permettant de faire tomber les particules lentement sur la tête rotative. Les particules

sont collectées dans des récipients attachés à la tête rotative et positionnés sous chaque vallée. Pour de meilleurs résultats, il est recommandé d'utiliser le séparateur sectoriel avec une vitesse de rotation constante et un débit d'alimentation lent, ce qui permet de minimiser l'erreur de ségrégation et de groupement. En effet, la trémie d'alimentation garantit un débit continu de particules vers le séparateur, ce qui équivaut à un échantillonnage en un très grand nombre d'incrément. Cette caractéristique du dispositif, couplée à la conception de la tête rotative, avec ses crêtes tranchantes permet également de d'éliminer les erreurs de délimitation et d'extraction (Gerlach & Nocerino, 2003). Ceci fait que la seule erreur d'échantillonnage qui persiste avec ce dispositif est l'erreur fondamentale.

Le montage utilisé (Figure 1.2) permet de créer à partir d'un lot, 8^n échantillons dans le cas où la tête de division est à huit secteurs et où n est le nombre de séparations effectuées. Cette technique fonctionne uniquement pour l'échantillonnage de matériau sec. Le schéma en Figure 1.2 donne une description visuelle du processus de séparation.

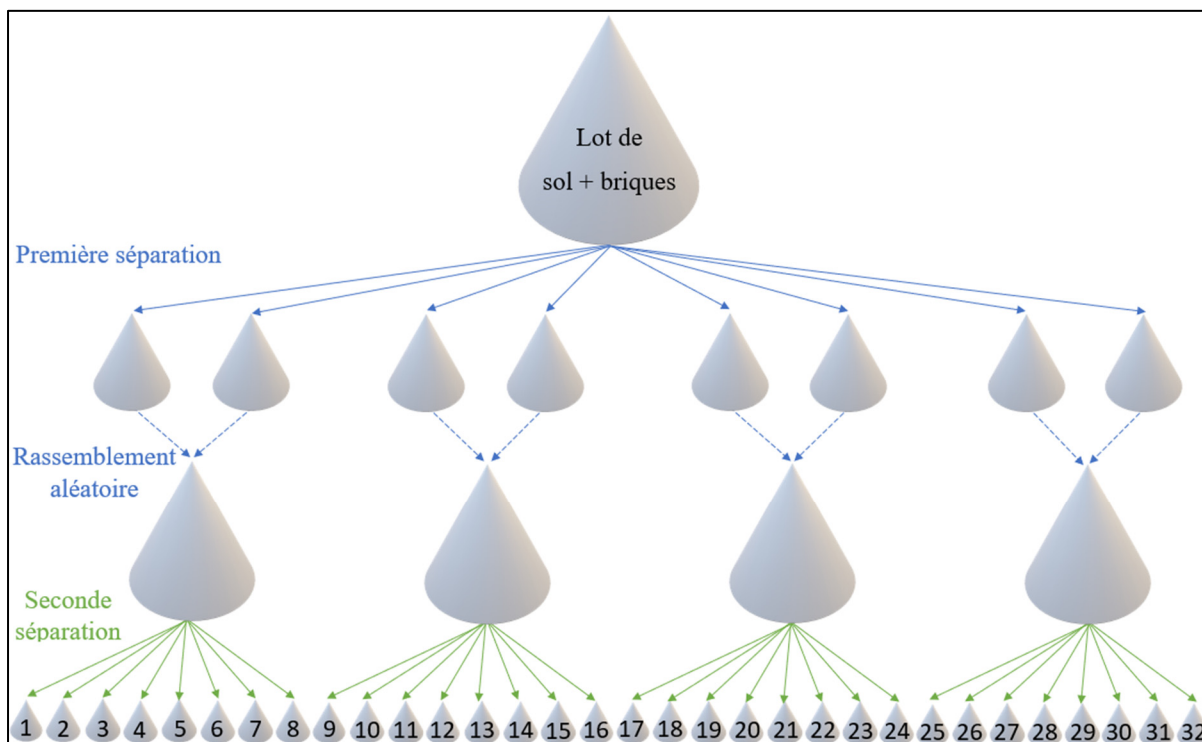


Figure 1.2 Schéma du processus d'échantillonnage par séparation rotative

Cette méthode est moins rapide que le grappillage, mais elle l'est tout de même plus que le pelletage fractionné (Gerlach & Nocerino, 2003).

1.2.4 Comparaison des différentes méthodes d'échantillonnage

Le Tableau 8 de Gerlach & Nocerino (2003) compile les travaux de plusieurs auteurs pour présenter une comparaison de treize méthodes d'échantillonnage (Tableau 1.1). La comparaison se fait sur cinq paramètres : la taille d'incrément, la susceptibilité de la méthode à produire du groupement et de la ségrégation, la possibilité d'appliquer la méthode sur un sol humide, si la méthode d'échantillonnage peut être considérée comme correcte ainsi que l'écart entre la variance liée à l'erreur fondamentale empirique et théorique. Un classement de ces treize méthodes d'échantillonnage de la meilleure à la moins recommandée est également réalisé à partir de ces différents paramètres.

Tableau 1.1 Classement des auteurs des méthodes d'échantillonnage (de la meilleure à la pire) avec N le nombre d'incrément, N.R. « non recommandé » et N.A. « non applicable »
Adapté de de Gerlach & Nocerino (2003, p.72)

Méthode	Taille typique d'incrément	Sensibilité à l'erreur de groupement et ségrégation	Présence de liquide	Échantillonnage correct possible	Proximité avec le s^2 (EFE) calculé	Commentaires
Séparateur sectoriel	Très petit	Faible	Sec	Oui	Très proche	Méthode de choix
Séparateur sectoriel à cône	Petit	Faible	Sec	Oui	Proche	Bonne méthode
Séparation par incréments	Petit à moyen	Modérée ; faible avec un grand nombre (>30) d'incrément	Sec à humide	Oui avec une pile	Proche	Biais toujours possible, diminue avec l'augmentation du nombre d'incrément et l'échantillonnage correcte
Séparateur à chute	Petit à moyen	Faible à modérée	Sec	Oui avec des précautions	Bon à correct	Pertes de fines possibles
Séparation alternée à la pelle	Moyen	Faible à modérée	Sec à humide	Oui avec des précautions	Bon à correct	Prends du temps à réaliser pour un grand lot
Pelletage fractionné	Moyen à grand	Modérée	Sec à humide	Oui avec des précautions	Correct	Performance liée à la masse du lot
Séparation de table	Moyen	Modérée à forte	Sec	Très difficilement, dépend du dimensionnement	Loin	Grande variabilité, non recommandée
Pelletage fractionné dégénéré	Moyen à grand	Modérée à forte	Sec à humide	Oui avec des précautions	Peu probable	Performance liée à la masse du lot, sujet aux biais, non recommandé
Roulement et quartage	Grand	Forte	Sec	Oui avec des précautions	Généralement loin	Très variable, non recommandé
Roulement et quartage	Grand	Forte	Sec	Oui avec des précautions	Généralement loin	Généralement biaisé, non recommandé
Mélangeur de type v	N.A.	Forte	Sec	N.A.	Très peu probable	Susceptible au groupement et séparation, non recommandé
Spatule vibrante	Petit	Très forte	Sec	Non	Loin	Susceptible au groupement et séparation, non recommandé
Grappillage	Variable	Très forte	Sec à humide	Non	Loin	Biaisé et variable, non recommandé

La séparation rotative se retrouve en première position et le pelletage fractionné en sixième position tandis que le grappillage se retrouve en dernière position. En effet, la séparation rotative s'avère être peu sensible à la ségrégation et au groupement, le pelletage fractionné modérément alors que le grappillage y est très sensible. D'un point de vue de la théorie d'échantillonnage, la séparation rotative est considérée correcte, le pelletage fractionné également à condition que l'opérateur soit précautionneux, cependant, le grappillage est considéré comme une méthode incorrecte. Le seul point pour lequel la séparation rotative est plus désavantageuse que les deux autres méthodes, est qu'elle ne peut pas être appliquée à un lot de sol humide mais qu'à des lots de sol sec.

L'article de Petersen et al. (2004) compare les performances d'une dizaine de méthodes d'échantillonnage selon la théorie de l'échantillonnage. Les méthodes de séparation sectorielles (séparation à chutes ou rotative) sont reconnues comme étant plus performantes que les méthodes par pelletage. Parmi ces dernières, le pelletage fractionné est généralement parmi les méthodes les plus performantes et le grappillage, la moins performante. D'autres études ont comparé le grappillage à d'autres méthodes d'échantillonnage en se basant sur différents paramètres tels que la reproductibilité (Boudreault, 2016; Sona & Dubé, 2021), les biais d'échantillonnage et la représentativité (Sona & Dubé, 2021). Les travaux de Boudreault et al. (2012) ainsi que de Sona & Dubé (2021) ne sont, une fois de plus, pas favorables à l'usage du grappillage. Ils soulignent l'importance de respecter les principes de la théorie de l'échantillonnage des matières particulières afin d'améliorer le degré de représentativité des échantillons produits.

1.3 Analyse granulométrique et quantification des matières résiduelles particulières

1.3.1 Analyse granulométrique

Au Québec, les lignes directrices à suivre pour réaliser une analyse granulométrique d'échantillon sont énoncées dans la méthode d'essai LC – 21 – 040 de Transport Québec datant de 2012 (Transport Québec, 2012) mais également dans la norme du Bureau de Normalisation

du Québec BNQ 2501-025 datant de 2013, qui est celle appliquée lors de cette étude. La procédure d'échantillonnage décrite en 1.2 permet d'obtenir 32 échantillons de 300 g environ par essai. Les tamis normalisés (ISO 3310-1) ne pouvant recevoir plus de 200 g individuellement au risque d'être colmatés (BNQ 2501-025), ce format d'échantillon est adapté aux procédures décrites dans cette norme au vu de la granulométrie du mélange de sol et briques (3.1.1).

En effet, cette norme fournit une description détaillée des étapes à suivre pour réaliser une analyse granulométrique complète. Elle couvre des aspects précis, adaptés au cas d'étude dans lequel cette recherche se place, tels que l'essai gravimétrique pour les sols pulvérulents (section 8) ainsi que la procédure de lavage des particules fines (section 7.2) en tant que préparation préalable nécessaire à l'analyse granulométrique elle-même. Cette norme détaille ces étapes spécifiquement pour différents types de sols, notamment ceux dont la taille est inférieure à 5 mm et qui contiennent des particules fines (section 7).

1.3.2 Teneur en matières résiduelles particulières

1.3.2.1 Quantification de la teneur en matières résiduelles particulières

La quantification de matières résiduelles particulières est essentielle, car au-delà d'une présence de ces dernières à un pourcentage de 50% dans un sol, ce sol est considéré comme une matière résiduelle à part entière et doit donc être traité comme tel d'après l'article 16, deuxième alinéa, du guide d'application du règlement concernant la valorisation de matières résiduelles fait par le Ministère de l'environnement et de lutte contre les changements climatiques (MELCC, 2021).

Le guide d'intervention pour la protection des sols et réhabilitation des terrains contaminés (Ministère de l'environnement et de lutte contre les changements climatiques, 2021) évoque deux méthodes visuelles permettant de mesurer le taux de matières résiduelles présentes dans les sols à partir d'échantillons. Les méthodes visuelles d'identification et de quantification de matière résiduelle consistent à extraire manuellement d'un échantillon, les matières résiduelles

reconnues par l'analyste et les classer par type avant de les peser. Les méthodes qui s'appliquent aux fractions granulométriques supérieures à 2,5 mm sont détaillées dans l'annexe II du Règlement concernant la valorisation de matières résiduelles (RVMR) et dans la méthode d'essai LC 21-901 du Laboratoire des chaussées du ministère des Transports du Québec (MTQ) (Transport Québec, 2019). Pour les fractions plus fines que 2,5 mm, le guide d'intervention suggère une évaluation visuelle à l'aide d'un stéréo-microscope.

L'annexe II du RVMR sur la détermination du contenu en impuretés concerne six catégories de particules. Le calcul du taux de matières résiduelles se fait par fraction granulométrique pour les fractions supérieures à 2,5 mm. Un exemple de calcul du pourcentage de matières résiduelles présentes dans la fraction entre 2,5 mm et 5 mm (P2,5-5 (%)) est donné à l'équation 3 :

$$P_{2,5-5} (\%) = \frac{(\text{masse de particules de MR retenue sur le tamis } 2,5 \text{ (g)})}{(\text{masse totale retenue sur le tamis de } 2,5 \text{ mm (g)})} \quad (1.6)$$

À quelques notations près, la méthode d'essai LC 21-901 indique le même processus. Seules les catégories de particules diffèrent légèrement, mais la méthode et les formules sont fondamentalement les mêmes.

1.3.2.2 Éléments d'évaluation de la représentativité d'un échantillonnage

La représentativité d'un échantillonnage peut se déterminer à partir de résultats de teneur en matière résiduelle.

Le biais relatif d'échantillonnage est l'erreur relative moyenne d'un ensemble de N mesures, soit :

$$m = \frac{\sum_{i=1}^N \varepsilon_s}{N} \quad (1.7)$$

L'erreur se calcule elle-même comme suit:

$$\varepsilon_s = \frac{a_s - a_L}{a_L} \quad (1.8)$$

Avec a_s qui est la teneur en analyte mesurée sur un échantillon provenant du lot, L, et a_L la teneur en analyte réelle du lot L (déterminée comme détaillé en 1.3.2.1 dans le cas où il s'agit de matière résiduelle).

Le degré de représentativité r^2 est une mesure de la justesse et de la reproductibilité des mesures tirées d'un échantillonnage, soit :

$$r^2 = m(ETE)^2 + s(ETE)^2 \quad (1.9)$$

Avec $m(ETE)^2$ qui est le biais relatif au carré et $s(ETE)^2$ qui est la variance relative des mesures.

1.3.3 Analyse granulométrique et particulière par analyse d'images

Les méthodes d'analyse d'images pour la caractérisation granulométrique des sols sont classées en deux groupes : les techniques directes et les techniques indirectes (Manashti, 2022). Les méthodes directes utilisent la segmentation d'image pour identifier le contour de chaque particule, déterminer leur taille et ainsi calculer la distribution granulométrique en utilisant des méthodes statistiques. WipFrag (Maerz et al., 1996) et Visiorock (Guyot et al., 2004) sont des exemples de logiciels commerciaux qui utilisent ces méthodes directes de segmentation d'image (Duhaime et al., 2021).

Les méthodes indirectes, quant à elles, extraient des caractéristiques texturales de l'image et les corrélient avec la taille des particules (Duhaime et al., 2021). Elles peuvent être traditionnelles (méthodes statistiques, de motifs d'image locaux et de domaine fréquentiel) ou modernes comme les réseaux de neurones artificiels ou convolutifs (Manashti, 2022). La présente recherche vise notamment à développer un réseau neuronal dit convolutif.

1.3.3.1 Réseau de neurones artificiels (ANN)

Réseau de neurones artificiels est une traduction de l'anglais *Artificial Neural Network* dit ANN. Il s'agit d'une approche algorithmique inspirée du fonctionnement du cerveau humain.

Un ANN est composé de plusieurs couches de neurones interconnectés qui traitent l'information et réalisent des tâches d'apprentissage et de prédiction.

Dans le cadre de l'analyse granulométrique d'un sol, un ANN utilise des paramètres texturaux en données d'entrée qui peuvent être extraits par différentes approches telles que la transformation en ondelettes de Haar (Hryciw et al., 2013) ou la détermination de l'entropie locale (Bouyim, 2022; Duhaime et al., 2021).

1.3.3.2 Réseau de neurones convolutif (CNN)

Un réseau de neurones convolutif est une traduction de l'anglais *Convolutional Neural Network* dit CNN. Il s'agit d'un type spécifique de réseau de neurones largement utilisé pour l'analyse d'images. Les CNN sont conçus pour extraire automatiquement les caractéristiques significatives des données d'entrée en utilisant des opérations de convolution, de sous-échantillonnage (Gonzales & Woods, 2018).

Un CNN se distingue d'un ANN par sa capacité à prendre directement des tableaux 2D (des images) en tant que données d'entrée et non pas des vecteurs (Goodfellow et al., 2016). Cependant, les calculs réalisés par ces deux types de réseaux sont très similaires (Gonzales & Woods, 2018). Un CNN peut être utilisé pour effectuer à la fois des tâches de classification d'objets, de détection d'objets, de segmentation d'images, ainsi que d'estimation de valeurs continues dans le cas d'une régression.

L'architecture classique d'un CNN (Figure 1.3) est représentée par l'architecture LeNet, qui a été proposée par Yann LeCun et ses collègues en 1998 (LeCun et al., 1998). Cette architecture a connu un grand succès dans la reconnaissance d'images, en particulier dans la reconnaissance de chiffres manuscrits du célèbre ensemble de données MNIST.

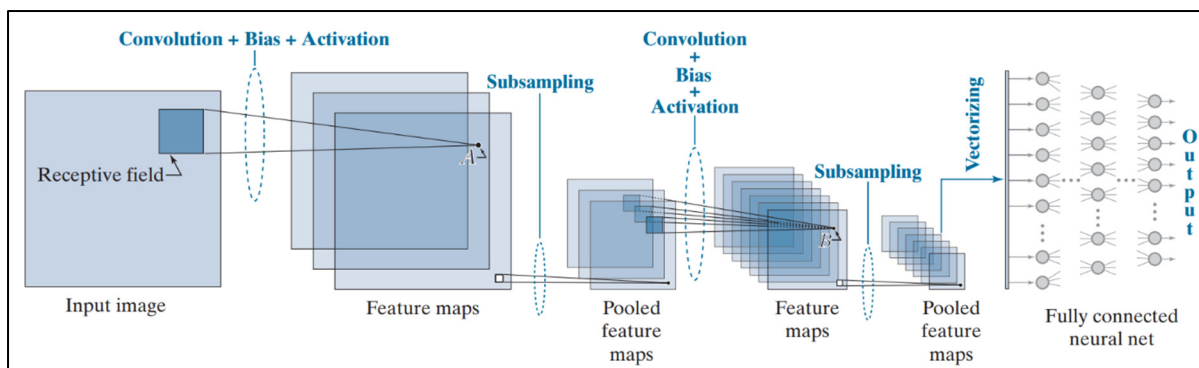


Figure 1.3 Architecture du CNN LeNet
Tirée de Gonzales & Woods (2018, p.965)

Les sept couches constituant le réseau LeNet présenté en Figure 1.3 ci-dessus sont ici décrites et illustrées de figures tirées de Gonzales & Woods (2018) :

- Couche de convolution (Convolutional Layer) : Cette couche est celle qui extrait en premier lieu des caractéristiques visuelles de l'image d'entrée. Elle applique des filtres (ou noyaux) convolutifs à l'image pour détecter des motifs tels que des bords, des textures et des formes. Chaque filtre produit un élément de la carte de caractéristiques obtenue en sortie, qui représente l'activation des caractéristiques détectées dans l'image (Gonzales & Woods, 2018; Goodfellow et al., 2016; LeCun et al., 1998).
- Couche de sous-échantillonnage (Pooling Layer) : Après la couche de convolution, une couche de sous-échantillonnage est généralement ajoutée pour réduire la dimension spatiale des cartes d'activation tout en conservant les caractéristiques les plus importantes (Gonzales & Woods, 2018; Goodfellow et al., 2016; LeCun et al., 1998).

La Figure 1.4 illustre ces deux premières couches du CNN LeNet pour une image d'entrée de taille 28×28 pixels et les effets qu'ont ces dernières sur cette image d'origine. La couche de convolution utilise des filtres de 5×5 pixels qui se déplacent d'un pas de 1×1 pixel, ce qui permet de constituer une carte de caractéristiques dite d'activation de taille 24×24 . La couche suivante de Pooling réduit la taille de cette carte d'un facteur deux en prenant tous les 2×2 pixels, les caractéristiques moyennes, ou maximales par exemple.

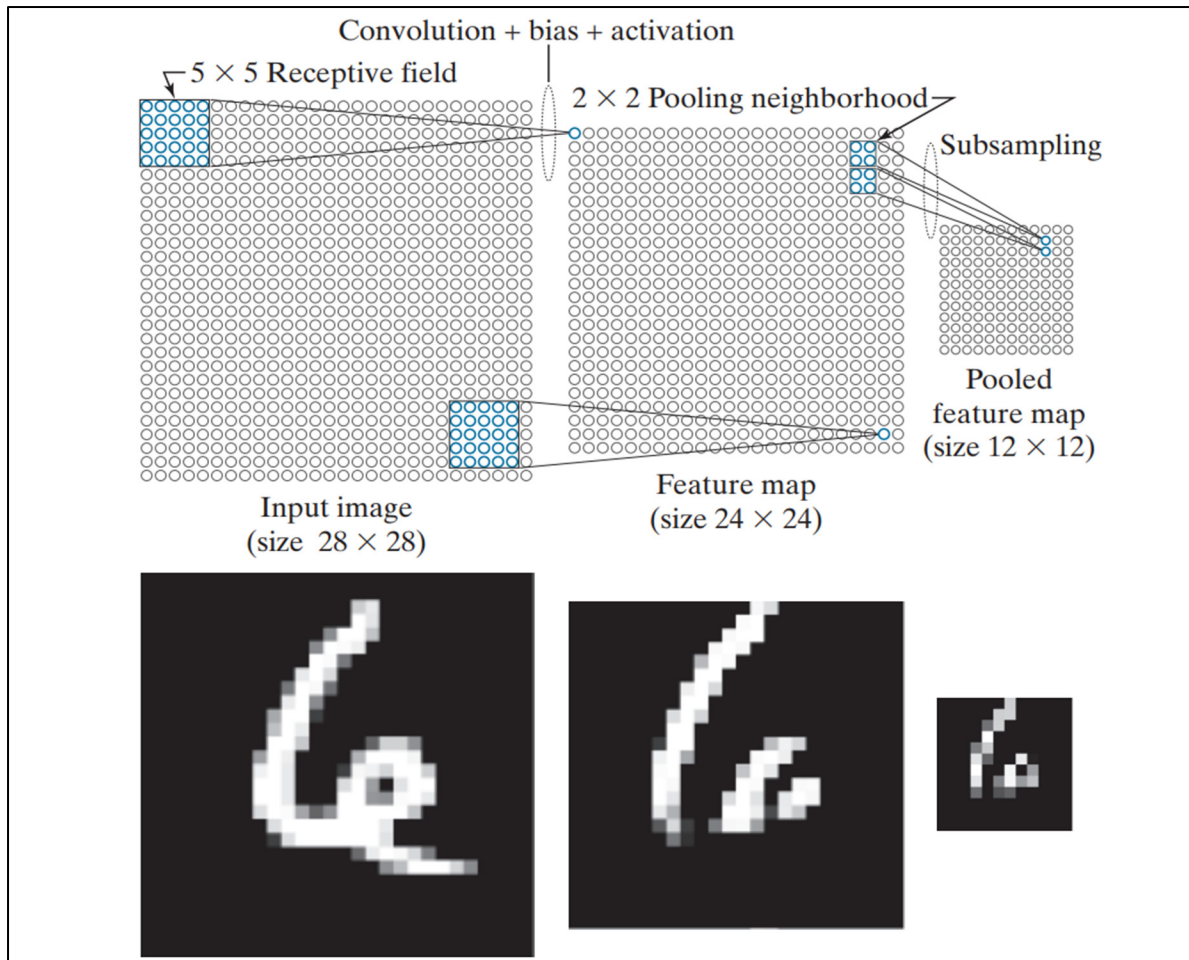


Figure 1.4 Illustration des premières couches de convolution et de Pooling du CNN LeNet ainsi que leur effet sur l'image d'entrée
Tirée de Gonzales & Woods (2018, p.969)

Ces deux premières couches sont suivies de deux autres couches de convolution puis de Pooling.

- Couche entièrement connectée (Fully Connected Layer) : Après les couches de convolution et de Pooling, les caractéristiques extraites sont aplaties en un vecteur pour être introduites dans une couche entièrement connectée. Cette couche est similaire à celle d'un réseau de neurones artificiel classique et est utilisée pour effectuer la classification finale des caractéristiques extraites (Gonzales & Woods, 2018; LeCun et al., 1998).
- Fonction d'activation : Entre chaque couche, une fonction d'activation est appliquée pour introduire de la non-linéarité dans le modèle. La fonction d'activation la plus couramment

utilisée est la fonction ReLU (Rectified Linear Unit), qui permet de conserver les activations positives et de supprimer les activations négatives (Gonzales & Woods, 2018; LeCun et al., 1998).

- Couche de sortie : La dernière couche du LeNet est une couche de sortie qui est utilisée pour la classification finale. La taille de cette couche dépend du nombre de classes que le modèle doit reconnaître (Gonzales & Woods, 2018; LeCun et al., 1998).

La Figure 1.5 propose un sommaire visuel du fonctionnement de cette architecture en reprenant l'exemple présenté en Figure 1.4.

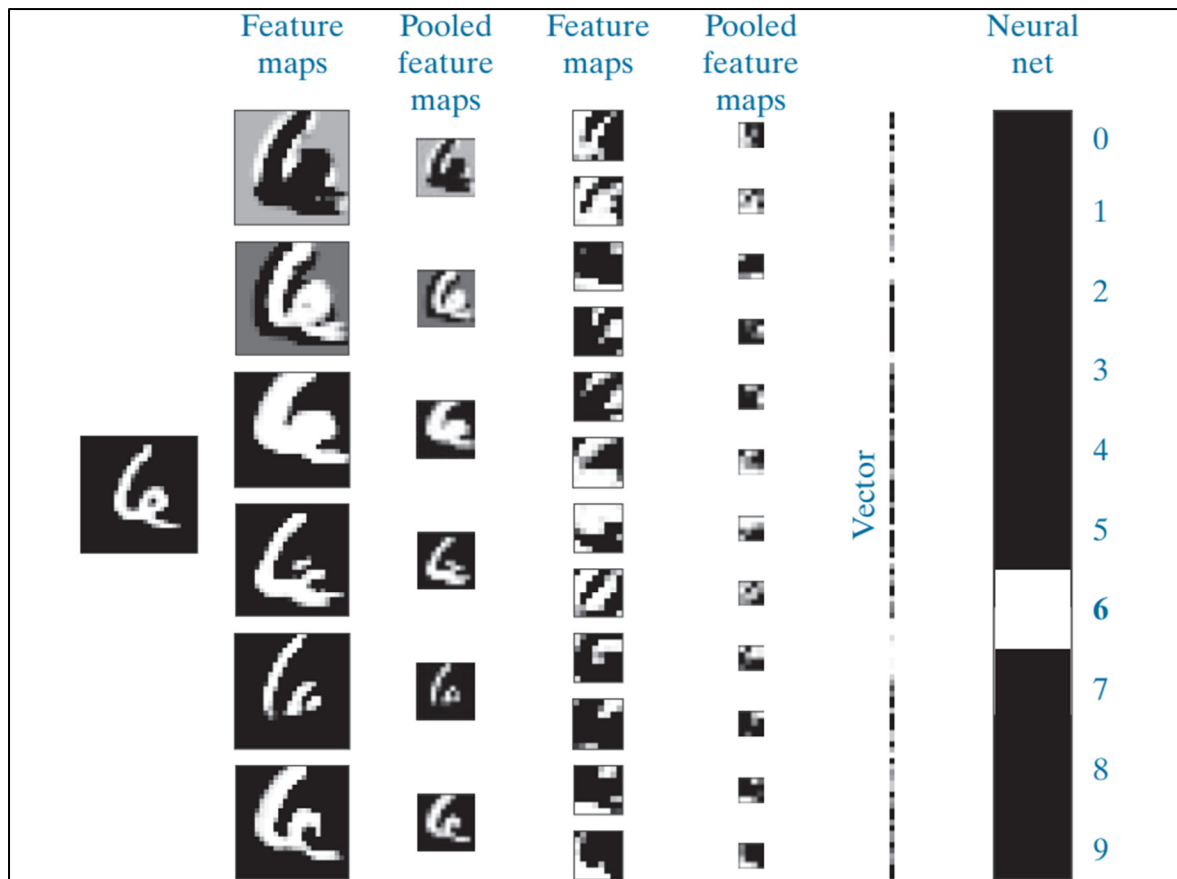


Figure 1.5 Résumé visuel de la propagation d'une image d'entrée à travers le CNN LeNet
Tirée de Gonzales & Woods (2018, p.971)

Cette architecture est utilisée pour des tâches de classifications dans le cadre du projet MNIST, cependant, elle peut également être adaptée à des tâches de régression en changeant la dernière

couche réservée à la classification en couche de régression. C'est à partir de cette dernière version qu'est élaboré le réseau neuronal convolutif développé au cours de cette recherche.

Dans le cas particulier de l'analyse de sol, des travaux ont permis de développer des réseaux neuronaux convolutifs pour l'analyse granulométrique de sédiments. En effet, les réseaux SediNet et GRAINet notamment (Buscombe, 2020; Lang et al., 2021), ont été développés pour œuvrer en ce sens et constituent des références en termes d'analyse de sol par réseau neuronal convolutif que ce soit par régression (SediNet et GRAINet) ou classification (SediNet).

1.3.3.3 Bases de données photographiques

L'une des clés fondamentales pour construire un réseau de neurones convolutif performant réside dans la composition et la quantité de photographies de la base de données utilisée. Pour obtenir un bon apprentissage et une bonne validation d'un CNN, il est essentiel d'utiliser une base de données de taille suffisante et équilibrée.

Les images utilisées dans les réseaux de neurones convolutifs peuvent être soit réelles, soit synthétiques. L'utilisation de photographies générées synthétiquement (Pirnia et al., 2018) peut être bénéfique pour augmenter facilement la taille des bases de données, à condition qu'elles ne créent pas de biais. Plusieurs études ont comparé les effets de l'utilisation de photographies artificielles sur des paramètres de performance tels que la RMSE (Duhaime et al., 2021; Manashti, 2022). Au cours de cette recherche, seules des photographies réelles sont employées.

CHAPITRE 2

DESCRIPTION MÉTHODOLOGIQUE

2.1 Description et préparation des matériaux utilisés

2.1.1 Présentation des essais effectués

Les essais réalisés ont permis d'étudier l'effet de trois paramètres sur la représentativité de l'analyse granulométrique et du pourcentage de matières résiduelles, soit :

- La présence de particules fines, notées seulement « fines » le long de ce mémoire ;
- La présence d'une phase liquide (eau ou phase organique) ;
- La méthode d'échantillonnage.

La brique particulière est la matière résiduelle choisie dans le cadre de cette recherche. Elle fait partie des débris provenant de la construction. (Ministère de l'environnement et de lutte contre les changements climatiques, 2021, partie 7.7). Elle a été choisie principalement pour sa non toxicité et la facilité de son identification qui est permise par sa couleur rouge vif pour sa quantification physique (1.3.2.1) et par analyse d'images (1.3.3). Se référer à la partie 2.1.3 pour plus d'information sur les particules de briques et leur préparation.

Pour étudier l'effet de la présence d'une phase organique sur l'échantillonnage, l'huile de canola a été choisie. Les produits pétroliers, comme le diesel, constituent la principale source de contamination des sols au Québec (1.1) mais sont toxiques donc il n'aurait pas été sécuritaire d'en manipuler. Les biodiesels étant inclus dans l'appellation de produits pétroliers d'après l'article 2 de la loi sur les produits pétroliers P-30.01, l'huile de canola s'est avérée être une bonne alternative non toxique au diesel dans le cadre de cette recherche. Le canola étant la matière première majoritairement utilisée dans plusieurs pays dont le Canada (et la France Figure 2.1 où colza = canola). Le Canada est également un important producteur de canola qui est une matière première particulièrement peu coûteuse et garantissant un bon rendement en huile (McKague & Brunke, 2023). De plus, la norme ASTM 6871 *Standard Specification for*

Natural (Vegetable Oil) Ester Fluids Used in Electrical Apparatus de 2008 qui définit les caractéristiques physiques requises pour qu'une huile puisse être utilisée dans certains appareils électriques, inclut notamment l'huile de canola. C'est pour toutes ces raisons que l'huile de canola brute est celle choisie en tant que contaminant liquide organique dans cette recherche.

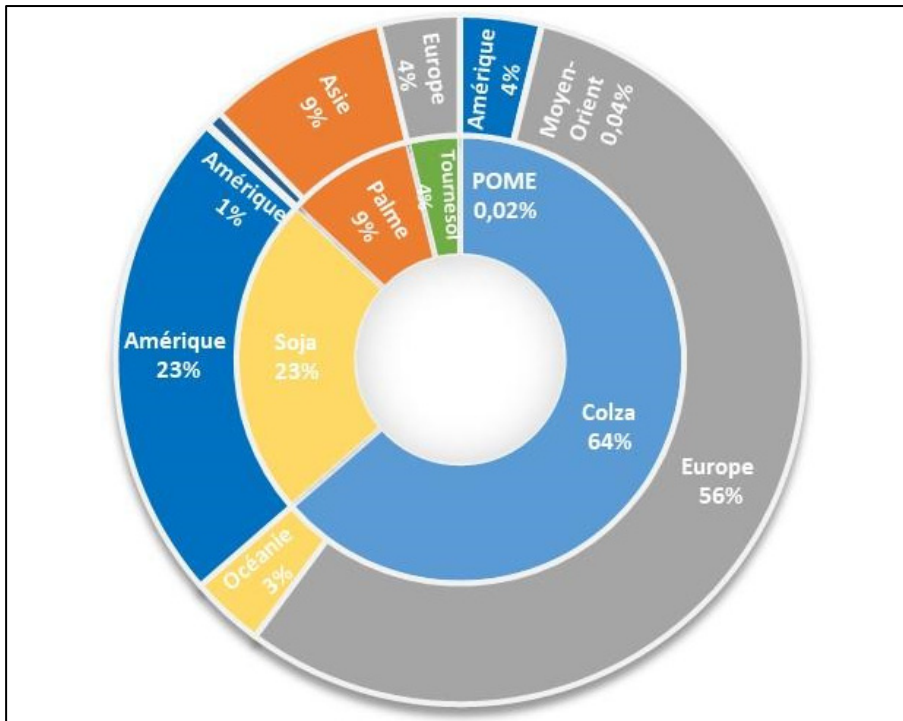


Figure 2.1 Origine des biocarburants mis sur le marché français en 2019
Tirée de Ministère Écologie Énergie Territoires (2019)

Le Tableau 2.1 présente les caractéristiques de chacun des dix essais.

Tableau 2.1 Caractéristiques des dix essais de la recherche

Essai	Particules fines	Phase liquide	Méthode d'échantillonnage
N°1	Avec fines	Non – essai sec	Séparation rotative
N°2	Avec fines	Non – essai sec	Grappillage
N°3	Avec fines	Non – essai sec	Pelletage fractionné
N°4	Sans fines	Non – essai sec	Grappillage
N°5	Sans fines	Non – essai sec	Séparation rotative
N°6	Sans fines	Non – essai sec	Pelletage fractionné
N°7	Avec fines	Oui – 10% d'eau	Grappillage
N°8	Avec fines	Oui – 10% d'eau	Pelletage fractionné
N°9	Avec fines	Oui – 10% d'huile	Grappillage
N°10	Avec fines	Oui – 10% d'huile	Pelletage fractionné

La séparation rotative a été seulement effectuée pour des essais secs, car elle n'est pas applicable aux sols humides. Par ailleurs, seuls des sols non lavés (avec fines c'est-à-dire avec des particules supérieures à 160 μm) ont été utilisés pour les essais humides. Premièrement, des résultats préliminaires ont montré que l'effet d'une phase liquide sur la caractérisation par analyse d'images est davantage mis en évidence avec des sols non lavés (Bouyim, 2022). Ensuite, deux autres études préalables et connexes à cette recherche (Dol, 2023; Zeryouh, 2023) se sont focalisées sur l'échantillonnage de sols lavés (sans fines), à l'état sec et humide.

Il a donc été nécessaire de constituer deux lots de sol soit, un lot de sol lavé et un lot de sol non lavé. Le sol utilisé pour constituer les deux lots ainsi que la procédure détaillée de la préparation de ces derniers est à retrouver en 2.1.2. Dans les deux cas, les sols ont été écrêtés à 5 mm (ce qui veut dire que seules les particules inférieures à 5 mm ont été gardées). Chaque lot contenait également 30% massique de particules de brique afin de représenter la présence d'une matière résiduelle. Au cours des projets précédents déjà évoqués, des teneurs de brique de 20% et 50% ont été utilisées (Dol, 2023; Zeryouh, 2023). Dans ces projets, l'obtention de

faux négatifs et de faux positifs a été étudiée, soit des échantillons contenant un pourcentage moins grand que 50% et provenant d'un lot à 50% (faux négatif) ou un pourcentage plus grand que 50% et provenant d'un lot à 20% (faux positif). Une teneur intermédiaire de 30% a donc été utilisée dans le présent projet afin de compléter ces études.

Pour déterminer la masse de sol nécessaire pour chacun des lots, le nombre d'échantillons par essai ainsi que leur masse individuelle ont d'abord été définis. La masse d'un échantillon a été établie à 300 g afin de respecter les exigences massiques minimales des normes BNQ 2501-025 de 2013 et LC 21-901 (Transport Québec, 2019). Pour chaque essai, 32 échantillons ont été prélevés afin d'avoir un ensemble de données suffisamment important pour que l'analyse statistique paramétrique soit robuste vis-à-vis de possibles violations des hypothèses de normalité et d'homoscédasticité (2.4.3).

Ainsi, une masse de 32×300 g, soit 9,6 kg, laquelle était composée de 6,72 kg de sol (lavé ou non) et de 2,88 kg de particules de briques était nécessaire par essai. Les deux lots utilisés pour l'ensemble des essais nécessitaient donc des masses de 67,14 kg pour le mélange sol et briques non lavé (7 essais d'après le Tableau 2.1) et de 28.80 kg pour le mélange sol et briques lavés (3 essais d'après le Tableau 2.1). Des quantités plus larges ont évidemment été préparées afin d'avoir une marge d'erreur.

Pour les essais humides, une masse d'eau distillée de 33,33 g a été ajoutée au mélange de sol et briques en début d'essai. Cette masse correspondait au pourcentage massique de 10% de l'échantillon de sol, briques et eau et à 11,11% de l'échantillon de sol et briques. Il a été fait de même avec l'huile pour les essais concernés. L'huile utilisée est de l'huile de canola de la marque Selection. Cette huile a été choisie car il s'agit de celle qui est majoritairement utilisée pour la fabrication de biocarburant en Amérique du Nord et elle est surtout non toxique à la manipulation (1.1.2).

2.1.2 Description et préparation du sol

Le sol utilisé était stocké au laboratoire dans des bidons comme ceux présentés en Figure 2.2. Il s'agissait de quantités excédentaires d'un sable utilisé pour la réalisation de précédents projets au LG2. Il était composé d'un large éventail de tailles de particules, de moins de $80\ \mu\text{m}$ à plus de $15\ \text{cm}$ (Figure 2.2). Le sol a tout d'abord été étalé dans de grandes pannes rectangulaires qui ont ensuite été disposées dans des étuves à 110°C pendant 24h. Ensuite, les plus grosses particules ont été enlevées à la main, puis avec un tamis $5\ \text{mm}$ pour écrêter à cette dimension. Le tamisage a été réalisé manuellement à l'aide d'un tamis d'un diamètre de $30\ \text{cm}$. Le passant $5\ \text{mm}$ a été conservé dans des sacs de polyéthylène.

Pour constituer sept lots identiques de sol avec fines nécessaires à la réalisation des sept essais associés (Tableau 2.1), un séparateur rotatif de grande capacité a été utilisé. Obtenir des lots identiques a permis de comparer les résultats d'un essai à l'autre. Le séparateur rotatif produisant huit échantillons à partir d'un lot, le huitième lot a été utilisé pour des analyses de référence, soit l'analyse granulométrique par tamisage et par sédimentométrie (ANNEXE I), la détermination du pourcentage de matière organique moyen ainsi que la détermination du taux d'humidité hygroscopique moyen (qui se détermine selon la procédure décrite en 2.2.3.1). Cette première analyse granulométrique a permis de constater que le sol contenait environ $16,5\%$ de particules fines inférieures à $160\ \mu\text{m}$.

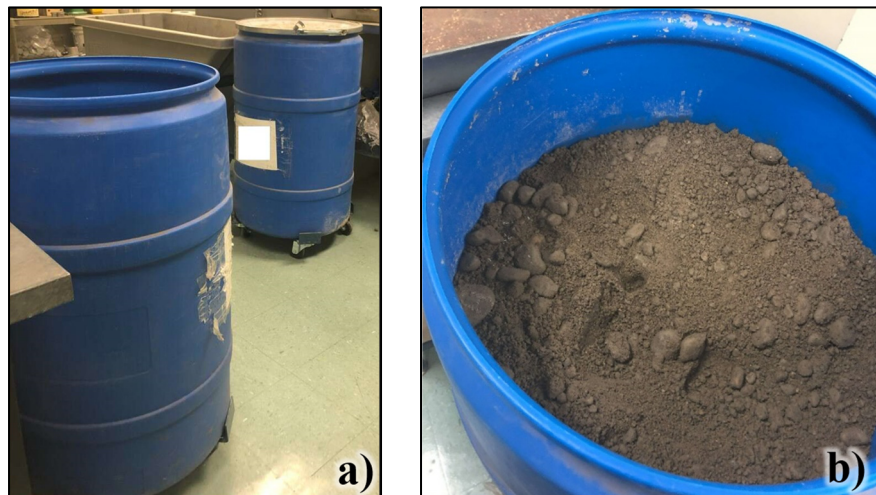


Figure 2.2 Photographies des bidons contenant le sol brut

Ainsi, pour constituer trois lots identiques de sol sans fines nécessaires à la réalisation des trois essais associés (Tableau 2.1), le séparateur rotatif de grande capacité a également été utilisé. Au préalable, il a fallu laver puis tamiser une quantité de sol suffisante pour en retirer les particules fines c'est-à-dire celles inférieures à 160 μm . Ceci a été réalisé avec le dispositif de lavage présenté en Figure 2.3 plutôt que selon la procédure décrite dans la norme BNQ 2501-025 de 2013, car plus il était plus adapté aux quantités à traiter.



Figure 2.3 Dispositif de lavage d'importantes quantités de sol

2.1.3 Description et préparation des fragments de brique ajoutés au sol

Comme décrit précédemment, des fragments de brique ont été incorporés aux dix lots associés aux dix essais (Tableau 2.1) de sorte à ce qu'ils constituent 30% de la masse totale des lots. La

masse de fragments de brique à ajouter aux différents lots de sol a été recalculée selon la masse réelle de ces derniers de sorte à ce qu'il y ait très exactement 30% de fragments de briques à chaque fois.

Afin que les fragments de brique s'intègrent à la granulométrie du sol tout en étant facilement identifiables à l'œil nu en vue d'une séparation physique, il a été choisi que leur taille soit comprise entre 2,5 mm et 5 mm.

Des briques entières, de marque BRS McAvoy red fine Mattex standard (SRFM) (Quebec size) de la société Montréal Brique & Pierre Inc (Figure 2.4), ont été utilisées. Elles ont été choisies pour la vivacité de leur couleur rouge qui devait permettre de faciliter leur identification visuelle mais également leur reconnaissance par le réseau de neurones par la suite.

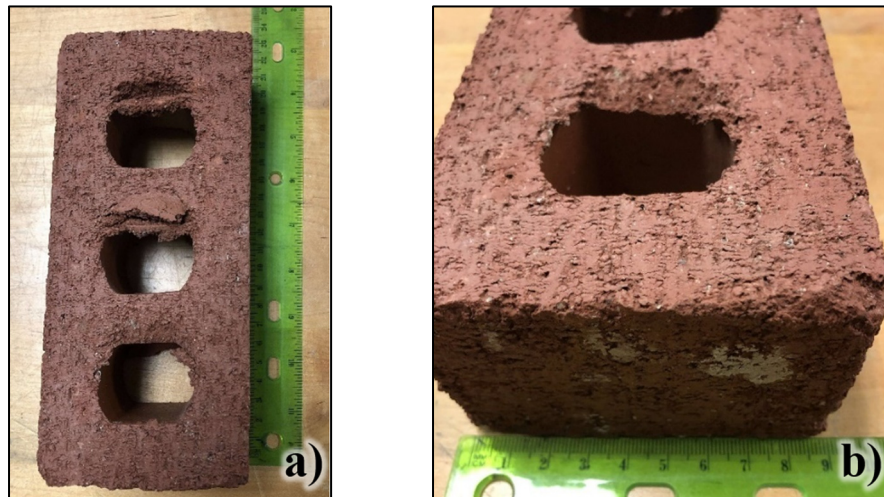


Figure 2.4 Brique entière ($\approx 20,5$ cm x 8,5 cm ; $\approx 1,92$ kg)

Les briques ont été concassées en deux étapes, étant donné qu'entières, les briques étaient trop grosses pour le concasseur mécanique utilisé pour atteindre leur taille finale. La première étape de concassage a donc consisté à casser les briques manuellement à l'aide d'un marteau pour pouvoir générer des morceaux de briques grossiers (Figure 2.5). La deuxième étape de concassage a été effectuée à l'aide du concasseur mécanique présenté en Figure 2.6.



Figure 2.5 Pré-concassage manuel



Figure 2.6 Concasseur mécanique

Après chaque concassage, les particules inférieures à 5 mm ont été réservées, alors que celles supérieures à 5 mm ont été renvoyées au concasseur jusqu'à ce qu'il n'y ait plus de particules supérieures à 5 mm.

À la fin du processus, les particules ainsi générées ont été tamisées manuellement à 2,5 mm. Le passant a été rejeté et le retenu, gardé pour les essais étant donné que la méthode d'analyse visuelle décrite en 1.3.2.1 ne permet que de quantifier des particules supérieures à 2,5 mm. Ces pertes de fragments inférieurs à 2,5 mm ont été estimées à 60%. Ainsi, afin d'avoir une marge confortable d'une dizaine de kilos, 100 kg de briques entières ont été concassés.

2.2 Déroulement des essais

Les essais se sont déroulés globalement tous de la même façon. En premier lieu, chaque lot a été échantillonné par la méthode d'échantillonnage choisie (Tableau 2.1). Ensuite, chaque échantillon obtenu a été photographié à l'aide du dispositif prévu à cet effet (Figure 2.10). Par la suite, si l'échantillon contenait une phase liquide, cette dernière était enlevée, soit par séchage pour l'eau, soit par perte au feu pour l'huile. Finalement, chaque échantillon a subi un tamisage mécanique afin de déterminer sa distribution granulométrique suivi d'une séparation manuelle des fractions supérieures à 1,25 mm pour déterminer leur teneur en fragments de brique. Dans la méthode d'analyse décrite en 1.3.2.1, il s'agit des fractions supérieures à 2,5 mm qui sont analysées mais puisque les particules de brique sont friables, une partie non négligeable aurait été omise si la fraction 1,25 – 2,5 mm n'avait pas été analysée. Les masses des échantillons ont été relevées en début et en fin de chaque étape de l'essai afin d'évaluer les pertes qu'elle générerait.

Le détail du déroulement des essais suit dans cette section 2.2.

2.2.1 Méthodes d'échantillonnage

2.2.1.1 Grappillage

Le grappillage est un mode de prélèvement qui repose sur le jugement du préleveur et qui est réalisé en prélevant un seul incrément à partir d'une partie accessible du lot, généralement sa surface. Ainsi, il n'existe pas de procédure normalisée, voire simplement balisée, pour ce mode de prélèvement.

Au cours de la présente recherche, le grappillage a été effectué en trois étapes. Tout d'abord le lot de mélange sol et briques a été homogénéisé à la main dans une panne; Ensuite, le mélange a été étalé sur toute la surface de la panne. Enfin, 32 échantillons ont été prélevés en un incrément chacun sans remise à l'aide d'une truelle d'une capacité de 300 g environ, puis disposé dans un sac en plastique. Les échantillons contenant une phase liquide ont été placés dans des sacs en plastique hermétiques.

Le grappillage est utilisé pour les essais n°2, 4, 7 et 9 (Tableau 2.1). La Figure 2.7 présente le grappillage d'un lot de matériel humide (essai n°7).



Figure 2.7 Grappillage d'un mélange humide

2.2.1.2 Pelletage fractionné

La méthode utilisée pour le pelletage fractionnée est celle décrite à la section 1.2.3.2. Pour cette méthode d'échantillonnage, le lot a été homogénéisé à la main, puis disposé en un tas de forme conique dans une panne. Ensuite, 32 échantillons ont été créés selon la méthode décrite à la section 1.2.3.2. Dans cette étude, chaque échantillon a été créé en combinant 20 incréments. L'outil de prélèvement utilisé, une cuillère ronde, avait une capacité de 15 g, de

manière à obtenir des échantillons de 300 g. Pour l'essai avec des lots contenant une phase liquide, les incréments ont été disposés dans des sacs hermétiques qui étaient systématiquement refermés entre chaque incrément pour limiter l'évaporation.

La Figure 2.8 présente un essai en cours de pelletage fractionné d'un lot sec (essai n°3) et d'un lot humide (essai n°8).



Figure 2.8 Pelletage fractionné sur des lots a) sec et b) humide

2.2.1.3 Séparation rotative

Le dispositif de séparation rotative était constitué du diviseur d'échantillons à cône rotatif ainsi que d'une trémie d'alimentation. La Figure 2.9 en présente les différents éléments.

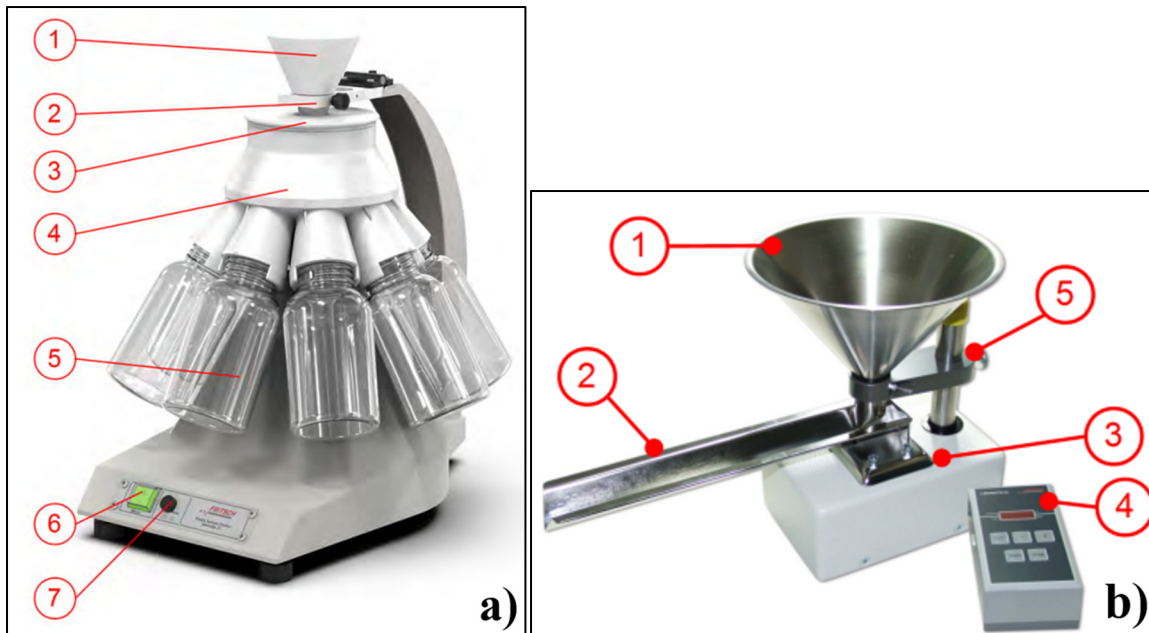


Figure 2.9 a) Diviseur d'échantillons à cône rotatif Laborette 27 de FRITSCH (1 – entonnoir ; 2 – support d'entonnoir ; 3 – couvercle de tête de division ; 4 – tête de division ; 5 – flacons d'échantillonnage ; 6 – interrupteur principal ; 7 – fusible) et b) trémie d'alimentation à vibrations Laborette 24 de FRITSCH (1 – entonnoir ; 2 – goulotte ; 3 – appareil de base ; 4 – unité d'entonnoir ; 5 – support d'entonnoir réglable en hauteur)

Tirées de FRITSCH (2019 et 2020)

Le mélange sol et briques, initialement contenu dans un grand saladier métallique, a été versé dans l'entonnoir de la trémie d'alimentation (Figure 2.9) à l'aide d'une truelle. Préalablement à la séparation elle-même, le lot de mélange sol et briques a été homogénéisé à l'aide du dispositif en le divisant et en le remélangeant deux fois.

Ensuite, afin d'obtenir 32 échantillons, il a été choisi de rassembler aléatoirement deux à deux les huit échantillons issus de la première séparation pour n'en avoir plus que quatre avant la seconde séparation. Grâce à cette étape supplémentaire, 32 échantillons ont été obtenus suite à une seconde séparation.

Un pinceau à poils doux a été utilisé pour l'essai avec fines afin de récupérer le plus de fines qui se seraient déposées dans le dispositif (dans un entonnoir ou la goulotte par exemple).

Enfin, les échantillons ont été disposés dans des sacs plastiques identifiés.

2.2.2 Photographie

L'objectif de la prise de photos est de constituer une base de données qui serait exploitée pour entraîner, valider et tester un réseau neuronal. Ce dernier serait donc probablement plus adapté à des photographies d'échantillons prises en laboratoire mais constituerait tout de même une préface à la mise en place de réseaux neuronaux convolutifs capables d'analyser des photographies prises directement sur le terrain.

Pour ce faire, un montage photo a été pensé pour photographier un échantillon étalé sur une plaque en verre afin de les prendre en photo par le haut ainsi que par le bas. L'utilité d'avoir ces deux vues est de capter toute la gamme granulométrique de l'échantillon photographié. En effet, les particules fines ont tendance à se déposer dans le fond du cadre tandis que les particules les plus grossières restent en surface.

2.2.2.1 Le montage photo et les réglages

Le montage pour la prise des photographies est constitué d'un cadre en aluminium (Figure 2.10) qui supporte un cadre en bois et en verre. Le cadre en aluminium inclut un système de réglage du positionnement de l'appareil photo du haut. L'appareil photo du bas est fixé sur un trépied en dessous de l'échantillon. La distance entre les appareils photo et l'échantillon a été définie telle qu'un pixel soit égal à environ 20 μm afin de distinguer les particules les plus fines.

Le cadre qui supporte l'échantillon est constitué d'une plaque en bois avec une ouverture centrale en verre de 250 \times 160 mm bordé d'un cadre en bois afin d'y déposer un échantillon. Ces dimensions ont été définies selon plusieurs critères :

- Le rapport largeur \times hauteur doit se rapprocher au mieux de celui des photos prises par l'appareil afin de limiter le traitement d'image (notamment le rognage) ;
- Le cadre doit être en mesure de recevoir des échantillons inférieurs à 2 kg et contenant des particules de maximum 5 mm pour correspondre au format d'échantillon de l'étude ;

- L'épaisseur de l'échantillon étalé doit y être de trois particules environ pour avoir des photographies les plus représentatives possible.

Les appareils photos utilisés sont des NIKON Z6 II de résolution 4024 x 6048 pixels. Ces appareils sont équipés d'objectifs macro et de télécommandes afin d'éviter au plus le dérèglement de la position des appareils lors de la prise de photos. Les appareils sont réglés en mode automatique et l'exposition est réglée à +1. Pour garantir exposition optimale, deux boîtes à lumières sont placées de sorte à éclairer, sans générer de reflet, le haut et le bas de l'échantillon.

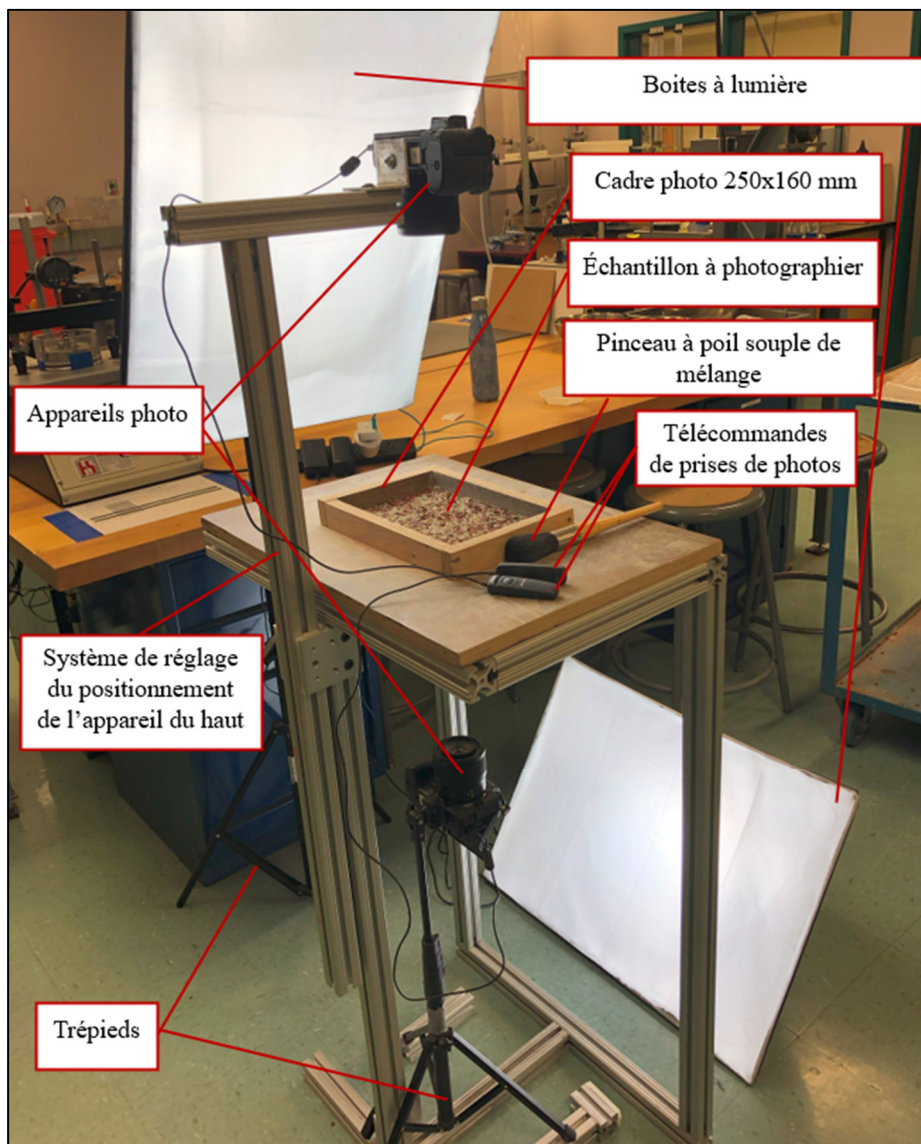


Figure 2.10 Montage pour la prise des photographies

2.2.2.2 Procédure de prise de photographie

Les photographies sont prises après l'étape de séparation, le même jour pour les échantillons contenant de l'eau afin d'en limiter l'assèchement préalable. La procédure est la même quel que soit l'essai duquel sont issus les échantillons photographiés :

- 1) Le montage photo est réglé au préalable : la distance entre l'appareil photo du haut et l'échantillon est ajustée avec le bras coulissant. Les appareils photo sont positionnés de sorte à bien cadrer l'échantillon. Ce réglage n'est fait qu'une fois par journée de prise de photos ;
- 2) L'échantillon est disposé dans le cadre et est étalé le plus homogènement possible à l'aide d'un pinceau à poils pour les échantillons secs, et d'une maryse pour les échantillons contenant une phase liquide. Pour rappel, l'épaisseur de l'échantillon étalé dans le cadre doit être environ de trois particules considérant la taille de plus grosse particule de l'échantillon ;
- 3) Le placement du cadre photo est ajusté à partir de ce que l'écran des appareils photos renvoie, le positionnement des appareils n'est en aucun cas touché et ce, tout au long du processus ;
- 4) L'échantillon est identifié à l'aide d'un papier contenant les informations nécessaires posé dessus, une règle est également ajoutée pour pouvoir mesurer au cas par cas la taille d'un pixel. Une première photographie est prise ainsi avant d'enlever ces éléments et de passer à la suite ;
- 5) L'échantillon est pris en photo avec les deux appareils en même temps avant d'être remanié à l'aide du pinceau ou de la maryse selon le cas. Le remaniement permet d'augmenter la taille de la base de données en prenant plusieurs photographies d'un même échantillon ;
- 6) L'étape précédente est répétée de sorte à produire quatre photographies du haut et du bas de l'échantillon en plus de la photo d'identification.

Pour chaque essai, la masse des échantillons est systématiquement mesurée avant et après la prise de photographies afin d'estimer les pertes générées lors de cette étape.

Au total, chaque échantillon est donc pris neuf fois en photos en comptant la photo d'identification de l'échantillon, quatre fois de haut et quatre fois par le bas. Pour les dix essais (Tableau 2.1) et les 32 échantillons par essai, 2560 photographies ont été prises.

2.2.3 Enlèvement et quantification de la phase liquide

2.2.3.1 Eau

La quantification et l'élimination de l'eau des échantillons, suite à la photographie, ont été réalisées selon de la méthode d'analyse décrite en 7.3 du MA. 100 – S.T. 1.1 du (Centre d'expertise en analyse environnementale du Québec, 2017). La procédure a consisté à peser l'échantillon le plus précisément possible, le laisser une nuit à l'étuve à 110°C et le peser le lendemain. La procédure du CEAEQ préconise un échantillon d'environ 10 g, cela dit, sous-échantillonner les échantillons d'environ 333 g n'aurait qu'engendré des pertes et des erreurs supplémentaires. Il a donc été choisi d'étaler les échantillons un à un dans une assiette et de la laisser 24h à l'étuve.

2.2.3.2 Huile

La quantification et l'élimination de l'huile des échantillons, suite à la photographie, ont été réalisées selon la méthode d'analyse décrite en 7.3 du MA. 100 – S.T. 1.1 du CEAEQ (2017). Cette fois-ci, la procédure a consisté à laisser les échantillons, l'un après l'autre, pendant environ une à deux heures, dans le four à moufle à 550°C présenté en Figure 2.11.



Figure 2.11 Photographie du four à moufle

La Figure 2.12 présente un échantillon de sol, briques et huile à différents stades de la combustion.

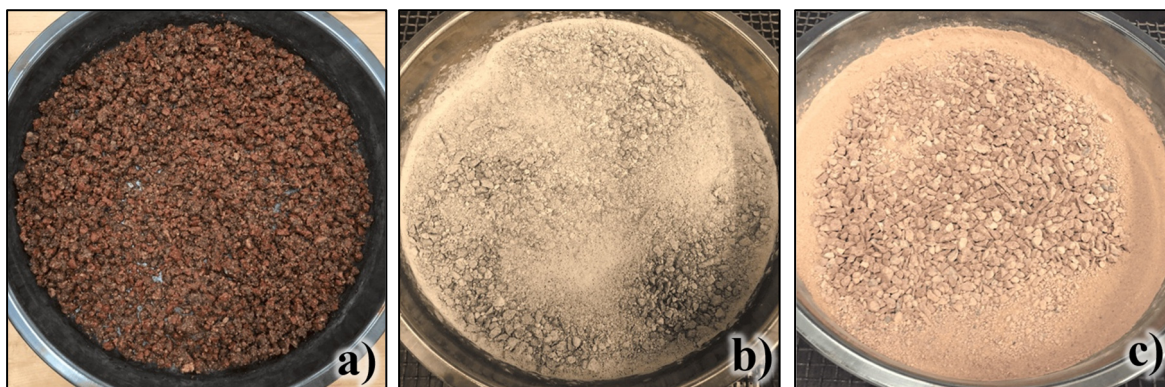


Figure 2.12 Échantillon de sol, briques et huile a) avant combustion, b) après combustion partielle et c) après combustion complète

Il est possible de constater, sur la deuxième photographie, que la combustion n'était pas complète puisqu'il restait des traces noires d'huile carbonisée. L'échantillon n'était retiré du four que lorsqu'il avait atteint l'aspect de la troisième photographie.

2.2.4 Granulométrie et séparation

Le matériel utilisé et la méthode suivie sont ceux décrits par la norme BNQ 2501-025 de 2013. Pour les échantillons contenant des particules fines, une première étape de lavage des particules inférieures à 80 μm a été requise. Pour ce faire, chaque échantillon avec fines a été pesé, puis lavé par rinçages successifs sur un tamis 80 μm selon la procédure indiquée dans la norme et ce, jusqu'à ce que l'eau de rinçage soit claire. En fin de processus, les particules retenues sur le tamis ont été récupérées et séchées une nuit à l'étuve à 110°C. L'échantillon a ensuite été pesé afin de déterminer la masse de particules inférieures à 80 μm éliminées lors du lavage.

L'échantillon a ensuite été placé dans une batterie de six tamis de 20 cm de diamètre à mailles carrées et normalisés (ISO 3310-1), soit ayant des ouvertures de 2,5 mm, 1,25 mm, 630 μm , 315 μm , 180 μm et 80 μm . Cette batterie a été agitée pendant vingt minutes dans un tamiseur mécanique comme recommandé la méthode d'essai LC 21 – 040 (Transport Québec, 2012), la norme BNQ 2501-025 de 2013 ne donnant pas d'indication particulière à ce sujet. En fin de tamisage, le contenu de chaque tamis a été méticuleusement récupéré à l'aide de brosses, puis pesé. Les retenus 2,5 mm, 1,25 mm, 630 μm , 315 μm , 180 μm puis 80 μm sont présentés respectivement pour un échantillon sans fines (Figure 2.13) et pour un échantillon avec fines (Figure 2.14) lavé et séché.

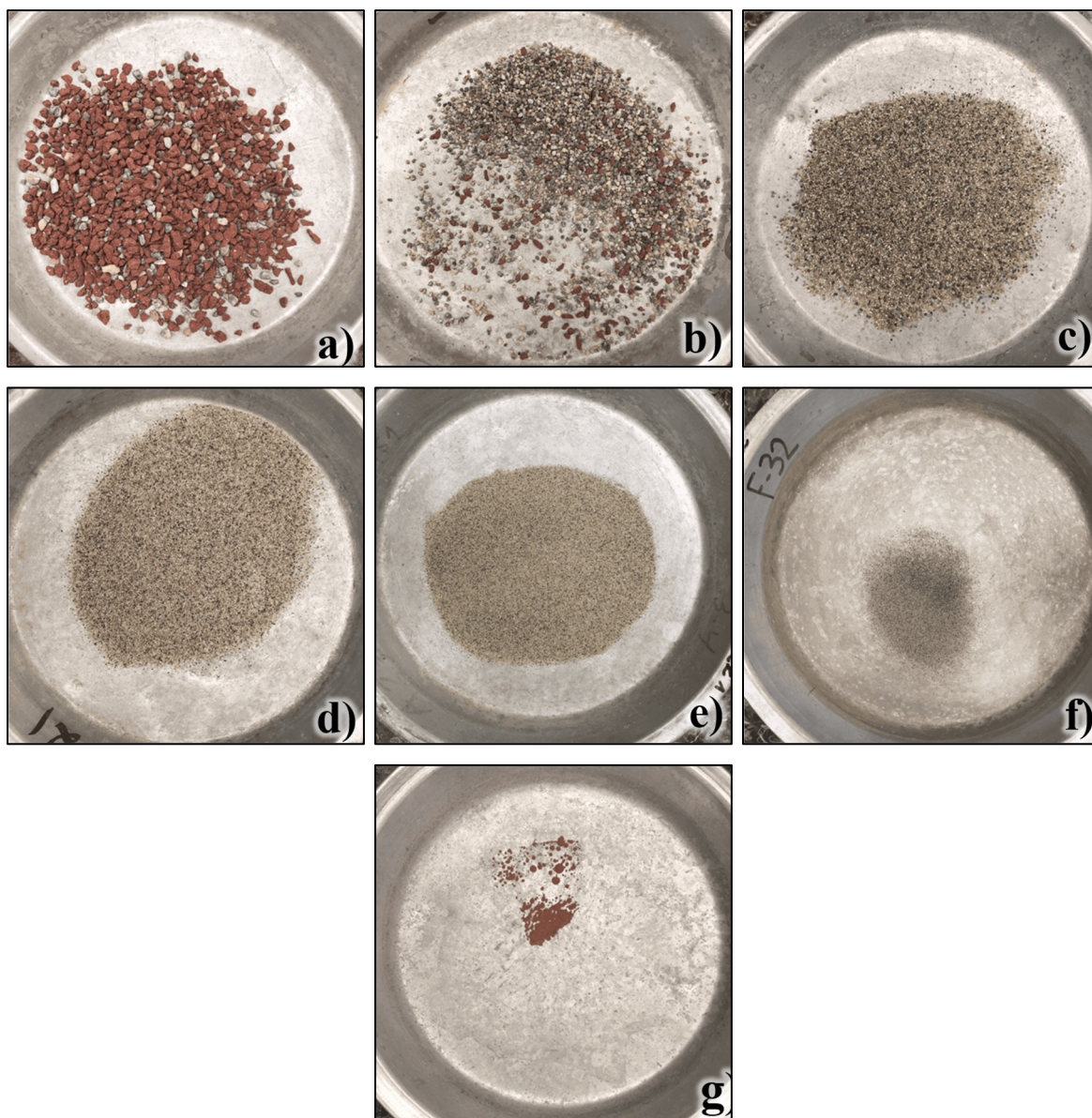


Figure 2.13 Retenus a) 2,5 mm, b) 1,25 mm, c) 630 μm , d) 315 μm , e) 180 μm , f) 80 μm et de g) la panne pour un échantillon sans fines

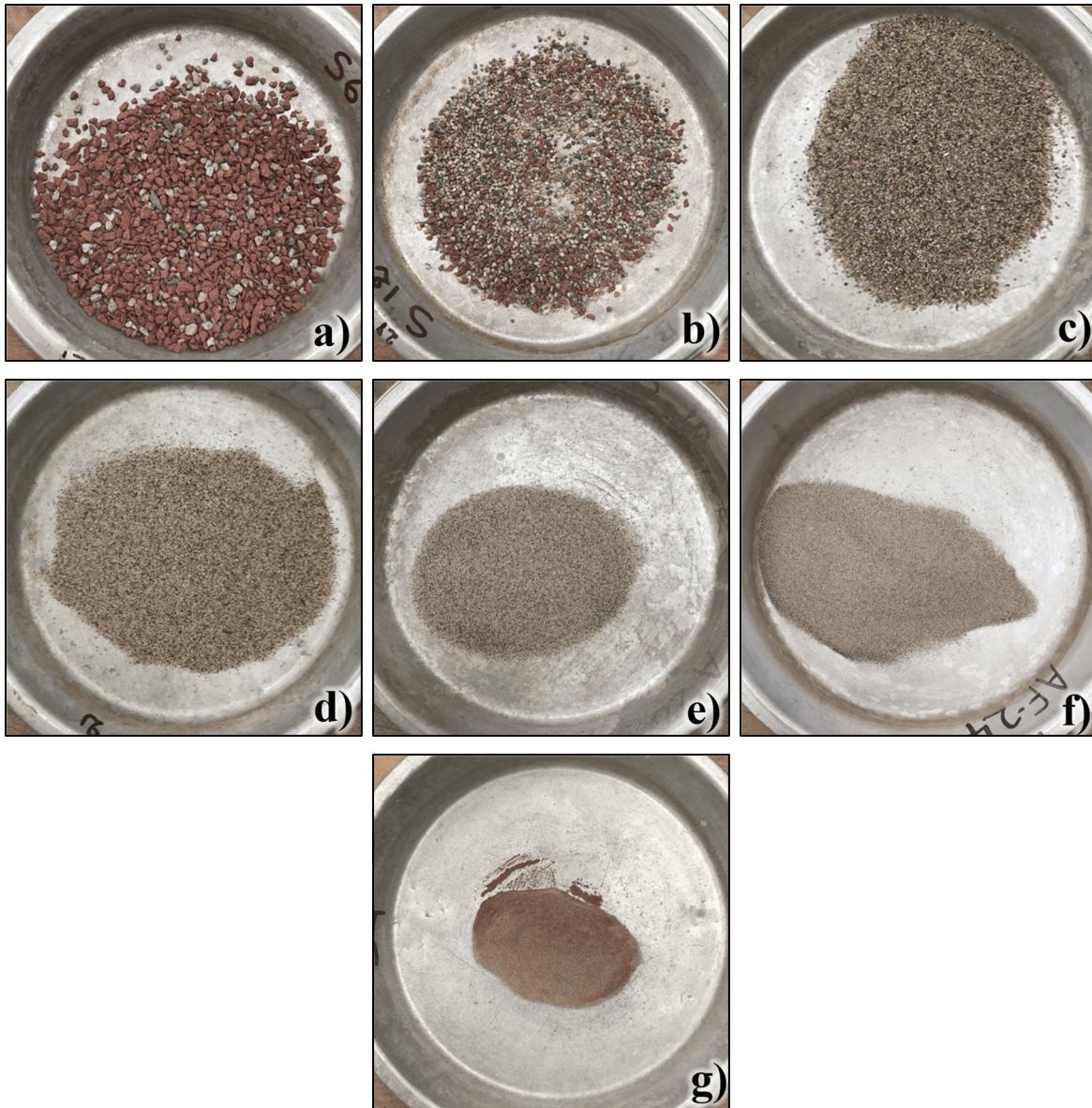


Figure 2.14 Retenus a) 2,5 mm, b) 1,25 mm, c) 630 μm , d) 315 μm , e) 180 μm , f) 80 μm et de g) la panne pour un échantillon avec fines

Ces deux figures permettent de constater que bien que les particules de briques se soient retrouvées essentiellement dans le retenu 2,5 mm tel que prévu, il y en avait également dans le retenu 1,25 mm, ainsi que dans la panne contenant les particules inférieures à 80 μm . Il y en avait également dans les autres retenus, mais en moins grande quantité. Cet élément est important, car la quantification manuelle de la teneur en fragments de brique a été réalisée sur

les retenus 1,25 mm et 2,5 mm seulement. Ceci implique que la teneur en fragments de brique a été légèrement sous-estimée.

Le triage des briques a été réalisé selon la méthode LC 21-901 (Transport Québec, 2019). Le pourcentage de fragments de brique d'un échantillon a été déterminé à l'aide de l'équation (2.1).

$$\%briques_{ech} = \frac{m_{briques_{1,25-2,5\text{ mm}}} + m_{briques_{2,5-5\text{ mm}}}}{m_{totale_{ech}}} \quad (2.1)$$

Avec :

- $\%briques_{ech}$: le pourcentage de fragments de brique d'un échantillon ;
- $m_{briques_{1,25-2,5\text{ mm}}}$: la masse de fragments de brique comprise dans la fraction 1,25-2,5 mm de l'échantillon (g) ;
- $m_{briques_{2,5-5\text{ mm}}}$: la masse de fragments de brique comprise dans la fraction 2,5-5 mm de l'échantillon (g) ;
- $m_{totale_{ech}}$: la masse totale de l'échantillon avant l'analyse granulométrique (g).

2.3 Analyse d'image

Pour l'analyse d'image, l'objectif a été d'entraîner et valider un réseau neuronal convolutif (CNN) de régression de sorte à ce qu'il soit capable de quantifier de manière la plus fiable possible, le pourcentage de briques contenu dans une photo d'échantillon de taille 512 x 810 pixels. Pour ce faire, une base de données photographiques équilibrée en termes de pourcentages de briques représentés a été mise en place parmi laquelle 80% des photos ont été consacrées à l'entraînement, 10% à la validation et 10% au test.

Il a été choisi d'implémenter le réseau en langage Python sur PyCharm qui est un environnement de développement intégré pour lequel il existe une version gratuite relativement performante au vu des besoins de la présente recherche.

2.3.1 Base de données photographique

2.3.1.1 Base de données photographique initiale (2072 photographies)

La base de données photographique regroupe un total de 4146 photographies d'échantillons de sol avec briques numérotées de 1 à 4146. Les 2072 premières photographies sont prises du haut. Les autres sont prises du bas. Parmi ces photographies, 2560 sont issues des essais réalisés au cours de ce projet (section 2.2.3) et 1586 sont issues des projets antérieurs (Dol, 2023; Zeryouh, 2023) pour lesquels le sol utilisé était similaire au sol de ce projet mais sans fines, les particules de briques étaient identiques, et la méthodologie d'échantillonnage et de prise de photos suivie a été la même que celle suivie au cours de cette recherche. Dans cette base de données, il est possible de retrouver des photographies d'échantillons avec différents pourcentages de briques, avec ou sans fines et avec ou sans phase liquide. La structure de la base de données est présentée dans le Tableau 2.2. La base de données fournit le pourcentage de briques ainsi que la granulométrie de chaque échantillon. Le Tableau 2.3 présente un extrait de la base de données.

Tableau 2.2 Description de la base de données photographique originale de sol avec briques

Référence	Détail sur les échantillons				N° des photos	
	%Briques du lot	Méthode d'échantillonnage	Particules fines	Phase liquide	Photos du haut	Photos du bas
Ce mémoire	30%	Grappillage	Sans fines	Sec	1 à 128	2073 à 2200
	30%	Séparation rotative	Sans fines	Sec	129 à 256	2201 à 2328
	30%	Pelletage fractionné	Sans fines	Sec	257 à 384	2329 à 2456
	30%	Grappillage	Avec fines	Sec	385 à 512	2457 à 2584
	30%	Séparation rotative	Avec fines	Sec	513 à 640	2585 à 2712
	30%	Pelletage fractionné	Avec fines	Sec	641 à 768	2713 à 2840
	30%	Grappillage	Avec fines	10% d'eau	769 à 896	2841 à 2968
	30%	Pelletage fractionné	Avec fines	10% d'eau	897 à 1024	2969 à 3096
	30%	Grappillage	Avec fines	10% d'huile	1025 à 1152	3097 à 3224
	30%	Pelletage fractionné	Avec fines	10% d'huile	1153 à 1280	3225 à 3352
(Zeryouh, 2023)	50%	Grappillage	Sans fines	6% d'eau	1281 à 1344	3353 à 3416
	50%	Grappillage	Sans fines	10% d'eau	1345 à 1408	3417 à 3480
	50%	Pelletage fractionné	Sans fines	6% d'eau	1409 à 1472	3481 à 3544
	50%	Pelletage fractionné	Sans fines	10% d'eau	1473 à 1535	3545 à 3607
	20%	Grappillage	Sans fines	6% d'eau	1536 à 1599	3608 à 3671
	20%	Pelletage fractionné	Sans fines	6% d'eau	1600 à 1662	3672 à 3734
(Dol, 2023)	20%	Grappillage	Sans fines	Sec	1663 à 1727	3735 à 3799
	50%	Grappillage	Sans fines	Sec	1728 à 1804	3800 à 3876
	20%	Pelletage fractionné	Sans fines	Sec	1805 à 1872	3877 à 3944
	50%	Pelletage fractionné	Sans fines	Sec	1873 à 1937	3945 à 4009
	20%	Séparation à chutes	Sans fines	Sec	1938 à 2008	4010 à 4080
	50%	Séparation à chutes	Sans fines	Sec	2009 à 2072	4081 à 4144

Tableau 2.3 Extrait de la base de données photographique de sol avec briques

N° photo	Autrice	Numéro d'échantillon	Présence de fines ?	Méthode d'échantillonnage	%briques lot	%d'eau	Haut/bas	Répétition	%briques éch - méthode visuelle
1	Clélia M	1	Sans fines	Grappillage	30%	0%	Haut	1	32,99%
2	Clélia M	1	Sans fines	Grappillage	30%	0%	Haut	2	32,99%
3	Clélia M	1	Sans fines	Grappillage	30%	0%	Haut	3	32,99%
4	Clélia M	1	Sans fines	Grappillage	30%	0%	Haut	4	32,99%
5	Clélia M	2	Sans fines	Grappillage	30%	0%	Haut	1	31,75%
6	Clélia M	2	Sans fines	Grappillage	30%	0%	Haut	2	31,75%
7	Clélia M	2	Sans fines	Grappillage	30%	0%	Haut	3	31,75%
8	Clélia M	2	Sans fines	Grappillage	30%	0%	Haut	4	31,75%

N° photo	% Passant 5 mm	% passant 2,5 mm	% passant 1,25 mm	% passant 630 µm	% passant 315 µm	% passant 160 µm	% passant 80 µm
1	100%	53,84%	35,80%	19,94%	7,28%	0,38%	0,03%
2	100%	53,84%	35,80%	19,94%	7,28%	0,38%	0,03%
3	100%	53,84%	35,80%	19,94%	7,28%	0,38%	0,03%
4	100%	53,84%	35,80%	19,94%	7,28%	0,38%	0,03%
5	100%	55,38%	38,38%	19,50%	6,67%	0,31%	0,06%
6	100%	55,38%	38,38%	19,50%	6,67%	0,31%	0,06%
7	100%	55,38%	38,38%	19,50%	6,67%	0,31%	0,06%
8	100%	55,38%	38,38%	19,50%	6,67%	0,31%	0,06%

La Figure 2.15, présente les photographies 257, 641, 897, 1153, 1805 et 1873. Cette figure met en avant la diversité d'aspects des photographies de la base de données. Les Tableau 2.4 et Tableau 2.5 donne les caractéristiques de chacun de ces échantillons.

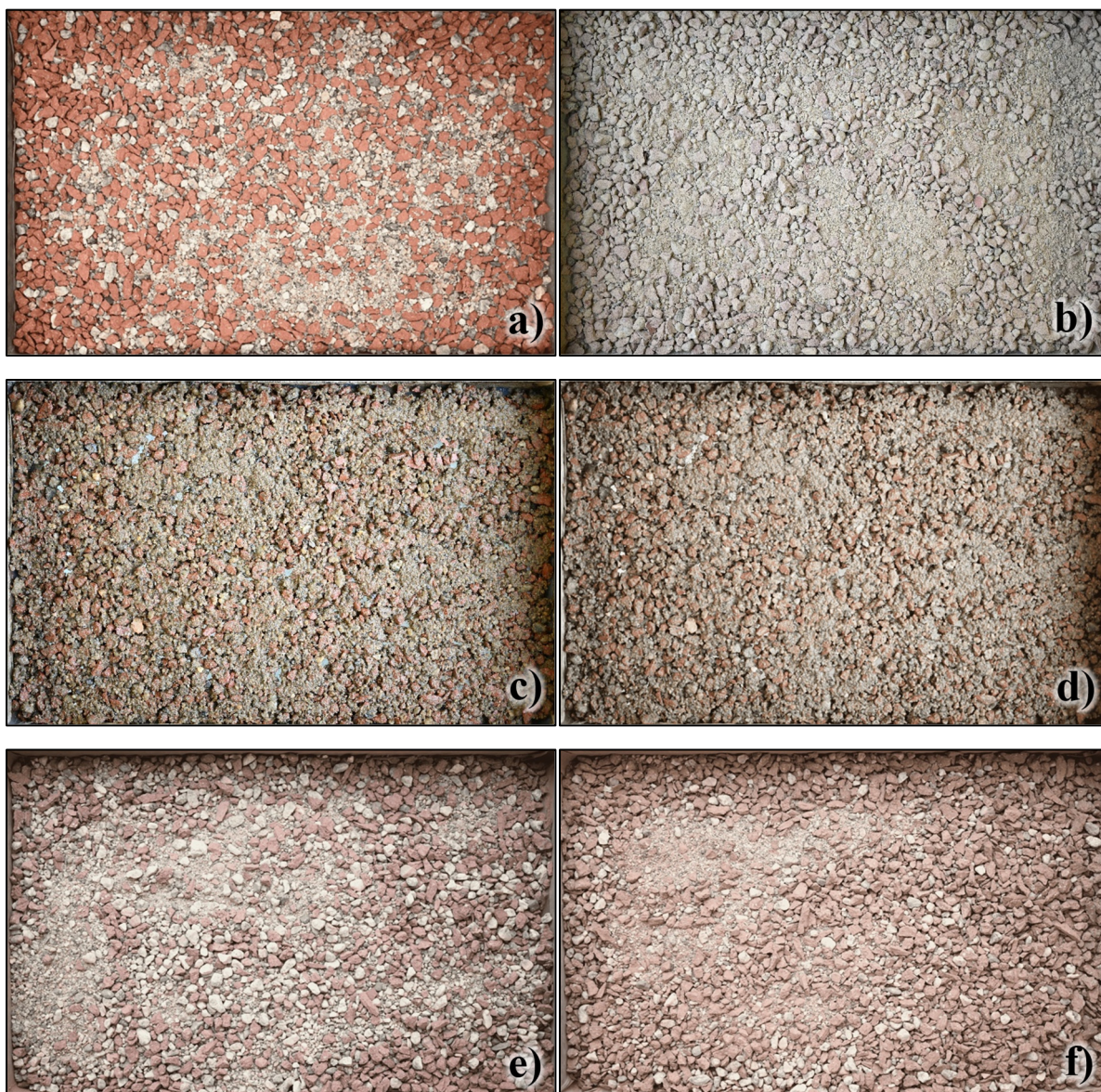


Figure 2.15 Photographies des échantillons a) 257, b) 641, c) 897, d) 1153, e) 1805 et f) 1873

Tableau 2.4 Information sur les essais réalisés et les pourcentages de briques des échantillons numéro 257, 641, 897, 1153, 1805 et 1873 présentés en Figure 2.15

Référence	N° photo	Essai	% briques du lot	% briques de l'échantillon
Ce mémoire	257	Sec, sans fines	30%	30,34%
Ce mémoire	641	Sec, avec fines	30%	30,98%
Ce mémoire	897	10%eau, avec fines	30%	30,55%
Ce mémoire	1153	10%huile, avec fines	30%	28,51%
(Dol, 2023)	1805	Sec, sans fines	20%	22,98%
(Dol, 2023)	1873	Sec, sans fines	50%	53,15%

Tableau 2.5 Information granulométriques des échantillons numéro 257, 641, 897, 1153, 1805 et 1873 présentés en Figure 2.15

Référence	N° photo	% passant 5 mm	% passant 2,5 mm	% passant 1,25 mm	% passant 630 µm	% passant 315 µm	% passant 160 µm	% passant 80 µm
Ce mémoire	257	100	60,33	43,13	22,40	7,90	0,50	0,18
Ce mémoire	641	100	57,41	41,43	28,12	17,62	10,98	7,37
Ce mémoire	897	100	60,64	43,06	28,98	18,05	11,30	7,66
Ce mémoire	1153	100	61,94	47,87	30,26	18,04	11,36	7,50
(Dol, 2023)	1805	100	54,77	38,98	-	-	0,35	-
(Dol, 2023)	1873	100	40,48	24,55	-	-	0,60	-

2.3.1.2 Base de données photographique rééquilibrée (3195 photographies)

La base de photographies du Tableau 2.2 contient 2072 photographies prises de haut, dont 1280 de sol à 30% de briques, 331 de sol à 20% de briques et 461 de sol à 50% de briques (Tableau 2.6). Pour l'entraînement des CNN, il est préférable d'utiliser une base de données équilibrées pour éviter les prédictions biaisées. Le réseau aura tendance à fournir des prédictions qui surreprésentent les pourcentages de briques les plus fréquents dans la base de données (Torralba & Efros, 2011). Pour pallier ce déséquilibre, il est possible d'utiliser une fonction de pertes personnalisée qui donne plus de poids aux valeurs extrêmes lors de l'entraînement. Il est également possible de rééquilibrer la base de données en appliquant des transformations simples aux photographies sous-représentées. Cette méthode de rééquilibrage a été choisie, car elle s'est avérée plus concluante que l'utilisation d'une fonction de pertes personnalisée.

Pour rééquilibrer la base de données et ainsi avoir approximativement le même nombre de photos d'échantillons issus de lots à 20%, 30% et 50% de particules de briques, il a été choisi d'augmenter le nombre de photographies de sol avec environ 20% de briques d'un facteur trois et d'augmenter le nombre de photographies de sol avec environ 50% de briques d'un facteur deux plutôt que de réduire le nombre de photos d'échantillons issus de lots à 30% de briques. Le nombre de photographies de la base rééquilibrée par pourcentage de briques est présenté au Tableau 2.6.

Tableau 2.6 Nombre de photographies de sol contenant 30%, 20% et 50% de briques dans la base de photos originale et rééquilibrée, facteur de rééquilibrage

	Base originale	Facteur de rééquilibrage	Base rééquilibrée
Nombre de photographies à 30% de briques	1280		1280
Nombre de photographies à 20% de briques	331	×3	993
Nombre de photographies à 50% de briques	461	×2	922
TOTAL	2072		3195

La base de données rééquilibrée contient 3195 photographies au total. Les opérations utilisées pour augmenter le nombre de photographies avec 20 et 50 % de brique sont des basculements par rapport à un axe ou *flip* en anglais. Ces opérations garantissent des formats homogènes dans la base de données, ce qui n'est pas été le cas avec un zoom ou une rotation des photographies. Pour respecter les facteurs de rééquilibrage, les photographies n°1281 à 2072 (correspondant à des sols à 20 et 50% de briques) de la base originale dont la structure est présentée en Tableau 2.2 ont été basculées selon l'axe de la plus petite dimension des images dans un premier temps. Dans un second temps, les photographies n°1536 à 1727, 1805 à 1872 et 1938 à 2008 (correspondant seulement à des sols à 20% de briques) ont été basculées selon l'axe de la plus grande dimension des images. Ces opérations ont été effectuées à l'aide d'un script Python. Ainsi, la structure de la nouvelle base de données correspond à celle de la base originale présentée en Tableau 2.2 à laquelle est ajouté le complément présenté en Tableau 2.7.

Tableau 2.7 Description de la base de données photographique de sol avec briques complémentaire constituante de la base de données rééquilibrée

Référence	Opération	%briques du lot	Méthode d'échantillonnage	Particules fines	Phase liquide	N° des photos (haut)
(Zeryouh, 2023)	<i>Flip selon x</i>	50%	Grappillage	Sans fines	6% d'eau	2073 à 2136
		50%	Grappillage	Sans fines	10% d'eau	2137 à 2200
		50%	Pelletage fractionné	Sans fines	6% d'eau	2201 à 2264
		50%	Pelletage fractionné	Sans fines	10% d'eau	2265 à 2327
		20%	Grappillage	Sans fines	6% d'eau	2328 à 2391
		20%	Pelletage fractionné	Sans fines	6% d'eau	2392 à 2454
(Dol, 2023)		20%	Grappillage	Sans fines	Sec	2455 à 2519
		50%	Grappillage	Sans fines	Sec	2520 à 2596
		20%	Pelletage fractionné	Sans fines	Sec	2597 à 2664
		50%	Pelletage fractionné	Sans fines	Sec	2665 à 2729
		20%	Séparation à chutes	Sans fines	Sec	2730 à 2800
		50%	Séparation à chutes	Sans fines	Sec	2801 à 2864
(Zeryouh, 2023)	<i>Flip selon y</i>	20%	Grappillage	Sans fines	6% d'eau	2865 à 2928
		20%	Pelletage fractionné	Sans fines	6% d'eau	2929 à 2991
(Dol, 2023)		20%	Grappillage	Sans fines	Sec	2992 à 3056
		20%	Pelletage fractionné	Sans fines	Sec	3057 à 3124
		20%	Séparation à chutes	Sans fines	Sec	3125 à 3195

2.3.1.3 Redimensionnement préalable des photographies (512×810 pixels)

Afin d'être en mesure d'entraîner le modèle avec les 3195 photographies de la base de données rééquilibrée, ces dernières ont été redimensionnées en des images RGB de 512×810 pixels. Ces dimensions ont été préférées à un redimensionnement carré comme appliqué lors de projets antérieurs (Dol, 2023; Zeryouh, 2023) car elles permettent de conserver les proportions des images originales qui sont de taille 6048×3824 pixels. Cela permet ainsi de limiter la perte d'information des images dans la direction x du plan cartésien par rapport à y .

Le redimensionnement permet d'une part d'uniformiser les tailles d'images qui pouvaient s'avérer différentes selon les éventuels rognages réalisés pour chaque photographie. Cette uniformisation est nécessaire, car le réseau est seulement capable de traiter des photographies de mêmes dimensions. Cette opération était également nécessaire d'autre part pour des raisons

de stockage mémoire et de temps de calcul. En effet, la base de données complète de 3195 avait une taille de 1,1 Go, auxquels s'ajoutent le stockage mobilisé par les opérations effectuées lors de la création, l'entraînement, la validation et le test du réseau.

2.3.2 Description du CNN

Le CNN créé dans le cadre de ce travail a été implémenté en Python en utilisant la bibliothèque TensorFlow.

2.3.2.1 Données d'entraînement, de validation et de test

Dans un premier temps, les 2072 photographies pour les simulations avec la base originale, et les 3195 photographies pour les simulations avec la base rééquilibrée, ont été réparties aléatoirement dans des dossiers d'entraînement, de validation et de test. La proportion de photographies utilisées pour l'entraînement est de 80% contre 10% pour la validation et 10% pour le test ce qui correspond à 2557 photographies d'entraînement, 319 photographies de validation et 319 photographies de test pour la base rééquilibrée.

Un script Python a permis dans un premier temps de répartir les photographies aléatoirement dans ces trois dossiers. La fonction « *sample* » qui provient de la bibliothèque Python appelée « *Random* » a été utilisée pour la sélection aléatoire tandis que la fonction « *rename* » de la bibliothèque Python « *os* » a déplacé les sélections aléatoires dans les trois dossiers.

Dans un second temps, un autre script a permis d'extraire les pourcentages de briques associés dans trois fichiers texte à partir d'un fichier texte, contenant tous les pourcentages, créé à partir de la base de données du Tableau 2.2. Le fichier texte de tous les pourcentages de briques a été converti en tableau à l'aide de la fonction « *fromstring* » de la bibliothèque Python « *Numpy* », la manipulation de tableaux étant plus simple que la manipulation de listes sous Python. Une fois répartis dans trois tableaux associés aux photographies d'entraînement, de validation et de test à l'aide d'une simple boucle *for*, les tableaux ont été reconvertis en liste avec la fonction « *tolist* ».

Une fois les dossiers de photographies et les fichiers texte associés d'entraînement, de validation et de test créés, ceux-ci restent inchangés durant tout le travail sur le réseau neuronal. Pour pouvoir les utiliser dans le CNN, ils sont simplement chargés en amont du script de ce dernier. Les fichiers texte sont chargés à l'aide de la fonction « *loadtxt* » de la bibliothèque « *numpy* » puis convertis en tenseurs avec la fonction « *convert_to_tensor* » de *TensorFlow*. Pour les photographies, trois matrices de zéros pouvant accueillir toutes les photographies associées à l'entraînement, à la validation et au test sont créées avec la fonction « *zeros* » de *Numpy*, puis remplies avec les photographies à l'aide d'une boucle *for*.

2.3.2.2 Description du modèle

Le modèle créé est un modèle séquentiel c'est-à-dire qu'il est constitué d'une pile linéaire de plusieurs couches. Plusieurs structures de réseau ont été essayées. Pour élaborer la structure de réseau la plus optimale, une structure similaire à celle élaborée lors du projet de Dol (2023) a été essayée en premier lieu puis modifiée et ajustée selon les performances (définies aux parties 2.4.4.1 Performance de l'entraînement et 2.4.4.2 Performance du test) qu'elle permettait d'obtenir avec la base de données de cette recherche (Tableau 2.7). La structure d'origine était donc un réseau neuronal convolutif de régression avec deux duos de couches de convolution et de sous-échantillonnage (*maxpooling*) suivies des couches de sortie classiques. Celles-ci ne permettant pas aux modèles générés de converger (c'est-à-dire, d'estimer des pourcentages de briques autres que constants et égaux à la moyenne des données d'entraînement), la complexité a dû être progressivement augmentée. Tout d'abord, le nombre de filtres de convolution a été progressivement augmenté, et comme cela ne permettait toujours pas au modèle de converger, il a été choisi d'ajouter un duo de couches de convolution et de *maxpooling*. De plus, le réseau neuronal GRAINet (Lang et al., 2021) élaboré pour l'analyse granulométrique de sédiments est composé de trois couches de convolution également. L'enchaînement de couches typique ayant permis aux modèles de converger était structuré comme suit :

- 1) **Une couche de normalisation des pixels d'entrée.** Pour une image, la valeur de chacun de ses pixels est divisée par 225 afin de leur attribuer une valeur comprise entre 0 et 1 ;
- 2) **Une couche de convolution 2D** (car les images correspondent à trois matrices de deux dimensions pour ses trois canaux RVB) qui, utilise quatre ou huit filtres de convolution de taille 4×4 ou 3×3 pixels qui se déplaceront d'un pas de 2×2 pixels sur l'image pour l'entraînement du modèle. Bien que dans la littérature récente, des filtres 3×3 sont régulièrement employés (Lang et al., 2021), des filtres de taille 4×4 pixels ont donné de meilleurs résultats dans cette recherche. Chaque filtre tire des caractéristiques de l'image. Ensuite, pour intégrer de la non-linéarité, la fonction d'activation « ReLu », notamment utilisée dans le réseau GRAINet, est appliquée à chaque neurone en sortie de couche ;
- 3) **Une couche *maxpooling* 2D** est ensuite intégrée pour tirer les caractéristiques les plus importantes de l'image. Dans le cas du réseau développé dans cette recherche, il s'agit des valeurs maximales tous les 2×2 pixels. Cette étape consiste également à réduire la taille de la sortie de la couche précédente d'un facteur 2 ;
- 4) **Une deuxième couche de convolution 2D** qui utilise cette fois-ci deux fois plus de filtres que la couche de convolution 2D précédente pour augmenter la complexité du modèle en lui permettant ainsi de mettre en lumière des motifs plus complexes et moins évidents à voir au premier abord ;
- 5) **Une deuxième couche *maxpooling* 2D** permettant d'extraire, une fois de plus, les caractéristiques maximales tous les 2×2 pixels et ainsi réduire l'entrée d'un facteur deux. Rajouter une couche de convolution ainsi que de *maxpooling* permet d'extraire des caractéristiques plus complexes et abstraites. Cela permet d'effectuer un niveau d'analyse de l'image plus poussé ;
- 6) **Une troisième et dernière couche de convolution 2D** qui utilise une fois de plus, deux fois plus de filtres que la couche de convolution 2D précédente pour les mêmes raisons. En effet, le modèle ne semblait pas converger avec seulement deux jeux de couches de convolution et de *maxpooling* combinées quelles que soient les augmentations du nombre de filtres essayées pour augmenter la complexité du modèle. L'ajout d'un quatrième jeu de couches de convolution et de *maxpooling* a été essayé mais a généré du surentraînement (défini en 2.4.4) ;

- 7) **Une troisième couche *maxpooling 2D*** permettant d'extraire les caractéristiques maximales tous les 4×4 pixels dans ce dernier cas, ce qui permet de réduire l'entrée d'un facteur quatre ;
- 8) **Une couche d'aplatissement dite *flatten*** afin de mettre toutes les caractéristiques issues des étapes précédentes en vecteur ligne. Cette opération est nécessaire pour la fonction de la couche suivante qui ne prend en argument que des vecteurs lignes.
- 9) **Une couche de régression dite *dense*** pour l'apprentissage d'informations complexes. Elle est constituée dans le cas présent de dix neurones et une fonction d'activation sigmoïde.
- 10) **Une couche de *Dropout* (éventuelle)** visant à limiter le surajustement qui a été un problème parfois lors des simulations effectuées. Ceci est permis par la désactivation d'un certain pourcentage de neurones aléatoirement lors des différentes étapes d'entraînement. Le pourcentage de désactivation le plus concluant dans les quelques situations de surajustement a été 30%.
- 11) **Une autre couche de régression dite *dense*** différente de la première puisqu'elle sert à ajuster les données de sorties de sorte à pouvoir faire des régressions. Elle est constituée d'un neurone et d'une fonction d'activation sigmoïde. La fonction d'activation linéaire a également été essayée mais n'a jamais permis au modèle de converger et générait généralement du sous-apprentissage (défini en 2.4.4).

2.3.2.3 Compilation et entraînement du modèle

Le modèle est ensuite compilé à l'aide de la fonction « *compile* » et la fonction de perte, l'optimiseur, et les métriques à surveiller sont précisés. Il s'agit respectivement de l'erreur quadratique moyenne pour la régression (MSE), de l'optimiseur Adam avec lequel est paramétré un des hyperparamètres qui est le taux d'apprentissage, et la racine carrée de l'erreur quadratique moyenne (RMSE) calculé à partir des pourcentages prédits et mesurés.

L'entraînement est effectué à l'aide de la fonction « *model.fit* » de la bibliothèque « *TensorFlow* » prenant en argument les données d'entraînement et de validation d'entrée et

de sortie en plus de deux hyperparamètres que sont le nombre d'époques d'entraînement et la taille du lot dite *batch size*.

Les hyperparamètres, comprenant le nombre d'époques ou *epoch*, la taille du lot ou du *batch* et le taux d'apprentissage, sont définis à ce stade du code donc. Le taux d'apprentissage définit la vitesse à laquelle le réseau apprend en jouant sur les poids du modèle d'une époque à l'autre. Le nombre d'époques est le nombre de fois où le réseau passe par toutes les données d'entraînement tandis que la taille de l'échantillon définit essentiellement la vitesse d'ajustement des poids (plus il est grand, plus ça sera rapide). Ces hyperparamètres ont été ajustés en fonction des différentes configurations de couches testées, dans le but de trouver les combinaisons d'hyperparamètres conduisant aux résultats les plus probants pour chaque modèle. Le nombre d'époques n'ayant pas grande influence sur le modèle, ce n'est pas le cas du taux d'apprentissage et de la taille du *batch*. Un taux d'apprentissage à 0,01 génèrait beaucoup d'instabilité qui se manifestait par des fluctuations des pertes d'entraînement et de validation au cours des époques d'entraînement sur les graphiques de performances générés à cette étape. Un taux d'apprentissage à 0,0001 permettait une convergence très stable et lisse du modèle au cours des époques d'entraînement, mais bien trop lente et donc pas suffisante. La taille du *batch* garantit une bonne convergence du modèle quand elle est optimale. Plus elle est grande, plus elle réduit de la durée de simulation mais elle nécessite plus de capacité en mémoire et peut générer du surapprentissage. Les valeurs qui ont généré les meilleurs résultats étaient généralement de 15 pour le nombre d'époques, de 0,001 pour le taux d'apprentissage, et de 16 ou 32 pour la taille du *batch* selon les configurations.

En résumé, le modèle ainsi créé est un réseau de régression, capable de prendre en entrée des images d'échantillons de taille 512×810 pixels et de prédire les % de briques qui s'y trouvent. Trois configurations de couches et d'hyperparamètres ont été retenues sur les dizaines d'essayées. Les scripts associés sont présentés en ANNEXE II.

2.4 Interprétation des résultats

2.4.1 Résultats granulométriques

L'interprétation des résultats granulométriques s'est déroulée en trois étapes. Tout d'abord, des courbes granulométriques ont été tracées, à partir desquelles des courbes d'erreur ont ensuite été générées. Enfin, une analyse statistique a complété l'étude des variances, rendue possible grâce aux courbes d'erreur.

Les méthodologies pour tracer les courbes granulométriques, les courbes d'erreur, et réaliser l'analyse statistique sont respectivement décrites en 2.4.1.1., 2.4.1.2 et 2.4.3.

2.4.1.1 Courbes granulométriques

Pour chaque échantillon produit comme décrit en 2.2.1, « le calcul et l'expression globale de l'analyse granulométrique » est effectué selon la dixième section du même nom de la norme BNQ 2501-025/2013.

Pour ce faire, le % retenu de chaque tamis est dans un premier calculé en divisant la masse qui y est retenue par la masse totale de l'échantillon avant lavage :

$$\%retenu_{tamis\ n} = \frac{m_{retenu_{tamis\ n}}}{m_{totale\ avant\ lavage}} \times 100 \quad (2.2)$$

Il est à noter que la mention « tamis n » est la notation associée à « tamis de taille n » et que $n \in (2,5\text{ mm} ; 1,25\text{ mm} ; 630\ \mu\text{m} ; 315\ \mu\text{m} ; 60\ \mu\text{m} ; 80\ \mu\text{m})$.

Le retenu cumulé est ensuite calculé. Ce dernier correspond pour chaque tamis de taille n, à la part totale de l'échantillon qui y serait retenu s'il n'y avait aucun autre tamis et est calculé selon la formule ci-dessous :

$$\%retenu\ cumulé_{tamis\ n} = \sum_{i=1}^n \%retenu_{tamis\ i} \quad (2.3)$$

Enfin, le pourcentage total passant de chaque fraction granulométrique peut être calculé en soustrayant le pourcentage retenu cumulé à 100% :

$$\%passant_{tamis\ n} = 100 - \%retenu\ cumulé_{tamis\ n} \quad (2.4)$$

C'est avec le % total passant qu'une courbe granulométrique peut être tracée sur un graphique semi logarithmique. En effet, il s'agit d'un graphique représentant le $\%passant_{tamis\ n}$ en fonction du $\log_{10}(tamis\ n)$ pour tout n (exemple en 3.1.1).

Afin de s'assurer de la validité de l'analyse granulométrique, il est vérifié que la masse totale après analyse granulométrique ne varie pas de la masse totale initiale de l'échantillon (c'est-à-dire avant lavage) de plus de 0,3% (point 7.3.4 de la norme BNQ 2501-025/2013).

2.4.1.2 Courbes d'erreur

Afin de comparer les résultats d'analyse granulométrique entre les différents essais, des courbes dites d'erreur sont calculées et tracées pour chaque essai.

Pour calculer les valeurs de la courbe d'erreur, la granulométrie moyenne (correspondant aux % passants moyens des différents tamis utilisés) des 32 échantillons est déterminée dans un premier temps, de même pour les écarts-types.

$$\mu\%passant_{tamis\ n} = \frac{\sum_{i=1}^{32} \%passant_{tamis\ n\ ech\ i}}{32} \quad (2.5)$$

$$\sigma\%passant_{tamis\ n} = \sqrt{\frac{\sum_{i=1}^{32} (\%passant_{tamis\ n\ ech\ i} - \mu\%passant_{tamis\ n})^2}{31}} \quad (2.6)$$

Les courbes d'erreurs d'un essai sont au nombre de deux et correspondent à la granulométrie moyenne moins trois fois l'écart-type et plus trois fois l'écart-type. L'écart entre ces courbes constitue l'intervalle de confiance à 99,7% puisque les données suivent globalement une loi normale (2.4.1). Elles se calculent selon la formule suivante :

Pour tout n :

$$\text{courbe d'erreur}^- = \mu\%passant_{tamis\ n} - 3 \times \sigma\%passant_{tamis\ n} \quad (2.7)$$

$$\text{courbe d'erreur}^+ = \mu\%passant_{tamis\ n} + 3 \times \sigma\%passant_{tamis\ n}$$

Juxtaposer les courbes d'erreur de différents essais sur un même graphique permet de comparer les variabilités granulométriques des échantillons produits entre ces différents essais, la granulométrie moyenne étant quasiment identiques pour tous les essais (2.1). Plus l'écart entre les courbes d'erreurs et la courbe granulométrique moyenne est faible, plus l'essai d'échantillonnage peut être considéré comme représentatif d'un point de vue granulométrique. Au contraire, plus cet écart est grand, moins l'essai peut être considéré comme représentatif. Ainsi, en juxtaposant les courbes d'erreur adéquates, il est possible de comparer différentes influences comme l'effet des différentes méthodes d'échantillonnage, ou de la présence d'une phase liquide, ou encore l'effet de la présence ou pas de particules fines sur la représentativité granulométrique.

2.4.2 Résultats de pourcentage de briques

L'interprétation des résultats de pourcentages de brique a été notamment rendue possible par l'analyse de variances relatives, la comparaison de coefficients de variation théoriques et empiriques, ainsi que par l'analyse statistique des variances qui a permis de compléter l'étude des éléments de statistiques descriptives.

Les méthodologies pour calculer la variance relative et autres grandeurs associées, pour calculer les coefficients de variation empiriques et théoriques, ainsi que pour réaliser l'analyse statistique sont respectivement décrites en 2.4.2.1, 0 et 2.4.3.

2.4.2.1 Étude des biais relatifs, variances relatives et degrés de représentativité

Le biais relatif se calcule à partir de l'erreur relative, elle-même calculée par échantillon n numéroté de 1 à 32 dans cette recherche, selon la formule (2.8) adaptée de la formule (1.8).

$$\varepsilon_{relative\ éch\ n} = \frac{\%briques_{éch\ n} - \%briques_{lot}}{\%briques_{lot}} \quad (2.8)$$

Le biais relatif m^2 et le degré de représentativité r^2 se calculent respectivement à partir des formules (1.7) et (1.9).

Le biais relatif étant une mesure de l'écart moyen à la moyenne. Puisque dans le cas des essais de cette recherche, l'entièreté des lots est systématiquement échantillonnée, les biais relatifs sont très faibles et non indicateurs de la représentativité. Les variances relatives, et par conséquent les degrés de représentativité, sont en revanche plus à même de donner une idée de la représentativité des essais.

2.4.2.2 Comparaison des coefficients de variations empiriques et théoriques

Les coefficients de variation théoriques ($CV_{\text{théo}}$) se calculent à partir de la variance relative due à l'erreur totale d'échantillonnage, cette dernière étant la somme des variances relatives dues aux différentes erreurs décrites à la section 1.2.2. Cependant, comme il n'est pas possible d'estimer a priori la variance due aux erreurs de délimitation, d'extraction et de préparation, le $CV_{\text{théo}}$ se calcule à partir de la somme des variances relatives dues aux erreurs fondamentales et de groupement et ségrégation. En ce sens, le $CV_{\text{théo}}$ correspond au coefficient de variation minimal si toutes les sources d'erreur, autres que ces deux dernières, sont éliminées. Le coefficient de variation s'exprime donc comme suit :

$$CV_{\text{théo}} \geq \sqrt{s^2(\text{ETE})} \times 100 = \sqrt{s^2(\text{EFE}) + s^2(\text{ESG})} \times 100 \quad (2.9)$$

Avec :

- $s^2(\text{ETE})$: la variance générée par l'erreur totale d'échantillonnage (1.2) ;
- $s^2(\text{EFE})$: la variance générée par l'erreur fondamentale d'échantillonnage (1.4) ;
- $s^2(\text{ESG})$: la variance générée par l'erreur de ségrégation et groupement (1.3).

Les coefficients de variation empiriques (CV_{emp}) sont calculés par essai à partir de la formule suivante :

$$CV_{\text{emp}}(\%) = \frac{\sigma}{\mu} \times 100 \quad (2.10)$$

Avec :

- CV_{emp} : le coefficient de variation empiriques de la série des pourcentages de briques, obtenue pour l'essai n (%) ;
- σ : l'écart type de la série des pourcentages de briques, obtenue pour l'essai n (%) ;
- μ : la moyenne de la série des pourcentages de briques, obtenue pour l'essai n (%) ;

Ces coefficients de variation empiriques ont ensuite pu être comparés aux coefficients de variation théoriques. Cette comparaison a permis d'établir l'écart de variabilité causé par les sources d'erreurs mal ou non contrôlées associées à chaque méthode d'échantillonnage ou à la présence d'une phase liquide (eau ou huile).

2.4.3 Analyse statistique des résultats granulométriques et de pourcentages de brique

L'ensemble de l'analyse statistique a été réalisée à l'aide du logiciel Minitab (Ryan, 2020). Au cours de l'analyse statistique des jeux de données granulométriques et de pourcentages de briques, des tests d'égalité de la variance ont été effectués afin d'évaluer la fiabilité et la robustesse des jeux de données ainsi que de les comparer entre eux. Pour juger de la normalité ou non des données, le test d'Anderson-Darling est réalisé en amont pour chaque jeu de données granulométriques et de pourcentages de briques.

Pour les jeux de données granulométriques, les tests de normalités ont donc été exécutés sur les jeux de données associés à chaque fraction granulométrique (des tamis utilisés pour l'analyse granulométrique) soit un total de $6 \times 7 = 42$ fois pour les essais avec fines (associés aux passants 2,5 mm, 1,25 mm, 630 μm , 315 μm , 160 μm et 80 μm pour chacun des sept essais avec fines) et de $4 \times 3 = 12$ fois pour les essais sans fines (associés aux passants 2,5 mm, 1,25 mm, 630 μm et 315 μm pour chacun des trois essais sans fines). Pour les jeux de données de pourcentages de briques, le test de normalité est exécuté dix fois, une fois pour chaque jeu de données associé à chacun des dix essais.

À l'occasion de chaque test de normalité, deux graphiques sont affichés à l'aide du logiciel Minitab : un graphique de probabilité normale ainsi qu'un graphique avec ajustement normal pour juger de la normalité des jeux de données étudiés. Parmi les jeux donnés granulométriques, 20% ne respectaient pas l'hypothèse de normalité contrairement aux jeux de données de pourcentages de briques qui suivaient globalement tous une distribution normale.

Ensuite, des tests d'égalité des variances appropriés ont été choisis pour les deux types de jeux de données. Le test de Bartlett est courant pour vérifier l'homogénéité ou non des variances mais est sensible à la violation de l'hypothèse de normalité contrairement au test non paramétrique de Levene, courant également (Ramos-Guajardo & Lubiano, 2012). Ainsi, le test de Levene semblait être plus adapté pour les jeux de données granulométriques tandis que le test de Bartlett était le plus adapté aux jeux de données de pourcentages de briques.

Pour les jeux de données granulométriques, les tests de normalités ont donc été exécutés sur les jeux de données associés à chaque fraction granulométrique (des tamis utilisés pour l'analyse granulométrique) soit un total de six fois pour les essais avec fines (associés aux passants 2,5 mm, 1,25 mm, 630 μm , 315 μm , 160 μm et 80 μm des sept essais avec fines) et de quatre fois pour les essais sans fines (associés aux passants 2,5 mm, 1,25 mm, 630 μm et 315 μm des trois essais sans fines). Pour les jeux de données de pourcentages de briques, le test de normalité est exécuté deux fois, une fois pour les jeux de données des essais avec fines et une autre fois pour les jeux de données des essais sans fines.

L'hypothèse nulle et alternative ainsi que le seuil de signification sont les mêmes pour les deux tests d'égalité des variances :

Tableau 2.8 Hypothèses nulle et alternative et seuil de signification des tests d'égalité des variances

Hypothèse nulle H_0	Toutes les variances sont égales
Hypothèse alternative	Au moins une variance est différente
Seuil de signification	$\alpha = 0,05$

Si :

- p-valeur $< \alpha$, alors H_0 est rejetée c'est-à-dire qu'il y a des preuves statistiques d'une différence significative des variances entre les groupes ;
- p-valeur $> \alpha$, alors H_0 ne peut être rejetée et les variances des groupes sont considérées comme statistiquement similaires.

Les graphiques générés à l'occasion des deux types de test d'égalité des variances sont fondamentalement les mêmes avec en amont des tests, le traçage d'un graphique en boîte à moustaches pour permettre une étude descriptive des variances puis un graphique des intervalles de comparaisons multiples pour l'écart-type. Le graphique des intervalles de comparaisons multiples présente lui aussi les intervalles de Bonferroni qui permettent de juger si les écarts types des pourcentages de briques associés aux dix configurations d'essais sont significativement différents. En effet, les intervalles de Bonferroni représentent les intervalles de confiance à 95% pour les écarts types de chaque groupe (associé aux configurations des différents essais) testé. Ainsi, si les intervalles de confiance de deux groupes se chevauchent, cela suggère que les écarts types de ces groupes ne sont pas statistiquement différents avec un niveau de confiance à 95%. Si les intervalles de confiance ne se chevauchent pas, cela indique que les écarts types de ces groupes sont statistiquement différents avec le niveau de confiance spécifié.

2.4.4 Analyse d'images

Afin d'évaluer la performance du modèle créé, entraîné, validé et testé, des analyses de performance sont implémentées à différentes étapes du code. Ces analyses permettent toutes de mettre en parallèle les résultats de pourcentage de briques obtenus au cours de l'analyse physique avec les pourcentages prédits par le modèle. Ainsi, grâce aux éléments décrits dans cette section, la lecture croisée de tous les résultats obtenus au cours des analyses physiques et d'images de la recherche sera permise.

2.4.4.1 Performance de l'entraînement

Premièrement, un graphique de performance de l'entraînement est tracé après l'étape d'entraînement. Ce graphique met en évidence deux courbes qui correspondent respectivement à l'évolution de la perte d'entraînement et de validation au cours des différentes époques d'entraînement. Une époque correspond à un passage du modèle sur toutes les données d'entraînement et la perte correspond à l'écart entre les données prédites par le modèle et les données réelles. La fonction utilisée pour répertorier ces pertes est la fonction *history* de la bibliothèque *TensorFlow*.

Ce graphique permet de vérifier deux éléments : le premier étant que le modèle converge bien ; le deuxième étant que le modèle ne soit pas surentraîné ou sous-entraîné. Le sous entraînement se manifeste lorsque les courbes de pertes ne diminuent pas assez voire pas du tout, et le surentraînement se manifeste sur le graphique de performance de l'entraînement lorsque la courbe de perte liée à la validation passe au-dessus de celle liée à l'entraînement (Goodfellow et al., 2016). Le sous entraînement et le surentraînement peuvent être dus à un choix non-approprié de fonction d'activation, des hyperparamètres (nombre d'époques, taux d'apprentissage et taille du lot dit *batch size*) (Goodfellow et al., 2016) mais également à la complexité du modèle (nombre de couches de convolution et *pooling*, nombre et taille des filtres de convolution).

2.4.4.2 Performance du test

Ensuite, après l'étape de prédiction des données tests par le modèle, un graphique de régression est tracé. A la manière du graphique d'entraînement, ce graphique sert à visualiser l'écart entre les pourcentages de briques prédits par le modèle et les pourcentages de briques réels. Cette fois, la visualisation se fait en traçant le nuage de points des pourcentages prédits par rapport aux pourcentages réels. Ce graphique est accompagné d'une droite de régression $y = x$ définissant une correspondance parfaite. Plus le nuage de points suit cette droite, plus le modèle peut être considéré comme performant sur la prédiction du jeu de données test.

Pour compléter l'étude visuelle permise par graphique de régression, la fonction Python « *evaluate* » du module Keras de la bibliothèque *TensorFlow* est appliquée au modèle. Elle permet d'en calculer des indicateurs de performances tels que la perte et la racine carrée de l'erreur quadratique moyenne (RMSE).

CHAPITRE 3

RÉSULTATS

3.1 Analyse granulométrique

3.1.1 Granulométrie générale

Cette partie présente les granulométries du sol seul et du mélange de sol et briques utilisés pour constituer les lots avec et sans fines. Afin de garantir d'estimer au mieux ces granulométries, ces dernières ont été réalisées à la fin des expériences en sommant les granulométries obtenues pour tous échantillons avec et sans fines obtenus lors des dix essais de cette recherche.

Pour réaliser les granulométries du sol seul (sans briques) avec et sans fines, les masses de briques ont été retranchées des masses de sol et briques. Cependant, puisque la brique n'a été quantifiée que dans les fractions 5-2,5 mm et 2,5-1,25 mm, il est possible que les pourcentages passants des fractions inférieures soient légèrement surestimés dans les granulométries de sol seul.

Le Tableau 3.1 présente la granulométrie du lot de sol seul avec et sans fines, pour lesquelles des courbes granulométriques ont été tracées en Figure 3.1 et Figure 3.2, respectivement.

Tableau 3.1 Granulométrie du lot de sol avec et sans fines

Tamis	%Passant du lot de sol avec fines	%passant du lot de sol sans fines
5 mm	100%	100%
2,5 mm	84,01%	80,72%
1,25 mm	64,14%	57,60%
630 µm	42,70%	32,19%
315 µm	26,09%	12,32%
160 µm	16,51%	0,73%
80 µm	11,06%	0,20%

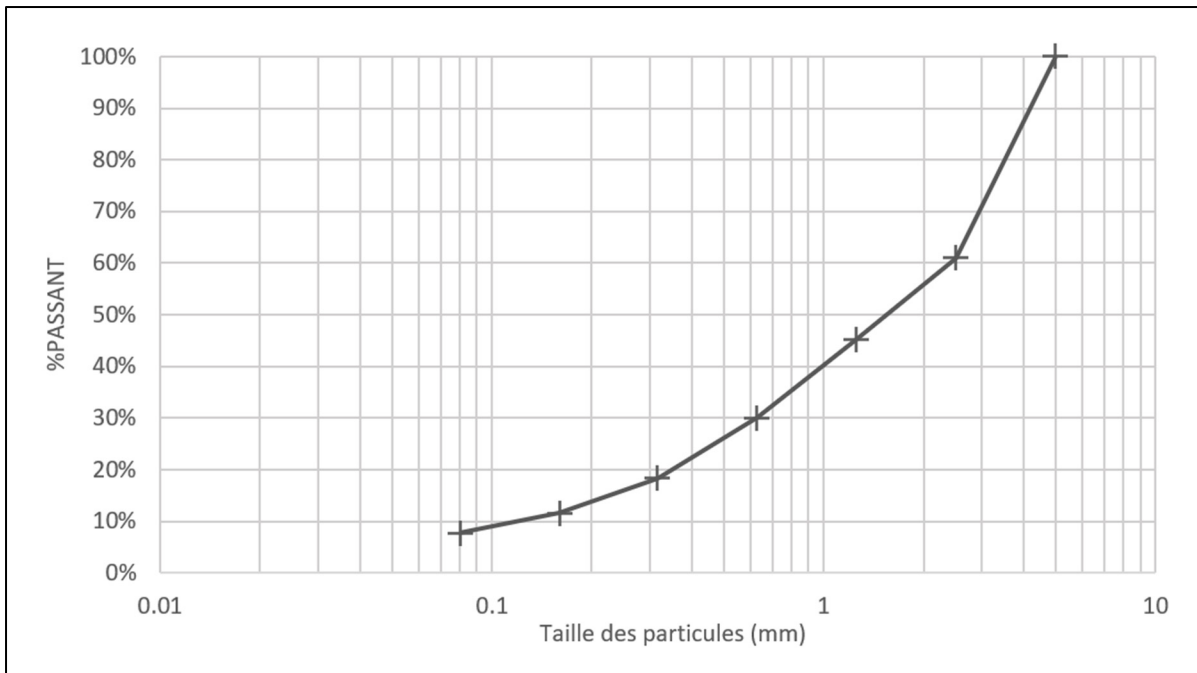


Figure 3.1 Courbe granulométrique du lot de sol avec fines

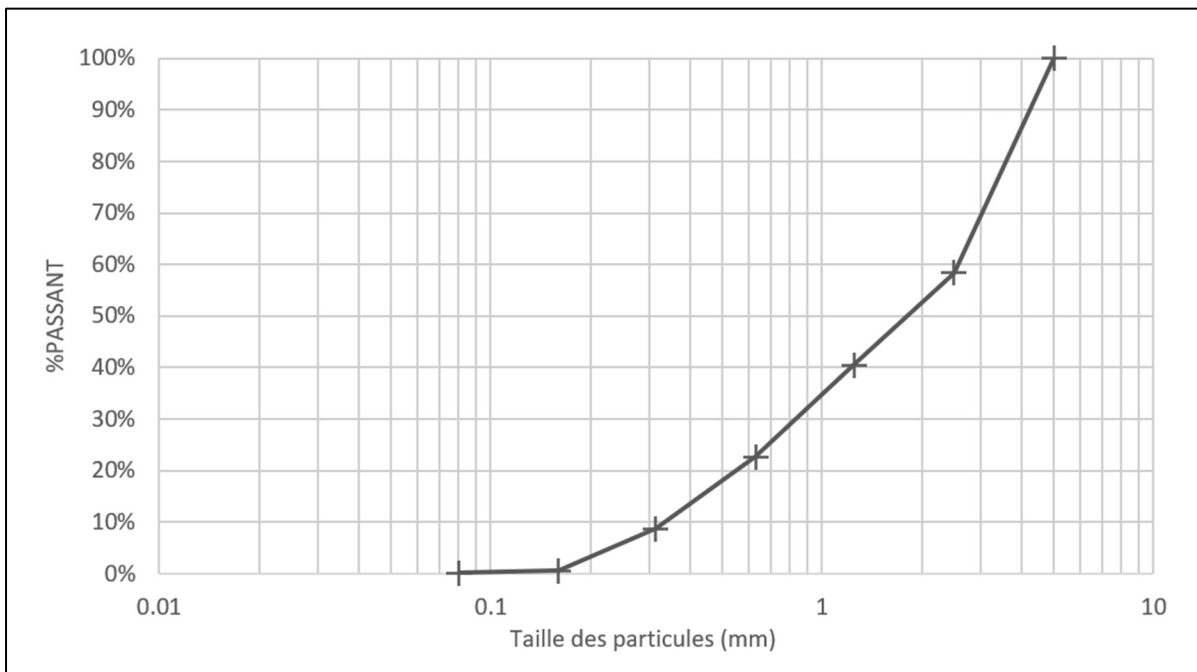


Figure 3.2 Courbe granulométrique du lot de sol sans fines

D'après le tableau 1 de la classification U.S.C.S (Portmann, 2023), les sol dits avec et sans fines contenant principalement des particules inférieures à 2 mm d'après les D50

(respectivement de 0,8 mm et 1 mm) et moins de 5% de particules inférieures à 0,06 mm, il s'agit de sable bien étalé classé SW.

Le Tableau 3.2 présente la granulométrie du lot de sol avec et sans fines avec particules de briques, pour lesquelles des courbes granulométriques ont été tracées en Figure 3.3 et Figure 3.4 (respectivement).

Tableau 3.2 Granulométrie du lot de sol et briques

Tamis	%Passant du lot de sol avec fines et briques	%passant du lot de sol sans fines et briques
5 mm	100,00%	100,00%
2,5 mm	61,07%	58,37%
1,25 mm	45,15%	40,45%
630 µm	30,04%	22,64%
315 µm	18,34%	8,71%
160 µm	11,60%	0,52%
80 µm	7,77%	0,14%

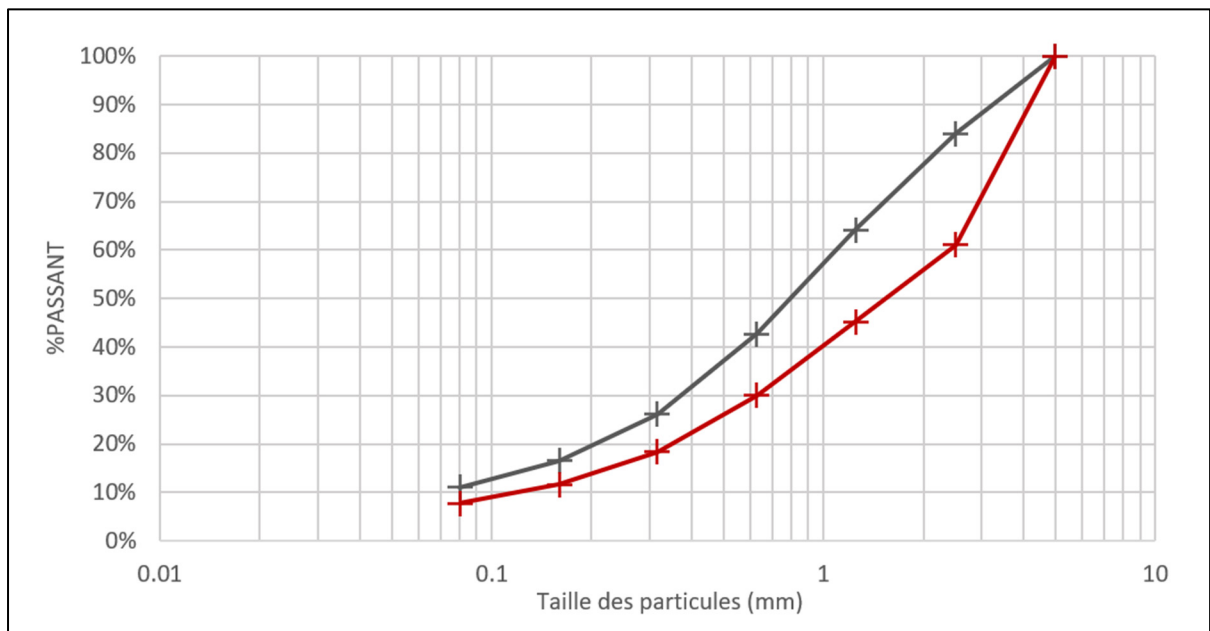


Figure 3.3 Courbe granulométrique du lot de sol avec fines sans briques (gris) et 30% de briques (rouge)

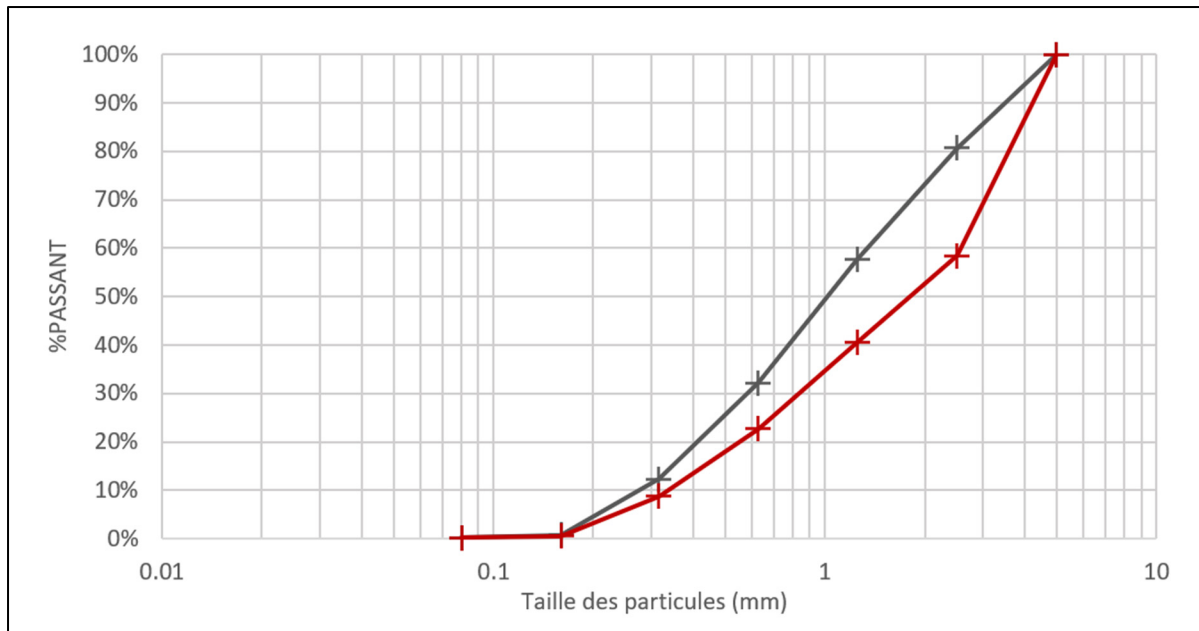


Figure 3.4 Courbe granulométrique du lot de sol avec fines sans briques (gris) et 30% de briques (rouge)

L'ajout de brique particulière a augmenté de 30% la masse totale principalement dans la fraction 5-2,5 mm et légèrement dans la fraction 1,25-2,5 mm. Par conséquent, le diamètre médian a été augmenté par rapport aux courbes précédentes et est passé de 0,8 mm à 1,11 mm pour le lot avec fines et de 1 mm à 1,57 mm pour le lot sans fines. La classe U.S.C.S (Portmann, 2023) reste inchangée par rapport au sol seul : il s'agit toujours de sable bien étalé classé SW.

3.1.2 Analyse descriptive des variances

L'étude de la représentativité (1.2.1) des analyses granulométriques se fait en étudiant l'effet de la présence de fines, puis de la présence d'une phase liquide (eau et huile) et enfin, de la méthode d'échantillonnage employée.

Dans un premier temps, les courbes granulométriques d'erreur associées sont tracées ainsi que les graphiques en boîte à moustaches par essai. Dans un second temps, les différences relevées au cours de cette première analyse descriptive sont étudiées plus précisément d'un point de vue statistique. Pour cette analyse, des graphiques de comparaison d'intervalles de confiances

à 95% sur l'écart-type générés grâce à des tests d'équivalence des variances sont tracés. Ces derniers types de graphiques permettent d'étudier si les écart-types, et donc les variances, sont statistiquement différents entre les configurations d'essai.

Des graphiques en boîte à moustaches sont tracés pour les données granulométriques, d'une part des essais avec fines, et d'autre part des essais sans fines. Six graphiques associés aux pourcentages passants des six fractions granulométriques sont donc tracés pour les pourcentages passants des essais de sol avec fines en Figure 3.5, et quatre autres pour les pourcentages passants des quatre fractions granulométriques les essais de sol sans fines en Figure 3.6. Combinés aux courbes d'erreur, ces graphiques permettent de décrire les effets de la présence de fines, de liquide et de la méthode d'échantillonnage employée sur la représentativité (1.2.1) des essais échantillonnages.

Quels que soient les graphiques, les astérisques représentent des valeurs aberrantes.

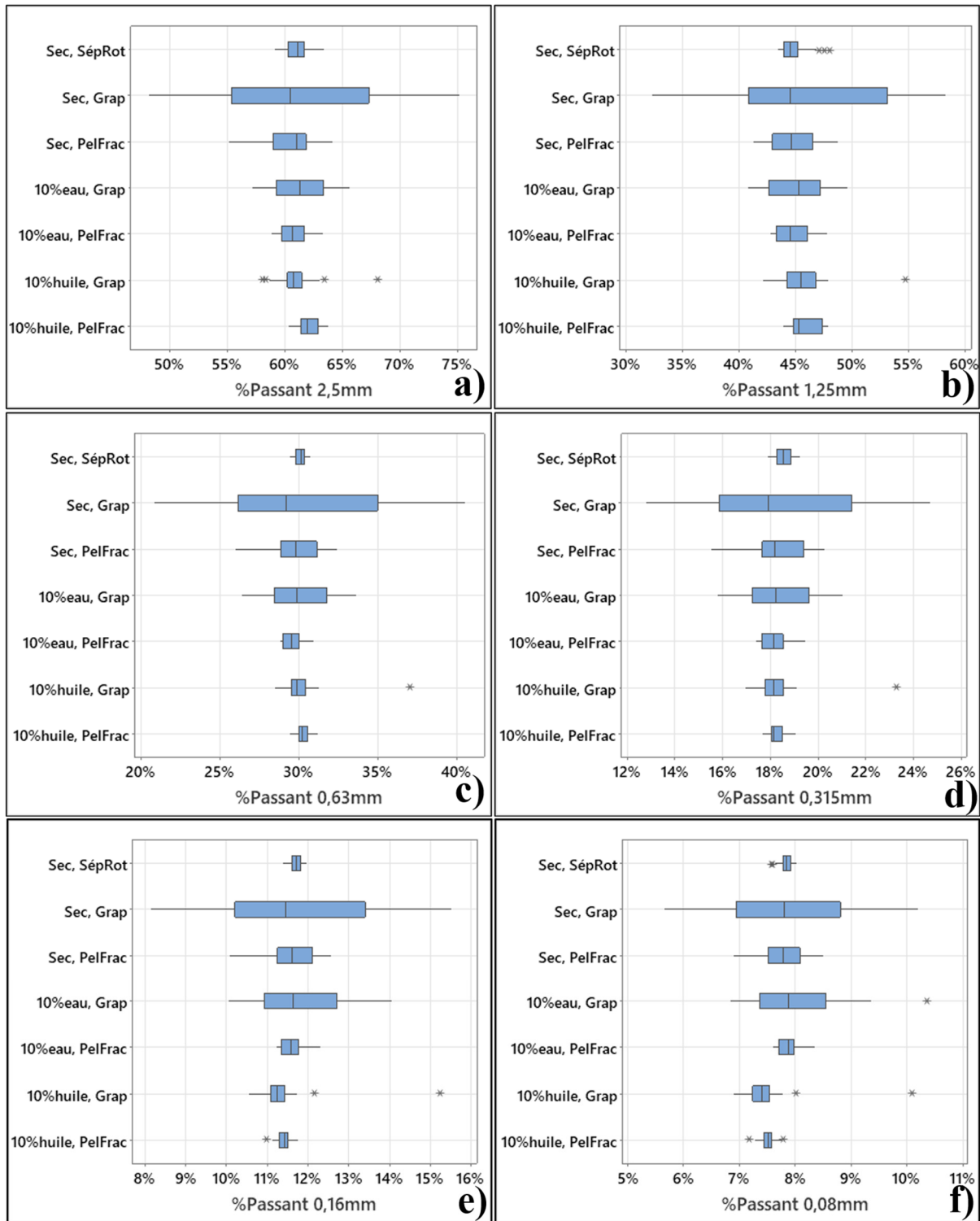


Figure 3.5 Graphiques en bo te   moustaches des pourcentages passants des sept essais avec fines des tamis a) 2,5 mm, b) 1,25 mm, c) 0,63 mm, d) 0,315 mm, e) 0,16 mm et f) 0,08 mm

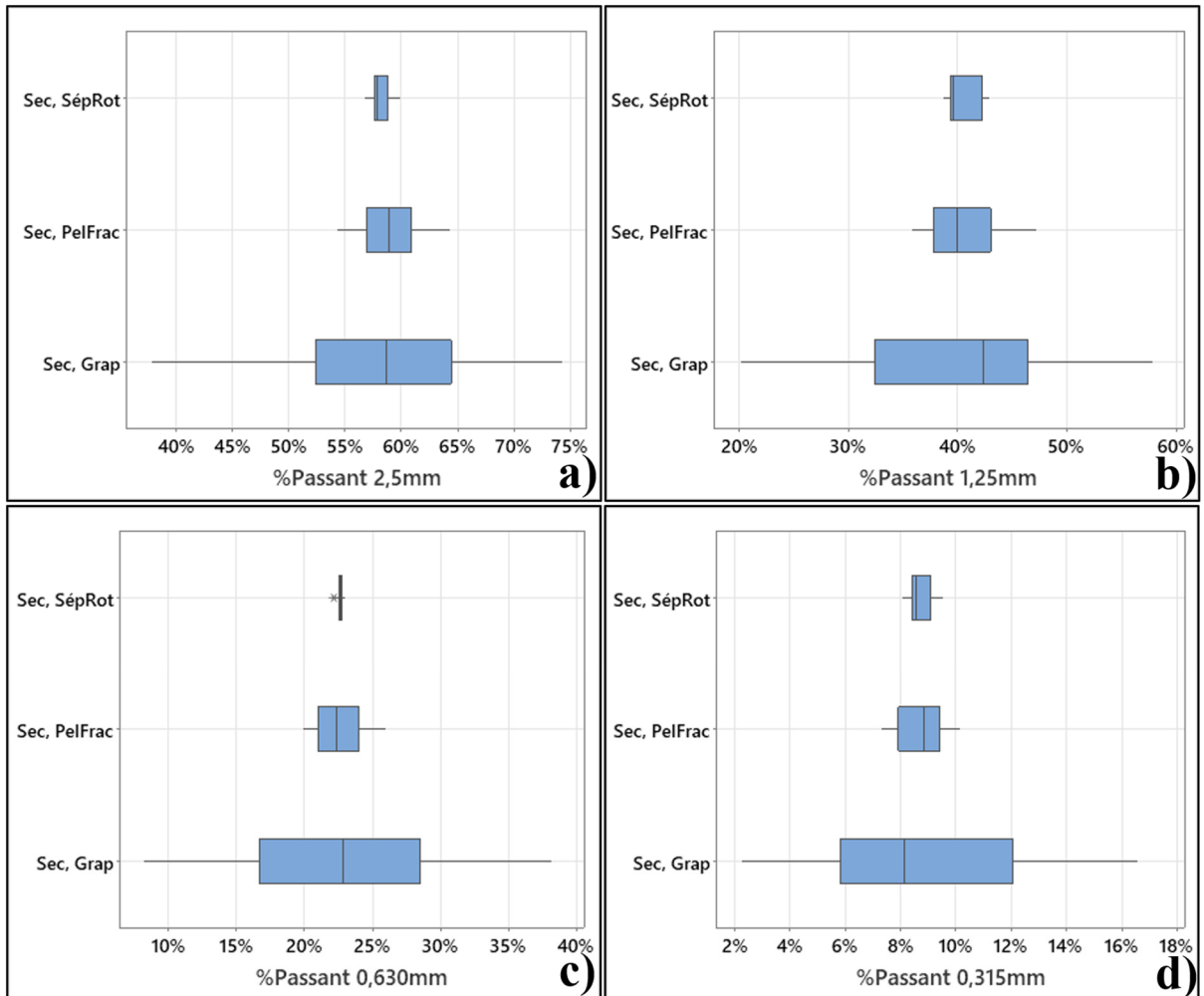


Figure 3.6 Graphiques en boîte à moustaches des pourcentages passants des trois essais avec fines des tamis a) 2,5 mm, b) 1,25 mm, c) 0,63 mm et d) 0,315 mm

Il est à noter et à retenir pour le reste de l'analyse des résultats granulométriques, qu'au vu de la Figure 3.5, un échantillon semble présenter une granulométrie aberrante pour l'essai de grappillage de sol avec fine et 10% d'huile. En effet, il se traduit par un astérisque très excentré sur la droite des boîtes à moustaches de l'essai en question. Cette valeur aberrante correspond à un échantillon pour lequel la granulométrie est très différente des autres échantillons de l'essai. Cela peut être dû soit à la procédure de grappillage qui peut générer des échantillons très différents, soit à une erreur dans la caractérisation granulométrique.

3.1.2.1 Effet de la présence de fines

Pour étudier l'effet que la présence de particules fines a eu sur les écarts de granulométries, ont été tracées les courbes granulométriques d'erreur des essais de sol sec avec fines puis sans fines échantillonnés par grappillage (n°2 et 4) en Figure 3.7 et Figure 3.8, puis pelletage fractionné (n°3 et 6) en Figure 3.9 et Figure 3.10 et enfin séparation rotative (n°1 et 5) en Figure 3.11 et Figure 3.12. Les courbes d'erreur ne pouvant être superposées sur un même graphique à la manière des graphiques des parties suivantes 3.1.2.2 et 3.1.2.3. Il s'agit de l'étalement des courbes d'erreurs par rapport à la courbe moyenne qui a été comparé et qui a permis d'évaluer l'effet de la présence de fines.

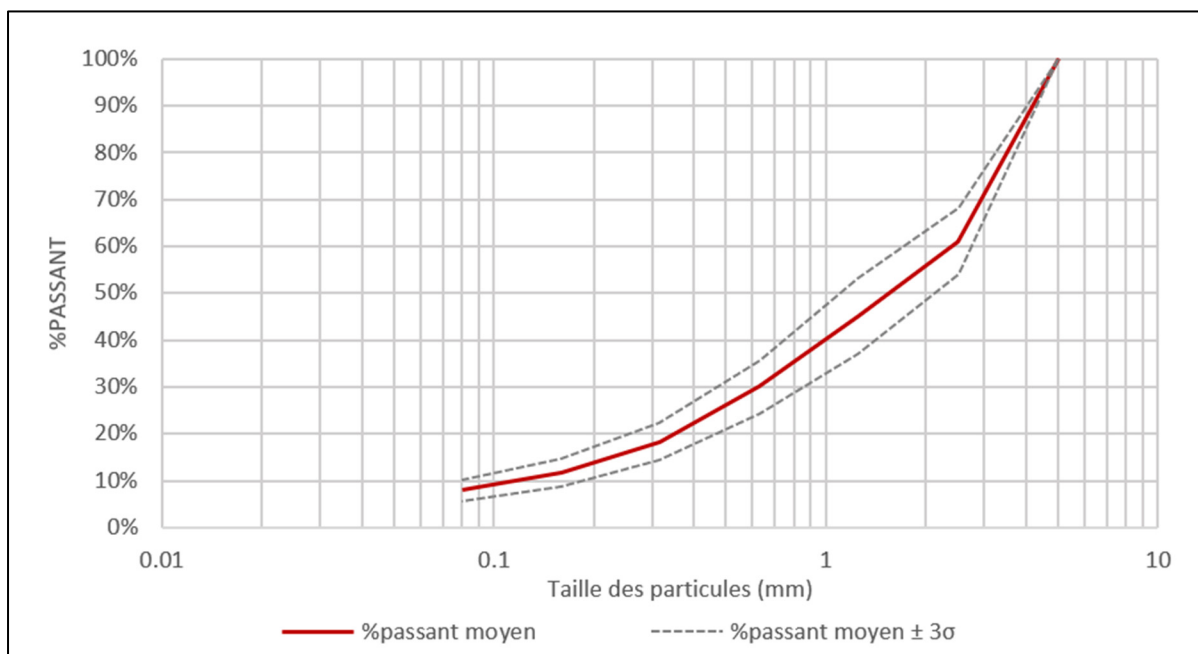


Figure 3.7 Granulométrie moyenne et intervalles de trois écart-types de l'essai n°2 de sol avec fines échantillonnés par grappillage

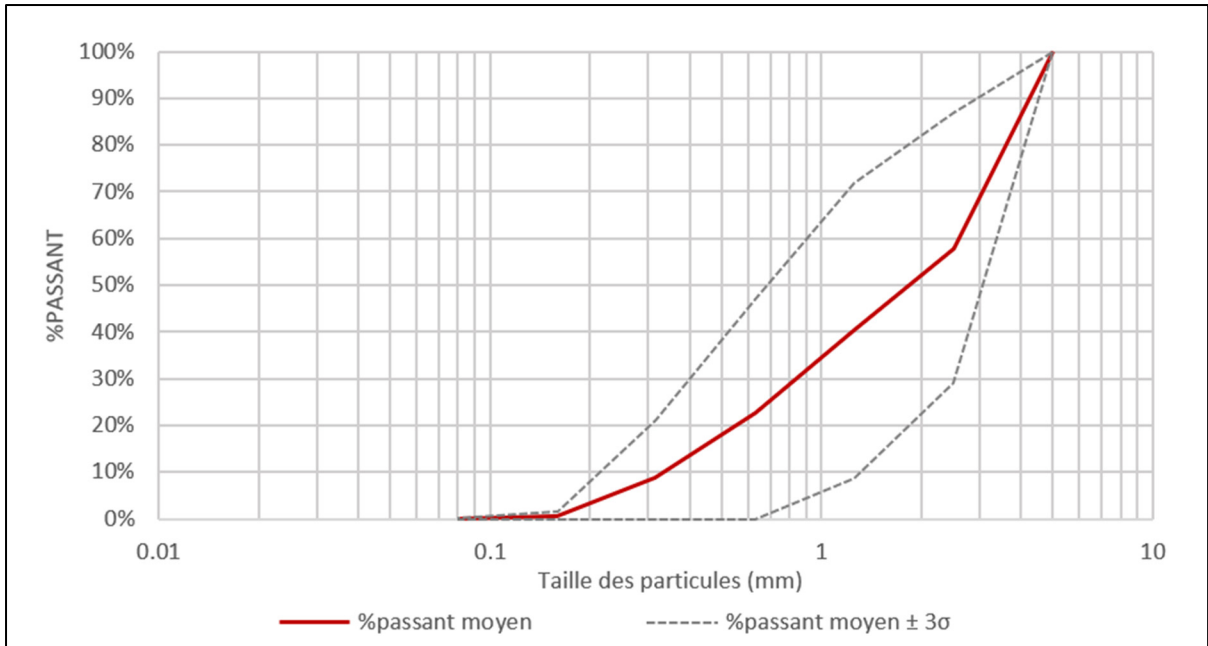


Figure 3.8 Granulométrie moyenne et intervalles de trois écart-types de l'essai n°4 de sol sans fines échantillonnés par grappillage

Les graphiques en Figure 3.7 et Figure 3.8 et les boîtes à moustaches associées en Figure 3.5 et Figure 3.6 ont mis en avant le fait que dans le cas du grappillage, l'absence de fines engendre des écarts de granulométrie beaucoup plus importants que pour l'échantillonnage de sol similaire en contenant.

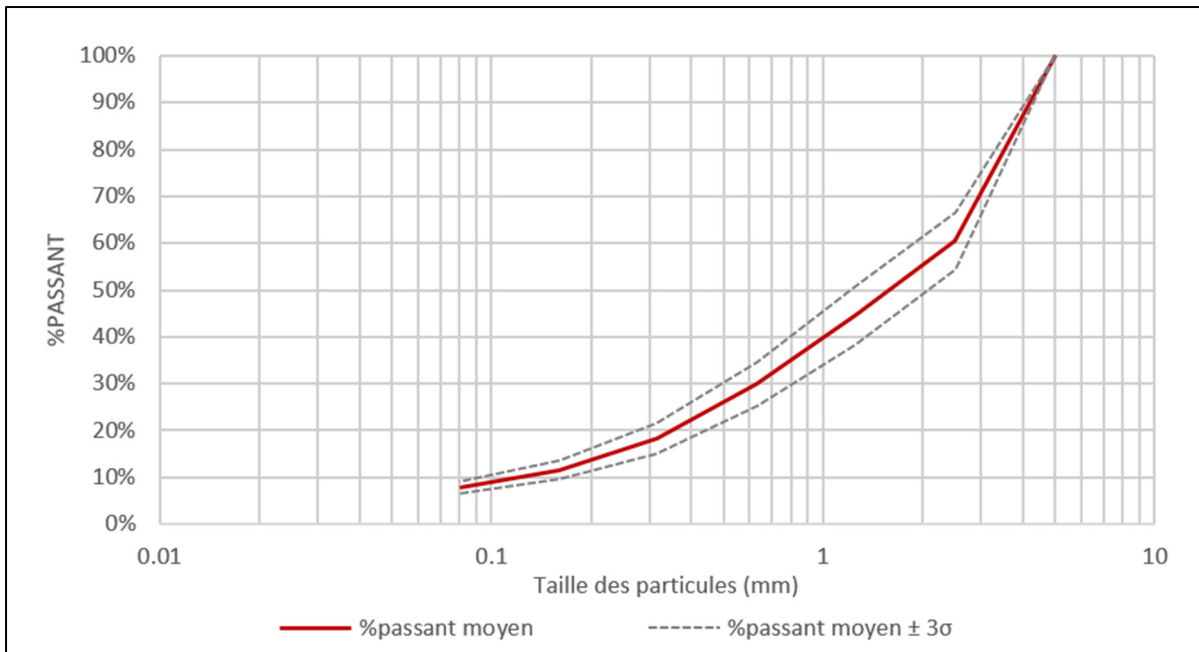


Figure 3.9 Granulométrie moyenne et intervalles de trois écart-types de l'essai n°3 de sol avec fines échantillonnés par pelletage fractionné

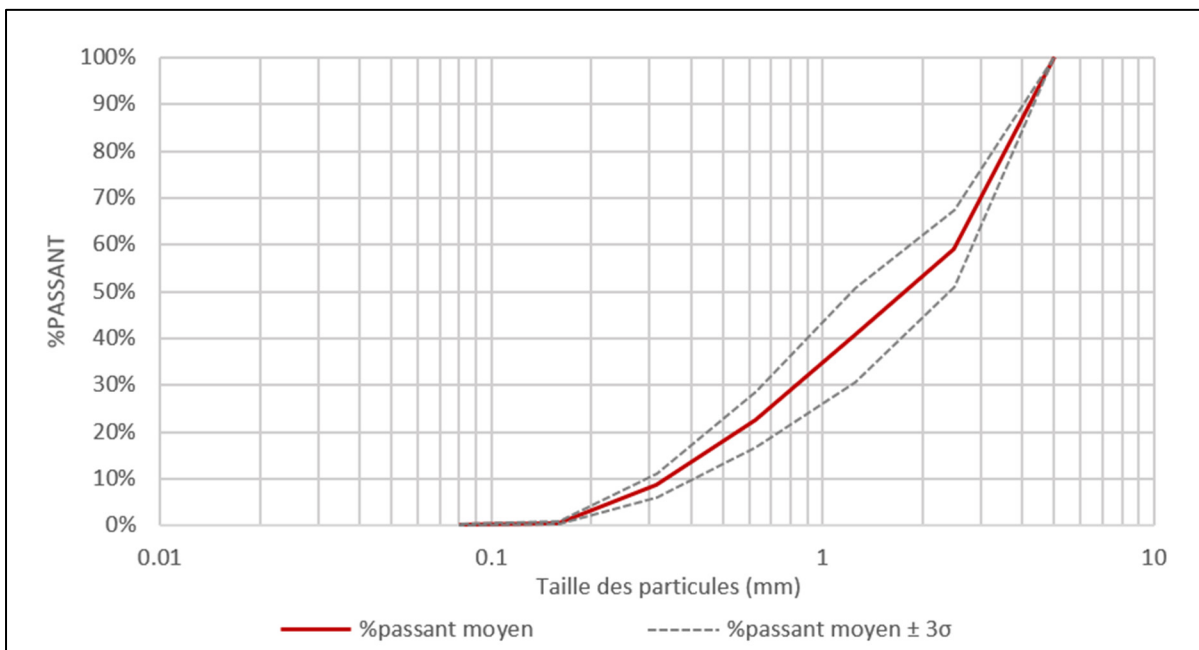


Figure 3.10 Granulométrie moyenne et intervalles de trois écart-types de l'essai n°6 de sol sans fines échantillonnés par pelletage fractionné

La Figure 3.9 et la Figure 3.10 et les boîtes à moustaches associées en Figure 3.5 et Figure 3.6 ont permis une fois de plus de constater que pour du sol sans fines, des écarts de granulométrie plus conséquents sont à prévoir dans le cas du pelletage fractionné.

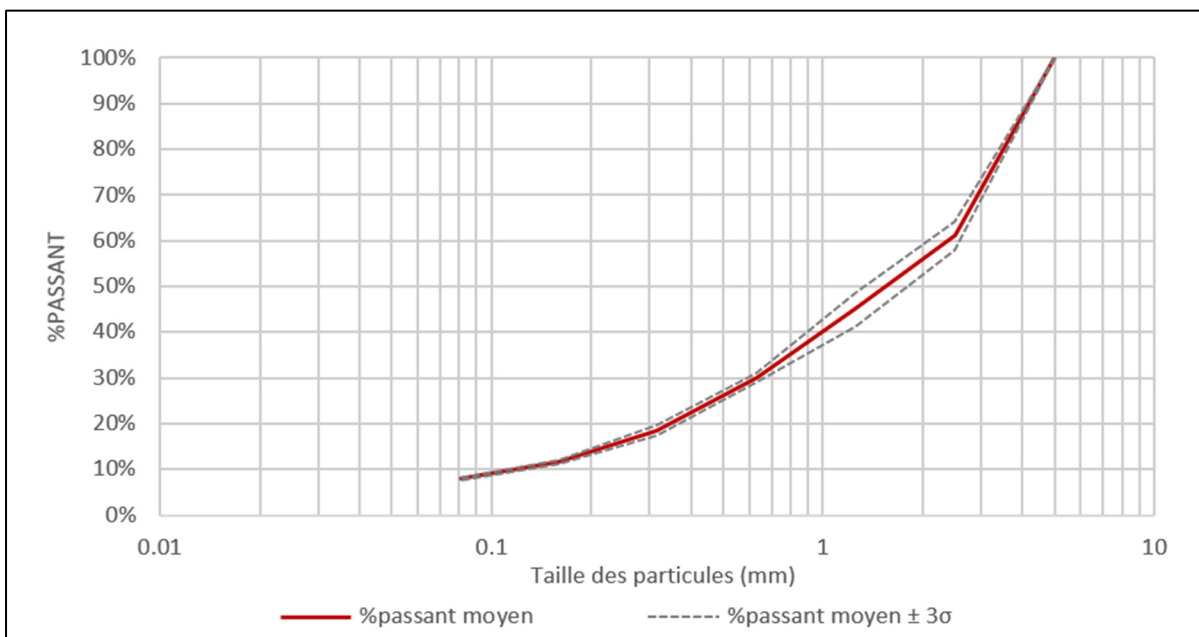


Figure 3.11 Granulométrie moyenne et intervalles de trois écart-types de l'essai n°1 de sol avec fines échantillonnés par séparation rotative

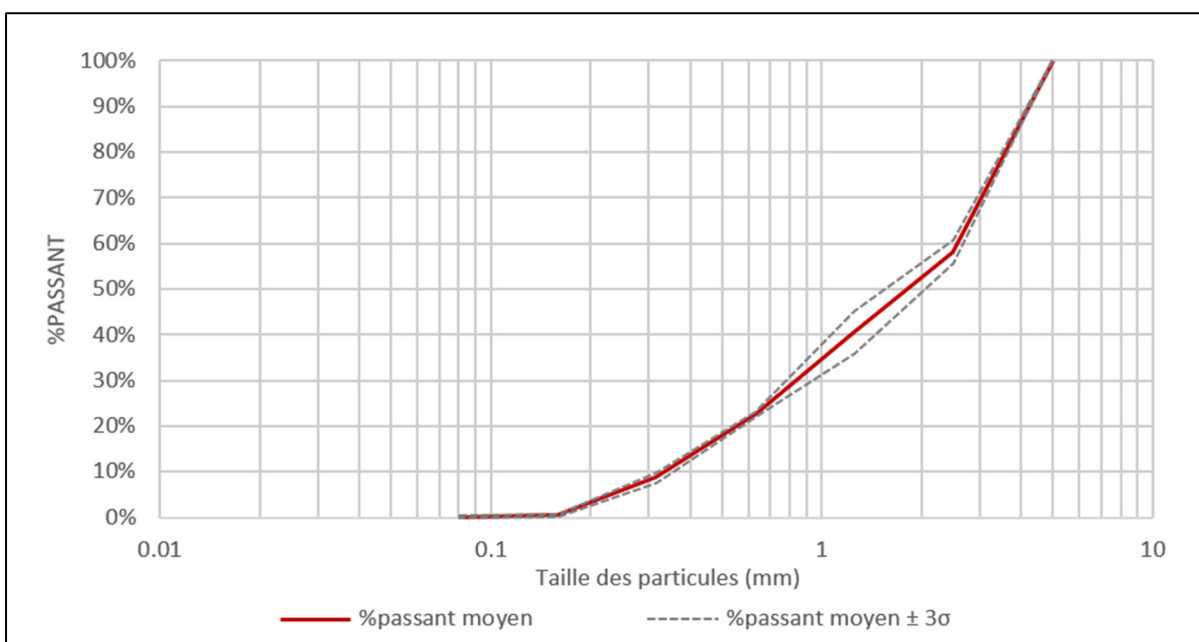


Figure 3.12 Granulométrie moyenne et intervalles de trois écart-types de l'essai n°5 de sol sans fines échantillonnés par séparation rotative

La Figure 3.11 et la Figure 3.12 et les boîtes à moustaches associées en Figure 3.5 et Figure 3.6 ont permis de constater que la séparation rotative a garanti une séparation représentative (1.2.1) du lot en termes de granulométrie et robuste à la présence ou l'absence de fines et ce, bien que le dispositif de trémie d'alimentation fut reconnu comme pouvant engendrer de la ségrégation en présence d'une large gamme granulométrique, notamment vis-à-vis des particules les plus fines (Gerlach & Nocerino, 2003).

En résumé, la présence de fines a eu tendance à améliorer la représentativité (1.2.1) granulométrique des échantillonnages par grappillage de manière très importante, et par pelletage fractionné de manière plus raisonnable. La séparation rotative quant à elle a semblé plutôt robuste à ce type de variation.

3.1.2.2 Effet de la présence d'une phase liquide

Dans cette partie, l'effet de la présence d'une phase liquide (eau ou huile) a été étudiée. Les essais choisis pour étudier cela sont les essais de sol avec fines de grappillage (n°2, 7 et 9) en Figure 3.13 et de pelletage fractionné (n°3, 8 et 10) en Figure 3.14. Les essais de séparation rotative sont exclus étant donné que cette méthode est seulement adaptée à l'échantillonnage de sol sec.

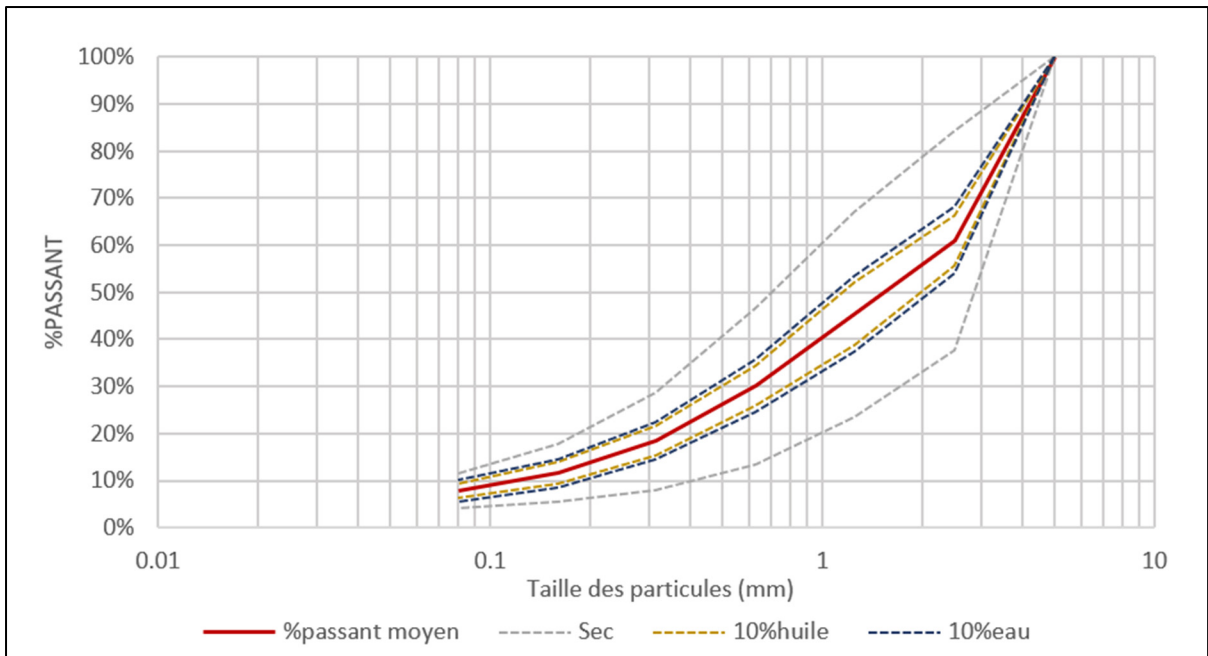


Figure 3.13 Comparaison des intervalles à trois écarts types centrés sur la moyenne pour les essais de sol avec fines échantillonnés par grappillage afin d'évaluer l'effet de la présence d'une phase liquide

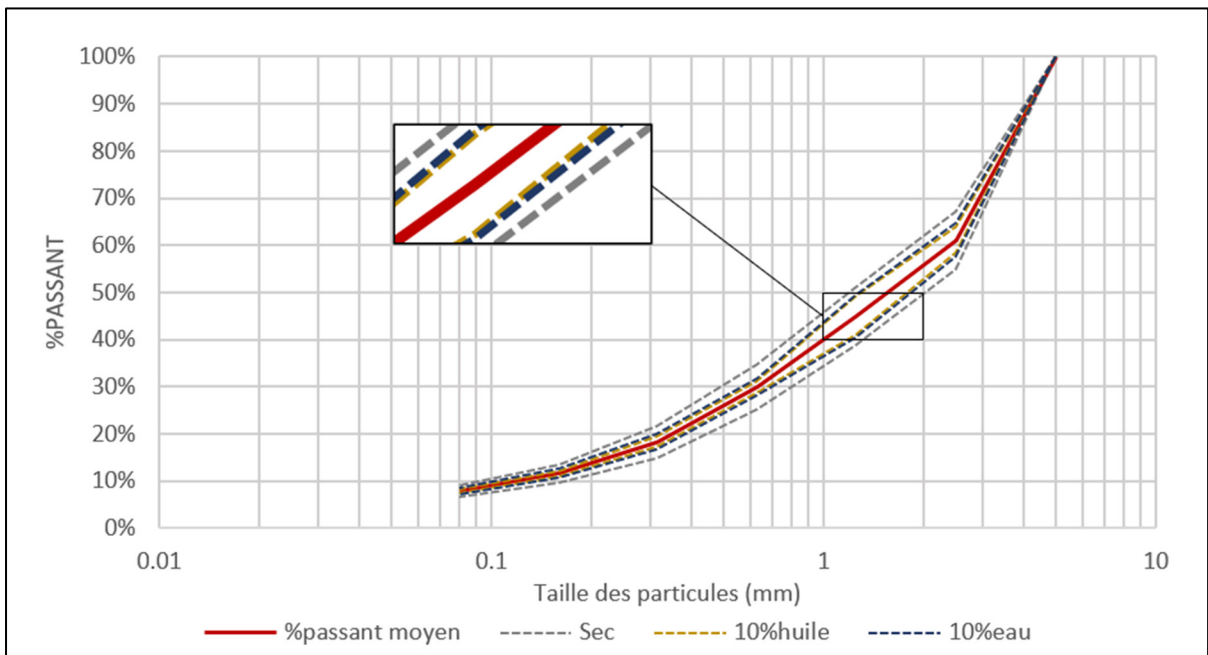


Figure 3.14 Comparaison des intervalles à trois écarts types centrés sur la moyenne pour les essais de sol avec fines échantillonnés par pelletage fractionné afin d'évaluer l'effet de la présence d'une phase liquide

Grâce à ces deux graphiques et les boîtes à moustaches associées en Figure 3.5 et Figure 3.6, il a été possible de constater que la présence d'une phase liquide améliore la représentativité (1.2.1) des échantillonnages, surtout pour le grappillage qui est une méthode plus approximative que le pelletage fractionné. En effet, la présence d'une phase liquide a permis d'augmenter la cohésion des particules de sol de manière plus homogène que pour un lot sec où les plus grosses particules se retrouvaient généralement en surface tandis que les plus petites particules avaient tendance à se déposer en fond.

Bien que la présence d'une phase liquide à une concentration massique de 10% a lissé grandement les écarts de représentativités entre les deux méthodes, le pelletage fractionné est resté tout de même plus représentatif que le grappillage.

La présence d'huile a permis d'améliorer légèrement la représentativité d'échantillonnage par rapport à l'eau. Cette amélioration est sûrement due à la différence de propriétés physiques de ces deux phases liquides telles que la viscosité.

3.1.2.3 Effet de la méthode d'échantillonnage employée

Dans cette partie, les différentes méthodes d'échantillonnage ont été comparées. Pour ce faire, les courbes d'erreur des essais de sol sec et avec fines (n°1, 2 et 3) en Figure 3.15, puis les essais de sol avec fines à 10% d'eau (n°7 et 8) en Figure 3.16 et enfin, les essais de sol avec fines à 10% d'huile (n°9 et 10) en Figure 3.17 ont été tracées.

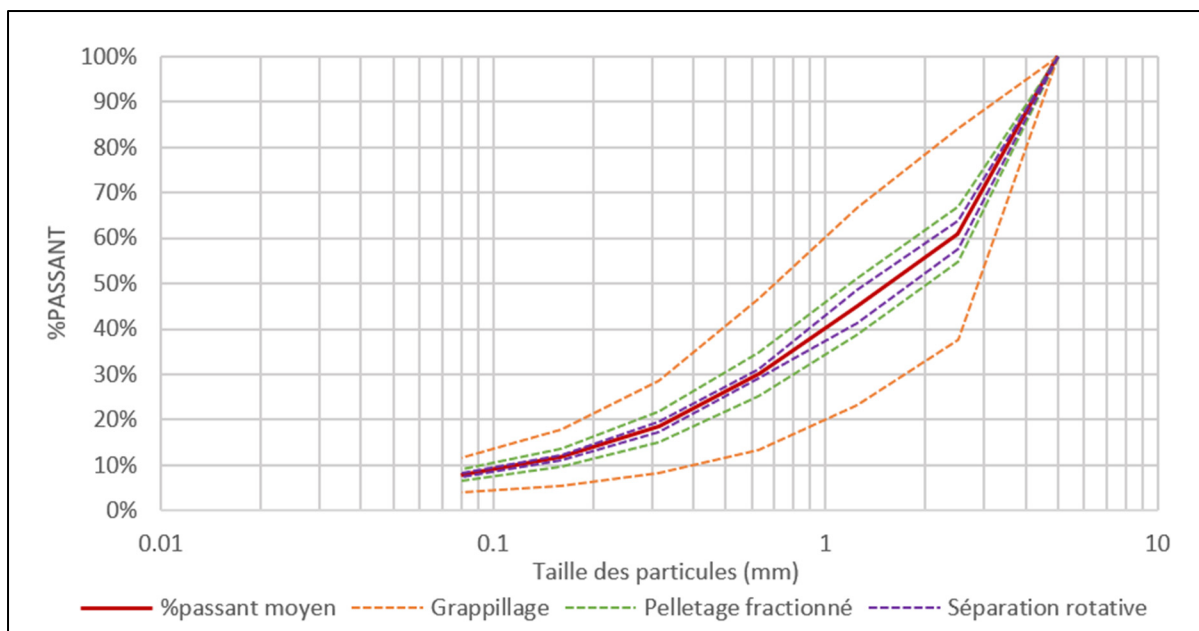


Figure 3.15 Comparaison des intervalles à trois écarts types centrés sur la moyenne pour les essais de sol sec avec fines de grappillage, pelletage fractionné et séparation rotative afin de comparer l'effet de la méthode d'échantillonnage employée

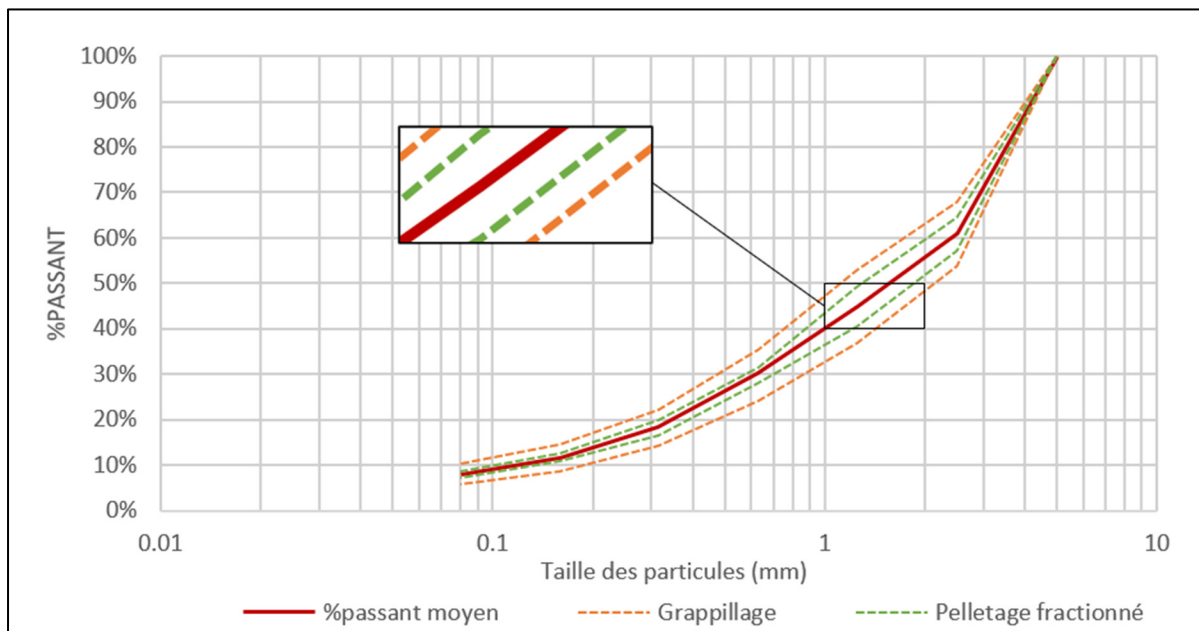


Figure 3.16 Comparaison des intervalles à trois écarts types centrés sur la moyenne pour les essais de sol avec fines et 10% d'eau de grappillage et pelletage fractionné afin de comparer l'effet de la méthode d'échantillonnage employée

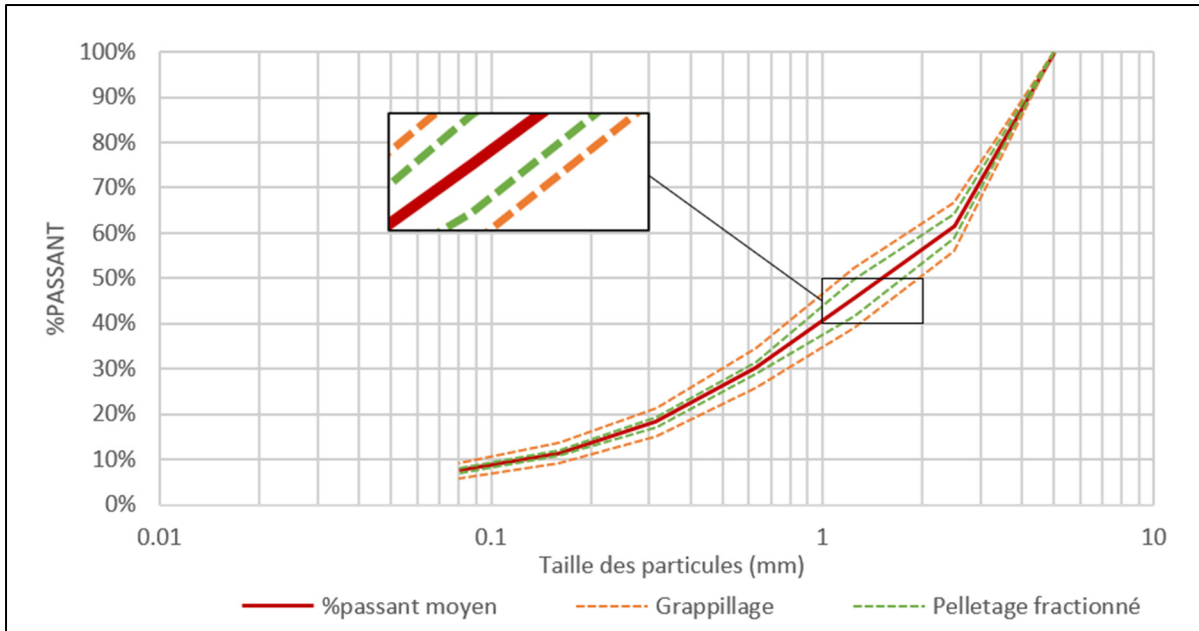


Figure 3.17 Comparaison des intervalles à trois écarts types centrés sur la moyenne pour les essais de sol avec fines et 10% d'huile de grappillage et pelletage fractionné afin de comparer l'effet de la méthode d'échantillonnage employée

Encore une fois, la hiérarchie dans la représentativité (1.2.1) des méthodes de grappillage, pelletage fractionné et de séparation rotative pour la Figure 3.15 est restée inchangée. En examinant la Figure 3.15, qui compare les courbes d'erreur pour les essais de sol sec et les boîtes à moustaches associées en Figure 3.5 et Figure 3.6, il est évident que le pelletage fractionné et la séparation rotative se sont positionnés au même niveau de performance, bien que cette dernière ait été légèrement plus performante. En revanche, les courbes d'erreur et les boîtes à moustaches du grappillage ont montré des écarts significatifs en termes de représentativité granulométrique. Cependant, l'ajout de 10% d'eau a effacé cette différence entre le grappillage et le pelletage fractionné bien que ce dernier soit toujours légèrement plus performant. Cette différence s'est atténuée encore davantage en présence d'huile, probablement pour des raisons similaires à celles évoquées précédemment en 3.1.2.2.

3.1.3 Analyse statistique de la variance

Pour rappel, dans le cadre de l'analyse statistique des résultats granulométriques, des tests préliminaires de normalité d'Anderson-Darling sont effectués pour les jeux de données de pourcentages passants associés à chaque fraction granulométrique de chacun des dix essais. Puisque presque 40% des jeux de données ne suivent pas une distribution normale, le test non paramétrique d'égalité des variances de Levene est donc préféré à celui de Bartlett qui est sensible à la violation de l'hypothèse de normalité.

Le test de Levene est effectué sur ces données granulométriques, d'une part pour les essais avec fines, et d'autre part pour les essais sans fines. Six tests d'égalité des variances de Levene associés aux six fractions granulométriques sont donc effectués pour les pourcentages passants des essais de sol avec fines (Figure 3.18), et quatre autres pour les essais de sol sans fines (Figure 3.19). Au total, ce sont donc 10 tests d'égalité des variances qui sont faits pour l'analyse des granulométries.

Les valeurs exactes des écarts types et bornes des intervalles de confiance à 95% associés aux écarts types des pourcentages passants sont calculées pour tous les graphiques des intervalles de comparaisons multiple générés. En ANNEXE IV, le Tableau-A IV-1 récapitule ces informations pour les tests de Levene effectués pour les configurations d'essais avec fines tandis que Tableau-A IV-2 les présente pour les essais sans fines.

3.1.3.1 Essais avec fines

La Figure 3.18 ci-dessous présente les graphiques de comparaisons multiples des six tests d'égalité des variances de Levene associés aux six fractions granulométriques effectués pour les pourcentages passants des essais de sol avec fines.

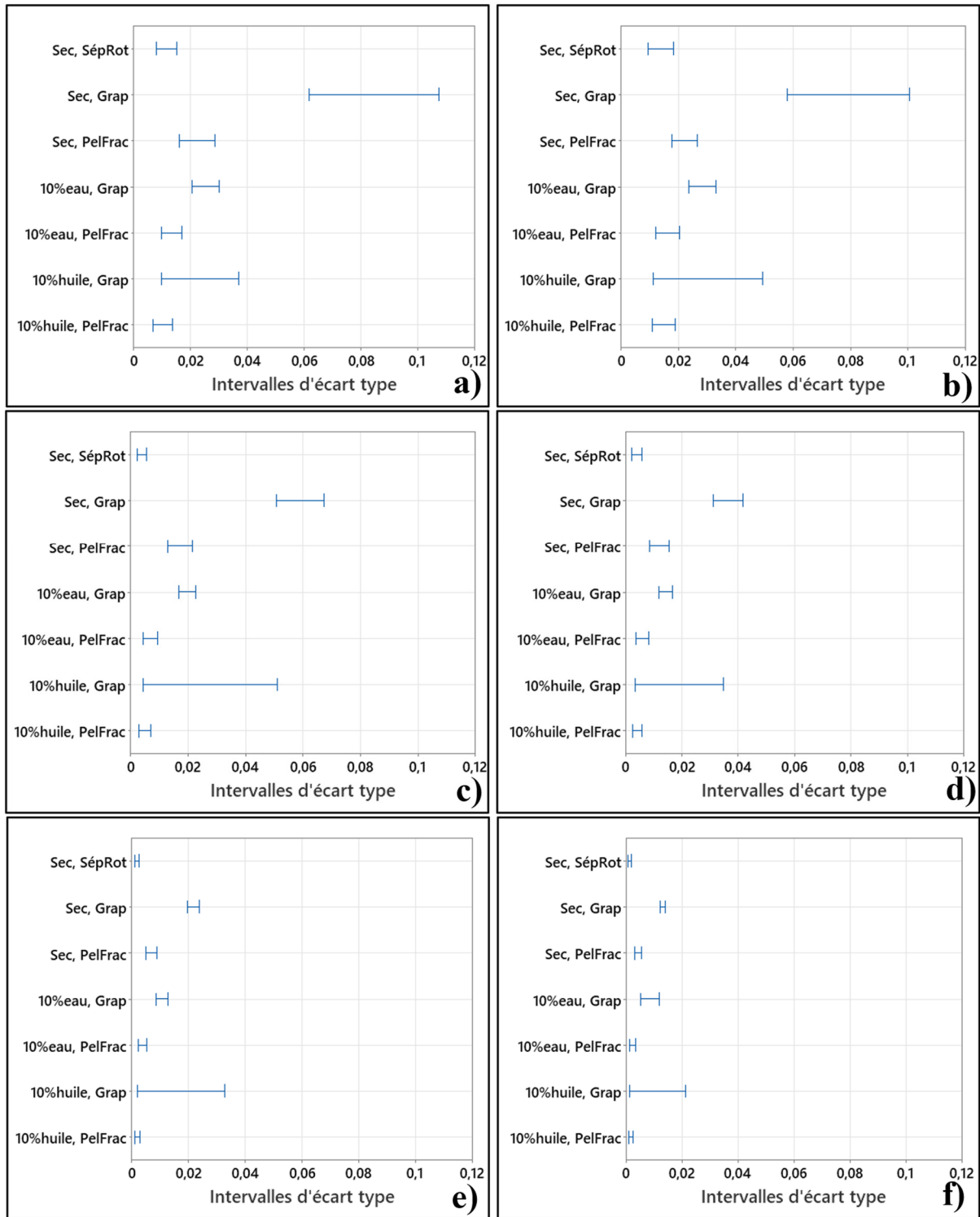


Figure 3.18 Graphique des intervalles de comparaisons multiples des écarts types pour les pourcentages passants des sept essais avec fines des tamis a) 2,5 mm, b) 1,25 mm, c) 0,63 mm, d) 0,315 mm, e) 0,16 mm et f) 0,08 mm

Les graphiques des intervalles multiples démontrent de manière cohérente, tant pour les essais à sec qu'avec de l'eau ou de l'huile, que les tailles et les ordres de grandeur des intervalles de Bonferroni pour les écarts types suivent systématiquement la même hiérarchie. Cette hiérarchie suit l'ordre croissant de la séparation rotative (uniquement dans le cas à sec), suivi du pelletage fractionné et enfin du grappillage. Ceci suit les constats faits en 3.1.2.3 sur l'analyse des masses et l'étude des courbes d'erreurs. Les écarts types liés au grappillage sur sol sec sont d'ailleurs compris dans un intervalle nettement plus grand que ceux des autres essais, au point même d'être considérés comme significativement différents d'eux pour les tests effectués sur les plus grandes fractions.

Pour le grappillage et le pelletage fractionné, la présence d'eau a permis pour toutes les fractions granulométriques de réduire la taille et l'ordre de grandeur des intervalles de Bonferroni par rapport à ceux des essais secs. Parfois, les améliorations sont telles que les intervalles associés au pelletage fractionné et au grappillage secs et avec eau ne se superposent pas, ce qui veut dire qu'ils sont significativement différents. Pour le pelletage fractionné, la présence d'huile a fait davantage diminuer la taille et l'ordre de grandeur des intervalles par rapport à ceux liés à la présence d'eau. La présence de phase liquide semble également lisser les écarts entre les différentes méthodes d'échantillonnage avec des intervalles d'écarts types qui sont toujours significativement différents entre les méthodes pour les essais secs, ce qui n'est plus le cas pour les essais avec phase liquide. Ces améliorations des essais avec phase liquide par rapport aux essais secs, et particulièrement des essais avec huile, suivent les constats faits en 3.1.2.2 sur l'analyse des masses et l'étude des courbes d'erreurs.

Pour le grappillage cependant, la présence d'une valeur aberrante (3.1.2, Figure 3.5) pour l'essai avec huile a, au contraire, fait grandement augmenter les intervalles de Bonferroni surtout pour les fractions les plus fines, ce qui diffère donc de ce qui était attendu. En effet, la présence d'huile était censée générer des intervalles du même ordre voir plus petits que pour l'essai avec eau. Ces dernières ont pu être à l'origine de ce résultat, puisque les étalements des boîtes à moustaches de l'essai de grappillage avec huile sont eux, cohérents. L'origine de ces valeurs aberrantes peut résider dans le manque de représentativité (1.2.1) connue de cette

méthode pouvant parfois mener à l'impossibilité de comparer les résultats de deux essais, même réalisés de manière identique (1.2.3.1). Pour appuyer ce fait, il est à constater que toutes les valeurs aberrantes mises en avant sur les six graphiques en boîte à moustache concernent des essais de grappillage.

3.1.3.2 Essais sans fines

Comme pour les essais avec fines précédemment, un graphique des intervalles de comparaisons multiples pour l'écart-type (avec $\alpha = 0,05$) a été fait (Figure 3.19) pour chacun des quatre tests d'égalité des variances sur les pourcentages passants des trois essais avec fines (n°4, 5 et 6).

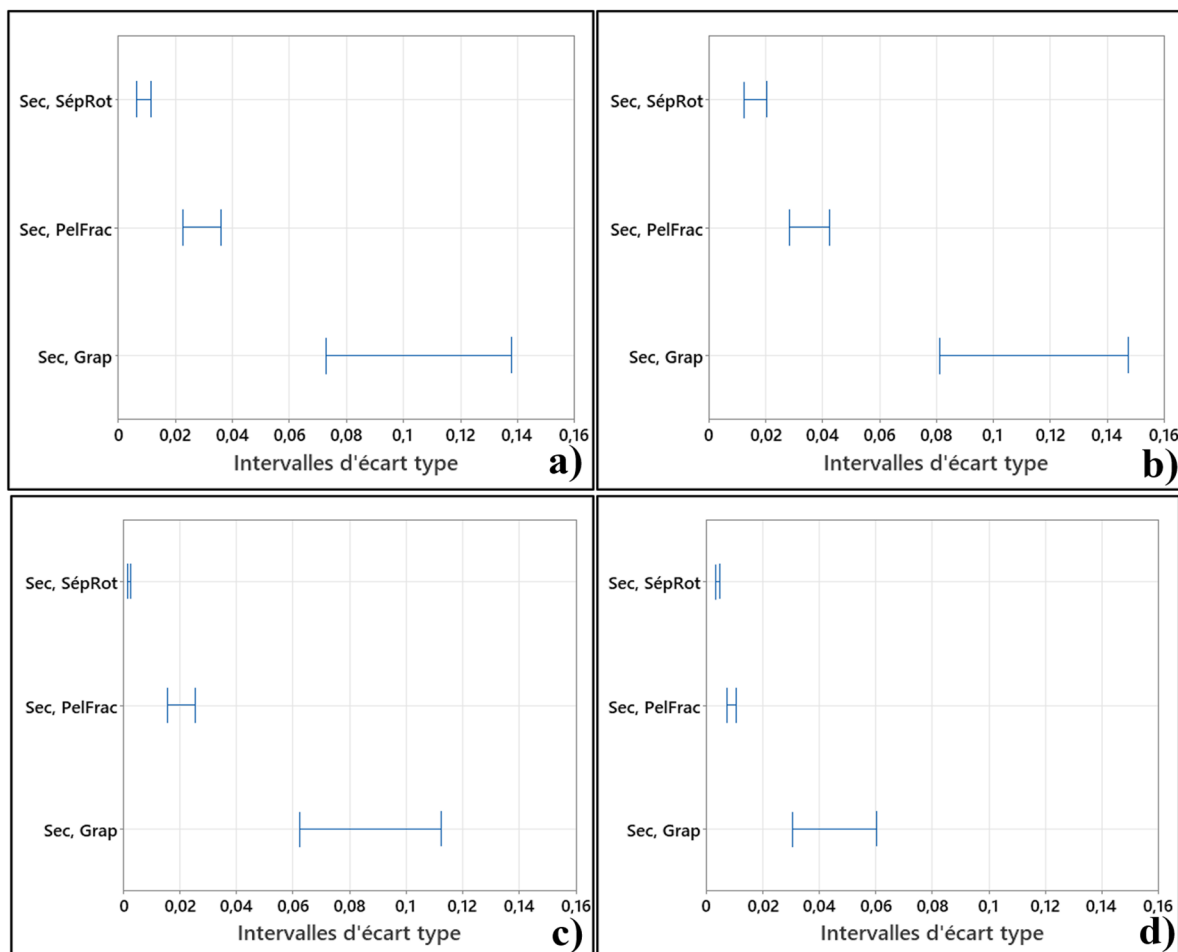


Figure 3.19 Graphique des intervalles de comparaisons multiples des écarts types pour les pourcentages passants des trois essais sans fines des tamis a) 2,5 mm, b) 1,25 mm, c) 0,63 mm et d) 0,315 mm

Comme pour les résultats des tests d'égalité des variances des essais de sol avec fines, les graphiques des intervalles multiples démontrent de manière cohérente, que les tailles et les ordres de grandeur des intervalles de Bonferroni pour les écarts types, suivent globalement la même hiérarchie de méthode d'échantillonnage qu'en 3.1.2.3 sur l'analyse des masses et l'étude des courbes d'erreurs. D'ailleurs, pour toutes les fractions granulométriques, il est à constater que les intervalles sont tous significativement différents.

L'intervalle de Bonferroni associé au pelletage fractionné est demeuré pratiquement inchangé quelle que soit la fraction granulométrique, ce qui n'est pas le cas de la séparation rotative. En effet, même si les intervalles de Bonferroni restent meilleurs que ceux du pelletage fractionné, ils se détériorent à mesure que la fraction granulométrique diminue. Cette observation peut être attribuée en grande partie à l'effet de ségrégation potentiellement induit par la trémie d'alimentation vibrante dans le dispositif de séparation rotative. (1.2.3.3).

3.2 Analyse des résultats de pourcentage de briques

3.2.1 Étude descriptive et statistique des variances

Afin d'être en mesure de décrire, dans un premier temps, la distribution des données de pourcentage de briques de chacune des dix configurations d'essais, un graphique en boîte à moustaches est réalisé et est présenté en Figure 3.22.

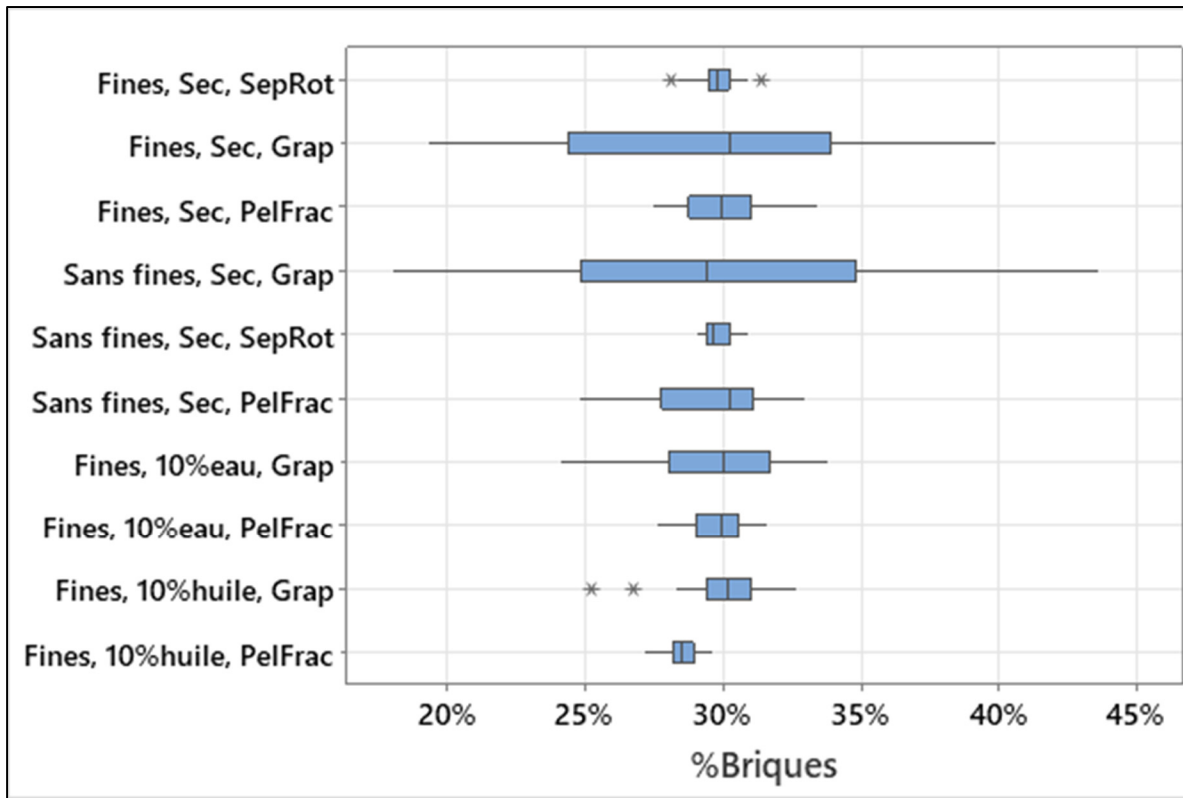


Figure 3.20 Graphique en boîtes à moustaches des pourcentages de briques des dix essais

Ce graphique en boîtes à moustaches met en avant le fait que le grappillage est la méthode générant le plus d'écart entre les pourcentages de briques des différents échantillons produits. Vient ensuite le pelletage fractionné et enfin, la séparation rotative. La présence de phase liquide diminue ces écarts, au point de rendre l'échantillonnage de sol avec 10% d'huile par pelletage fractionné aussi performant que l'échantillonnage de sol sec par séparation rotative. Comme pour les résultats granulométriques, des valeurs aberrantes sont à déplorer pour l'essai de grappillage de sol avec fines et avec huile. Cette valeur aberrante correspond à un échantillon pour lequel la granulométrie est très différente des autres échantillons de l'essai. Une analyse statistique peut maintenant confirmer ces ressemblances et ces différences observées entre les essais sont significatives. Elle permet également d'observer d'autres tendances non visibles ou peu mises en évidence avec ce seul graphique en boîtes à moustaches.

Dans le cadre de l'analyse statistique des résultats de pourcentages de briques, des tests préliminaires de normalité d'Anderson-Darling sont effectués pour les jeux de données associés à chacun des dix essais. L'ensemble des jeux de données suivent une distribution normale ainsi, le test Bartlett est préféré pour la suite de l'analyse statistique. Un test de Bartlett a donc été réalisé pour les dix jeux de 32 pourcentages de briques associés aux dix essais réalisés lors de cette recherche. Ce test d'égalité des variances a permis d'apporter un point de vue statistique dans la comparaison des résultats de pourcentage de briques associés aux différentes configurations d'essai. Au cours de ce test, un graphique des intervalles de comparaisons multiples pour l'écart-type est généré et est présenté en Figure 3.21. Pour plus de détails sur les lignes directrices d'interprétation de ces graphiques, se référer à la partie 2.4.3. En plus du graphique des intervalles de comparaisons multiples, les valeurs exactes des écarts types et des bornes des intervalles de confiance à 95% sont calculées lors du test. Le Tableau-A VI-1 en ANNEXE IV récapitule ces informations pour les différentes configurations d'essais.

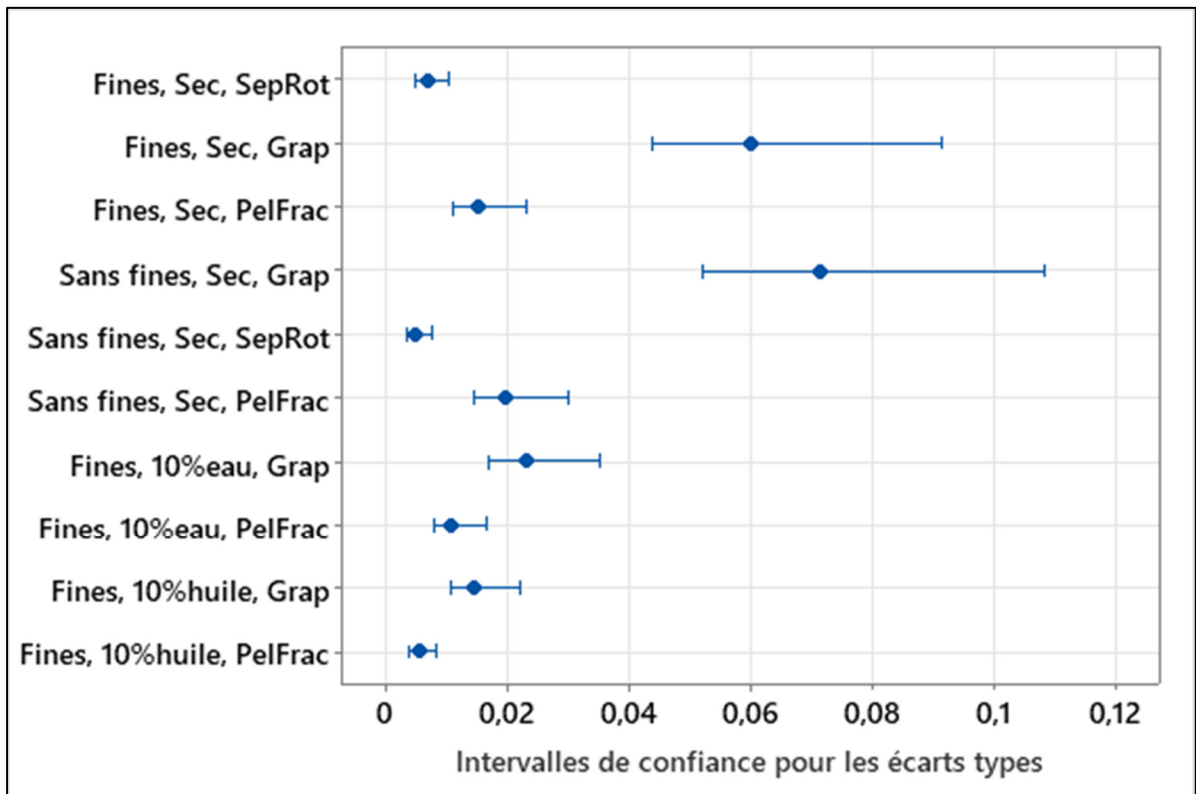


Figure 3.21 Graphique des intervalles de confiance de Bonferroni à 95 % pour les écarts types des pourcentages de briques des dix essais

Pour les essais secs, les écarts types liés aux essais de grappillage sont supérieurs, plus variables, et significativement différents des intervalles liés au pelletage fractionnés et à la séparation rotative. Les intervalles liés à ces deux dernières méthodes sont semblables mais la séparation rotative reste tout de même plus performante.

La présence d'une phase liquide fait grandement diminuer la taille et l'étalement des intervalles de Bonferroni liés au grappillage, les rendant plus semblables à ceux du pelletage fractionné bien que toujours moins performants. L'huile apporte une amélioration notable par rapport à l'eau. Cependant, tant en présence d'eau que d'huile, les intervalles liés aux essais de grappillage et de pelletage fractionné sont significativement différents.

La présence de fines permet de générer des intervalles de Bonferroni sur les écarts types plus petits et légèrement moins étalés que leurs homologues sans fines pour le grappillage et le pelletage fractionné mais ne sont pas significativement différents. La situation contraire est observée pour l'essai de séparation rotative. En effet, l'absence de fines permet de diminuer la taille et l'étalement de l'intervalle de Bonferroni étant donné que cette méthode n'est quasiment soumise qu'à l'erreur fondamentale qui est diminuée avec la gamme granulométrique (explications plus détaillées à venir en partie 3.2.3).

3.2.2 Étude des biais relatifs, variances relatives et degrés de représentativité

Le Tableau 3.3 présente les biais d'échantillonnage (m^2 exprimé à l'équation (1.7)) des dix essais. Ces biais sont directement associés à la perte de brique particulière au cours de l'échantillonnage (étudiée en 4.1 plus bas) ainsi qu'à l'omission des fragments compris dans les fractions inférieures à 1,25 mm. Ce tableau présente également les variances relatives (s_{relative}^2) et les degrés de représentativité (r^2 exprimé à l'équation (1.9)) associés.

Tableau 3.3 Biais d'échantillonnage, variance relative et degré de représentativité des résultats de pourcentage de briques des dix essais

	Séparateur rotatif	Grappillage	Pelletage fractionné
Sol sec sans fines	Essai n°5 $m^2 = 5,55 \times 10^{-5}$ $s_{\text{relative}}^2 = 2,94 \times 10^{-4}$ $r^2 = 3,49 \times 10^{-4}$	Essai n°4 $m^2 = 6,50 \times 10^{-6}$ $s_{\text{relative}}^2 = 5,67 \times 10^{-2}$ $r^2 = 5,67 \times 10^{-2}$	Essai n°6 $m^2 = 1,66 \times 10^{-4}$ $s_{\text{relative}}^2 = 4,36 \times 10^{-3}$ $r^2 = 4,53 \times 10^{-3}$
Sol sec avec fines	Essai n°1 $m^2 = 4,81 \times 10^{-5}$ $s_{\text{relative}}^2 = 5,42 \times 10^{-4}$ $r^2 = 5,90 \times 10^{-4}$	Essai n°2 $m^2 = 2,99 \times 10^{-4}$ $s_{\text{relative}}^2 = 4,03 \times 10^{-2}$ $r^2 = 4,06 \times 10^{-2}$	Essai n°3 $m^2 = 1,92 \times 10^{-5}$ $s_{\text{relative}}^2 = 2,60 \times 10^{-3}$ $r^2 = 2,62 \times 10^{-3}$
Sol à 10%eau (avec fines)	-	Essai n°7 $m^2 = 3,43 \times 10^{-5}$ $s_{\text{relative}}^2 = 6,00 \times 10^{-3}$ $r^2 = 6,03 \times 10^{-3}$	Essai n°8 $m^2 = 1,92 \times 10^{-5}$ $s_{\text{relative}}^2 = 1,34 \times 10^{-3}$ $r^2 = 1,43 \times 10^{-3}$
Sol à 10%huile (avec fines)	-	Essai n°9 $m^2 = 1,90 \times 10^{-6}$ $s_{\text{relative}}^2 = 2,41 \times 10^{-3}$ $r^2 = 2,41 \times 10^{-3}$	Essai n°10 $m^2 = 2,44 \times 10^{-3}$ $s_{\text{relative}}^2 = 3,51 \times 10^{-4}$ $r^2 = 2,79 \times 10^{-3}$

Les biais semblent indépendants de la méthode d'échantillonnage et de la présence ou l'absence de phase liquide quelle qu'elle soit malgré les tendances observées précédemment pour les pertes de sol et briques cumulées (étudiées en 4.1 plus bas). Ceci prouve que les biais sont principalement dus à l'omission des fragments de particules de briques présentes dans les fractions inférieures à 1,25 mm plus qu'à la perte de particules au cours de la séparation et la prise photographique qui concerne surtout les particules les plus fines et donc pas les particules de brique. Ces dernières ne sont effritées qu'à partir de l'étape de tamisage à priori.

Cela dit, puisqu'au cours des essais, les lots de sol et briques étaient échantillonnés en totalité, il est normal que les biais relatifs soient faibles. Le biais relatif est la moyenne des erreurs relatives d'échantillonnage et celles-ci auront tendance à s'annuler lorsqu'un lot est échantillonné en totalité. La variance relative est donc plus appropriée pour étudier l'effet de la présence de liquide, de la méthode d'échantillonnage employée et de la présence de fines. Les constats sont les mêmes que ceux faits pour l'étude des variances (3.2.2). En effet, la hiérarchie des méthodes d'échantillonnage est la même avec les variances relatives associées

aux essais de séparation rotative qui sont les plus faibles, suivies par celles des essais de pelletage fractionné et enfin de grappillage. La présence d'une phase liquide permet de réduire les variances relatives d'un ou plusieurs ordres de grandeur mais la hiérarchie est conservée entre les méthodes d'échantillonnages applicables. La présence de fines permet de diminuer très légèrement les variances relatives pour les essais secs de grappillage et de pelletage fractionné, principalement influencés par les erreurs autres que fondamentale. Pour les essais de séparation rotative, la situation contraire est observée, la variance liée à l'erreur totale étant principalement influencée par la variance liée l'erreur fondamentale.

Le degré de représentativité suit la tendance dictée par la variance relative, les biais relatifs étant très faibles en raison de l'échantillonnage complet des lots.

3.2.3 Coefficient de variation

3.2.3.1 Coefficient de variation théorique $CV_{théo}$

Comme détaillé en 2.4.2.2, le coefficient de variation théorique correspond à la racine carrée de la variance relative de l'erreur totale d'échantillonnage qui correspondant elle, à la somme des variances relatives liées à l'erreur fondamentale et à l'erreur de ségrégation et groupement (1.3). La variance liée à l'erreur fondamentale se calcule à partir de l'invariant d'hétérogénéité d'après la formule (1.4). Ce dernier est exprimé en formule (1.5). La variance liée à l'erreur de ségrégation et groupement quant à elle se calcule selon l'expression (1.3) avec les nombres d'incrément constitutifs d'un échantillon selon les trois méthodes d'échantillonnage.

Il a donc d'abord fallu calculer l'invariant d'hétérogénéité (Tableau 3.4 et Tableau 3.5), puis la variance liée à l'erreur fondamentale (Tableau 3.6), suivie de la variance liée à l'erreur de ségrégation et groupement (1.3) ce qui a permis de déterminer la variance liée à l'erreur totale et finalement (1.2), le coefficient de variation théorique (Tableau 3.7).

Tous ces éléments énumérés ont été calculés par essai afin de garantir une estimation du coefficient de variation théorique la plus précise possible dans l'optique de les comparer avec

les valeurs de coefficient de variation empiriques dans le Tableau 3.9. Cette comparaison permet d'estimer le plus justement possible la variance liée aux erreurs autres que fondamentale et de ségrégation et groupement et ainsi de comparer les différentes configurations d'essai en fonction.

Le Tableau 3.4 récapitule les éléments nécessaires au calcul du terme $d_{\alpha}^3 \frac{m_{L\alpha}}{m_L} \frac{(a_{\alpha} - a_L)^2}{a_L^2}$ de l'invariant d'hétérogénéité qui sera noté A par soucis de simplification. Le Tableau 3.5 qui suit, présente d'une part ces termes calculés par fraction granulométrique pour chaque essai, d'autre part les diamètres moyens des fractions granulométriques. Le facteur de forme f usuel pour les sol sablonneux est de 0,5 (Dol, 2023; Sona & Dubé, 2021) tandis que la masse volumique des particules est de 2,7 g/cm³.

Tableau 3.5 Calculs intermédiaires de l' invariant d' hétérogénéité, diamètre moyen des particules et invariant d' hétérogénéité

d_α (cm)	$A = d_\alpha^3 \frac{m_{L_\alpha}}{m_L} \frac{(a_\alpha - a_L)^2}{a_L^2}$ (cm ³)									
0,375	5,04×10 ⁻²	4,94×10 ⁻²	4,87×10 ⁻²	4,22×10 ⁻²	4,26×10 ⁻²	4,51×10 ⁻²	5,08×10 ⁻²	4,96×10 ⁻²	5,05×10 ⁻²	5,37×10 ⁻²
0,1875	3,21×10 ⁻⁴	3,89×10 ⁻⁴	4,98×10 ⁻⁴	6,16×10 ⁻⁴	5,71×10 ⁻⁴	4,70×10 ⁻⁴	3,23×10 ⁻⁴	3,75×10 ⁻⁴	3,40×10 ⁻⁴	3,39×10 ⁻⁴
0,094	1,23×10 ⁻⁴	1,25×10 ⁻⁴	1,22×10 ⁻⁴	1,48×10 ⁻⁴	1,49×10 ⁻⁴	1,49×10 ⁻⁴	1,26×10 ⁻⁴	1,25×10 ⁻⁴	1,29×10 ⁻⁴	1,28×10 ⁻⁴
0,04725	1,22×10 ⁻⁵	1,23×10 ⁻⁵	1,21×10 ⁻⁵	1,47×10 ⁻⁵	1,47×10 ⁻⁵	1,48×10 ⁻⁵	1,23×10 ⁻⁵	1,21×10 ⁻⁵	1,26×10 ⁻⁵	1,27×10 ⁻⁵
0,02375	9,17×10 ⁻⁷	9,01×10 ⁻⁷	9,03×10 ⁻⁷	1,09×10 ⁻⁶	1,10×10 ⁻⁶	1,08×10 ⁻⁶	8,84×10 ⁻⁶	8,76×10 ⁻⁶	9,24×10 ⁻⁶	9,17×10 ⁻⁶
0,012	6,69×10 ⁻⁸	6,61×10 ⁻⁸	6,58×10 ⁻⁸	6,33×10 ⁻⁸	6,24×10 ⁻⁸	6,60×10 ⁻⁸	6,57×10 ⁻⁸	6,45×10 ⁻⁸	6,79×10 ⁻⁸	6,74×10 ⁻⁸
HIL=fp_{Σα} A	6,86×10⁻²	6,74×10⁻²	6,66×10⁻²	5,80×10⁻²	5,85×10⁻²	6,18×10⁻²	6,93×10⁻²	6,76×10⁻²	6,88×10⁻²	7,31×10⁻²

Le Tableau 3.6 récapitule pour chaque essai, les variances théoriques liées à l'erreur fondamentale ($s^2(\text{EFE})$), celles liées à l'erreur de ségrégation et groupement ($s^2(\text{ESG})$) ainsi que les variances théoriques liées à l'erreur totale ($s^2(\text{ÉTÉ})$).

Tableau 3.6 Récapitulatif des variances théoriques liées à l'erreur fondamentale, de groupement et ségrégation, et totale pour les différentes configurations d'essai calculées selon les équations (1.4), (1.3) et (1.2) respectivement

	Séparation rotative	Grappillage	Pelletage fractionné
Sec (sol sans fines)	Essai n°5	Essai n°4	Essai n°6
	$s^2(\text{EFE}) = 1,90 \times 10^{-4}$	$s^2(\text{EFE}) = 1,85 \times 10^{-4}$	$s^2(\text{EFE}) = 1,96 \times 10^{-4}$
	$s^2(\text{ESG}) = 0$	$s^2(\text{ESG}) = 1,85 \times 10^{-4}$	$s^2(\text{ESG}) = 9,80 \times 10^{-6}$
	$s^2(\text{ETE}) = 1,90 \times 10^{-4}$	$s^2(\text{ETE}) = 3,70 \times 10^{-4}$	$s^2(\text{ETE}) = 2,06 \times 10^{-4}$
Sec (sol avec fines)	Essai n°1	Essai n°2	Essai n°3
	$s^2(\text{EFE}) = 2,32 \times 10^{-4}$	$s^2(\text{EFE}) = 2,14 \times 10^{-4}$	$s^2(\text{EFE}) = 2,11 \times 10^{-4}$
	$s^2(\text{ESG}) = 0$	$s^2(\text{ESG}) = 2,14 \times 10^{-4}$	$s^2(\text{ESG}) = 1,06 \times 10^{-5}$
	$s^2(\text{ETE}) = 2,32 \times 10^{-4}$	$s^2(\text{ETE}) = 4,28 \times 10^{-4}$	$s^2(\text{ETE}) = 2,22 \times 10^{-4}$
10%eau (sol avec fines)	-	Essai n°7	Essai n°8
		$s^2(\text{EFE}) = 2,19 \times 10^{-4}$	$s^2(\text{EFE}) = 2,20 \times 10^{-4}$
		$s^2(\text{ESG}) = 2,19 \times 10^{-4}$	$s^2(\text{ESG}) = 1,10 \times 10^{-5}$
		$s^2(\text{ETE}) = 4,38 \times 10^{-4}$	$s^2(\text{ETE}) = 2,31 \times 10^{-4}$
10%huile (sol avec fines)	-	Essai n°9	Essai n°10
		$s^2(\text{EFE}) = 2,34 \times 10^{-4}$	$s^2(\text{EFE}) = 2,38 \times 10^{-4}$
		$s^2(\text{ESG}) = 2,34 \times 10^{-4}$	$s^2(\text{ESG}) = 1,19 \times 10^{-5}$
		$s^2(\text{ETE}) = 4,68 \times 10^{-4}$	$s^2(\text{ETE}) = 2,50 \times 10^{-4}$

La variance théorique liée à l'erreur fondamentale n'est pas fonction de la méthode d'échantillonnage employée mais seulement de la constitution du lot et de ses différentes fractions de densité et granulométriques. C'est pour cela que la seule différence de variance notable réside entre celles associées aux essais sans fines et celles associées aux essais avec fines. En effet, la suppression des particules fines de sol pour les essais n°4, 5 et 6 réduit l'hétérogénéité et la variance liée à l'erreur fondamentale avec une moyenne de $2,24 \times 10^{-4}g$ pour les essais avec fines et $1,90 \times 10^{-4}g$ pour les essais sans fines.

Le nombre d'incrément nécessaire à la constitution d'un échantillon pour les trois méthodes d'échantillonnage étudiées lors de la présente recherche servent à calculer les variances théoriques liées à l'erreur de ségrégation et de groupement selon la formule (1.3). Le nombre d'incrément est considéré comme égal à l'infini pour la séparation rotative, de un pour le grappillage, et de 20 pour le pelletage fractionné. Bien que l'erreur de ségrégation et

groupement soit fonction de l'erreur fondamentale, elle est également contrôlée notamment par le nombre d'incrémentes prélevés pour constituer chaque échantillon, lui-même en lien avec la méthode d'échantillonnage employée. Plus ce nombre est élevé, moins la variance due à l'erreur de ségrégation et groupement sera grande. Le grappillage consiste à prélever des échantillons en un unique incrément. Il d'agit donc de la méthode exerçant le moins de contrôle sur la variance due à l'erreur de ségrégation et groupement, suivie par le pelletage fractionné et ses 20 incrémentes. La variance due à l'erreur de ségrégation et groupement tend vers zéro avec la séparation rotative, les échantillons étant constitués par un flux de particules constant et donc, un très grand nombre d'incrémentes.

Les variances théoriques liées à l'erreur totale pour les différentes configurations d'essai sont obtenues en sommant les variances théoriques liées à l'erreur fondamentale et de ségrégation et groupement. Pour la séparation rotative, l'erreur totale est donc uniquement due à l'erreur fondamentale, pour le pelletage fractionné, elle est principalement influencée par l'erreur fondamentale également tandis que pour le grappillage, en théorie, elle est autant influencée par l'erreur fondamentale que par celle de ségrégation et groupement.

Le Tableau 3.7 présente les coefficients de variation théoriques associés aux variances théoriques liées à l'erreur totale du Tableau 3.6 pour les différentes configurations d'essai.

Tableau 3.7 Récapitulatif des coefficients de variation théoriques pour les différentes configurations d'essai calculés selon l'équation (2.9)

	Séparateur rotatif	Grappillage	Pelletage fractionné
Sec (sol sans fines)	Essai n°5 1,38%	Essai n°4 1,92%	Essai n°6 1,43%
Sec (sol avec fines)	Essai n°1 1,52%	Essai n°2 2,07%	Essai n°3 1,49%
10%eau (sol avec fines)	-	Essai n°7 2,09%	Essai n°8 1,52%
10%huile (sol avec fines)	-	Essai n°9 2,16%	Essai n°10 1,58%

Comme le montre ce Tableau 3.7, les coefficients de variation théoriques sont tous du même ordre de grandeur. Cependant, les coefficients de variation théoriques sont tout de même un

légèrement plus élevés pour le grappillage que pour les deux autres méthodes. Comme déjà évoqué en 3.1.2.3 au sujet des courbes d'erreur, pour un échantillonnage sec, la performance du pelletage fractionné à 20 incréments se rapproche plus de celle de la séparation rotative que du grappillage. Il est important de mentionner que ces variances théoriques et leurs CV associés représentent un minimum pour une granulométrie donnée, car ils ne quantifient pas la variance due aux erreurs autres que les erreurs fondamentales, de ségrégation et de groupement. Ainsi, bien que les variance et CV théoriques soient du même ordre de grandeur pour les trois méthodes d'échantillonnage, les variances et CV empiriques, qui mesurent la variabilité due à l'ensemble des sources d'erreurs, sont très différents comme il est possible de le constater ci-dessous en 3.2.3.2.

3.2.3.2 Coefficient de variation empirique CV_{emp}

Le Tableau 3.8 présente par essai, les pourcentages de briques moyens, la variance et les coefficients de variations obtenus en fonction. Il renseigne également les pourcentages de briques minimum et maximum relevés dans les 32 échantillons de chaque essai.

Tableau 3.8 Pourcentage de briques moyen, minimal et maximal, variance et coefficient de variation empirique associés à chaque configuration d'essai

	Séparateur rotatif	Grappillage	Pelletage fractionné
Sol sec sans fines	Essai n°5 Moyenne = 29,78% Min = 29,04% Max = 30,87% $s^2 = 2,65 \times 10^{-5}$ CV_{emp} = 1,73%	Essai n°4 Moyenne = 29,92% Min = 17,98% Max = 43,64% $s^2 = 5,01 \times 10^{-3}$ CV_{emp} = 23,86%	Essai n°6 Moyenne = 29,61% Min = 24,80% Max = 32,96% $s^2 = 3,92 \times 10^{-4}$ CV_{emp} = 6,69%
Sol sec avec fines	Essai n°1 Moyenne = 29,79% Min = 28,11% Max = 31,37% $s^2 = 4,88 \times 10^{-5}$ CV_{emp} = 2,34%	Essai n°2 Moyenne = 29,48% Min = 19,31% Max = 39,86% $s^2 = 3,62 \times 10^{-3}$ CV_{emp} = 20,42%	Essai n°3 Moyenne = 29,87% Min = 27,49% Max = 33,37% $s^2 = 2,34 \times 10^{-4}$ CV_{emp} = 5,12%
Sol à 10%eau (avec fines)	-	Essai n°7 Moyenne = 29,82% Min = 27,80% Max = 33,84% $s^2 = 5,40 \times 10^{-4}$ CV_{emp} = 7,79%	Essai n°8 Moyenne = 29,72% Min = 27,61% Max = 31,55% $s^2 = 1,21 \times 10^{-4}$ CV_{emp} = 3,70%
Sol à 10%huile (avec fines)	-	Essai n°9 Moyenne = 30,04% Min = 25,17% Max = 32,69% $s^2 = 2,18 \times 10^{-4}$ CV_{emp} = 4,90%	Essai n°10 Moyenne = 28,52% Min = 27,16% Max = 29,60% $s^2 = 3,16 \times 10^{-5}$ CV_{emp} = 1,97%

Le Tableau 3.9 vient compléter le Tableau 3.8 ci-dessus en apportant des éléments de compréhension de ce premier. Il permet également de conclure sur ce qui a été vu au cours de cette partie 3.2.3.

Tableau 3.9 Comparaison des coefficients de variation théoriques et empiriques des différentes configurations d'essai

	Séparateur rotatif	Grappillage	Pelletage fractionné	
Sec (sol sans fines)	Essai n°5 $CV_{théo} = 1,38\%$ $CV_{emp} = 1,73\%$ $\Delta CV_{emp-théo} = 0,35\%$	Essai n°4 $CV_{théo} = 1,92\%$ $CV_{emp} = 23,86\%$ $\Delta CV_{emp-théo} = 21,94\%$	Essai n°6 $CV_{théo} = 1,43\%$ $CV_{emp} = 6,69\%$ $\Delta CV_{emp-théo} = 5,26\%$	
	Sec (sol avec fines)	Essai n°1 $CV_{théo} = 1,52\%$ $CV_{emp} = 2,34\%$ $\Delta CV_{emp-théo} = 0,82\%$	Essai n°2 $CV_{théo} = 2,07\%$ $CV_{emp} = 20,42\%$ $\Delta CV_{emp-théo} = 18,35\%$	Essai n°3 $CV_{théo} = 1,49\%$ $CV_{emp} = 5,12\%$ $\Delta CV_{emp-théo} = 3,63\%$
		10% eau (sol avec fines)	-	Essai n°7 $CV_{théo} = 2,09\%$ $CV_{emp} = 7,79\%$ $\Delta CV_{emp-théo} = 5,70\%$
10% huile (sol avec fines)			-	Essai n°9 $CV_{théo} = 2,16\%$ $CV_{emp} = 4,90\%$ $\Delta CV_{emp-théo} = 2,74\%$

À la vue du Tableau 3.8, il peut être constaté que les coefficients de variation varient beaucoup selon la méthode d'échantillonnage et sont diminués en présence de phase liquide.

En comparant les méthodes d'échantillonnage, le constat est toujours le même que précédemment : le grappillage est la méthode d'échantillonnage générant les coefficients de variations empirique les plus grands, suivi du pelletage fractionné et enfin de la séparation rotative (pour les essais secs seulement). Le pelletage fractionné, bien que beaucoup moins performant que la séparation rotative, est tout de même plus proche de cette dernière que du grappillage en termes de coefficient de variation.

La présence d'une phase liquide améliore la représentativité (1.2.1) des échantillonnages tant par pelletage fractionné que par grappillage. L'huile permet d'avoir des coefficients de variation empiriques presque divisés par deux par rapport à l'eau. En effet, les différences de propriétés physiques de ces deux phases liquides, notamment la viscosité, peuvent être à l'origine de ce phénomène en améliorant la cohésion des particules. Autrement, il peut être

également constaté que la présence d'une phase liquide diminue l'écart de représentativité d'échantillonnage entre les méthodes étudiées d'un point de vue du coefficient de variation.

Autrement, la différence de tendance qui peut être observée réside dans la comparaison des coefficients de variations des essais secs avec et sans fines. En effet, pour la séparation rotative, la présence de fines fait augmenter le coefficient de variation tandis que pour le pelletage fractionné et le grappillage, la présence de fines semble faire diminuer le coefficient de variation. En réalité, comme l'erreur de ségrégation et de groupement pour la séparation rotative est très faible voire nulle (Tableau 3.9), les coefficients de variation des essais de séparation rotative sont quasiment essentiellement influencés par l'erreur fondamentale qui, comme expliqué plus haut dans cette partie, est plus importante pour les essais avec fines que sans fines. Les coefficients de variation empiriques ne suivent pas cette tendance pour les essais de grappillage et de pelletage fractionné car ils sont plus influencés par d'autres sources d'erreurs, dont principalement l'erreur de ségrégation et de groupement ou de délimitation et d'extraction pour le grappillage notamment, plus que par l'erreur fondamentale. L'explication de ce qui est avancé peut se voir dans les écarts entre les coefficients de variation théoriques et empirique présentés dans le Tableau 3.9. En effet, pour les essais de séparation rotative, ces écarts sont relativement faibles, ce qui signifie que le coefficient de variation empirique est principalement dû à l'erreur fondamentale dont la variance est calculée dans le Tableau 3.6. Pour les essais de grappillage et pelletage fractionné en revanche, l'écart est plus conséquent, ce qui est sûrement dû à la sous-estimation de l'erreur de ségrégation et de groupement dans le Tableau 3.6, voire à l'omission d'autres sources d'erreurs.

En résumé, la séparation rotative constitue la méthode d'échantillonnage la plus représentative possible car les sources d'erreurs autres que fondamentales sont presque toutes minimisées. Ceci n'est pas le cas du pelletage fractionné et du grappillage, grandement influencés par des phénomènes de ségrégation et groupement. Cela dit, le pelletage fractionné pour un nombre de 20 incréments est tout de même bien meilleur que la séparation rotative, surtout en présence d'huile. En effet, elle permet une meilleure cohésion du sol ce qui minimise, l'effet des phénomènes de ségrégation et de groupement sous l'action de la gravité.

3.3 Analyse d'image

Comme détaillé dans la section 2.4.4, des dizaines de configurations de réseaux neuronaux ont été essayées. Les performances d'entraînement et de validation ainsi que les performances du réseau sur les données de test ont été évaluées à chaque simulation. En effet, après l'entraînement, un graphique de performance est généré (comme expliqué en 2.4.4.1). Ensuite, après avoir fait des prédictions avec le modèle sur les données de test, un graphique de régression est fait (comme expliqué en 2.4.4.2). Parmi les dizaines de réseaux neuronaux testés, les trois configurations générant les meilleures performances lors de l'entraînement, la validation et surtout lors du test, ont été retenues.

Le script de la première configuration de réseau neuronal est à retrouver dans Algorithme-A II-1 de l'ANNEXE II. Cette configuration possède trois couches de convolutions et de *maxpooling* alternées, avec notamment une première couche de convolution à huit filtres de 4×4 pixels, une deuxième à 16 filtres de 4×4 pixels et une troisième à 32 filtres de 4×4 pixels. Quant aux hyperparamètres, le nombre d'époques est de 15 et la taille du *batch* est de 16.

Le script de la deuxième configuration de réseau neuronal est à retrouver dans Algorithme-A II-2 de l'ANNEXE II. Cette configuration possède également trois couches de convolutions et de *maxpooling* alternées, avec les mêmes nombre et tailles de filtres. La différence avec la première est la taille du *batch* qui est de 32.

Le script de la troisième configuration de réseau neuronal est à retrouver dans Algorithme-A II-3 de l'ANNEXE II. Dans cette configuration, il y a également trois couches de convolutions et de *maxpooling* alternées, avec cependant, une première couche de convolution à 16 filtres de 4×4 pixels et une troisième à 32 filtres de 4×4 pixels et une troisième à 64 filtres de 4×4 pixels. La complexité du modèle est de ce fait augmentée dans cette version par rapport aux deux autres. La taille du *batch* est aussi de 32.

Pour chacune de ces trois configurations, six simulations ont été lancées successivement afin d'évaluer la variabilité de modèles que chaque configuration génère. Autrement dit, pour chacune des trois configurations de couches et d'hyperparamètres, six modèles ont été créés,

entraînés, validés et testés avec les mêmes jeux de données afin de voir s'ils aboutissaient aux mêmes performances d'une simulation à l'autre. Les graphiques de performance et validation ainsi que les graphiques de régression et les performances lors du test sont présentés dans l'ANNEXE III en Tableau-A III-1 pour la configuration de réseau neuronal n°1, en Tableau-A III-2 pour la configuration de réseau neuronal n°2, et en Tableau-A III-3 pour la configuration de réseau neuronal n°3. Les points des graphiques de régression ont été colorés par la même occasion afin d'évaluer visuellement si les pourcentages de briques étaient plus ou moins bien prédits selon l'aspect des échantillons (incluant les échantillons secs sans fines, secs avec fines, humides ou avec huile) ou les méthodes d'échantillonnages employées pour les générer (grappillage, pelletage fractionné, séparation rotative et séparation à chutes).

Les simulations générant les meilleures performances parmi les six simulations sont la n°1 pour la première et la deuxième configuration, et la n°6 pour la troisième configuration. Les graphiques de performance d'entraînement et de régression sont à retrouver respectivement en Figure 3.22, Figure 3.23 et Figure 3.24.

Pour rappel, 1280 photos sont utilisées pour l'entraînement des modèles, 319 pour la validation et 319 pour le test.

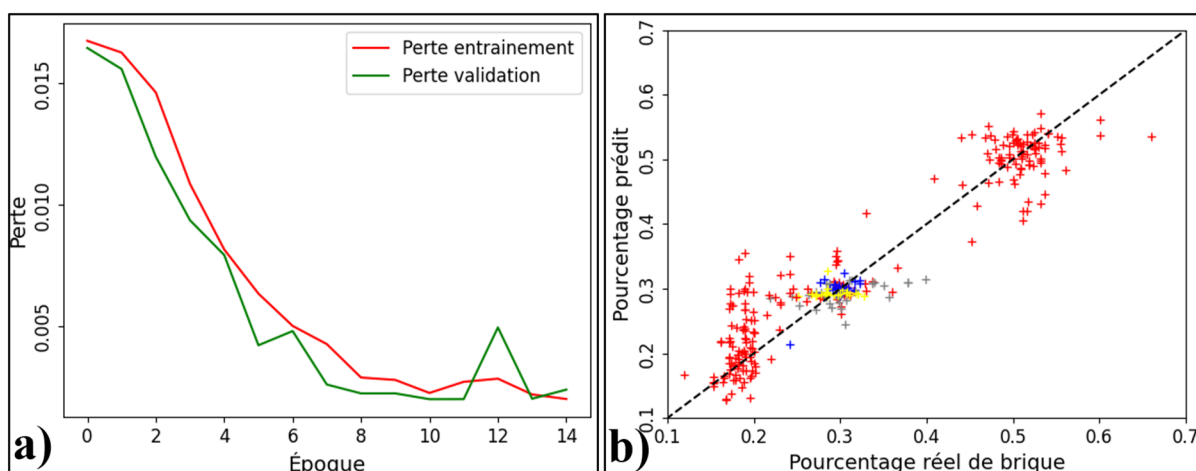


Figure 3.22 a) Graphique des pertes d'entraînement et de validation au fil des époques et b) graphique de régression lors du test et performances du test pour la simulation n°1 effectuée avec la configuration de réseau n°1 - les points rouges représentent les échantillons secs sans fines, en gris les secs avec fines, en bleu les humides, en jaune ceux avec huile

Pour cette simulation n°1 du réseau neuronal n°1, la perte obtenue lors du test est de 0,0011 tandis que l'erreur quadratique moyenne est de 0,0337.

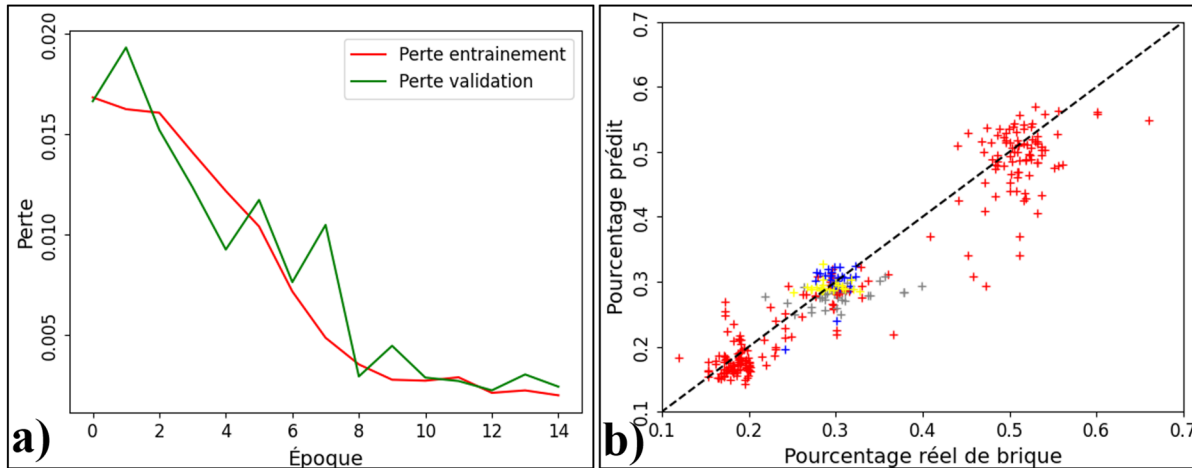


Figure 3.23 a) Graphique des pertes d'entraînement et de validation au fil des époques et b) graphique de régression lors du test et performances du test pour la simulation n°1 effectuée avec la configuration de réseau n°2 - les points rouges représentent les échantillons secs sans fines, en gris les secs avec fines, en bleu les humides, en jaune ceux avec huile

Pour cette simulation n°1 du réseau neuronal n°2, la perte obtenue lors du test est de 0,0016 tandis que l'erreur quadratique moyenne est de 0,0401.

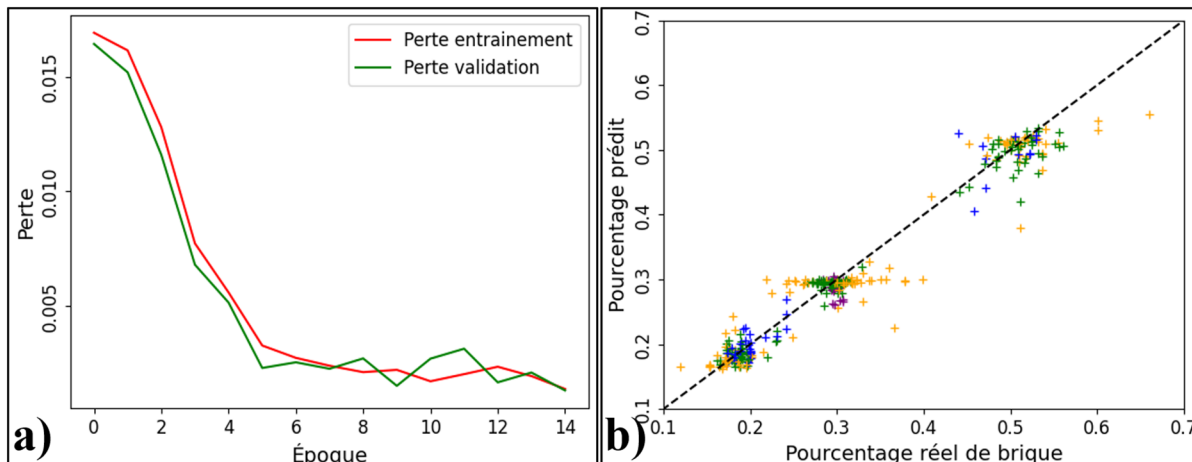


Figure 3.24 a) Graphique des pertes d'entraînement et de validation au fil des et b) graphique de régression lors du test et performances du test pour la simulation n°6 effectuée avec la configuration de réseau n°3 - les points jaunes sont associés au grappillage, les points verts au pelletage fractionné, les points violets à la séparation rotative et les points bleus à la séparation à chutes

Pour cette simulation n°6 du réseau neuronal n°3, la perte obtenue lors du test est de 0,000838 tandis que l'erreur quadratique moyenne est de 0,0290. Cette dernière simulation est celle garantissant les meilleures prédictions entre toutes les simulations, toutes configurations de réseaux confondues.

Les graphiques de performance d'entraînement et validation, les graphiques de régression et les performances lors du test sont présentés dans l'ANNEXE III pour les trois simulations. Le réseau n°3 étant le plus performant, neuf simulations supplémentaires ont été réalisées afin de mieux l'étudier ultérieurement. Les performances associées sont également présentées en ANNEXE III. Les graphiques de performances permettent de constater différents éléments :

- 1) Il y a des cas où le réseau neuronal prédit un même pourcentage de briques (égal à la moyenne soit à 32,42%) pour toutes les photographies du test, ce qui fait apparaître les points du graphique de régression alignés horizontalement. Ce cas a concerné les trois réseaux neuronaux à travers la simulation n°5 pour le premier (Tableau-A III-1), la simulation n°6 pour le second (Tableau-A III-2) et les simulations n°3, 8, 9 et 14 pour le troisième (Tableau-A III-3). Cette situation indique que le réseau ne converge pas.

Les graphiques de régression utilisant le code couleur mettant en avant les points associés aux différents aspects de photographies (les points rouges représentant les échantillons secs sans fines, en gris les secs avec fines, en bleu les humides et en jaune ceux avec huile) ont permis de constater que les pourcentages de briques des photographies d'échantillons de sol sec avec fines sont généralement mal prédits. En effet, les réseaux leur prédit un pourcentage correspondant environ au pourcentage moyen comme il est le cas dans

l'exemple en Figure 3.25. Cette estimation semble être celle faite par défaut quand le réseau ne parvient pas à détecter les particules de briques.

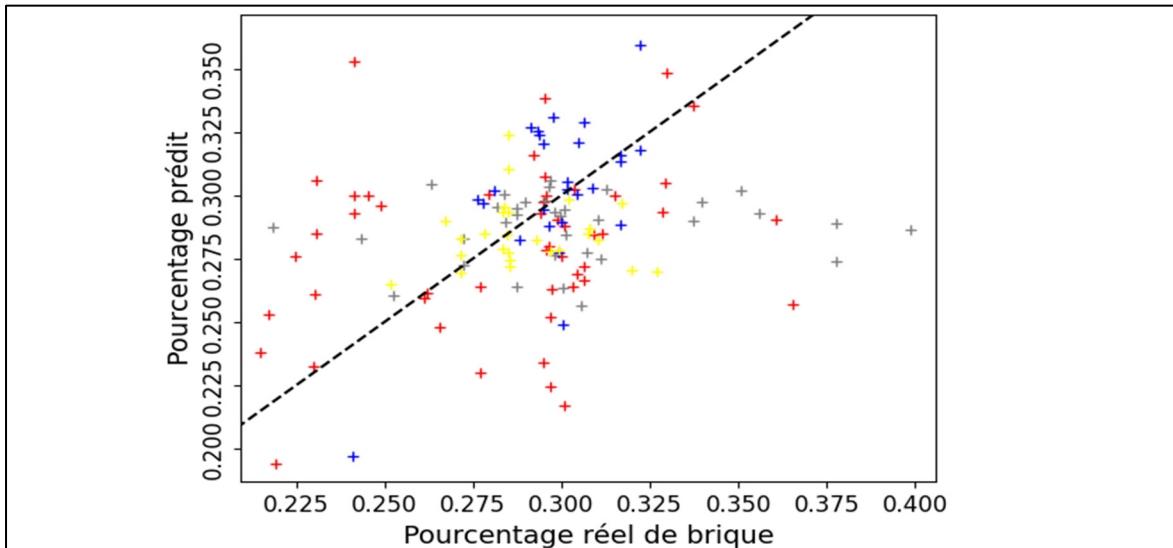


Figure 3.25 Zoom sur le graphique de régression présenté dans le Tableau-A III-2 associé à la simulation n°3 du réseau neuronal n°2

- 2) Sur beaucoup de graphiques, il est constaté que les réseaux donnent des prédictions moins précises pour les échantillons contenant autour de 50% de briques. En effet, sur ces graphiques, les deux nuages de points associés aux échantillons issus des lots à 20 et 30% de briques sont plus proches de la courbe de régression que pour les points associés aux échantillons issus du lots à 50% de briques qui sont plus dispersés. Ce cas a concerné toutes les simulations pour lesquelles les réseaux convergeaient, allant de la simulation n°5 du réseau n°3 dans le pire des cas à la simulation n°6 du réseau n°3 dans le meilleur des cas (Figure 3.24).

En résumé, bien que les deux premières configurations de réseau neuronal aient donné des modèles qui convergent de manière performante, les meilleurs modèles sont générés par le réseau n°3 du fait de sa complexité supérieure adaptée à une telle variété et quantité de photographies. Sa capacité à générer des modèles identiques d'une simulation à l'autre constitue néanmoins son principal point faible. Cela dit, puisque la bibliothèque *Keras* intégrée

à *TensorFlow* permet de sauvegarder les modèles entraînés à l'aide de la fonction *model.save* et de les réutiliser plus tard avec la fonction *load_model*, il suffit donc de générer un bon modèle.

3.4 Lien entre les résultats d'échantillonnage et d'analyse d'images

3.4.1 Biais relatifs, variances relatives et degrés de représentativité liés à l'échantillonnage et au réseau

Dans l'optique de faire le lien entre l'échantillonnage et l'analyse d'images et de répondre à l'objectif principal de cette recherche, une étude des biais, variances et degrés de représentativité d'échantillonnage et d'analyse d'images est effectuée (Tableau 3.10). Les biais, variances et degrés de représentativité d'échantillonnage caractérisent les erreurs dues à l'échantillonnage, soient les erreurs entre les pourcentages de briques prédits dans échantillons grâce à la procédure d'échantillonnage ($\%Briques_{éch}$), par rapport à la réalité du lot ($\%Briques_{lot}$). Les biais variances et degrés de représentativité de l'analyse d'images caractérisent les erreurs d'estimation du réseau neuronal, soient les erreurs entre les pourcentages de briques prédits dans échantillons grâce à la procédure d'échantillonnage ($\%Briques_{éch}$) et les pourcentages prédits par le réseau ($\%Briques_{réseau}$).

Ces biais, variances et degrés de représentativité sont calculés à partir des données de pourcentages de briques $\%Briques_{lot}$, $\%Briques_{éch}$ et $\%Briques_{réseau}$ associés aux 319 photographies utilisées pour le test des modèles. Ces biais, variances et degrés de représentativité d'échantillonnage et d'analyse d'images couplés caractérisent ainsi l'erreur d'estimation du réseau neuronal par rapport à la réalité du lot.

Le Tableau 3.10 présente également la variance relative associées aux erreurs relatives d'échantillonnage, ainsi que le degré de représentativité, qui a pour avantage de combiner le biais et la variance.

Tableau 3.10 Biais relatifs, variances relatives et degrés de représentativité associés aux erreurs entre les %Briques_{lot} et %Briques_{éch} puis entre les %Briques_{éch} et %Briques_{réseau} pour différentes simulations du réseau neuronal n°3

	Simulation	Biais relatif	Variance relative	Degré de représentativité
Échantillon par rapport au lot		$6,09 \times 10^{-4}$	$1,18 \times 10^{-2}$	$1,24 \times 10^{-2}$
Prédictions du réseau par rapport à l'échantillon	N°4	$3,35 \times 10^{-3}$	$2,63 \times 10^{-2}$	$2,96 \times 10^{-2}$
	N°5	$8,73 \times 10^{-3}$	$1,87 \times 10^{-2}$	$2,74 \times 10^{-2}$
	N°6	$1,25 \times 10^{-6}$	$1,48 \times 10^{-2}$	$1,48 \times 10^{-2}$
	N°7	$8,41 \times 10^{-7}$	$1,93 \times 10^{-2}$	$1,93 \times 10^{-2}$
	N°10	$4,31 \times 10^{-3}$	$2,04 \times 10^{-2}$	$2,47 \times 10^{-2}$
	N°11	$3,24 \times 10^{-3}$	$1,69 \times 10^{-2}$	$2,02 \times 10^{-2}$
	N°12	$4,72 \times 10^{-6}$	$2,01 \times 10^{-2}$	$2,01 \times 10^{-2}$
	N°13	$3,10 \times 10^{-4}$	$1,63 \times 10^{-2}$	$1,67 \times 10^{-2}$
	N°15	$2,74 \times 10^{-3}$	$2,65 \times 10^{-2}$	$2,92 \times 10^{-2}$

Le degré de représentativité lié à l'échantillonnage est de $1,24 \times 10^{-2}$. Les degrés de liberté liés au réseau varient d'une simulation à l'autre. Le plus petit enregistré est de $1,48 \times 10^{-2}$ pour la simulation n°6 effectuée avec la structure de réseau retenue, indiquant qu'il s'agit de la simulation qui garantit les meilleures prédictions. Les degrés de liberté associés à l'échantillonnage et au réseau sont donc du même ordre de grandeur. La lecture des pourcentages prédits doit se faire en considérant les marges d'erreurs en conséquent.

3.4.2 Nombre de faux positifs et négatifs générés lors de l'échantillonnage et l'analyse d'images

Pour rappel, la notion de faux positifs et faux négatifs a été évoquée pour la première fois en 2.2.4. Sont appelés des faux positifs, des échantillons analysés comme contenant plus de 50% de particules de brique mais pourtant issus d'un lot à 20 ou 30% de briques. Sont appelés des faux négatifs, des échantillons analysés comme contenant moins de 50% de particules de brique mais pourtant issus d'un lot à 50% de briques. Leur étude est essentielle car l'enjeu d'une bonne classification du sol est de taille comme expliqué en introduction.

Le Tableau 3.11 présente le nombre de faux négatifs et positifs générés par l'échantillonnage et l'analyse d'images pour les 319 photographies testées dans le réseau neuronal. Aucun faux positif n'a été à déplorer que ce soit pour l'échantillonnage que pour l'analyse d'images.

Tableau 3.11 Nombre de faux négatifs générés par l'échantillonnage et l'analyse d'images

Nombre de faux négatifs	
Échantillonnage	32
Simulation n°4	28
Simulation n°5	81
Simulation n°6	37
Simulation n°7	28
Simulation n°10	21
Simulation n°11	28
Simulation n°12	42
Simulation n°13	50
Simulation n°15	34

Aucun faux négatif n'est produit pour les échantillons issus de lot à 20 et 30%, mais des faux positifs ont été générés pour ceux issus des lots à 50%. L'échantillonnage a permis d'estimer ce nombre à 32 contre 37 pour la simulation n°6 effectuée avec le réseau neuronal. Ce nombre étant parmi les plus proches des 32 de référence, ce qui confirme l'efficacité du modèle créé lors de cette simulation par rapport aux autres.

CHAPITRE 4

DISCUSSION

4.1 Discussion sur l'analyse granulométrique

Les résultats présentés en 3.1 peuvent être nuancés. En effet tout d'abord, des pertes de sol et briques particulières ont été à déplorer au cours des essais. Par ailleurs, en raison de la nature friable des particules de brique, des fragments de taille inférieure à 1,25 mm ont été générés à partir des particules de taille initiale comprise entre 2,5 et 5 mm introduites en début d'essai. L'identification des briques se faisant seulement dans les fractions 1,25 – 2,5 mm et 2,5 – 5 mm (2.2.4), les fragments inférieurs à 1,25 mm sont donc omis de la quantification.

4.1.1 Pertes de sol et briques au cours de l'échantillonnage

Le Tableau 4.1 présente les pourcentages de perte totale de sol et briques au cours des dix essais. Ces pourcentages ont été calculés en divisant la masse sèche totale des échantillons obtenus en bout de ligne de chaque essai par la masse sèche initiale du lot.

Tableau 4.1 Perte totale de sol et briques au cours des essais

	Séparateur rotatif	Grappillage	Pelletage fractionné
Sol sec sans fines	Essai n°5 0,07%	Essai n°4 0,07%	Essai n°6 0,17%
Sol sec avec fines	Essai n°1 0,27%	Essai n°2 0,28%	Essai n°3 0,27%
Sol à 10%eau (avec fines)	-	Essai n°7 0,48%	Essai n°8 0,46%
Sol à 10%huile (avec fines)	-	Essai n°9 1,08%	Essai n°10 2,85%

Il peut être constaté la perte est croissante entre les essais de sol sec sans fines, secs avec fines, avec eau et avec huile. En effet, l'échantillonnage de sol avec fines générerait une importante quantité de poussières lors des étapes de séparation et de prise photographique, ce qui n'était

pas le cas lors des essais de sol sans fines. Ainsi, les granulométries sous-estiment probablement la part des particules fines, notamment celles inférieures à 80 µm. Le Tableau 4.1 met également en évidence le fait que la présence d'une phase liquide augmente les pertes de sol et briques. En effet, plus de matière se perd sur les outils et contenants d'échantillonnage et de prise de photos quand elle est humide. La présence d'huile semble d'autant plus accentuer ce type de pertes par rapport à l'eau. Cette tendance en partie due à l'élimination de la matière organique et l'humidité présentes dans le mélange de sol et briques conjointement à l'élimination de l'huile dans le four à moufles à 550°C. Le pourcentage d'humidité et de matière organique présentes dans le mélange de sol et briques sec a été estimé à environ 0,05% et 0,4% respectivement selon la procédure d'estimation décrite dans la méthode d'analyse MA. 100 – S.T. 1.1 appliquée à cinq échantillons représentatifs (Centre d'expertise en analyse environnementale du Québec, 2017). Ainsi, conjointement à l'élimination de l'huile, une perte de masse de 0,45% d'humidité et de matières organiques est attendue.

Le Tableau 4.2 présente les pourcentages de perte générée par les étapes de séparation, de prise de photos et d'analyse granulométrique pour les dix essais. Les pourcentages présentés dans le tableau ont été calculés en divisant la masse totale perdue lors de l'étape indiquée en tête de colonne par la masse initiale du lot. Pour les essais avec phase liquide, ce ratio a été fait avec des masses humides, puisque les pertes comprenaient également des pertes de phase liquide sans que ces dernières puissent être quantifiées précisément. C'est pour cette raison que les pourcentages de pertes associés ont été soulignés.

Tableau 4.2 Pourcentages de perte lors des différentes étapes (souligné quand le ratio est fait avec des masses humides)

Essai	Séparation	Prise de photos	Granulométrie
N°1 - sol sec avec fines, séparation rotative	0,14%	0,11%	0,01%
N°2 - sol sec avec fines, grappillage	0,12%	0,14%	0,02%
N°3 - sol sec avec fines, pelletage fractionné	0,09%	0,16%	0,02%
N°4 - sol sec sans fines, grappillage	0,02%	0,03%	0,01%
N°5 - sol sec sans fines, séparation rotative	0,05%	0,01%	0,01%
N°6 - sol sec sans fines, pelletage fractionné	0,09%	0,05%	0,03%
N°7 - sol 10%eau avec fines, grappillage	<u>0,41%</u>	<u>0,70%</u>	<u>0,03%</u>
N°8 - sol 10%eau avec fines, pelletage fractionné	<u>0,71%</u>	<u>0,85%</u>	<u>0,03%</u>
N°9 - sol 10%huile avec fines, grappillage	<u>0,31%</u>	<u>0,38%</u>	<u>0,03%</u>
N°10 - sol 10%huile avec fines, pelletage fractionné	<u>1,77%</u>	<u>0,43%</u>	<u>0,03%</u>

Le Tableau 4.2 permet de constater que les étapes de séparation et de prise de photos génèrent approximativement les mêmes pertes par essai. Il permet également de constater que les pertes générées lors de l'analyse granulométrique étaient systématiquement inférieures ou égales à 0,03%, ce qui est bien en dessous de la limite des 0,3% exigée par la norme BNQ 2501-025/2013. Le Tableau 4.2 fait également remarquer que la perte par étape est plus importante pour les essais de sol sec avec fines par rapport à leurs équivalents sans fines, et qu'elle est également plus importante pour les essais de sol humides par rapport aux essais de sol secs.

Le Tableau 4.3 présente quant à lui les pourcentages de perte spécifique des phases liquides au cours des essais.

Tableau 4.3 Pourcentages de perte de phase liquide au cours des essais (souligné quand le ratio est fait avec des masses humides)

	% Perte de phase liquide
N°7 - sol 10%eau avec fines, grappillage	<u>7,08%</u>
N°8 - sol 10%eau avec fines, pelletage fractionné	<u>11,72%</u>
N°9 - sol 10%huile avec fines, grappillage	<u>1,50%</u>
N°10 - sol 10%huile avec fines, pelletage fractionné	<u>0,64%</u>

Ce tableau permet de constater qu'il y a des pertes de phase liquide plus importantes à déplorer pour les essais avec eau. Ce fait s'explique probablement par les effets d'évaporation. En effet,

les pertes sont d'autant plus conséquentes pour l'essai de pelletage fractionné qui dure plus longtemps que le grappillage avec 2h40 pour le pelletage contre 30 minutes pour le grappillage.

4.1.2 Pertes de brique particulière pendant l'échantillonnage ainsi que dans les fractions inférieures à 1,25 mm

Le Tableau 3.3 présente les biais d'échantillonnage (m^2) des dix essais. De par le processus d'échantillonnage total des lots lors des essais échantillonnages réalisés au cours de cette recherche, l'étude des biais est assimilable à une étude des pertes et des omissions des fragments de briques compris dans les fractions inférieures à 1,25 mm. Le Tableau 4.4 présente les pourcentages de pertes et omissions confondues de particules de briques au cours des dix essais réalisés au cours de cette recherche.

Tableau 4.4 Pertes et omissions de particules de briques au cours des essais en pourcentage massique (souligné quand le ratio est fait avec des masses humides)

Essai	Pourcentage de particules de brique perdues
N°1 - sol sec avec fines, séparation rotative	0,93%
N°2 - sol sec avec fines, grappillage	0,81%
N°3 - sol sec avec fines, pelletage fractionné	0,71%
N°4 - sol sec sans fines, grappillage	0,72%
N°5 - sol sec sans fines, séparation rotative	0,81%
N°6 - sol sec sans fines, pelletage fractionné	1,47%
N°7 - sol 10%eau avec fines, grappillage	<u>1,23%</u>
N°8 - sol 10%eau avec fines, pelletage fractionné	<u>1,35%</u>
N°9 - sol 10%huile avec fines, grappillage	<u>1,16%</u>
N°10 - sol 10%huile avec fines, pelletage fractionné	<u>1,90%</u>

Le principal constat permis par le Tableau 4.4 est que plus de pertes sont à déplorer pour les essais de sol avec phase liquide. En effet, au cours de ces essais, plus de sol et briques se perdaient sur les outils d'échantillonnage et contenants.

4.2 Discussion sur l'analyse d'images

Les résultats présentés en 3.3 mettent en avant le fait qu'il peut être difficile pour les réseaux d'estimer précisément les valeurs de pourcentages de briques extrêmes bien que la base de données photographique ait été rééquilibrée. En effet, les pourcentages liés aux lots à 50% de briques constituent les valeurs les plus extrêmes des jeux de données d'entraînement, de validation et de test à côté des échantillons issus de lots à 20 et 30% qui sont des valeurs plutôt proches. Un meilleur rééquilibrage aurait été d'ajouter des photographies d'échantillons à environ 40% de briques en nombre similaire à celui des photographies d'échantillons à environ 20, 30 et 50% de briques pour pallier le manque entre les photographiques de sols à 30 et 50%.

Les modèles produits au cours de cette recherche convergent, mais leurs performances semblent atteindre une même limite quelles que soient les configurations essayées. En effet, les meilleurs modèles avaient des pertes autour de 0,012 et des racines carrées d'erreurs quadratiques moyennes à 0,035 pour des graphiques de régression corrects mais qui semblent être améliorables. L'augmentation de la complexité du réseau et/ou l'augmentation et l'équilibrage de la base de données photographique sont des pistes d'amélioration. Cependant, la complexification, que ce soit par l'ajout de couches et de filtres et l'augmentation de la taille du *batch* peuvent rapidement faire basculer l'entraînement du modèle en surentraînement sans passer par la génération d'un modèle optimal. En effet, l'étude des variabilités en 3.3 démontre ceci à travers les différentes simulations réalisées pour trois différentes configurations de réseau neuronal. En plus de cela, ces mécanismes de complexification ainsi que l'utilisation d'une base plus conséquente nécessitent tous deux une grande capacité en mémoire.

4.3 Lecture croisée des résultats d'analyses granulométriques et d'images

L'étude des biais présentée en 3.4 permet de constater que les prédictions du réseau neuronal par rapport aux prédictions d'échantillonnage sont correctes en comparaison du biais d'échantillonnage déjà présent. L'étude du nombre de faux positifs et faux négatifs générés par l'échantillonnage et l'analyse d'images permet de constater que le réseau neuronal identifie globalement le même nombre de faux positifs et négatifs que l'échantillonnage.

Cependant, en contexte réel, il est rare que l'ensemble d'un lot soit traité afin de le caractériser. Une poignée d'échantillons seulement est analysée. Il suffit donc que certains échantillons soient modérément représentatifs (1.2.1) à l'origine pour diverses raisons comme l'emploi d'une méthode de prélèvement approximative, pour que l'erreur liée au réseau, aussi minimisée soit-elle, fausse la classification du sol. L'analyse d'images est très prometteuse, mais pour l'instant, ses biais d'interprétations s'ajoutent à ceux liés à l'échantillonnage plutôt que les réduire, pouvant ainsi compromettre la bonne caractérisation du sol étudié.

CONCLUSION

1) Conclusion sur l'échantillonnage

Les résultats d'échantillonnage de cette recherche ont permis dans un premier temps de confirmer la hiérarchie de représentativité (1.2.1) déjà connue (1.2.4), tant à travers les résultats granulométriques que de pourcentage de briques, des trois méthodes d'échantillonnage qui y ont été traitées. Ces dernières sont le grappillage, le pelletage fractionné et la séparation rotative (utilisée pour les essais secs seulement) dans l'ordre croissant de représentativité. Cependant, un apport quant à la quantification de ces différences de représentativités est tout de même à constater. En effet, pour un pelletage fractionné en 20 incréments, la représentativité granulométrique et de pourcentage de briques obtenue entre les échantillons générés est plus de l'ordre de celle de la séparation rotative que du grappillage bien que la séparation rotative reste tout de même meilleure. En effet, la séparation rotative avec son flux constant de particules arrivant sur le cône rotatif du dispositif de séparation, garantit une séparation du lot presque particule par particule, minimisant ainsi les erreurs de ségrégation et de groupement pour ne laisser place qu'à l'erreur fondamentale (3.2.3). Le pelletage fractionné étant réalisé en 20 incréments sur des échantillons de 300 g de masse sèche, chaque pelletée contenait 15 g de particules, ce qui est relativement faible. C'est le fait que les pelletées soient petites qui limite ainsi le nombre de particules par incréments, et qui a permis de rapprocher les résultats des essais de pelletage fractionné de ceux de séparation rotative. Les écarts qui demeuraient étaient principalement dus aux erreurs de ségrégation et groupement, elles même causées par le fait que l'outil d'échantillonnage était incorrect et qu'il permettait d'échantillonner petites masses par petites masses, certes, mais que ces dernières contenaient un nombre de particules supérieur à ce que le processus de séparation rotative permet avec son flux continu.

La présence d'une phase liquide améliore considérablement la représentativité des échantillonnages, que ce soit en termes de granulométrie qu'en termes de pourcentage de briques. En effet, la représentativité a été en particulier améliorée dans le cas du grappillage, qui est une méthode bien moins précise que le pelletage fractionné. La présence de phase liquide a donc considérablement atténué les disparités de représentativité entre les deux

méthodes, bien que le pelletage fractionné soit resté tout de même plus représentatif que le grappillage. En effet, l'introduction d'une phase liquide a engendré une meilleure cohésion des particules de sol, rendant leur répartition plus uniforme par rapport à un échantillon sec. Dans un échantillon sec, les plus grandes particules avaient tendance à se retrouver en surface tandis que les plus petites se déposaient au fond. L'ajout d'une phase liquide à une concentration de 10% en masse a permis de limiter cette ségrégation des particules dans le lot, sans non plus en créer d'autre puisqu'à une telle concentration, il n'y avait pas d'eau ou d'huile libre. De plus, l'introduction d'huile a légèrement amélioré la représentativité de l'échantillonnage par rapport à l'eau. Cette amélioration peut être attribuée aux différences de propriétés physiques entre ces deux phases liquides, notamment leur viscosité.

La présence de fines impacte essentiellement l'erreur fondamentale. En effet dans la théorie, l'erreur fondamentale n'est pas fonction de la méthode d'échantillonnage employée mais seulement de la constitution du lot et de ses différentes fractions de densité et granulométriques. Le fait d'avoir une gamme granulométrique moins importantes pour les lots des essais sans fines réduit donc l'hétérogénéité et par conséquent, la variance liée à l'erreur fondamentale également (3.2.3). Cela dit, cet effet est minime à côté de celui causé par les erreurs de ségrégation et de groupement quand il y en a. C'est pour cela que l'effet de la présence de fines se constate uniquement entre les essais de séparations rotative, qui est une méthode d'échantillonnage minimisant ces sources d'erreur.

Cependant, au cours des étapes de séparation et de prise de photos durant les dix essais d'échantillonnage, des pertes de sol notamment des particules les plus fines ont pu légèrement biaiser les résultats granulométriques présentés en 3.1. Ces biais ont pu affecter les résultats des essais secs mais surtout des essais humides. En effet, pour les essais secs avec fines, les particules inférieures à 80 μm ont été sous-estimées tandis que pour les essais humides, ce sont les particules fines inférieures à 160 μm qui ont sûrement été sous-estimées. En plus de cela dans le cas de l'huile, l'élimination d'approximativement 0,45% de matière organique et d'humidité est venue s'ajouter et se confondre avec les pertes de sol et briques. Les pertes de

particules de briques quant à elles sont essentiellement dues à leur omission suite à leur fragmentation dans des fractions non analysées au cours du tamisage.

2) Conclusion sur le réseau neuronal et la lecture croisée

Dans un premier temps, les résultats liés au réseau neuronal de cette recherche a permis de mettre en lumière deux choses. La première étant qu'une base de données photographique déséquilibrée menait à des mauvaises prédictions par le réseau, des pourcentages de briques des photos d'échantillon avec les pourcentages les plus extrêmes de la base. Il a ensuite été constaté qu'un rééquilibrage par des processus d'amplification était indispensable pour la représentativité des données extrêmes et était plus concluant que l'attribution de poids aux photographies sous-représentées lors de l'entraînement via l'ajustement de la fonction de perte.

Cette recherche a surtout permis de constater que plus une base de données photographique est diversifiée, plus la complexité du réseau neuronal convolutif doit être importante afin de lui permettre de converger et de prédire correctement toutes les valeurs, même les plus extrêmes qui ont tendances à être laissées pour compte. La complexité peut être augmentée via la multiplication des nombres de filtres de convolution, et peut être augmentée de manière encore plus conséquente en ajoutant des couches de convolution et de *pooling* associées. Dans le cas de la base de données utilisée au cours de ce travail, un nombre de trois couches de convolutions et de *maxpooling* s'est avéré optimal. Il a été constaté que l'ajustement des hyperparamètres ont eu des effets spécifiques selon les configurations de réseau testées. L'augmentation raisonnée de la taille du *batch* garantissait notamment une convergence des plus optimales mais pouvait parfois générer du surentrainement. Le nombre d'époques n'avait pas d'incidence notable sur les performances des configurations de réseaux testées, ce qui n'était pas le cas du taux d'apprentissage dont la valeur optimale était 0,001 pour la plupart des configurations. Malgré cela, les valeurs les plus extrêmes restaient tout de même moins bien prédites que les autres.

L'étude des biais présentée en 3.4 permet de constater qu'entre les prédictions du réseau neuronal et les prédictions d'échantillonnage, le degré de représentativité est, dans le meilleur

des cas, du même ordre de grandeur que celui entre les prédictions d'échantillonnage et la réalité du lot échantillonné. L'étude du nombre de faux positifs et faux négatifs générés par l'échantillonnage qui a été réalisée, tant les résultats d'échantillonnage que pour prédictions du réseau, permet de constater que le réseau neuronal identifie globalement les mêmes nombres de faux positifs et négatifs que l'échantillonnage, ce qui ajoute une confiance supplémentaire quant en les résultats de ce dernier.

Cependant, dans des situations réelles, il est peu courant que l'intégralité d'un lot soit traitée dans le but de le caractériser. Seule une poignée d'échantillons est en général analysée. Ainsi, à une représentativité modérée au départ de certains échantillons analysés due à un prélèvement par grappillage par exemple, peut venir s'ajouter l'erreur intrinsèque d'identification du réseau. Ceci peut ainsi engendrer une erreur de classification du sol étudié, dans le cas d'une caractérisation du taux de matières résiduelles. Plutôt que cette erreur de caractérisation intrinsèque vienne s'additionner aux erreurs liées à l'échantillonnage, l'analyse d'images par intelligence artificielle pourrait tenter de les réduire.

RECOMMANDATIONS

La présente section propose des recommandations suite aux travaux effectués lors de cette recherche. Elles concernent l'analyse d'image principalement.

Concernant la base de données d'entraînement et de validation d'un réseau neuronal convolutif régressif, dans le cas de l'analyse d'images tout particulièrement, il a été vu qu'il est essentiel que ces dernières soient parfaitement équilibrées autant vis-à-vis du ou des paramètres recherchés par le réseau, que sur leur aspect. Pour pallier les éventuels manques, l'utilisation de photographies synthétiques peut être une solution rapide et prometteuse à condition seulement que leur nombre soit équilibré avec celui des photographies réelles. En effet, l'utilisation de différents types de photographies peut générer des biais (Duhaime et al., 2021; Manashti, 2022).

L'objectif du développement de l'analyse d'images de sol par intelligence artificielle est éventuellement à terme, d'analyser des photos directement prises sur le terrain. En effet, cela permettrait un gain de temps et d'argent important tant pour les analystes que les clients. Dans cette optique, il pourrait être intéressant de voir si un réseau entraîné et validé avec une base de données de photos prises en laboratoire et/ou des photos artificielles est capable d'estimer correctement des photos prises en contexte réel.

L'étude des réseaux neuronaux artificiels et convolutifs (par classification et régression) ayant été réalisée au cours de plusieurs travaux antérieurs et actuels à l'ÉTS (Bouyim, 2022; Dol, 2023; Zeryouh, 2023), des techniques de *machine learning* autres que les réseaux neuronaux pourraient également être explorées. Le *machine learning* est un domaine de l'informatique qui englobe les réseaux de neurones. Il correspond à l'élaboration de méthodes et d'algorithmes permettant aux ordinateurs d'apprendre à partir de données via le développement de modèles informatiques qui peuvent s'entraîner sur des données, extraire des schémas et des informations à partir de ces données, puis utiliser ces informations pour prendre des décisions ou effectuer des prédictions sur de nouvelles données. Le *machine learning* englobe diverses techniques, y

compris les réseaux de neurones donc, mais également les SVM (machines à vecteurs de support), la régression logistique, les arbres ou forêts de décision et bien d'autres. Dans le cadre d'analyse d'images, l'utilisation d'un réseau neuronal convolutif est très courante et pratique puisqu'il permet d'utiliser directement des photographies brutes en données d'entrée. Cependant, la régression logistique et les SVM peuvent être de bonnes pistes à explorer. La régression logistique et les SVM ont déjà été employées pour réaliser des tâches de classification de sols de manière générale pour les SVM (Barman & Choudhury, 2020; Shivhare & Cecil, 2021) et plus précisément pour réaliser des cartographies dans le cas de la régression logistiques (Abbaszadeh Afshar et al., 2018; Piccini et al., 2019). Ces deux méthodes peuvent également être employées pour des tâches de régression et dans le cadre de l'analyse d'images de sol, cet aspect reste à développer. Cela dit, ces deux méthodes ne permettent pas d'utiliser des images brutes en entrée mais des caractéristiques comme les couleurs ou les paramètres texturaux.

ANNEXE I

ANALYSE SÉDIMENTOMÉTRIQUE D'UN ÉCHANTILLON

Les procédures, formules et notations de cette annexe sont celles décrites dans la section 8 sur l'analyse granulométrique par sédimentation (sédimentométrie) de la norme BNQ 2501-025 de 2013 sur l'analyse granulométrique des sols inorganiques. La sédimentométrie consiste à déterminer la granulométrie des particules de sol inférieure à 80 µm. La formule pour calculer les diamètres équivalents et les pourcentages de particules d'un diamètre inférieur au diamètre équivalent ci-dessous.

$$D = \sqrt{\frac{3\eta}{g(D_{rs} - D_{ri})\rho_w} \left(\frac{H'}{t}\right)} \quad (\text{A1.1})$$

Avec :

- D , le diamètre équivalent d'une particule de sol (mm) ;
- η , la viscosité de l'eau (P.s) ;
- g , la constante gravitationnelle (= 9,8 m/s²) ;
- D_{rs} , densité des particules solides (= 2,7 g/cm³ déterminé pour les briques et sol par pycnomètre à gaz dans le cas d'étude) ;
- D_{ri} , la densité du liquide de suspension ($D_{ri} \approx 1,0$ g/cm³) ;
- ρ_w , la masse volumique de l'eau (= 1 g/cm³) ;
- H , la profondeur hydromètre (cm) ;
- t , le temps de sédimentation (min).

Pour un hydromètre 152H :

$$P = \frac{R_c \times F \times a}{m_4} \times 100 \quad (\text{A1.2})$$

Avec :

- P , pourcentage de particules d' un diamètre inférieur au diamètre équivalent (%);
- R_c , lecture corrigée de l' hydromètre (g/L) avec $R_c = R_a - C$ avec R_a correspondant à la lecture de l' hydromètre dans l' échantillon de l' essai et C la lecture de l' hydromètre dans l' échantillon blanc ;
- F , le facteur de correction qui se détermine selon l' équation 8.4.1.2 b) de la norme (\emptyset) ;
- a , le facteur de correction provenant du tableau 4 de la norme ;
- m_4 , masse sèche de l' échantillon (= 81,33 g dans le cas présent).

Tableau-A I-1 Détermination de la sédimentométrie du sol utilisé pour la recherche

Temps de sédimentation t (sec)	Température (°c)		Lecture hydromètre (g/L)			H (mm)	η (Pa.s)	D (μ m)	P
	Essai	Blanc	Essai	Ra	Blanc C				
75	20,5	20,2	14,0	7	7,0	136,772	0,000990	18,03	7,91%
150	20,5	20,2	13,0	7	6,0	166,886	0,000990	14,08	6,78%
300	20,0	20,2	12,0	7	5,0	197,000	0,001002	10,89	5,65%
915	20,1	20,5	12,0	7	5,0	197,000	0,000980	6,17	5,65%
1800	20,0	20,0	11,5	7	4,5	212,057	0,001002	4,61	5,09%
3620	19,8	19,9	11,0	7	4,0	227,114	0,001022	3,40	4,52%
7200	19,8	19,9	11,0	7	4,0	227,114	0,001022	2,41	4,52%
15360	19,8	19,9	10,0	7	3,0	257,228	0,001022	1,76	3,39%
28800	19,9	19,9	10,0	7	3,0	257,228	0,001025	1,28	3,39%
86400	19,8	19,9	9,0	7	2,0	287,342	0,001022	0,78	2,26%

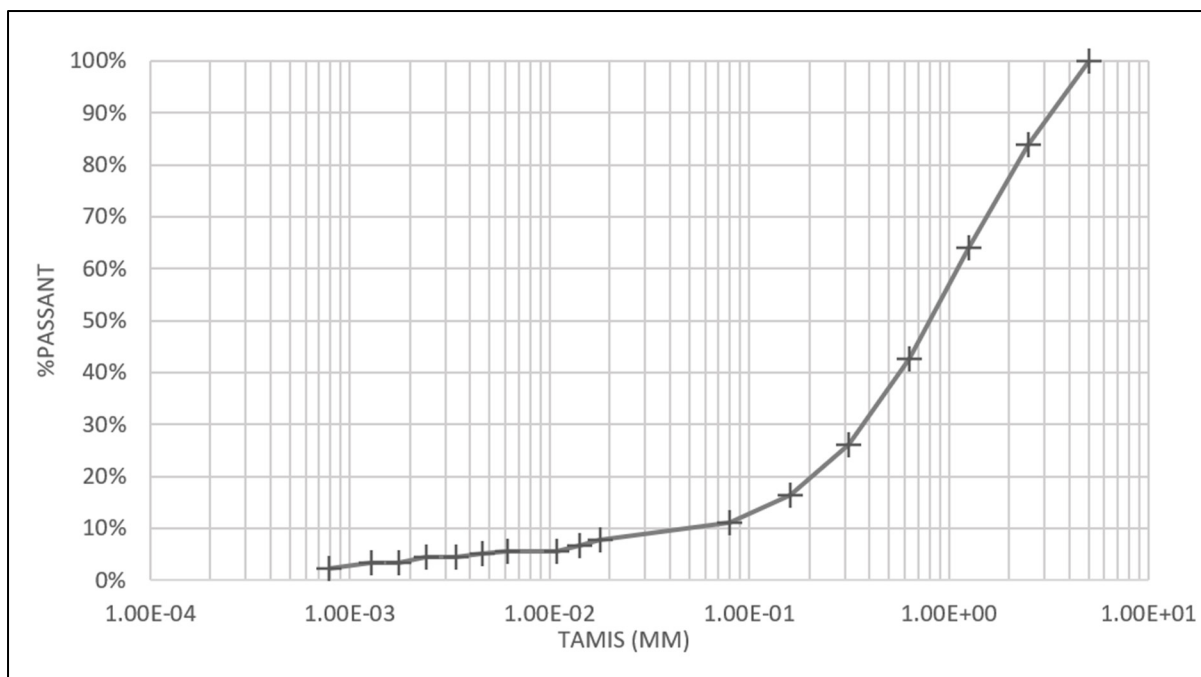


Figure-A I-1 Granulométrie complète du sol seul comprenant les résultats granulométriques et de sédimentométrie

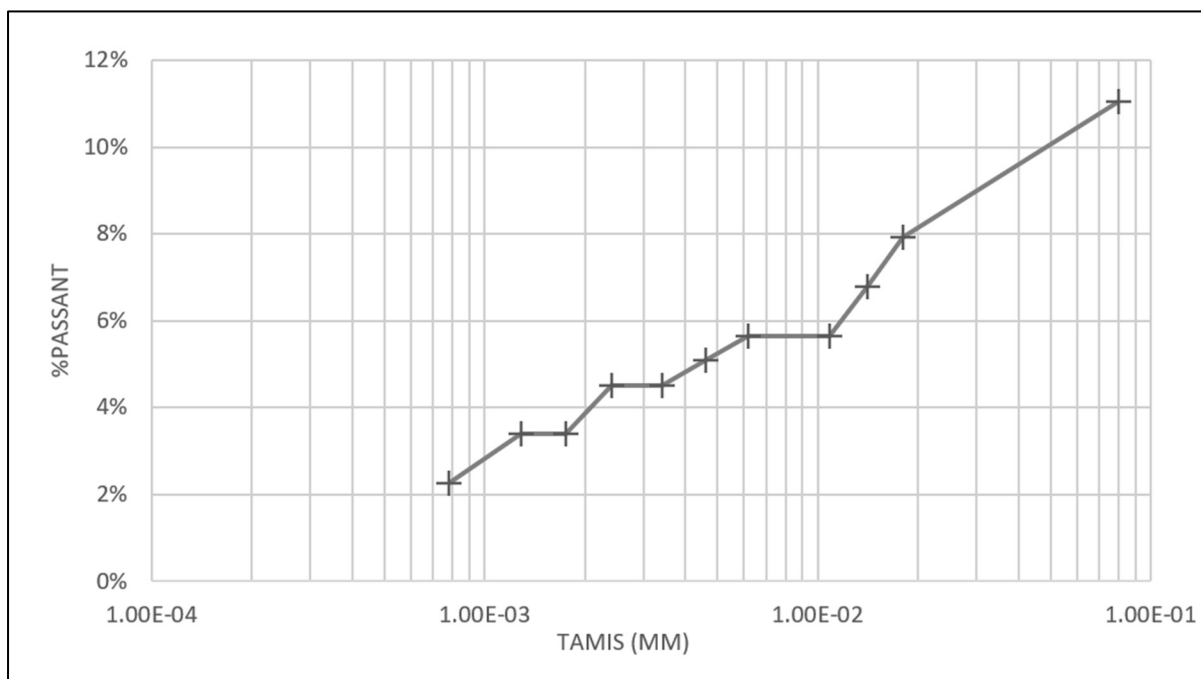


Figure-A I-2 Sédimentométrie du sol seul

ANNEXE II

CONFIGURATIONS DE RÉSEAU NEURONAL

Algorithme-A II-1 CNN n°1 – première Conv2D à 4 filtres, la deuxième à 8 filtres et la troisième à 16 filtres, taille du batch à 16

```
# CREATION DU RESEAU DE NEURONES
model=Sequential([
    tf.keras.layers.experimental.preprocessing.Rescaling(1.0/255.0),
    Conv2D(8,(4,4),strides=(2,2),
    ativation='relu',padding='same',input_shape=(512,810,3)),
    MaxPooling2D(pool_size=(2,2)),
    Conv2D(16,(4,4),strides=(2,2),activation='relu',padding='same'),
    MaxPooling2D(pool_size=(2,2)),
    Conv2D(32,(4,4),strides=(2,2),activation='relu',padding='same'),
    MaxPooling2D(pool_size=(4,4)),
    Flatten(),
    Dense(10,activation='sigmoid'),
    Dense(1,activation='sigmoid')])

model.compile(loss=tf.losses.mean_squared_error,
optimizer=tf.keras.optimizers.Adam(learning_rate=0.001),metrics=['RootMeanSquaredEr
ror'])

# ENTRAINEMENT
History=model.fit(input_train,output_train,epochs=15,
validation_data=(input_val,output_val),batch_size=16)
```

Algorithme-A II-2 CNN n°2 – première Conv2D à 4 filtres, la deuxième à 8 filtres et la troisième à 16 filtres, taille du *batch* à 32

```
# CREATION DU RESEAU DE NEURONES
model=Sequential([
    tf.keras.layers.experimental.preprocessing.Rescaling(1.0/255.0),
    Conv2D(8,(4,4),strides=(2,2),
    activation='relu',padding='same',input_shape=(512,810,3)),
    MaxPooling2D(pool_size=(2,2)),
    Conv2D(16,(4,4),strides=(2,2),activation='relu',padding='same'),
    MaxPooling2D(pool_size=(2,2)),
    Conv2D(32,(4,4),strides=(2,2),activation='relu',padding='same'),
    MaxPooling2D(pool_size=(4,4)),
    Flatten(),
    Dense(10,activation='sigmoid'),
    Dense(1,activation='sigmoid')])

model.compile(loss=tf.losses.mean_squared_error,
optimizer=tf.keras.optimizers.Adam(learning_rate=0.001),metrics=['RootMeanSquaredError'])

# ENTRAINEMENT
History=model.fit(input_train,output_train,epochs=15,
validation_data=(input_val,output_val),batch_size=32)
```

Algorithme-A II-3 CNN n°3 – première Conv2D à 16 filtres, la deuxième à 32 et la troisième à 64, taille du *batch* à 32

```
# CREATION DU RESEAU DE NEURONES
model=Sequential([
    tf.keras.layers.experimental.preprocessing.Rescaling(1.0/255.0),
    Conv2D(16,(4,4),strides=(2,2),
    ativation='relu',padding='same',input_shape=(512,810,3)),
    MaxPooling2D(pool_size=(2,2)),
    Conv2D(32,(4,4),strides=(2,2),activation='relu',padding='same'),
    MaxPooling2D(pool_size=(2,2)),
    Conv2D(64,(4,4),strides=(2,2),activation='relu',padding='same'),
    MaxPooling2D(pool_size=(4,4)),
    Flatten(),
    Dense(10,activation='sigmoid'),
    Dense(1,activation='sigmoid')])

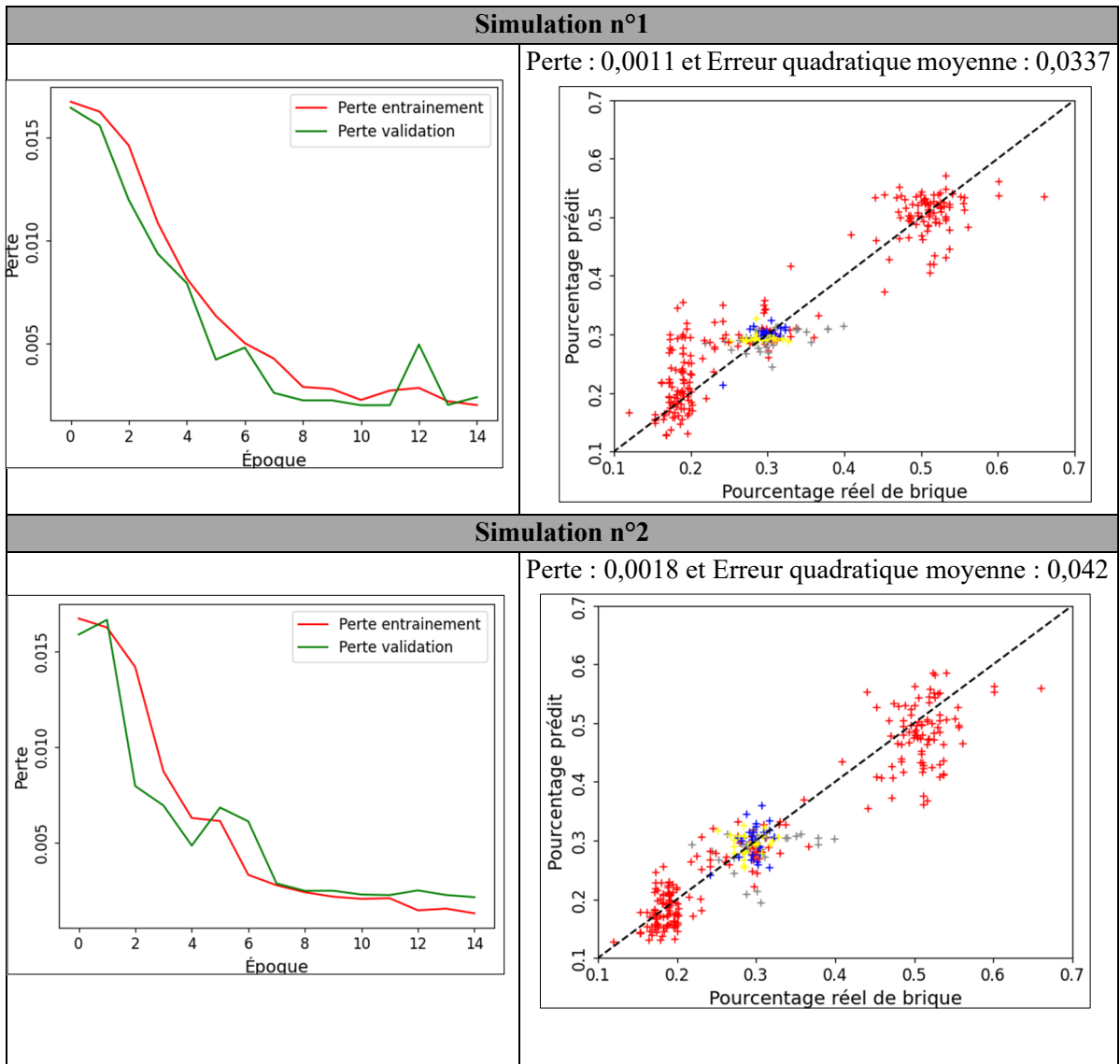
model.compile(loss=tf.losses.mean_squared_error,
optimizer=tf.keras.optimizers.Adam(learning_rate=0.001),metrics=['RootMeanSquaredError'])

# ENTRAINEMENT
History=model.fit(input_train,output_train,epochs=15,
validation_data=(input_val,output_val),batch_size=32)
```


ANNEXE III

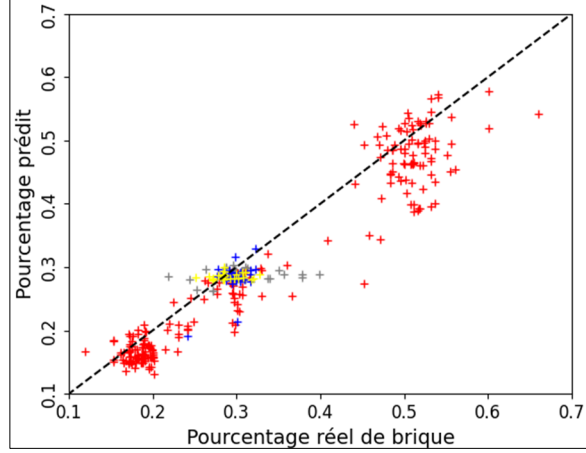
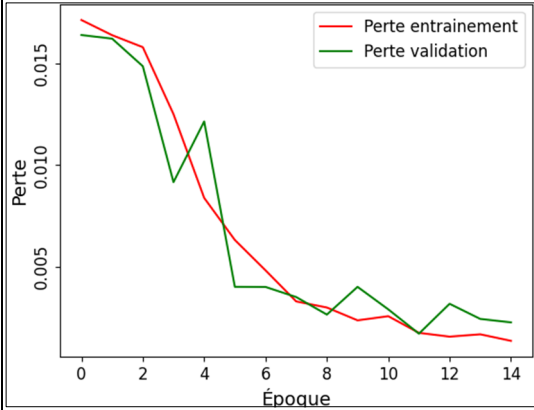
TEST DE VARIABILITÉ DES CONFIGURATIONS DE RÉSEAUX NEURONAUX

Tableau-A III-1 Graphique des pertes d'entraînement et de validation au fil des époques (à gauche) et graphique de régression lors du test et performances du test (à droite) pour les six simulations effectuées avec la configuration de réseau n°1 - des simulations n°1 à 3 incluses, les points rouges représentent les échantillons secs sans fines, en gris les secs avec fines, en bleu les humides, en jaune ceux avec huile, des simulations n°4 à 6 incluses, les points jaunes sont associés au grappillage, les points verts au pelletage fractionné, les points violets à la séparation rotative et les points bleus à la séparation à chutes



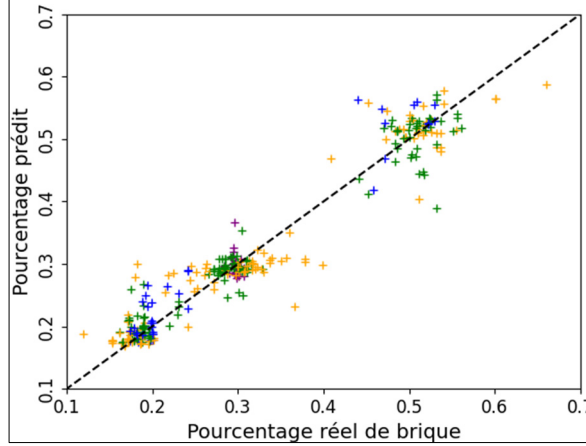
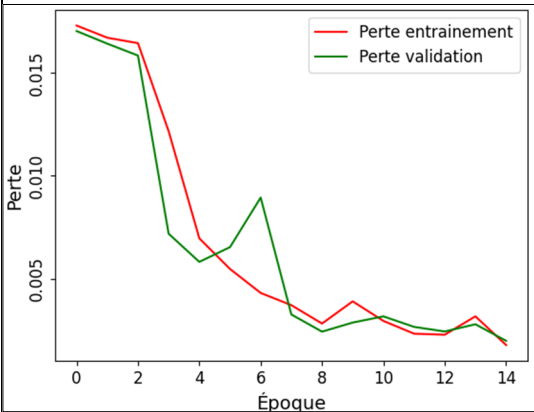
Simulation n°3

Perte : 0,0020 et Erreur quadratique moyenne : 0,0442



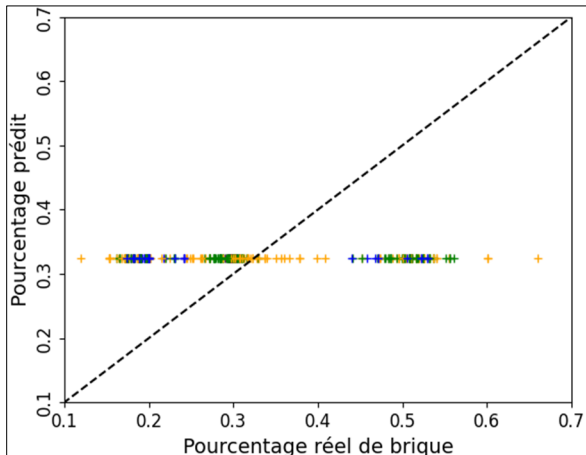
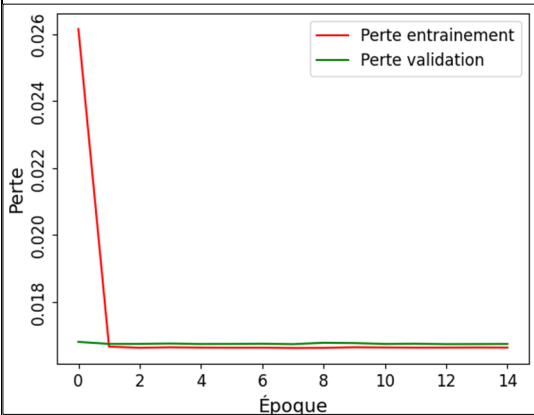
Simulation n°4

Perte : 0,0011 et Erreur quadratique moyenne : 0,0338



Simulation n°5

Perte : 0,0172 et Erreur quadratique moyenne : 0,1310



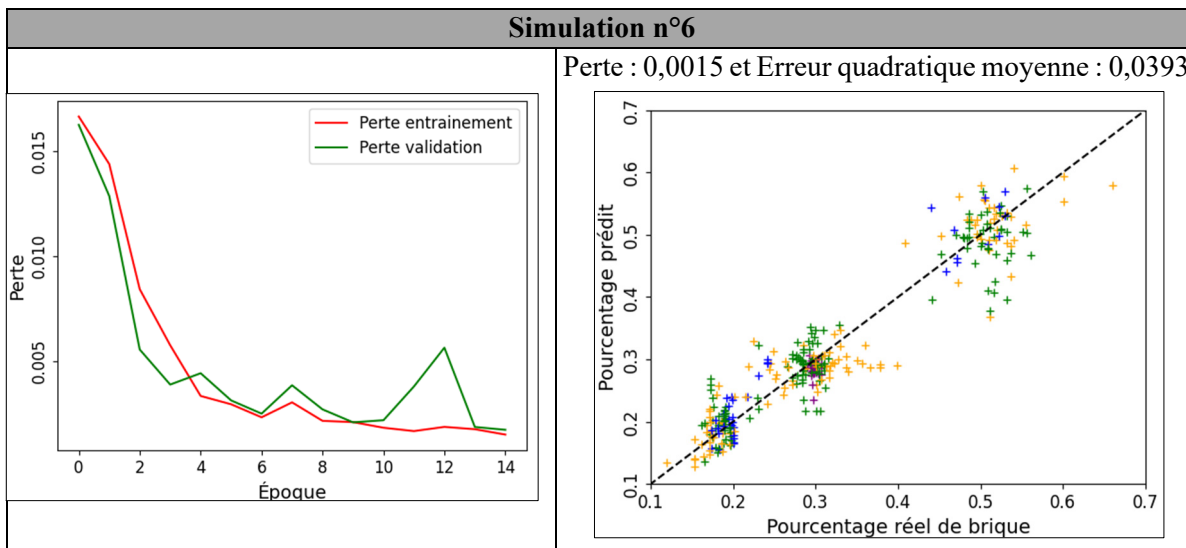
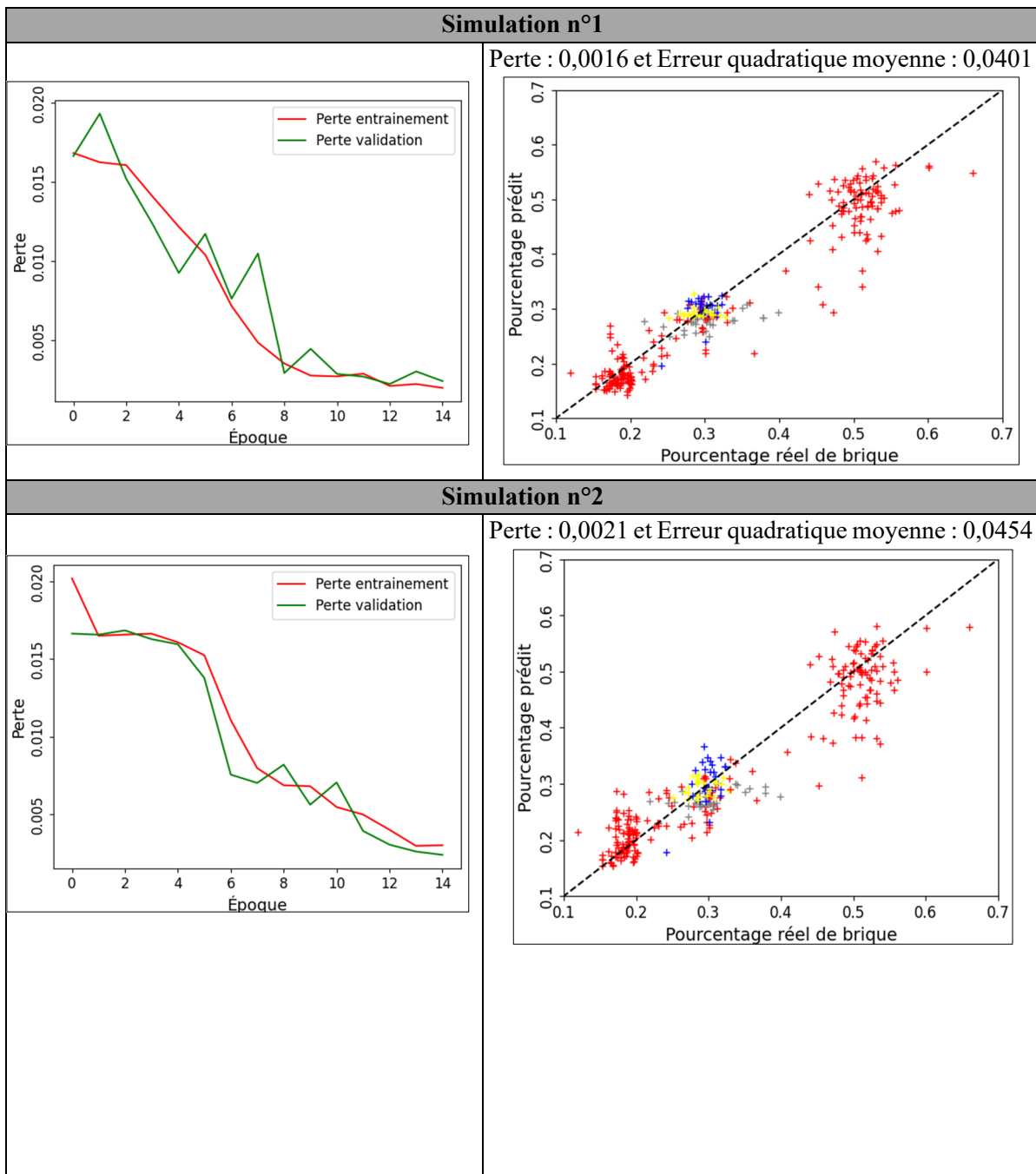
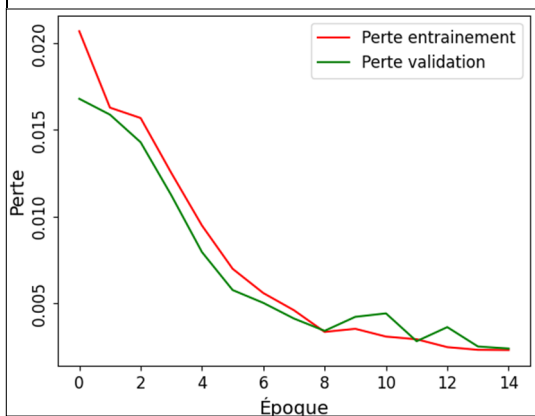


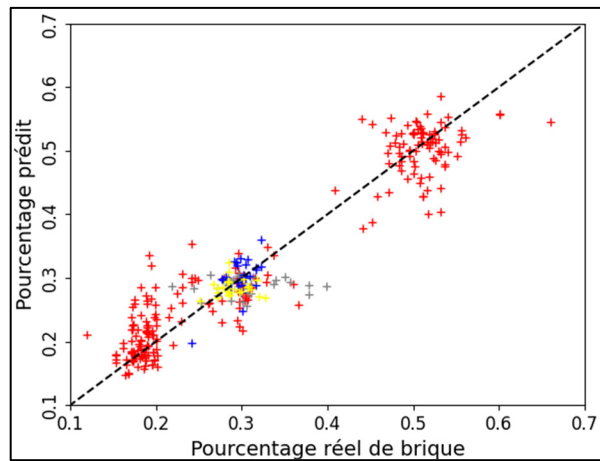
Tableau-A III-2 Graphique des pertes d'entraînement et de validation au fil des époques (à gauche) et graphique de régression lors du test et performances du test (à droite) pour les six simulations effectuées avec la configuration de réseau n°2 - des simulations n°1 à 3 incluses, les points rouges représentent les échantillons secs sans fines, en gris les secs avec fines, en bleu les humides, en jaune ceux avec huile, des simulations n°4 à 6 incluses, les points jaunes sont associés au grappillage, les points verts au pelletage fractionné, les points violets à la séparation rotative et les points bleus à la séparation à chutes



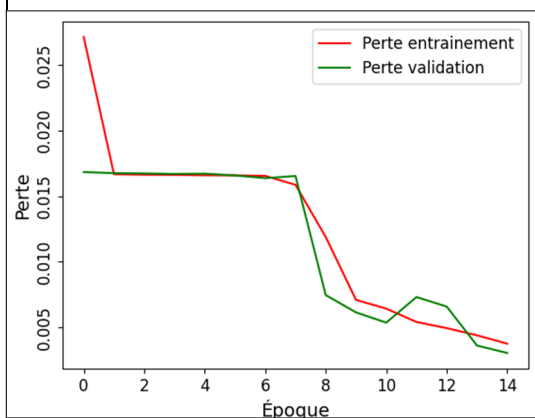
Simulation n°3



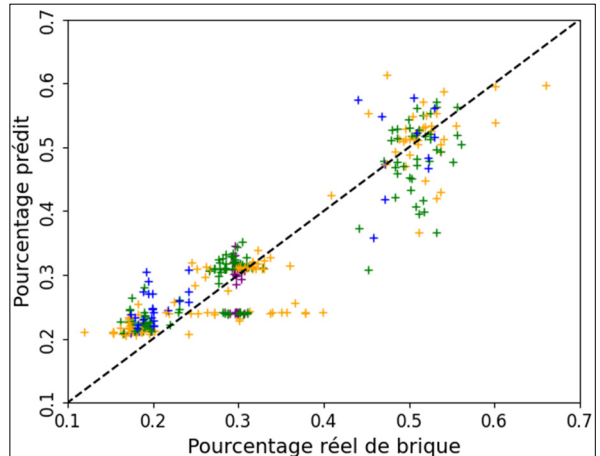
Perte : 0,0017 et Erreur quadratique moyenne : 0,0409



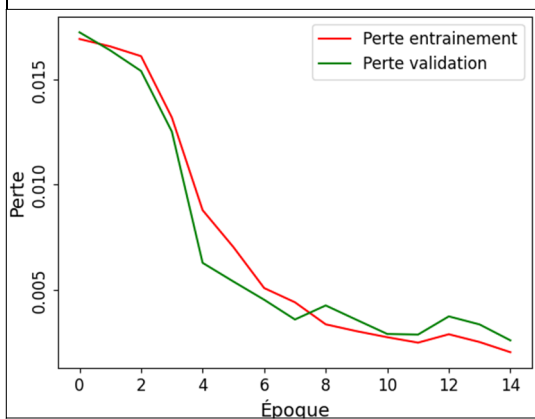
Simulation n°4



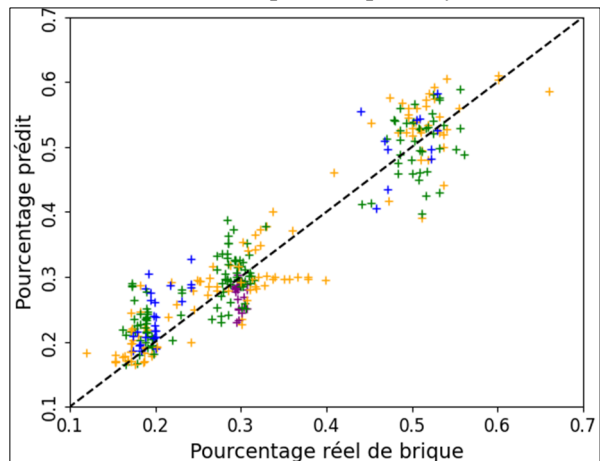
Perte : 0,0026 et Erreur quadratique moyenne : 0,0510



Simulation n°5



Perte : 0,0020 et Erreur quadratique moyenne : 0,0447



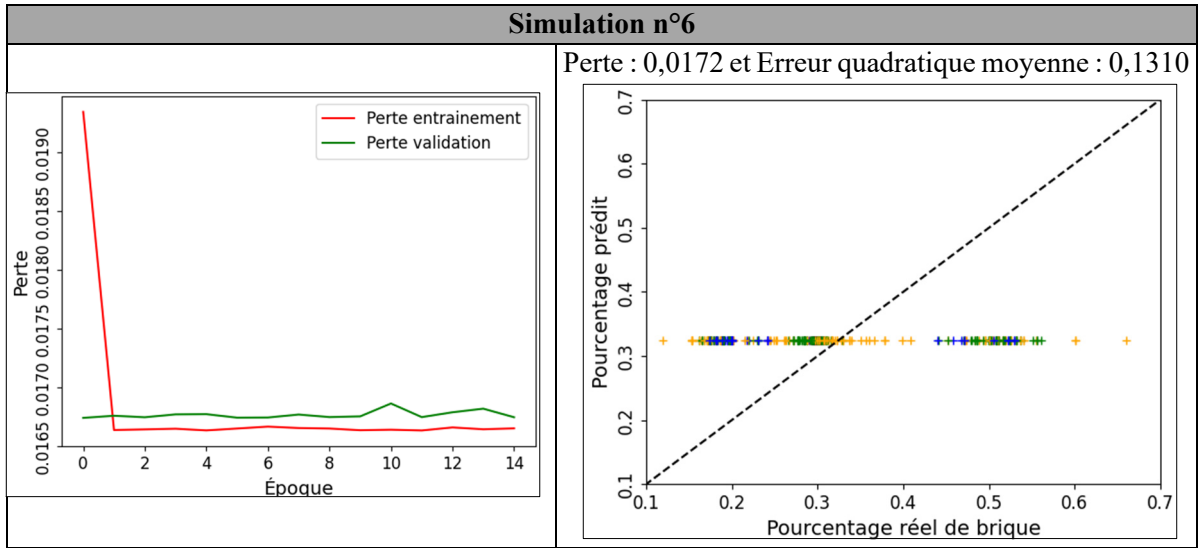
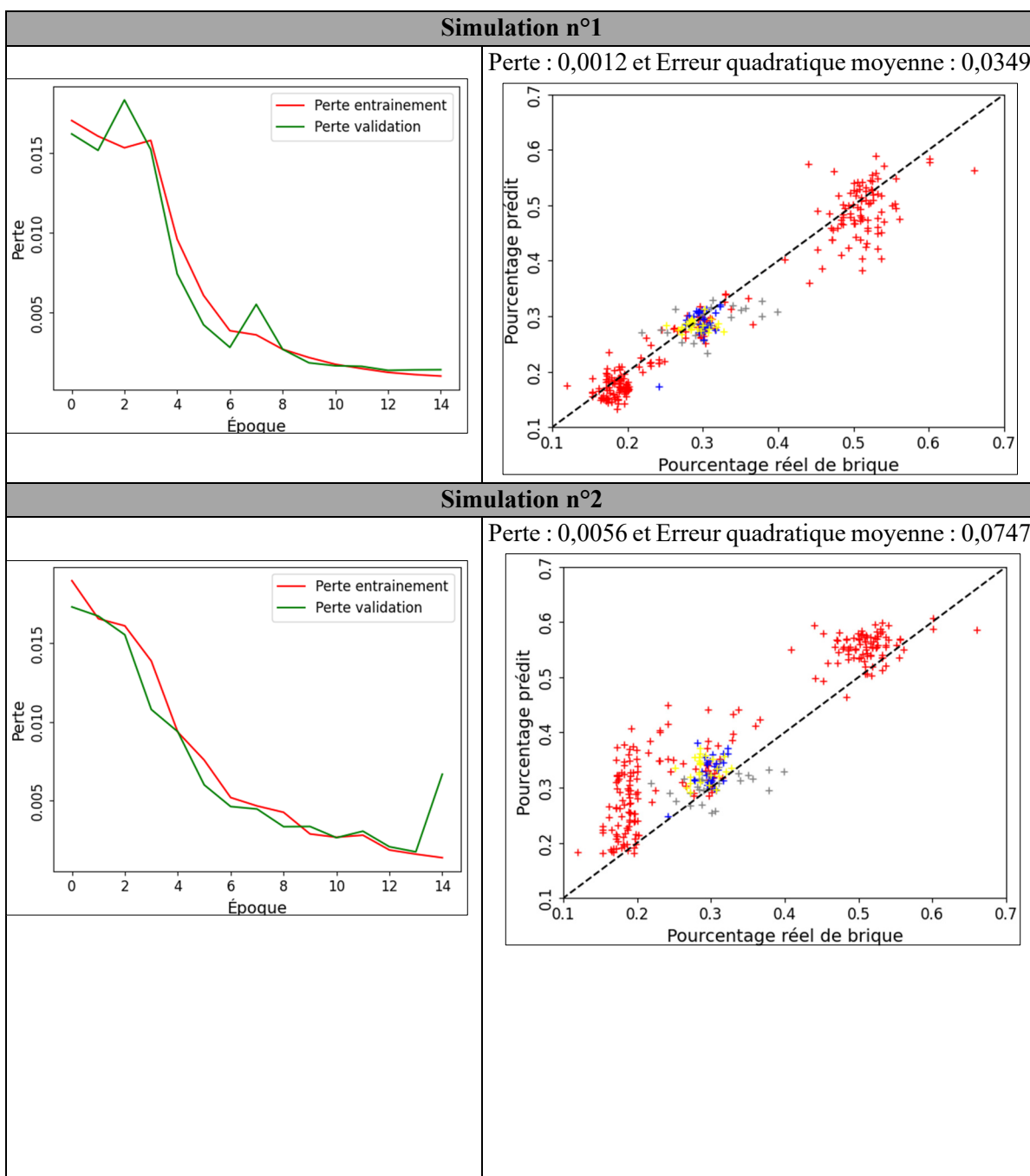
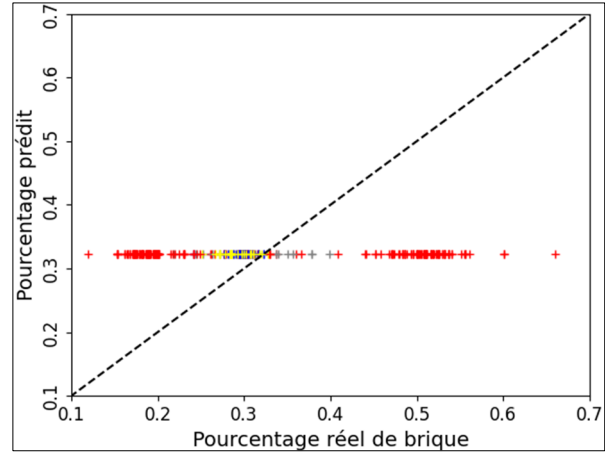
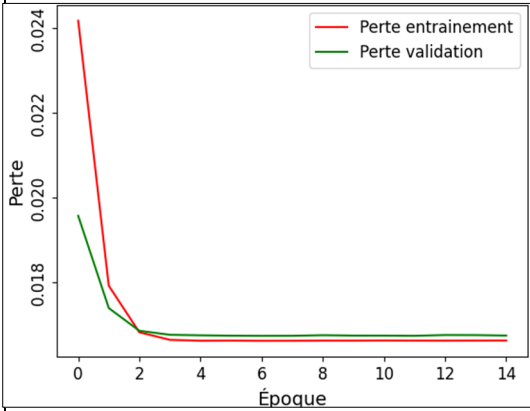


Tableau-A III-3 Graphique des pertes d'entraînement et de validation au fil des époques (à gauche) et graphique de régression lors du test et performances du test (à droite) pour les 15 simulations effectuées avec la configuration de réseau n°3 - des simulations n°1 à 3 et 7 à 15 incluses, les points rouges représentent les échantillons secs sans fines, en gris les secs avec fines, en bleu les humides, en jaune ceux avec huile, des simulations n°4 à 6 incluses, les points jaunes sont associés au grappillage, les points verts au pelletage fractionné, les points violets à la séparation rotative et les points bleus à la séparation à chutes



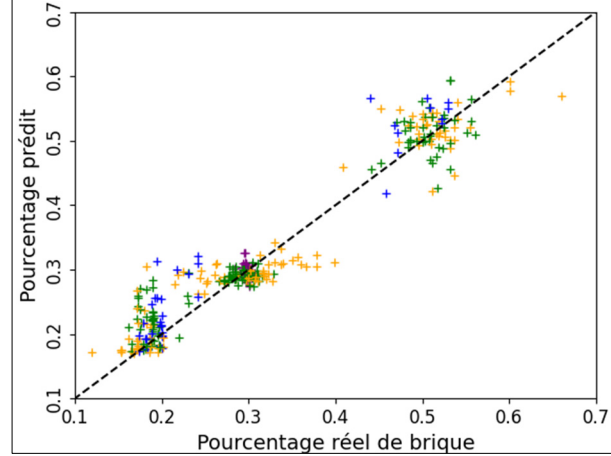
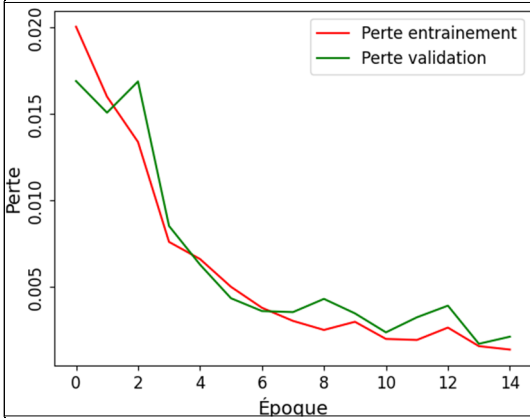
Simulation n°3

Perte : 0,0173 et Erreur quadratique moyenne : 0,1310



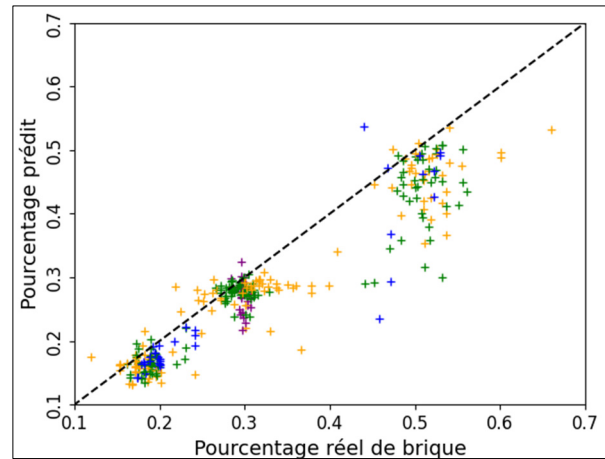
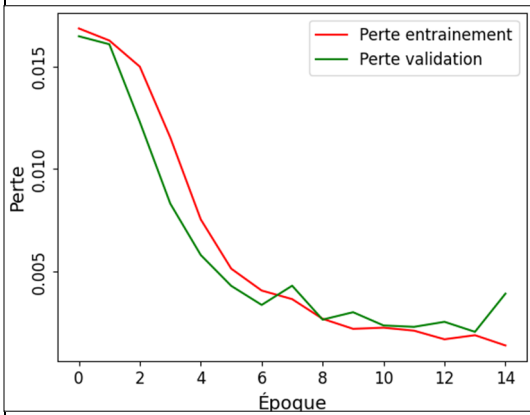
Simulation n°4

Perte : 0,0013 et Erreur quadratique moyenne : 0,0367

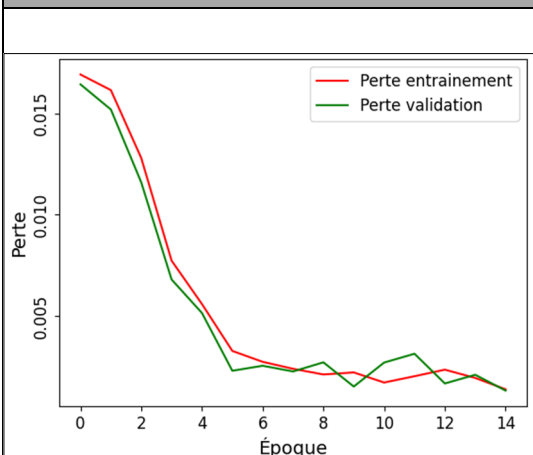


Simulation n°5

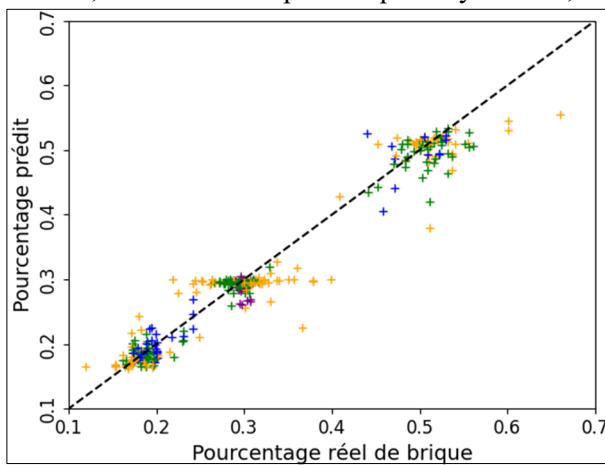
Perte : 0,0032 et Erreur quadratique moyenne : 0,0562



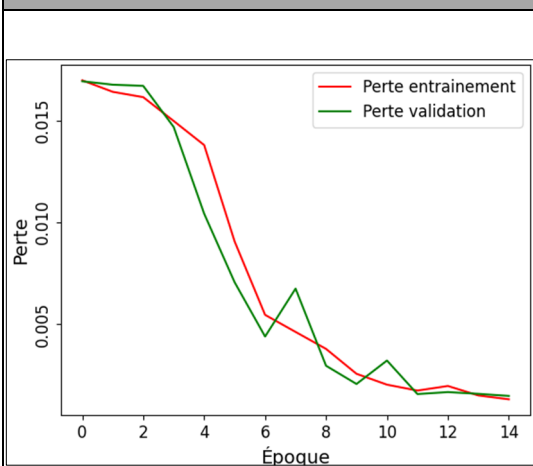
Simulation n°6



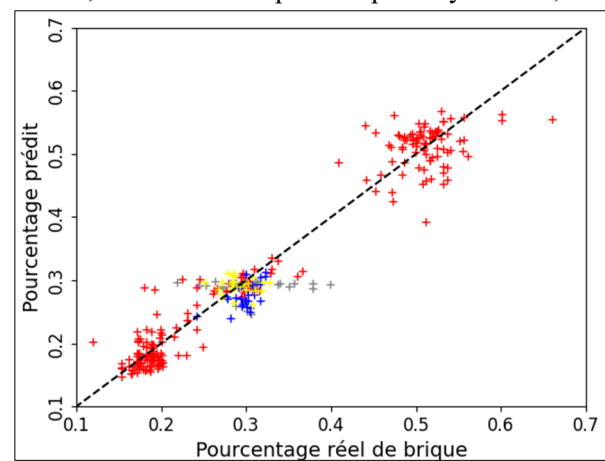
Perte : 0,000838 Erreur quadratique moyenne : 0,029



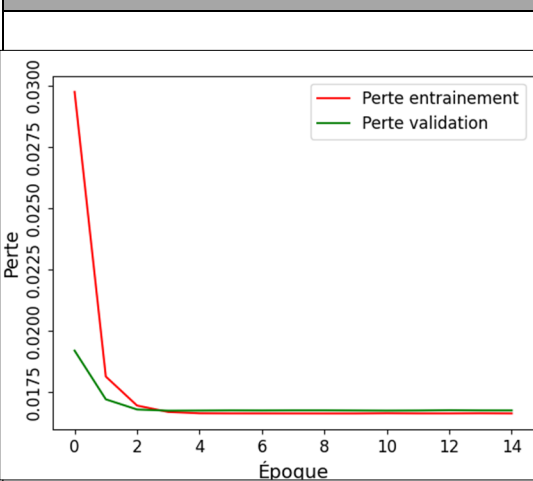
Simulation n°7



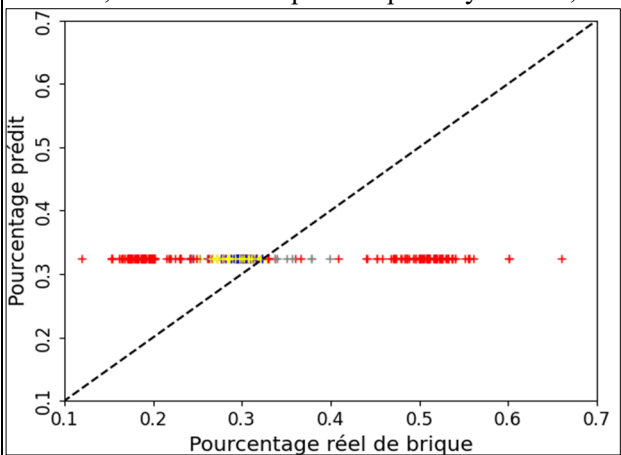
Perte : 0,0011 et Erreur quadratique moyenne : 0,0329



Simulation n°8

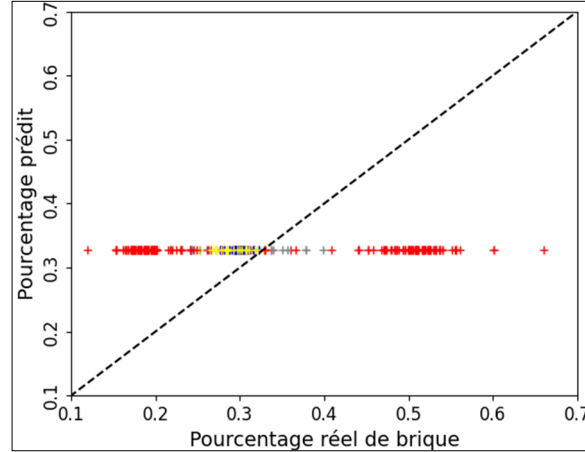
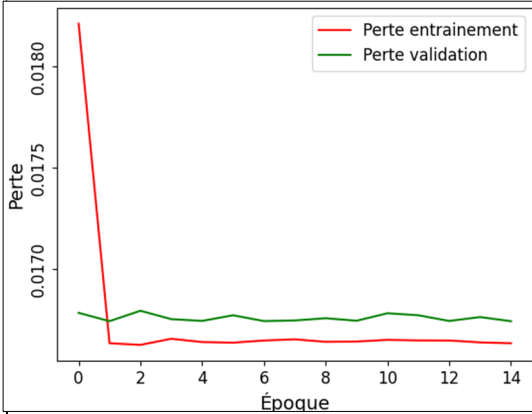


Perte : 0,0172 et Erreur quadratique moyenne : 0,1310



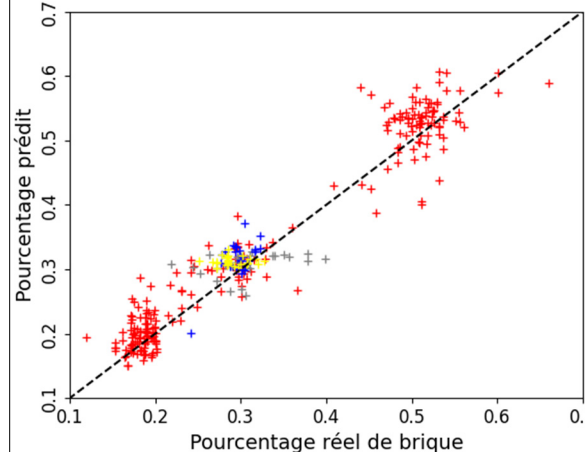
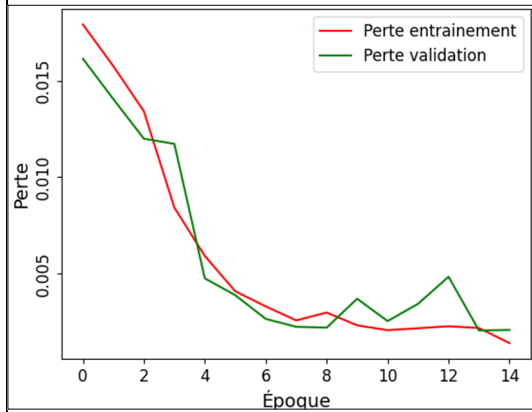
Simulation n°9

Perte : 0,0172 et Erreur quadratique moyenne : 0,1311



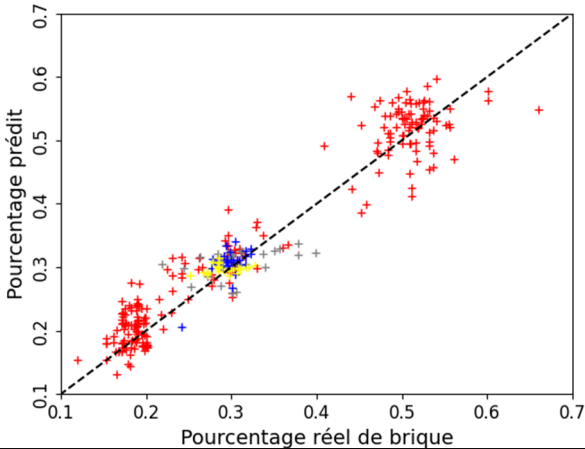
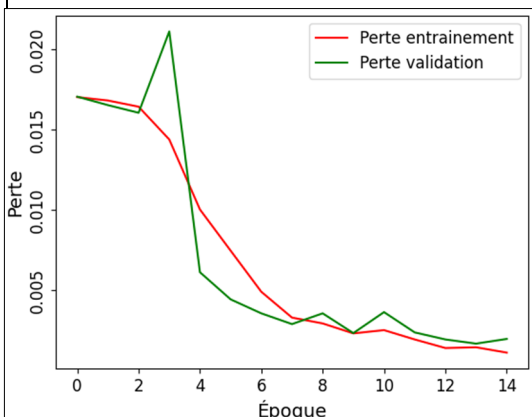
Simulation n°10

Perte : 0,0014 et Erreur quadratique moyenne : 0,037

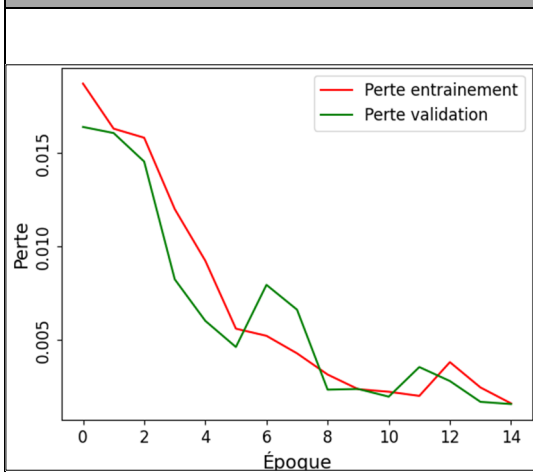


Simulation n°11

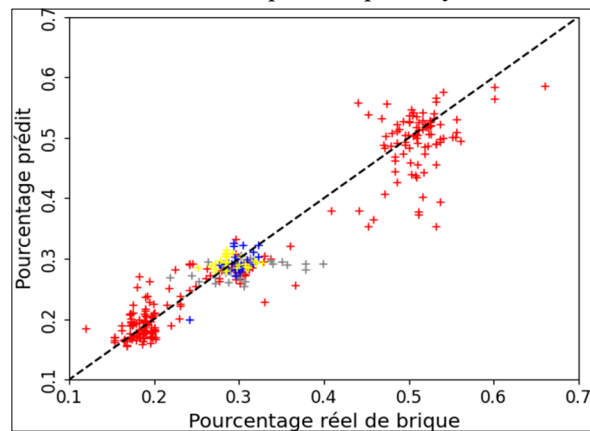
Perte : 0,0013 et Erreur quadratique moyenne : 0,037



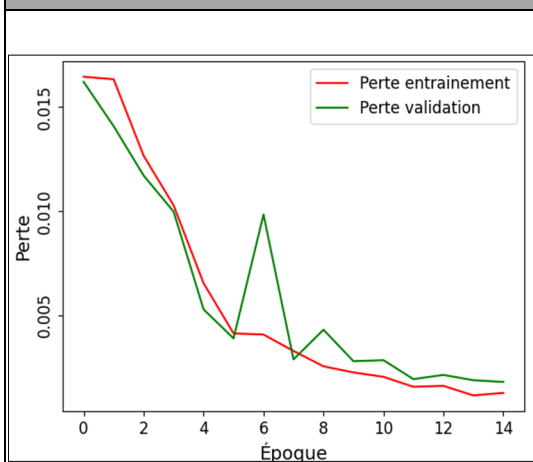
Simulation n°12



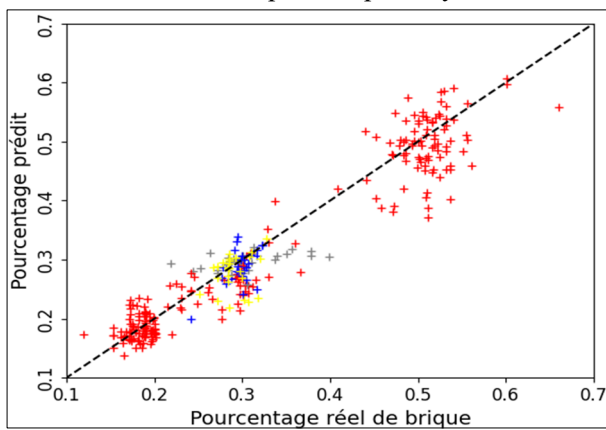
Perte : 0,0014 et Erreur quadratique moyenne : 0,0375



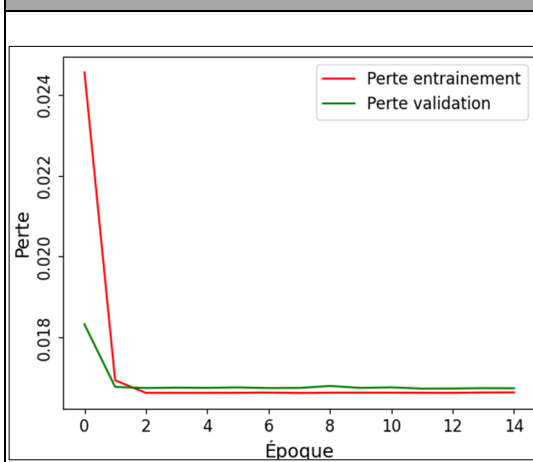
Simulation n°13



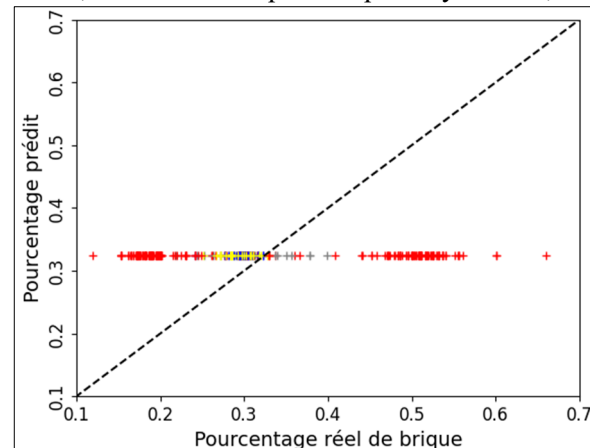
Perte : 0,0014 et Erreur quadratique moyenne : 0,0381



Simulation n°14

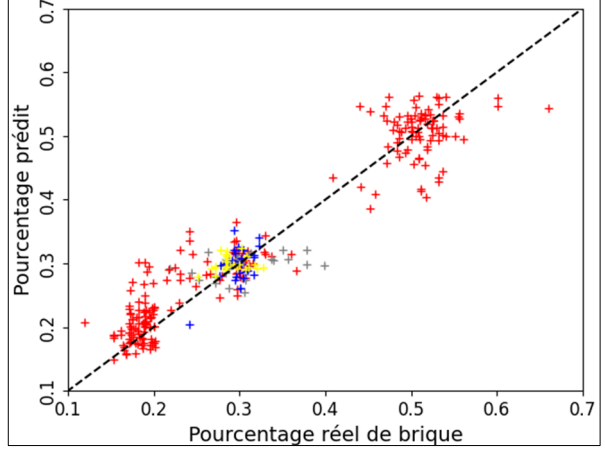
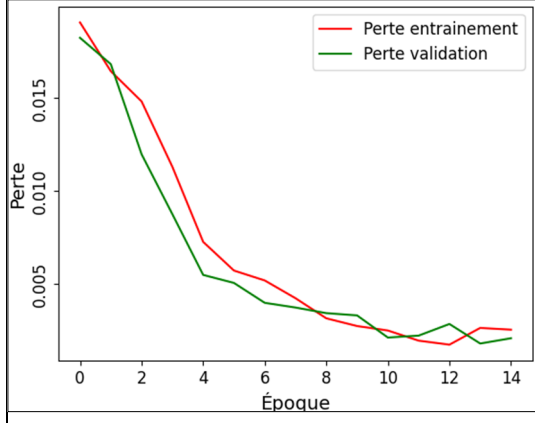


Perte : 0,0172 et Erreur quadratique moyenne : 0,1310



Simulation n°15

Perte : 0,0015 et Erreur quadratique moyenne : 0,039



ANNEXE IV

RÉSULTATS DES TESTS D'ÉGALITÉ DES VARIANCES SUR LES RÉSULTATS GRANULOMÉTRIQUES

Tableau-A IV-1 : Écartis types et intervalles de confiance à 95% associés aux écarts types des %passants - essais avec fines

Jeu de données	Essai n°1 :	Essai n°2 :	Essai n°3 :	Essai n°7 :	Essai n°8 :	Essai n°9 :	Essai n°10 :
	sec, séparation rotative	sec, grappillage	sec, pelletage fractionné	eau, grappillage	eau, pelletage fractionné	huile, grappillage	huile, pelletage fractionné
N	32	32	32	32	32	32	32
Écart-type	0,0103415	0,0774409	0,0202846	0,0236565	0,0120698	0,0178906	0,0090861
Intervalle de confiance	(0,0073335 ; 0,015922)	(0,0611558 ; 0,107063)	(0,0144115 ; 0,031172)	(0,0187477 ; 0,032590)	(0,0094512 ; 0,016829)	(0,0084462 ; 0,041374)	(0,0070217 ; 0,012837)
Écart-type	0,0124673	0,0726903	0,0205967	0,026616	0,0147539	0,0221793	0,0135332
Intervalle de confiance	(0,0083175 ; 0,020403)	(0,0576092 ; 0,100138)	(0,0164062 ; 0,028231)	(0,0217139 ; 0,035619)	(0,0115606 ; 0,020557)	(0,0100265 ; 0,053565)	(0,0106648 ; 0,018749)
Écart-type	0,0033229	0,0555447	0,0158599	0,0185046	0,0060054	0,0141562	0,0042171
Intervalle de confiance	(0,0025349 ; 0,0047556)	(0,0438793 ; 0,0767646)	(0,0116915 ; 0,0234893)	(0,0145184 ; 0,0257500)	(0,0045003 ; 0,0087495)	(0,0045033 ; 0,0485843)	(0,0030850 ; 0,0062936)
Écart-type	0,00353	0,0342863	0,0110638	0,0133225	0,0054258	0,0104568	0,0036046
Intervalle de confiance	(0,0027727 ; 0,0049066)	(0,0270459 ; 0,0474544)	(0,0078542 ; 0,0170155)	(0,0103498 ; 0,0187229)	(0,0039103 ; 0,0082197)	(0,0034582 ; 0,0345213)	(0,0026491 ; 0,0053549)
Écart-type	0,0015345	0,0204629	0,0063677	0,0099588	0,0031637	0,0077727	0,0017253
Intervalle de confiance	(0,0011100 ; 0,0023160)	(0,0158799 ; 0,0287888)	(0,0047096 ; 0,0093997)	(0,0073203 ; 0,0147918)	(0,0022292 ; 0,0049022)	(0,0023086 ; 0,0285707)	(0,0012311 ; 0,0026397)
Écart-type	0,0011846	0,0125106	0,0041026	0,007536	0,0021637	0,0053412	0,0014594
Intervalle de confiance	(0,0008344 ; 0,0018362)	(0,0096838 ; 0,0176458)	(0,0031345 ; 0,0058627)	(0,0047320 ; 0,0131032)	(0,0015157 ; 0,0033721)	(0,0016214 ; 0,0192099)	(0,0010435 ; 0,0022282)

Tableau-A IV-2 : Écartis types et intervalles de confiance à 95% associés aux écarts types des %passants - essais sans fines

Jeux de données		Essai n°4 : sec, grappillage	Essai n°5 : sec, séparation rotative	Essai n°6 : sec, pelletage fractionné
N		32	32	32
Écart-type		0,008251	0,0273024	0,0963384
Intervalle de confiance		(0,0064291 ; 0,011445)	(0,0215585 ; 0,037373)	(0,0745036 ; 0,134645)
Écart-type		0,015362	0,033468	0,105394
Intervalle de confiance		(0,0132717 ; 0,0192220)	(0,0264285 ; 0,045810)	(0,0836860 ; 0,143465)
Écart-type		0,0018904	0,0194358	0,0807124
Intervalle de confiance		(0,0013061 ; 0,002957)	(0,0158759 ; 0,025718)	(0,0638625 ; 0,110257)
Écart-type		0,0039367	0,008534	0,041395
Intervalle de confiance		(0,0032885 ; 0,0050938)	(0,0071244 ; 0,0110491)	(0,0332248 ; 0,0557447)
Écart-type		0,0007699	0,0008466	0,0031562
Intervalle de confiance		(0,0003109 ; 0,0020607)	(0,0005803 ; 0,0013348)	(0,0016261 ; 0,0066215)
Écart-type		0,000799	0,0003228	0,0003892
Intervalle de confiance		(0,0003265 ; 0,0021134)	(0,0001951 ; 0,0005773)	(0,0002837 ; 0,0005772)

%Passant 0,080 mm
 %Passant 0,160 mm
 %Passant 0,315 mm
 %Passant 0,63 mm
 %Passant 1,25 mm
 %Passant 2,5 mm

ANNEXE V

RÉSULTATS DES TESTS DE NORMALITÉ SUR LES POURCENTAGES DE BRIQUES

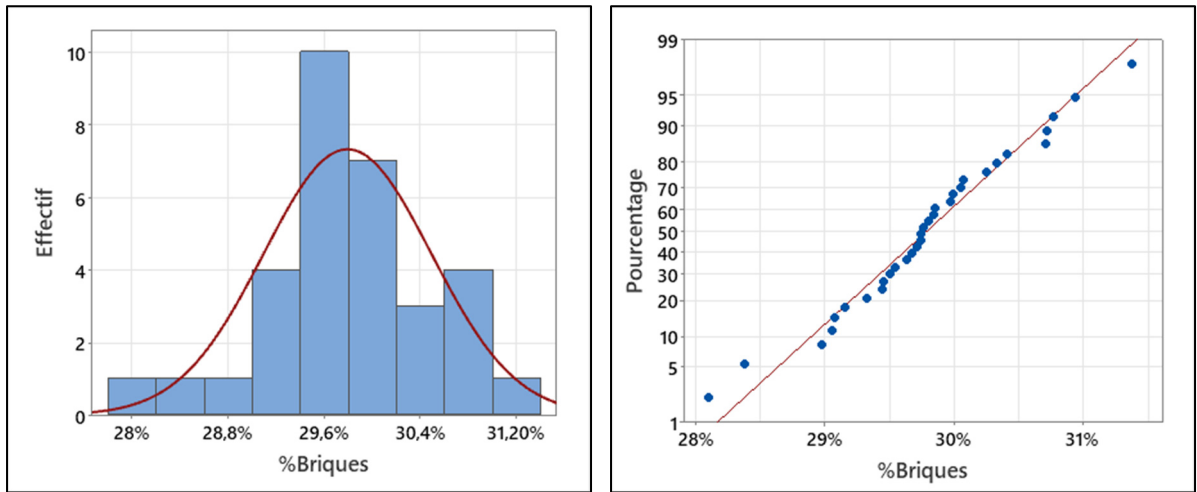


Figure-A V-1 Histogramme avec ajustement normale (à gauche) et graphique de probabilité normale (à droite) pour les résultats granulométriques de l'essai n°1 de sol sec avec fines échantillonné par séparation rotative

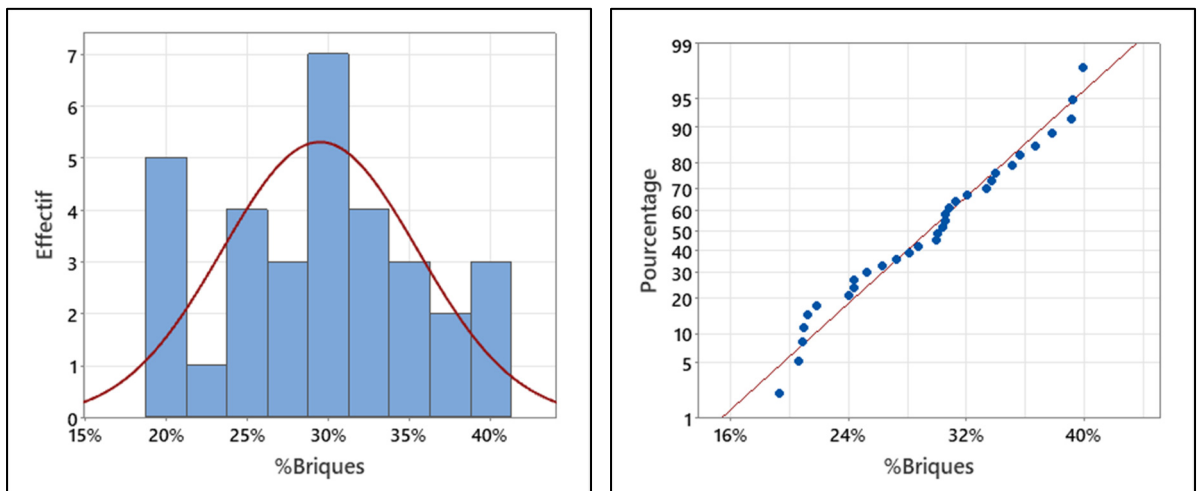


Figure-A V-2 Histogramme avec ajustement normale (à gauche) et graphique de probabilité normale (à droite) pour les résultats granulométriques de l'essai n°2 de sol sec avec fines échantillonné par grappillage

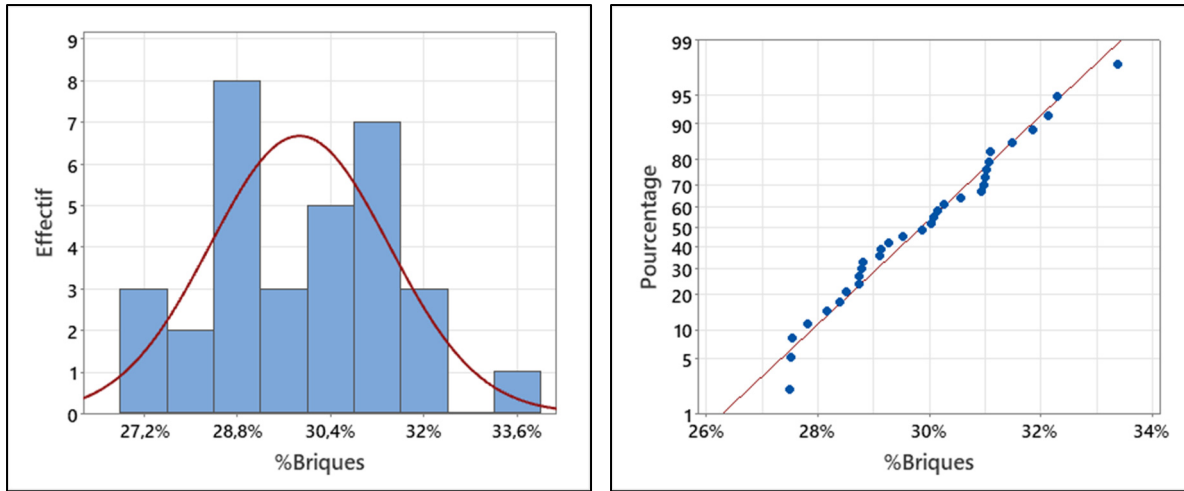


Figure-A V-3 Histogramme avec ajustement normale (à gauche) et graphique de probabilité normale (à droite) pour les résultats granulométriques de l'essai n°3 de sol sec avec fines échantillonné par pelletage fractionné

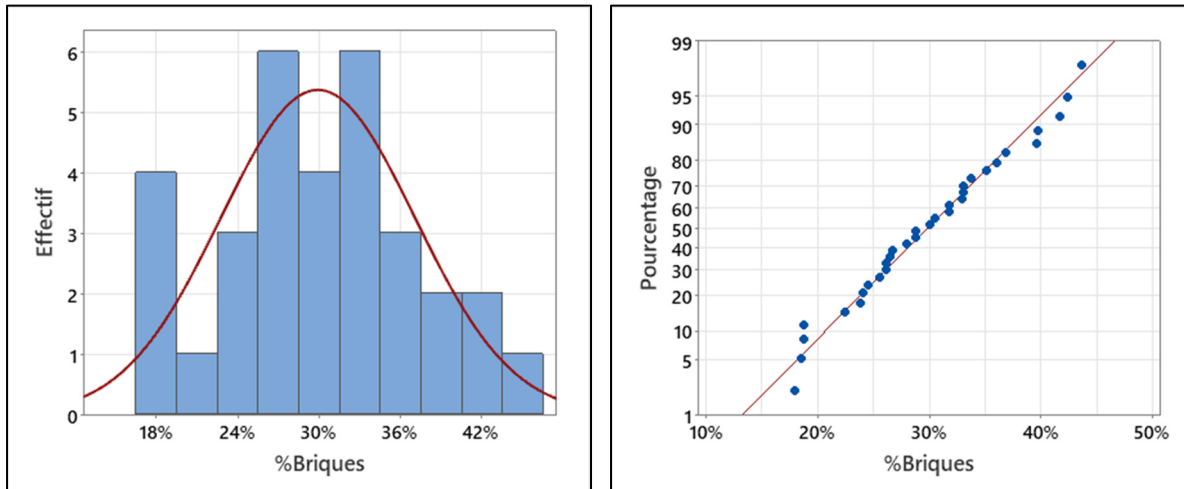


Figure-A V-4 Histogramme avec ajustement normale (à gauche) et graphique de probabilité normale (à droite) pour les résultats granulométriques de l'essai n°4 de sol sec sans fines échantillonné par grappillage

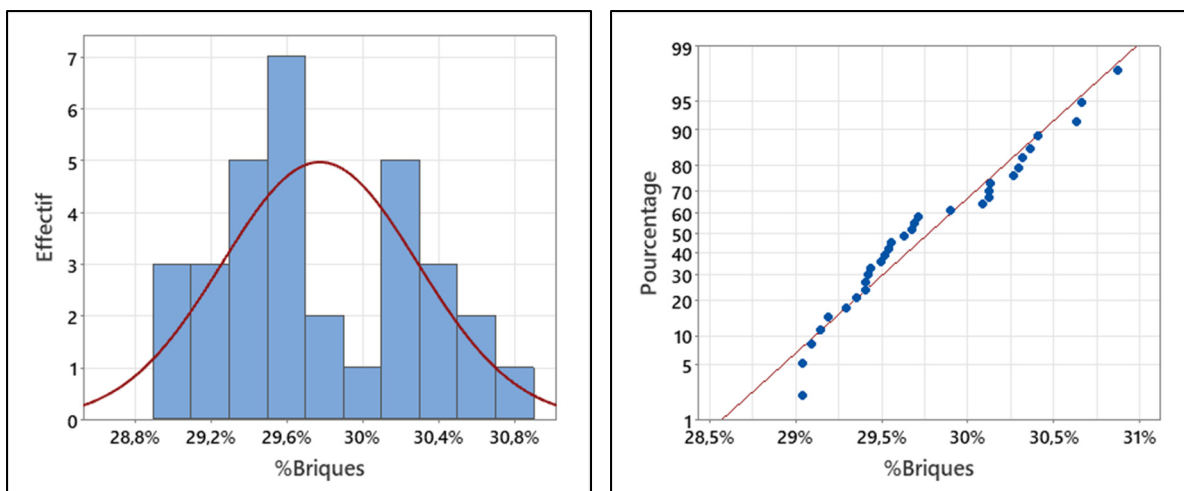


Figure-A V-5 Histogramme avec ajustement normale (à gauche) et graphique de probabilité normale (à droite) pour les résultats granulométriques de l'essai n°5 de sol sec sans fines échantillonné par séparation rotative

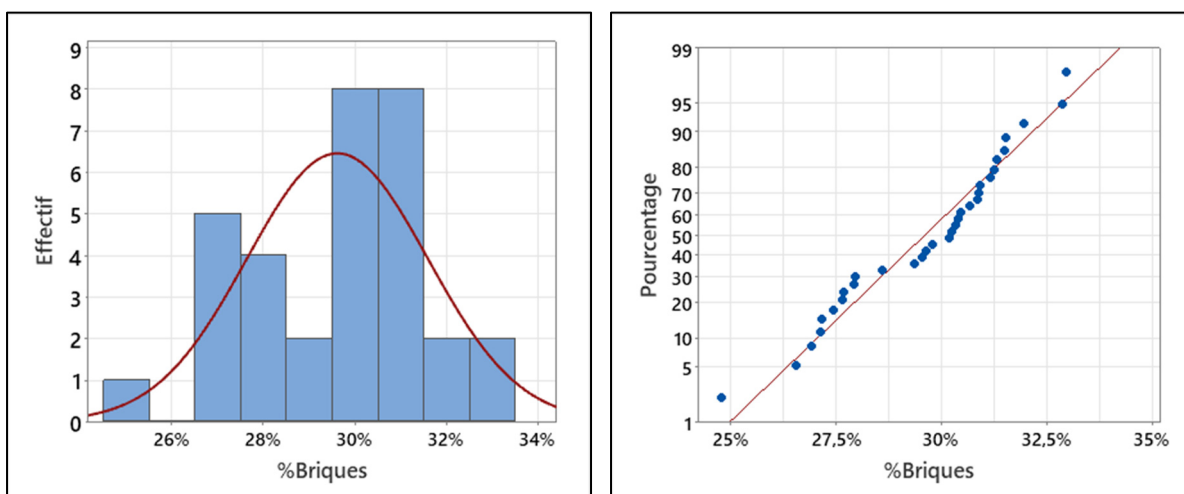


Figure-A V-6 Histogramme avec ajustement normale (à gauche) et graphique de probabilité normale (à droite) pour les résultats granulométriques de l'essai n°6 de sol sec sans fines échantillonné par pelletage fractionné

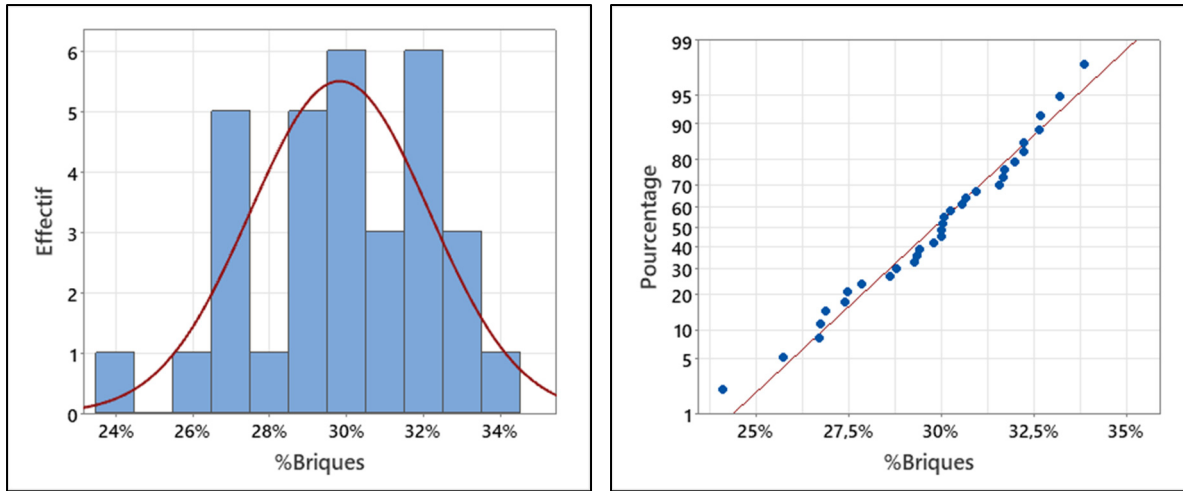


Figure-A V-7 Histogramme avec ajustement normale (à gauche) et graphique de probabilité normale (à droite) pour les résultats granulométriques de l'essai n°7 de sol avec 10% d'eau et fines échantillonné par grappillage

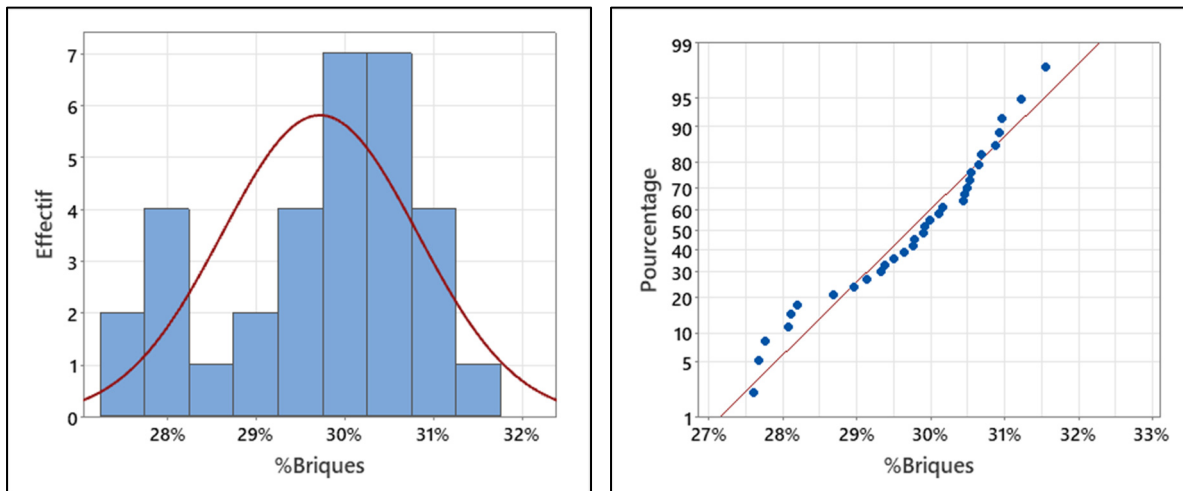


Figure-A V-8 Histogramme avec ajustement normale (à gauche) et graphique de probabilité normale (à droite) pour les résultats granulométriques de l'essai n°8 de sol avec 10% d'eau et fines échantillonné par pelletage fractionné

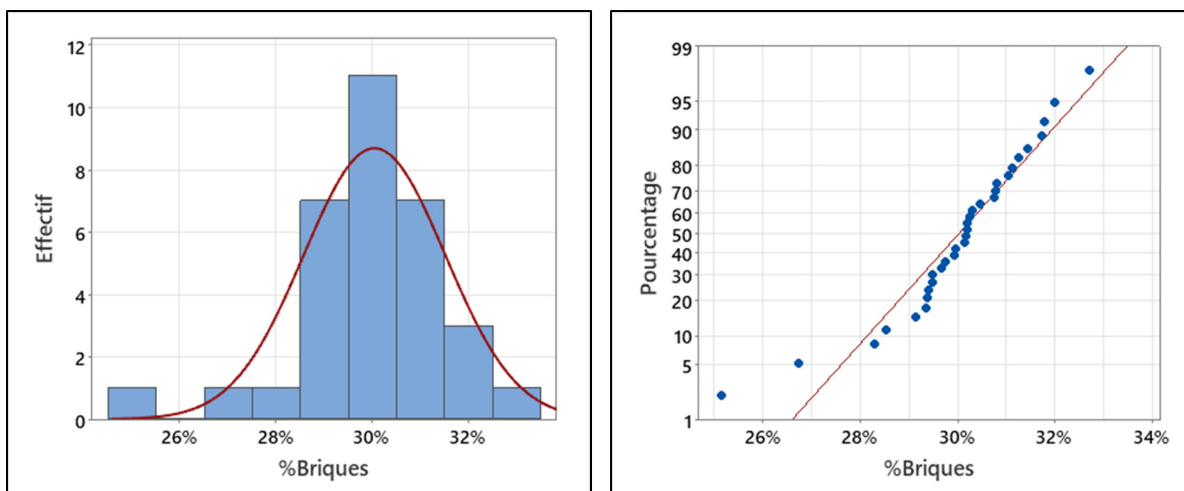


Figure-A V-9 Histogramme avec ajustement normale (à gauche) et graphique de probabilité normale (à droite) pour les résultats granulométriques de l'essai n°9 de sol avec 10% d'huile et fines échantillonné par grappillage

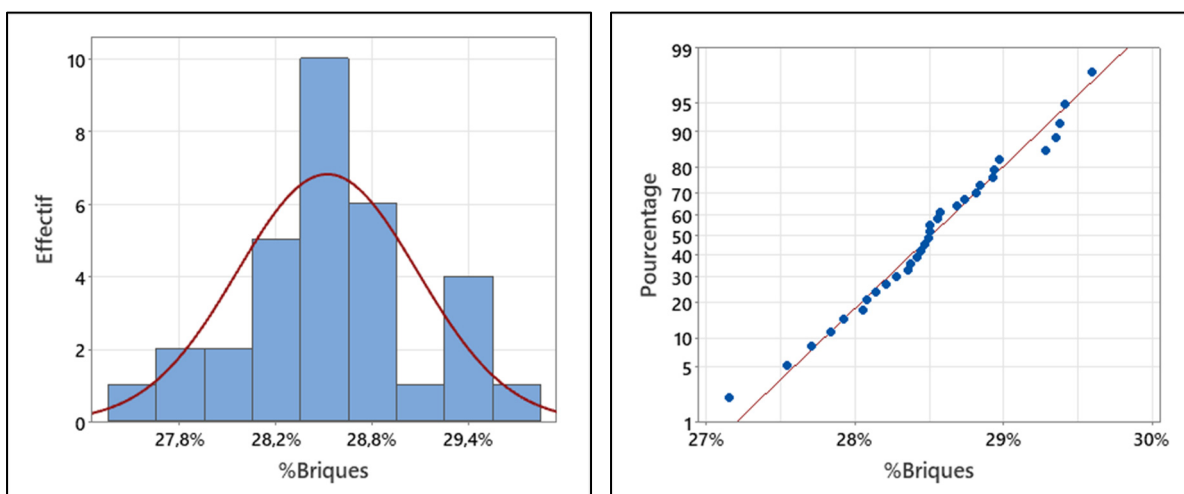


Figure-A V-10 Histogramme avec ajustement normale (à gauche) et graphique de probabilité normale (à droite) pour les résultats granulométriques de l'essai n°10 de sol avec 10% d'huile et fines échantillonné par pelletage fractionné

ANNEXE VI

RÉSULTATS DES TESTS D'ÉGALITÉ DES VARIANCES SUR LES RÉSULTATS DE POURCENTAGE DE BRIQUES

Tableau-A VI-1 Écarts types et intervalles de confiance à 95% associés aux écarts types par fractions granulométriques des différents essais de sol sans fines

Essai	N	Écart-type	Intervalle de confiance
N°1 - sol avec fines, sec, séparation rotative	32	0,006986	(0,0051210 ; 0,010623)
N°2 - sol avec fines, sec, grappillage	32	0,0602022	(0,0441300 ; 0,091543)
N°3 - sol avec fines, sec, pelletage fractionné	32	0,015299	(0,0112146 ; 0,023263)
N°4 - sol sans fines, sec, grappillage	32	0,0714072	(0,0523436 ; 0,108581)
N°5 - sol sans fines, sec, séparation rotative	32	0,0051434	(0,0037702 ; 0,007821)
N°6 - sol sans fines, sec, pelletage fractionné	32	0,0198083	(0,0145201 ; 0,030120)
N°7 - sol avec fines, 10%eau, grappillage	32	0,0232283	(0,0170270 ; 0,035321)
N°8 - sol avec fines, 10%eau, pelletage fractionné	32	0,0109911	(0,0080568 ; 0,016713)
N°9 - sol avec fines, 10%huile, grappillage	32	0,0147188	(0,0107893 ; 0,022381)
N°10 - sol avec fines, 10%huile, pelletage fractionné	32	0,0056244	(0,0041229 ; 0,008552)

LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES

- Abbaszadeh Afshar, F., Ayoubi, S., & Jafari, A. (2018). The extrapolation of soil great groups using multinomial logistic regression at regional scale in arid regions of Iran. *Geoderma*, 315, 36-48. <https://doi.org/10.1016/j.geoderma.2017.11.030>
- Barman, U., & Choudhury, R. D. (2020). Soil texture classification using multi class support vector machine. *Information Processing in Agriculture*, 7(2), 318-332. <https://doi.org/10.1016/j.inpa.2019.08.001>
- Boudreault, J.-P. (2016). *Contrôle de l'incertitude par une approche multidisciplinaire lors de la caractérisation environnementale des sites urbains contaminés* [D.Eng., Ecole de Technologie Supérieure (Canada)]. <https://www.proquest.com/pqdtglobal/docview/1891353065/abstract/8360A87BF28F4EF9PQ/1>
- Boudreault, J.-P., Dube, J.-S., Sona, M., & Hardy, E. (2012). Analysis of procedures for sampling contaminated soil using Gy's Sampling Theory and Practice. *Science of the Total Environment*, 425, 199-207. <https://doi.org/10.1016/j.scitotenv.2012.03.013>
- Bouyim, F.-R. (2022). *Influence de la teneur en eau sur les analyses granulométriques par photographies*. École de Technologie Supérieure.
- Buscombe, D. (2020). SediNet : A configurable deep learning model for mixed qualitative and quantitative optical granulometry. *Earth Surface Processes and Landforms*, 45(3), 638-651. <https://doi.org/10.1002/esp.4760>

- Centre d'expertise en analyse environnementale du Québec. (2008). *Guide d'échantillonnage à des fins d'analyses environnementales—Cahier 5: Échantillonnage des sols*.
<https://www.ceaeq.gouv.qc.ca/documents/publications/echantillonnage.htm>
- Centre d'expertise en analyse environnementale du Québec. (2017). *Méthode d'analyse—Détermination des solides totaux et des solides totaux volatils: Méthode gravimétrique*. https://www.ceaeq.gouv.qc.ca/methodes/methode_numer.htm
- Centre d'expertise en analyse environnementale du Québec. (2023). *Guide d'échantillonnage à des fins d'analyses environnementales—Cahier 1: Généralités*.
<https://www.ceaeq.gouv.qc.ca/documents/publications/echantillonnage.htm>
- Dol, M. (2023). *Influence de la méthode d'échantillonnage sur l'estimation du pourcentage de matière résiduelle dans le sol et analyse d'images par réseau de neurones*. École de Technologie Supérieure.
- Dubé, J.-S., & Esbensen, K. H. (2022). Revisiting Pierre Gy's formula (TOS) – A return to size-density classes for applications to contaminated soils, coated particular aggregates and mixed material systems. *Analytica Chimica Acta*, 1193, 339227.
<https://doi.org/10.1016/j.aca.2021.339227>
- Duhaime, F., Pirnia, P., Manashti, M., Temimi, M., Dubé, J.-S., & Toews, M. (2021, août 11). *Particle size distribution from photographs: Comparison of synthetic and real granular material images*. GeoNiagara 2021.
- FRITSCH. (2019). *Manuel d'utilisation—Diviseur d'échantillon à cône rotatif*.
- FRITSCH. (2020). *Manuel d'utilisation—Trémie d'alimentation à vibrations*.

- Gerlach, R. W., & Nocerino, J. M. (2003). *Guidance for obtaining representative laboratory analytical subsamples from particulate laboratory samples*. US Environmental Protection Agency, Office of Research and Development
- Gonzales, R. C., & Woods, R. E. (2018). *Digital Image Processing* (2^e éd.).
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
<http://www.deeplearningbook.org>
- Guyot, O., Monredon, T., LaRosa, D., & Broussaud, A. (2004). *VisioRock, an integrated vision technology for advanced control of comminution circuits*.
- Gy, P. M. (1992). *Sampling of Heterogeneous and Dynamic Material Systems : Theories of Heterogeneity, Sampling and Homogenizing*. Elsevier.
- Hébert, J., & Bernard, J. (2013). *Bilan sur la gestion des terrains contaminés*. Ministère du Développement Durable, de l'Environnement, de la Faune et des Parcs.
<https://www.environnement.gouv.qc.ca/sol/terrains/index.htm#bilans>
- Hryciw, R., Ohm, H., & Zhou, J. (2013). Theoretical Basis for Optical Granulometry by Wavelet Transformation. *Journal of Computing in Civil Engineering*, 29, 04014050.
[https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000345](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000345)
- ISO. (2023). *Échantillonnage*. DemarcheISO17025.
<https://www.demarcheiso17025.com/echantillonnage.html>
- Lang, N., Irniger, A., Rozniak, A., Hunziker, R., Wegner, J. D., & Schindler, K. (2021). GRAINet: Mapping grain size distributions in river beds from UAV images with convolutional neural networks. *Hydrology and Earth System Sciences*, 25(5), 2567-2597.

- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
- Maerz, N. H., Palangio, T. C., & Franklin, J. A. (1996). *WipFrag image based granulometry system* [Software]. <https://wipware.com/fr/products/wipfrag-image-analysis-software/>
- Manashti, M. (2022). *Grain size analyses of granular material based on image analysis techniques and machine learning* [Phd, École de technologie supérieure].
<https://espace.etsmtl.ca/id/eprint/3076/>
- McKague, K., & Brunke, R. (2023). *Production de biodiesel à la ferme | ontario.ca*.
<http://www.ontario.ca/fr/page/production-de-biodiesel-la-ferme>
- MELCC. (2021). *Règlement concernant la valorisation de matières résiduelles fait par le Ministère de l'environnement et de lutte contre les changements climatiques—Guide d'application*.
<https://www.legisquebec.gouv.qc.ca/fr/document/rc/Q-2,%20r.%2049%20/>
- MELCC. (2023a). *Loi sur les produits pétroliers*. Ministère du Travail, de l'Emploi et de la Solidarité. <https://www.legisquebec.gouv.qc.ca/fr/document/lc/P-30.01>
- MELCC. (2023b). *Répertoire des terrains contaminés*. Ministère de l'Environnement et de la Lutte contre les changements climatiques.
<https://www.environnement.gouv.qc.ca/sol/terrains/terrains-contamines/recherche.asp>
- Ministère de l'environnement et de lutte contre les changements climatiques. (2021). *Guide d'intervention—Protection des sols et réhabilitation des terrains contaminés*.
<https://www.environnement.gouv.qc.ca/sol/terrains/guide-intervention/>
- Ministères Écologie Énergie Territoires. (2019). *Biocarburants*. Ministères Écologie Énergie Territoires. <https://www.ecologie.gouv.fr/biocarburants>

- MTESS. (2023). *Règlement sur la protection et la réhabilitation des terrains*. Ministère du Travail, de l'Emploi et de la Solidarité.
<https://www.legisquebec.gouv.qc.ca/fr/document/rc/q-2,%20r.%2037>
- Petersen, L., Dahl, C. K., & Esbensen, K. H. (2004). Representative mass reduction in sampling—A critical survey of techniques and hardware. *Chemometrics and Intelligent Laboratory Systems*, 74(1), 95-114. <https://doi.org/10.1016/j.chemolab.2004.03.020>
- Piccini, C., Marchetti, A., Riviaccio, R., & Napoli, R. (2019). Multinomial logistic regression with soil diagnostic features and land surface parameters for soil mapping of Latium (Central Italy). *Geoderma*, 352, 385-394.
<https://doi.org/10.1016/j.geoderma.2018.09.037>
- Pirnia, P., Duhaime, F., & Manashti, J. (2018). Machine learning algorithms for applications in geotechnical engineering. *Proc. GeoEdmonton*.
- Pitard, F. F. (2019). *Theory of Sampling and Sampling Practice, Third Edition (3^e éd.)*. Chapman and Hall/CRC. <https://doi.org/10.1201/9781351105934>
- Portmann, J. P. (2023). *Nouvelle classification des sédiments meubles : Classification U.S.C.S.* 58.
- Ramos-Guajardo, A. B., & Lubiano, M. A. (2012). K-sample tests for equality of variances of random fuzzy sets. *Computational Statistics & Data Analysis*, 56(4), 956-966.
<https://doi.org/10.1016/j.csda.2010.11.025>
- Ryan, B. (2020). *MINITAB* [Software]. Minitab, Inc. <http://www.minitab.com/en-US/products/minitab/>
- Shivhare, S., & Cecil, K. (2021). Automatic Soil Classification by using Gabor Wavelet Support Vector Machine in Digital Image Processing. *3rd International Conference on*

Inventive Research in Computing Applications, ICIRCA 2021, September 2, 2021 - September 4, 2021, 1738-1743. <https://doi.org/10.1109/ICIRCA51532.2021.9544897>

Sona, M., & Dubé, J.-S. (2021). Sampling particulate matter for analysis – Controlling uncertainty and bias using the theory of sampling. *Analytica Chimica Acta*, 1185, 338982. <https://doi.org/10.1016/j.aca.2021.338982>

Torralba, A., & Efros, A. A. (2011). Unbiased look at dataset bias. *CVPR 2011*, 1521-1528. <https://doi.org/10.1109/CVPR.2011.5995347>

Transport Québec. (2012). *Méthode d'essai LC-21-040—Analyse granulométrique*.

Transport Québec. (2019). *Méthode d'essai LC 21-901—Détermination de la composition d'un matériau recyclé contenant des résidus d'enrobé et de béton*.

Zeryouh, S. (2023). *Influence de la teneur en eau sur l'échantillonnage des sols à des fins d'estimation du pourcentage des matières résiduelles et d'analyse d'image par réseau neuronal*. École de Technologie Supérieure.