

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

ÉCOLE NATIONALE POLYTECHNIQUE



Génie des Procédés et Environnement

Mémoire de projet de fin d'études

Pour l'obtention du Diplôme d'Ingénieur d'Etat en Génie des Procédés et
Environnement

Optimisation du système de séparation Verti-G par
apprentissage automatique pour le traitement des déblais
de forage

DJEFFAL Khalil & BALAMANE Islam

Sous la direction de :

Promoteur : Mr. BENAMIRA Elias ENP

Co-promoteur : Mr. HACENE Mouhoub ADFC

Présenté et soutenu publiquement le 02/06/2025

Composition du jury :

Président :	Mr. CHERGUI Abdelmalek	Professeur, ENP
Examineur :	Mr. KERCHICH Yacine	Professeur, ENP
Promoteur :	Mr. BENAMIRA Elias	MCB, ENP
Co-Promoteur :	Mr. HACENE Mouhoub	Responsable IT, ADFC

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

ÉCOLE NATIONALE POLYTECHNIQUE



Génie des Procédés et Environnement

Mémoire de projet de fin d'études

Pour l'obtention du Diplôme d'Ingénieur d'Etat en Génie des Procédés et
Environnement

Optimisation du système de séparation Verti-G par
apprentissage automatique pour le traitement des déblais
de forage

DJEFFAL Khalil & BALAMANE Islam

Sous la direction de :

Promoteur : Mr. BENAMIRA Elias ENP

Co-promoteur : Mr. HACENE Mouhoub ADFC

Présenté et soutenu publiquement le 02/06/2025

Composition du jury :

Président : Mr. CHERGUI Abdelmalek Professeur, ENP

Examineur : Mr. KERCHICH Yacine Professeur, ENP

Promoteur : Mr. BENAMIRA Elias MCB, ENP

Co-Promoteur : Mr. HACENE Mouhoub Responsable IT, ADFC

Dédicace

" À toi, **maman**,
douce étoile dans mes nuits d'incertitude,
dont l'amour inconditionnel a guidé chacun de mes pas.

À toi, **papa**,
pilier silencieux et fort,
dont le regard discret m'a toujours porté plus loin.

À toi, **ma sœur**,
complice de cœur et d'âme,
pour tes sourires, ton écoute et ta lumière dans les jours gris.

À **mes enseignants**,
artisans de l'esprit,
qui m'ont transmis bien plus que du savoir : une voie.

À **mes amis**,
présents dans les doutes, dans les rires, dans le silence,
merci d'avoir été le souffle et l'équilibre dans cette traversée.

À vous tous,
cette réussite porte l'empreinte de votre amour, de votre foi et de votre bienveillance."

Khalil

Dédicace

À ma source de tendresse, à la femme la plus patiente et la plus forte que je connaisse, à celle dont le courage, la dignité et l'amour inconditionnel m'ont porté jusqu'ici, à ma très chère mère, en témoignage de mon affection profonde, de mon respect éternel et de toute ma reconnaissance, que les mots ne suffiront jamais à exprimer.

À mon repère et mon idéal, à l'homme généreux et discret, dont les sacrifices silencieux et les valeurs transmises m'ont guidé sans relâche, à mon très cher père, qui m'a tout donné sans jamais rien attendre en retour, et qui, par ses conseils, sa dignité et son sens du devoir, m'a inspiré le courage et l'amour du travail, que ce travail soit le reflet de sa confiance et de ses espérances.

À mon seul et unique frère, à Souhib, dont la présence m'est si précieuse, je dédie ce travail en lui souhaitant de tout cœur une vie pleine de réussite, d'accomplissements et de sérénité.

À tous les enseignants qui ont contribué, de près ou de loin, à ma formation, je vous exprime ma profonde gratitude pour la richesse de votre enseignement, la rigueur de votre accompagnement et la passion que vous avez su transmettre.

À tous mes amis et camarades, en témoignage de mes plus profondes amitiés et de ma sincère sympathie, merci pour les moments partagés, le soutien, les sourires et la solidarité tout au long de ce parcours.

À tous ceux qui ont contribué, de près ou de loin, à la réalisation de ce travail, je vous adresse mes remerciements les plus sincères. Votre aide, votre soutien ou même vos encouragements discrets ont tous compté.

Islam

المخلص

تولد عمليات حفر الابار النفطية نفايات صلبة (Cuttings) ملوثة بسوائل الحفر الزيتية (OBM)، مما يشكل تحديات بيئية كبيرة. يهدف هذا العمل، المنجز بالتعاون مع شركة ADFC، إلى تحسين أداء نظام Verti-G، وهو فاصل ميكانيكي (جهاز طرد مركزي) مصمم لتقليل محتوى الزيت المتبقي (OOC) على هذه النفايات.

تم الاعتماد على بيانات صناعية واقعية لاختبار عدة نماذج تنبؤية Random Forest، SVR، LightGBM، CatBoost، حيث تم تحسينها باستخدام مكتبة Optuna، ثم دمجها بتقنية Stacking، مما مكن من تحقيق معامل تحديد (R^2) قدره 0.889 وخطأ جذري متوسط (RMSE) قدره 0.294. في المرحلة الأخيرة، تم تنفيذ تحسين عكسي لإعدادات جهاز Verti-G، حيث أوصى النموذج بإعدادات قللت نسبة الزيت المتوقعة من 7.41% إلى 5.66% مع الحفاظ على القيم ضمن النطاقات التشغيلية الواقعية.

تظهر هذه النتائج مدى فعالية تقنيات التعلم الآلي في تحسين الأداء الصناعي والحد من الأثر البيئي لنفايات الحفر النفطية.

الكلمات المفتاحية : الحفر – سوائل الحفر – نفايات الحفر – Verti-G – التحسين – التعلم الآلي – الذكاء الاصطناعي.

Abstract

Oil drilling operations generate cuttings contaminated with oil-based mud (OBM), posing significant environmental challenges. This study, conducted in collaboration with ADFC, aims to optimize the VertiG system, a mechanical separator designed to reduce the residual oil content (OOC) on cuttings.

Using industrial data, several regression models (Random Forest, SVR, LightGBM, CatBoost) were tested, optimized with Optuna, and combined through stacking, achieving a maximum R^2 of 0.889 and an RMSE of 0.294. Finally, an inverse optimization of the Verti-G settings was performed, recommending parameters that reduced the predicted OOC from 7.41% to 5.66%, while remaining within operationally feasible ranges.

These results highlight the relevance of Machine Learning techniques to enhance industrial performance and mitigate the environmental impact of petroleum waste.

Keywords : Drilling – drilling fluids – cuttings – Verti-G – optimization – Machine Learning – Artificial Intelligence (AI).

Résumé

Les opérations de forage pétrolier génèrent des déblais (cuttings) contaminés par des boues à base d'huile (OBM), posant des défis environnementaux majeurs. Ce travail, en collaboration avec la société ADFC, vise à optimiser le système Verti-G, un séparateur mécanique destiné à réduire la teneur en huile résiduelle (OOC) des cuttings.

À partir de données industrielles, plusieurs modèles de régression ont été testés (Random Forest, SVR, LightGBM, CatBoost), optimisés via Optuna, puis combinés par stacking, atteignant un R^2 maximal de 0.889 et un RMSE de 0.294. Enfin, une optimisation inverse des réglages Verti-G a permis de recommander des paramètres réduisant l'OOC prédite de 7.41% à 5.66 %, tout en restant dans des plages industrielles réalistes.

Ces résultats démontrent l'intérêt des techniques de Machine Learning pour améliorer les performances industrielles et limiter l'impact environnemental des déchets pétroliers.

Mots clés : Forage – fluides de forage – déblais de forage – Verti-G – optimisation – Machine Learning – Intelligence Artificielle (IA).

Remerciements

Nous remercions avant tout Allah, Le Tout-Puissant, pour nous avoir guidés, soutenus et permis d'achever ce travail dans les meilleures conditions.

Nous exprimons notre profonde reconnaissance à Monsieur Mouhoub HACENE, notre promoteur, pour son encadrement rigoureux, ses conseils pertinents et sa disponibilité constante tout au long de ce projet.

Nos remerciements les plus sincères vont également à Docteur BENAMIRA Elias, notre co-promoteur, pour son accompagnement technique, sa bienveillance et son soutien tout au long de cette expérience.

Nous tenons à exprimer notre gratitude à Monsieur le Professeur CHERGUI Abdelmalek, Président du jury, ainsi qu'à Monsieur le Professeur KERCHICH Yacine, Examineur, pour l'honneur qu'ils nous ont fait en acceptant d'évaluer ce travail, et pour l'intérêt qu'ils y ont porté.

Enfin, nous adressons nos vifs remerciements à toute l'équipe de l'ADFC (Algerian Drilling Fluids Company) pour leur accueil chaleureux, leur accompagnement et les moyens mis à notre disposition pour la réussite de ce projet.

Table des matières

Liste des tableaux

Table des figures

Liste des acronymes

Introduction générale	14
1 Le forage pétrolier	17
1.1 Introduction	17
1.2 Le forage pétrolier	17
1.3 le système de forage	18
1.4 Préparation du site de forage	19
1.5 Les méthodes de forage	20
1.6 Classification des appareils de forage pétrolier	22
1.7 Gestion des Opérations de Forage et Organisation du Personnel	24
2 Généralités sur les fluides de forage	25
2.1 Introduction	25
2.2 Définition d'un fluide de forage	25
2.3 Rôles des fluides de forage	26
2.4 Les types de fluide de forage	27
2.5 Les propriétés des fluides de forage	29
2.6 Le choix des boues de forage	29
2.7 Conclusion	30
3 Les déblais de forage (cuttings)	31

3.1	Introduction	31
3.2	Définition et origine des déblais de forage	31
3.3	Types de déblais de forage (cuttings)	32
3.4	Déblais de Forage : Propriétés, Usages et Enjeux Opérationnels	33
3.5	Sources de contamination des déblais de forage	34
3.6	Impacts des déblais de forage sur l'environnement	37
3.7	Législation et normes	39
4	La gestion des déblais de forage	41
4.1	Introduction	41
4.2	Les méthodes de traitement physique des cuttings	42
4.3	Les méthodes de traitement chimiques des déblais de forage	48
4.4	Traitement thermique	49
4.5	Traitement biologique	52
5	Présentation du processus Verti-G et des données industrielles	55
5.1	Introduction générale à la partie pratique	55
5.2	Présentation des données industrielles	56
6	Nettoyage, traitement et analyse exploratoire des données	60
6.1	traitement des valeurs manquantes (NaN)	60
6.2	Méthodes de gestion des valeurs manquantes	63
6.3	Mise en place du prétraitement	63
6.4	Distribution de la cible : Teneur en huile résiduelle sur les cuttings secs (% OOC)	63
6.5	Heatmap de la matrice de corrélation	65
6.6	Importance des variables	66
6.7	Conclusion	69
7	Modélisation prédictive de la teneur en huile résiduelle	71
7.1	Sélection et comparaison des modèles de Machine Learning	71
7.2	Optimisation des hyperparamètres avec Optuna	74
7.3	Assemblage de modèles (stacking)	76
7.4	Construction de l'assembleur	78

7.5	Évaluation des performances du stacking	79
7.6	Sélection du meilleur modèle et discussion des résultats	80
7.7	Validation du modèle sélectionné : RandomF_SVR_LGBM	83
8	Optimisation inverse des réglages Verti-G	87
8.1	Formulation du problème inverse	87
8.2	Cas pratique : exemple sur un mélange réel	89
8.3	Discussion sur la faisabilité opérationnelle des réglages	91
	Conclusion générale	93
	Bibliographie	94

Liste des tableaux

1.1	Types d'appareils de forage et leurs caractéristiques	24
3.1	Composition chimique, quantité et toxicité des fluides de forage utilisés à Hassi Messaoud	36
3.2	Valeurs limites pour la mise en décharge des déblais de forage selon le standard européen	40
6.1	Pourcentage de valeurs manquantes (NaN) par variable	62
7.1	Résultats comparatifs des performances des modèles de régression	73
7.2	Comparaison des performances (R^2 , RMSE) des modèles optimisés par Grid Search et Optuna	75
7.3	Comparaison des performances des modèles SVR, Random Forest et Stacking (RF + SVR)	79
7.4	Performances comparées des modèles individuels et empilés (stacking) selon cinq métriques d'évaluation	82

Table des figures

1.1	Etapes de forage	18
1.2	Principaux éléments des systèmes de forage actuels (Bobo et Hoch, 1957).	19
1.3	Une installation de forage par câble conventionnelle.	21
1.4	Une installaUne installation de forage rotatif conventionnelle	22
1.5	Types d'installations pétrolières : Onshore et Offshore	22
2.1	Circuit de circulation du fluide de forage	26
2.2	Les différents composants des fluides de forage (les boues à base d'huile et les boues à base d'eau)	28
3.1	Zone de stockage des déblais de forage (bourbier)	32
3.2	Interaction entre les fluides de forage et la formation géologique	35
3.3	Sauvetage d'un chameau piège dans un bournier	38
3.4	Préparation du terrain	39
4.1	Schéma du traitement des déblais de forage à base d'huile	41
4.2	Vibrateur de schiste typique	43
4.3	Un dégazeur à vide typique	44
4.4	Principe de fonctionnement des hydrocyclones	45
4.5	Un désableur typique	46
4.6	Un désiltateur typique	46
4.7	Un nettoyeur de boue	47
4.8	Une centrifugeuse de décantation typique	48
4.9	Schéma simplifié de l'unité de traitement par désorption thermique.	50
4.10	Thermal Cuttings Cleaner (TCC)	51
4.11	Thermal Phase Separation (TPS)	52

5.1	Vue extérieure du Verti-G	56
5.2	Vue en coupe détaillée du Verti-G	57
5.3	Échantillon des données industrielles brutes	58
5.4	Extrait du tableau récapitulatif des statistiques descriptives des variables	59
6.1	Heatmap des valeurs manquantes dans la base de données	61
6.2	Distribution de la teneur en huile résiduelle sur les cuttings secs (% OOC) . . .	64
6.3	Matrice de corrélation des variables	65
6.4	Importance des variables (barplot)	67
6.5	Visualisation des valeurs SHAP pour l'importance des variables	68
6.6	Flowchart pour choisir un estimateur en Machine Learning	70
7.1	Comparaison des modèles de régression selon le score R^2	73
7.2	Schéma de fonctionnement d'un modèle de stacking	77
7.3	Courbe d'apprentissage du modèle <code>RandomF_SVR_LGBM</code>	84
7.4	Distribution des résidus du modèle <code>RandomF_SVR_LGBM</code>	85
8.1	Définition des paramètres d'entrée du mélange réel (ligne 59 de la base de données)	90
8.2	Réglages Verti-G optimaux proposés par Optuna et prédiction de la teneur en huile	91

Liste des acronymes

- **A.I.** : Artificial Intelligence (Intelligence Artificielle)
- **ADFC** : Algerian Drilling Fluids Company
- **API** : American Petroleum Institute
- **BOP** : Blowout Preventer
- **CO₂** : Dioxyde de carbone
- **CV** : Cross Validation
- **HP** : High Pressure
- **LC50** : Lethal Concentration 50 – Concentration létale pour 50% des individus
- **LD50** : Lethal Dose 50 – Dose létale pour 50% des individus
- **LGBM** : Light Gradient Boosting Machine
- **ML** : Machine Learning
- **MSE** : Mean Squared Error
- **RMSE** : Root Mean Squared Error
- **OBM** : Oil-Based Mud
- **OOC** : Oil On Cuttings
- **RPM** : Revolutions Per Minute
- **SG** : Specific Gravity (masse volumique relative)
- **SHAP** : SHapley Additive exPlanations
- **SVM** : Support Vector Machine
- **SVR** : Support Vector Regression
- **TCC** : Thermal Cuttings Cleaner
- **TPE** : Tree-structured Parzen Estimator
- **TPS** : Thermal Phase Separation
- **VG** : Verti-G (séparateur mécanique pour le traitement des cuttings)
- **WBM** : Water-Based Mud
- **WOB** : Weight On Bit
- **XGBoost** : Extreme Gradient Boosting

Introduction générale

L'exploitation des ressources pétrolières demeure, depuis plusieurs décennies, un pilier fondamental du développement économique mondial. En Algérie, ce secteur joue un rôle stratégique, contribuant de manière significative au produit intérieur brut et représentant la majeure partie des recettes d'exportation. Toutefois, derrière cette importance économique se cachent des défis environnementaux et opérationnels considérables, en particulier liés aux déchets générés par les opérations de forage. Parmi ces déchets figurent les déblais de forage, ou cuttings, qui constituent des fragments de roche imprégnés de fluides de forage, notamment à base d'huile, et qui représentent une source majeure de pollution potentielle des sols, des eaux souterraines et de l'atmosphère.

La gestion et le traitement de ces déchets sont donc devenus des enjeux cruciaux pour concilier exploitation pétrolière et préservation de l'environnement. Les normes environnementales de plus en plus strictes imposent aux opérateurs de mettre en œuvre des procédés de traitement efficaces, capables de réduire la charge polluante des cuttings tout en permettant la récupération et la valorisation des hydrocarbures résiduels. Dans ce contexte, le procédé Verti-G, développé et exploité par la société ADFC, se distingue comme une solution technique incontournable. Ce séparateur mécanique est conçu pour réduire la teneur en huile des cuttings avant leur élimination ou leur valorisation, contribuant ainsi à la durabilité et à la rentabilité des opérations de forage.

Cependant, l'efficacité du Verti-G repose sur un ensemble complexe de paramètres opérationnels qui interagissent entre eux de façon non triviale. Une mauvaise configuration peut entraîner une séparation inefficace, une consommation excessive d'énergie ou un rejet non conforme aux normes environnementales. Ainsi, l'optimisation de ces réglages devient un objectif prioritaire pour garantir un fonctionnement optimal, minimiser les impacts écologiques et améliorer la rentabilité globale.

Dans ce cadre, les avancées récentes en intelligence artificielle, et plus particulièrement en apprentissage automatique (Machine Learning), offrent de nouvelles perspectives. Ces outils permettent d'exploiter les vastes volumes de données opérationnelles collectées sur le terrain pour développer des modèles prédictifs capables de capturer la complexité des interactions entre les paramètres et de proposer des réglages optimaux.

Le présent projet de fin d'études s'inscrit dans cette dynamique et se donne pour objectif d'étudier, d'analyser et d'optimiser le fonctionnement du système Verti-G. Plus précisément, ce travail vise à développer un modèle prédictif robuste capable d'estimer la teneur en huile résiduelle sur les cuttings secs en fonction des paramètres opérationnels mesurés. Ce modèle servira ensuite de base pour une optimisation inverse, permettant d'identifier les combinaisons de réglages les plus performantes afin de réduire l'impact environnemental tout en préservant l'efficacité industrielle.

Pour atteindre cet objectif, une démarche rigoureuse a été adoptée, articulant plusieurs étapes complémentaires : une étude bibliographique des procédés de traitement existants, une analyse exploratoire des données industrielles fournies par ADFC, la conception et l'évaluation de modèles d'apprentissage automatique, l'optimisation des hyperparamètres par des techniques avancées comme Optuna, et enfin l'exploitation du modèle pour guider le réglage optimal du Verti-G. Cette approche, à la fois technique et environnementale, illustre la convergence entre les exigences de durabilité et les opportunités offertes par la transformation numérique de l'industrie pétrolière.

En somme, ce mémoire ambitionne de démontrer comment l'intégration des méthodes d'intelligence artificielle dans la gestion des déchets pétroliers peut contribuer à améliorer les performances industrielles tout en réduisant l'empreinte écologique. Au-delà de la dimension technique, il met également en lumière la nécessité de repenser les pratiques industrielles dans une optique de développement durable, en faisant de la donnée un levier stratégique pour optimiser les processus et répondre aux défis environnementaux contemporains.

Partie Théorique

Chapitre 1

Le forage pétrolier

1.1 Introduction

L'exploration et l'exploitation des ressources pétrolières constituent des enjeux majeurs pour l'industrie énergétique mondiale. Au cœur de ce processus se trouve une étape technique essentielle : le forage pétrolier. Cette opération complexe permet d'atteindre les réservoirs souterrains d'hydrocarbures, souvent situés à des milliers de mètres de profondeur. En alliant ingénierie de pointe, logistique rigoureuse et technologies innovantes, le forage joue un rôle déterminant dans la mise en valeur des gisements pétroliers, qu'ils soient terrestres ou offshore. Ce chapitre vise à présenter les principes fondamentaux du forage pétrolier, en détaillant les équipements utilisés, les différentes méthodes mises en œuvre, ainsi que les types de puits forés selon les objectifs poursuivis. Il met également en lumière l'organisation des opérations sur site et les avancées en matière de forage directionnel, soulignant ainsi l'évolution constante de cette discipline au service de l'industrie des hydrocarbures.

1.2 Le forage pétrolier

Le forage pétrolier désigne l'ensemble des opérations permettant d'atteindre les couches souterraines poreuses et perméables susceptibles de contenir du pétrole ou du gaz naturel. Cette technique consiste à installer un dispositif de forage à la surface, équipé d'un derrick, une structure métallique d'environ 30 mètres de haut, utilisée pour insérer verticalement les tiges de forage. Ces tiges, constituées de tubes vissés les uns aux autres, forment une garniture de forage à l'extrémité de laquelle se trouve le trépan, l'outil chargé de percer le sous-sol (voir Figure 1.1a). À mesure que le forage progresse, de nouveaux tubes sont ajoutés pour approfondir le puits. Pendant cette opération, un fluide de forage est injecté en continu afin de remonter les débris, maintenir la stabilité du puits, contrôler la pression et refroidir le trépan (voir Figure 1.1b). [1]

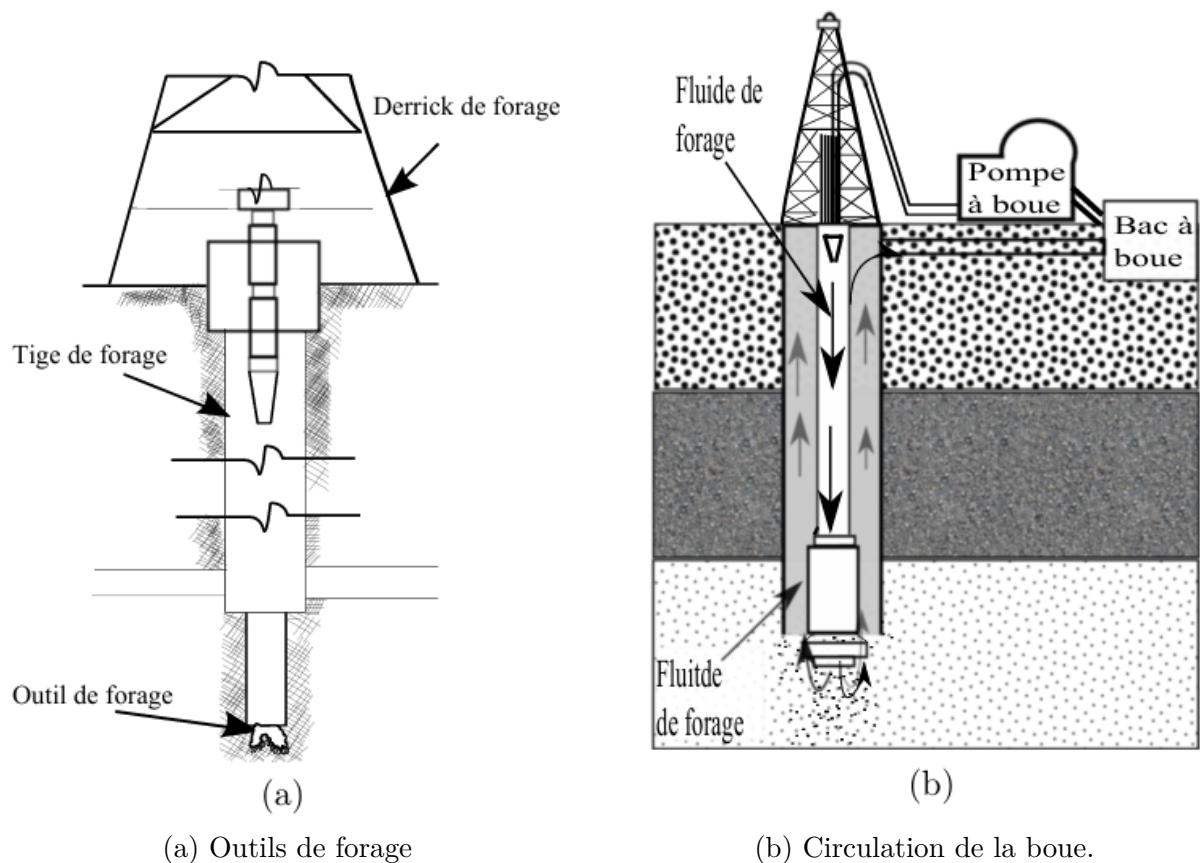


FIGURE 1.1 – Etapes de forage

1.3 le système de forage

Les principaux éléments des systèmes de forage actuels sont illustrés à la Figure 1.2. La tête de forage (*drill head*) broie la roche au fond du trou. Dans la plupart des cas, il s'agit d'un trépan (*drill bit*) qui fracture la roche par action mécanique ou mécano-hydraulique.

La vitesse de fragmentation dépend de plusieurs facteurs : la conception du trépan (efficacité de coupe et résistance à l'usure), le type de roche, la température, la pression, ainsi que la procédure d'exploitation (l'expérience et les compétences de l'opérateur peuvent fortement influencer la performance du forage).

Le trépan est entraîné soit depuis la surface via une tige de transmission (*drive string*), soit par un moteur en fond de trou (*downhole motor*).

Après fragmentation, les débris rocheux doivent être évacués vers la surface. Dans les tunnels, cela se fait souvent par des systèmes mécaniques complexes. En forage pétrolier et géothermique, l'évacuation se fait généralement à l'aide d'un fluide de forage (*mud*) ou d'air comprimé, comme illustré dans la figure 1.2 [2].

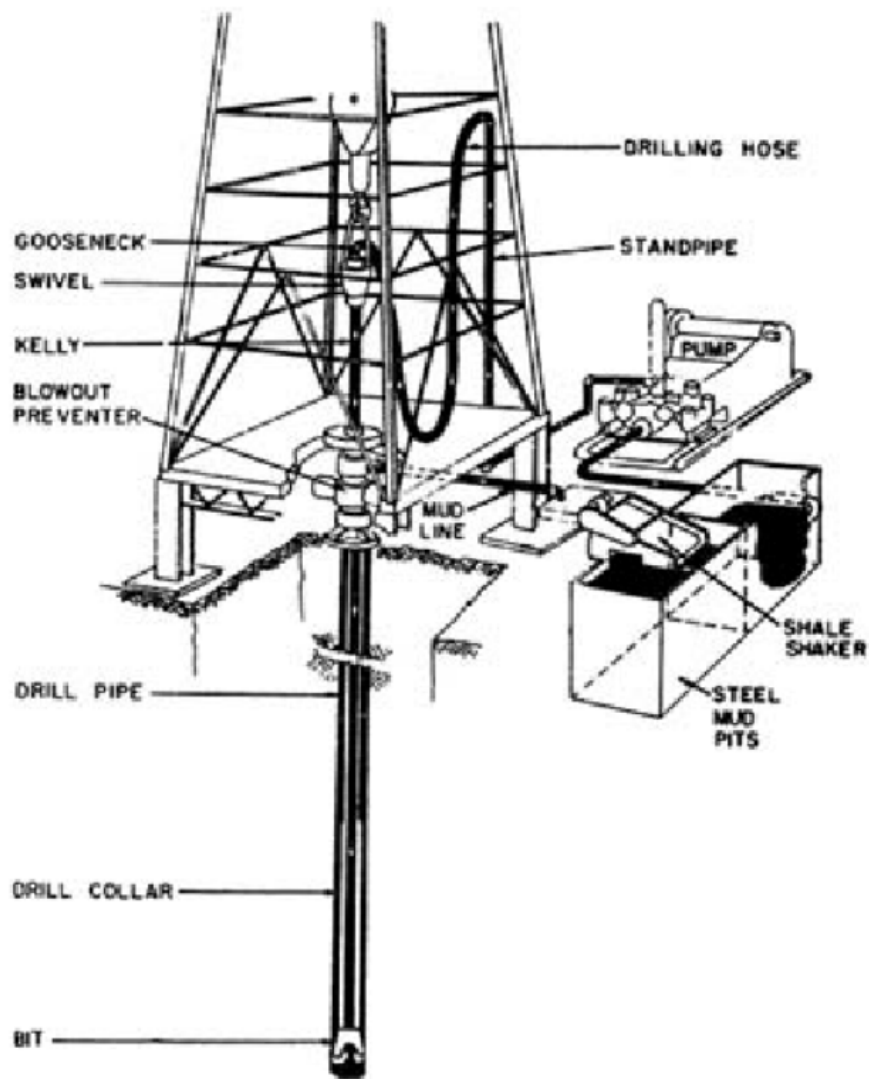


FIGURE 1.2 – Principaux éléments des systèmes de forage actuels (Bobo et Hoch, 1957).

1.4 Préparation du site de forage

La préparation d'un site de forage débute par le relevé topographique de l'emplacement du puits, permettant de définir précisément sa position et son altitude. Un plan est ensuite soumis aux autorités compétentes. Un bulldozer aménage une route d'accès, une zone de retournement et aplanit la surface du site (plateforme de forage), qui couvre en général entre 4 et 6 acres (soit 0,016 à 0,024 km²). En terrain humide, des planches ou des madriers (souvent 7,5 × 30 cm) peuvent être posés pour stabiliser la surface et favoriser le drainage. Un bassin de décantation (reserve pit) est creusé à proximité de l'appareil de forage, doublé d'un revêtement plastique, pour recueillir les boues usagées, les déblais et autres déchets. L'approvisionnement en eau est assuré par un puits d'eau ou une canalisation. La profondeur du puits détermine la taille et la puissance de l'équipement. Les forages peu profonds (<1000 m) sont réalisés avec un appareil monté sur camion (portable rig). Pour les puits plus profonds, le matériel est transporté en modules sur plusieurs camions et assemblé sur place. Pour les puits profonds, un cellier rectangulaire est creusé sous le plancher du forage pour accueillir les blowout preventers (BOPs). En zones très reculées, des helirigs (appareils transportables par hélicoptère) sont utilisés. L'installation du matériel est appelée rig up, son démontage rig down. Le début du forage, appelé spudding in, commence souvent avec un appareil léger qui fore un trou conducteur

peu profond (7 à 30 m de profondeur, diamètre 50 cm). Ce trou est ensuite équipé d'un tubage conducteur (conductor casing), cimenté ou enfoncé directement dans les sols meubles, afin de stabiliser l'ouverture du puits et fixer les BOPs, notamment en présence de gaz peu profonds.[3]

1.5 Les méthodes de forage

Il existe deux principales techniques de forage utilisées pour l'extraction des hydrocarbures[4] :

- le forage par câble (Cable Tool Drilling),
- le forage rotatif (Rotary Drilling).

1.5.1 Le forage par câble

Le forage par câble repose sur l'utilisation d'un trépan lourd fixé à un câble, qui est levé puis relâché de façon répétée afin de broyer la roche par percussion. Ce mouvement vertical est généré par un bras oscillant (walking beam), actionné par un moteur à travers un ensemble de poulies et de roues. À intervalles réguliers, le train de tiges est retiré pour permettre l'évacuation des déblais à l'aide d'un baquet (bailer).

L'équipement typique comprend : un moteur, une chaudière, une tour de forage (derrick), un système de câbles, des roues motrices (bull wheel), une roue de sable (sand wheel), le trépan et d'autres outils auxiliaires (Figure 1.3). Des tubes de cuvelage sont utilisés pour maintenir la stabilité du puits.

Bien que progressivement remplacée par des techniques plus modernes, cette méthode reste utile pour les forages de grands diamètres ou dans les zones rurales. Elle présente toutefois certains inconvénients : lenteur, efficacité limitée dans les terrains meubles, interruptions fréquentes pour nettoyer le fond du puits, et faible contrôle des pressions souterraines, comme illustré dans la figure 1.3 [4]

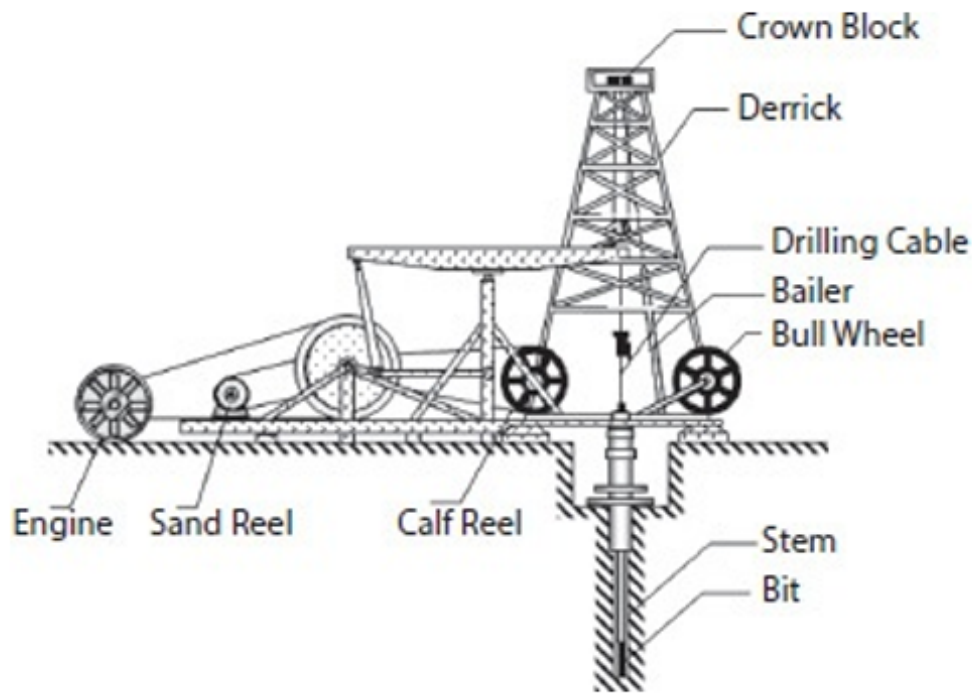


FIGURE 1.3 – Une installation de forage par câble conventionnelle.

1.5.2 Le forage rotatif

Le forage rotatif repose sur l'application d'un mouvement de rotation sur un trépan situé à l'extrémité d'un assemblage de fond de trou. Ce système est relié à une colonne de forage composée de tiges et de colliers de forage. À mesure que le forage progresse, de nouvelles sections de tiges sont ajoutées.

Un fluide de forage (ou boue) est injecté sous pression à travers les tiges, traverse le trépan, et remonte en surface en entraînant les débris rocheux. Ce fluide est ensuite traité et réutilisé en circuit fermé, assurant à la fois le nettoyage du puits, le refroidissement du trépan, la stabilisation des parois et le contrôle de la pression.

Cette technique, plus récente et largement adoptée dans l'industrie, permet d'atteindre des profondeurs importantes, jusqu'à 9.000 mètres. Elle est adaptée à divers types de formations géologiques. Les configurations typiques de forage rotatif est illustrée à la Figure 1.4 :[4]



FIGURE 1.4 – Une installaUne installation de forage rotatif conventionnelle

1.6 Classification des appareils de forage pétrolier

Il existe différents types d'appareils de forage pour extraire le pétrole, que ce soit sur terre ou en mer selon le processus de forage requis.[5]



(a) Plateforme pétrolière Onshore [5]



(b) Plateforme pétrolière Offshore [5]

FIGURE 1.5 – Types d'installations pétrolières : Onshore et Offshore

1.6.1 Appareil Leger

La plateforme de forage pétrolier léger est une unité de petite taille utilisée pour les opérations de forage à faible profondeur. Comprend un mât, un système de levage, une table rotative et une unité de puissance telle qu'un moteur diesel ou un moteur électrique. La plateforme est généralement utilisée pour forer des puits de petit diamètre et est portable, souvent montée sur un camion ou une remorque.[4]

1.6.2 Appareil moyen

La plateforme de forage pétrolier moyen est une unité de taille moyenne utilisée pour les opérations de forage peu profondes et profondes. Il comprend un mât, un système de levage, une table rotative et une unité d'alimentation comme une plate-forme légère. Cependant, il a une plus grande capacité de creusement et peut forer à de plus grandes profondeurs. Il est souvent monté sur un patin ou une remorque pour un transport facile et peut être utilisé pour le forage à terre et en mer.[4]

1.6.3 Appareil lourd

La plateforme de forage pétrolier lourd est une unité à grande échelle conçue pour les opérations de forage profond, capable de forer à des profondeurs de plusieurs milliers de mètres. Il comprend un mât, un système de levage, une table rotative, une unité motrice et des équipements auxiliaires tels que des pompes à boue, des obturateurs anti-éruptions et des démenageurs aériens. Il est généralement utilisé pour les opérations de forage complexes et est souvent monté sur une plate-forme en acier ou un navire de forage pour les opérations de forage en mer.[4]

1.6.4 Appareil super

lourd L'appareil de forage pétrolier super lourd est l'appareil de forage le plus grand et le plus puissant utilisé dans l'industrie pétrolière et gazière, conçu pour des opérations de forage très profondes jusqu'à 12 000 mètres ou plus. Il comprend un énorme mât ou tour de levage, un système de levage puissant, une grande plaque tournante, une unité motrice de grande capacité et une gamme d'équipements auxiliaires. Ce type de plate-forme est généralement utilisé dans les opérations de forage en eau très profonde et est monté sur une grande plate-forme de forage. Sa construction et son exploitation nécessitent des investissements importants en ressources et constituent l'une des plates-formes de forage les plus sophistiquées et les plus avancées de l'industrie. [4]

D'après le tableau 1.1, les types d'appareils de forage peuvent être classés comme suit [6] :

TABLE 1.1 – Types d’appareils de forage et leurs caractéristiques

Types d’appareils	Distance d’excavation (m)	Capacité
Appareil léger	1500 m – 2000 m	650 HP
Appareil moyen	3500 m	1300 HP
Appareil lourd	6000 m	2000 HP
Appareil super lourd	8000 m – 10000 m	3000 HP

1.7 Gestion des Opérations de Forage et Organisation du Personnel

L’exploitation d’un appareil de forage est coûteuse et se fait en continu, 24h/24, généralement en trois tours de 8 heures ou deux de 12 heures, selon le site. Le foreur (driller) est responsable de chaque tour et opère les équipements depuis une console, en surveillant plusieurs paramètres clés comme la profondeur, la pression, la vitesse de rotation et la vitesse de pénétration.

L’appareil est exploité par un entrepreneur de forage, tandis que le chef de chantier (tool pusher) supervise l’ensemble des opérations sur site. L’opérateur, c’est-à-dire la compagnie pétrolière, finance le projet et désigne un représentant pour collaborer avec le chef de chantier. Ensemble, ils rédigent un rapport quotidien détaillant les activités, la progression du forage et les données techniques.

Le personnel comprend aussi un opérateur de derrick, responsable des équipements de circulation de la boue, et des aides-foreurs qui assurent la manipulation et l’entretien du matériel. Sur les grandes installations, un opérateur moteur gère les moteurs principaux.[\[3\]](#)

Chapitre 2

Généralités sur les fluides de forage

2.1 Introduction

Dans les opérations de forage pétrolier et gazier, le fluide de forage (boue de forage) joue un rôle essentiel à toutes les étapes du processus. Véritable outil multifonction, il intervient non seulement dans le transport des déblais et le refroidissement des équipements, mais aussi dans la stabilisation du puits, le contrôle de la pression et la préservation des formations géologiques traversées. La diversité des environnements souterrains et la complexité croissante des forages exigent une formulation rigoureuse et adaptée de ces fluides, qu'ils soient à base d'eau, d'huile ou de gaz. Ainsi, la sélection, la préparation, le contrôle et la gestion des boues de forage constituent des éléments techniques cruciaux pour garantir la sécurité, la performance et la durabilité des opérations.

2.2 Définition d'un fluide de forage

Les fluides de forage, également appelés boues de forage, sont des mélanges complexes de liquides utilisés dans les opérations de forage pétrolier et gazier. Leur rôle est fondamental pour assurer le bon déroulement des travaux, en garantissant à la fois la sécurité, la stabilité du puits et l'efficacité du processus de forage. Ces fluides sont généralement constitués d'eau, d'huile, ou d'un mélange des deux, auxquels sont incorporés divers additifs chimiques destinés à répondre aux exigences spécifiques de chaque opération (contrôle de la pression, stabilisation des parois du puits, refroidissement et lubrification de l'outil de forage, etc.).[7]

La préparation du fluide s'effectue dans des bacs appelés bacs à boues, avant son injection dans la colonne de forage. Il circule ensuite jusqu'à l'outil de forage, puis remonte dans l'espace annulaire, chargé des déblais (cuttings) issus de la désagrégation des formations géologiques traversées. Avant sa réutilisation, le fluide doit être traité en surface afin d'éliminer les particules solides qu'il a transportées. Ce traitement inclut des étapes telles que le tamisage, la dilution et l'ajout de produits chimiques, dans le but de restaurer les propriétés physico-chimiques du fluide et d'assurer son efficacité pour les cycles de forage suivants. [8]

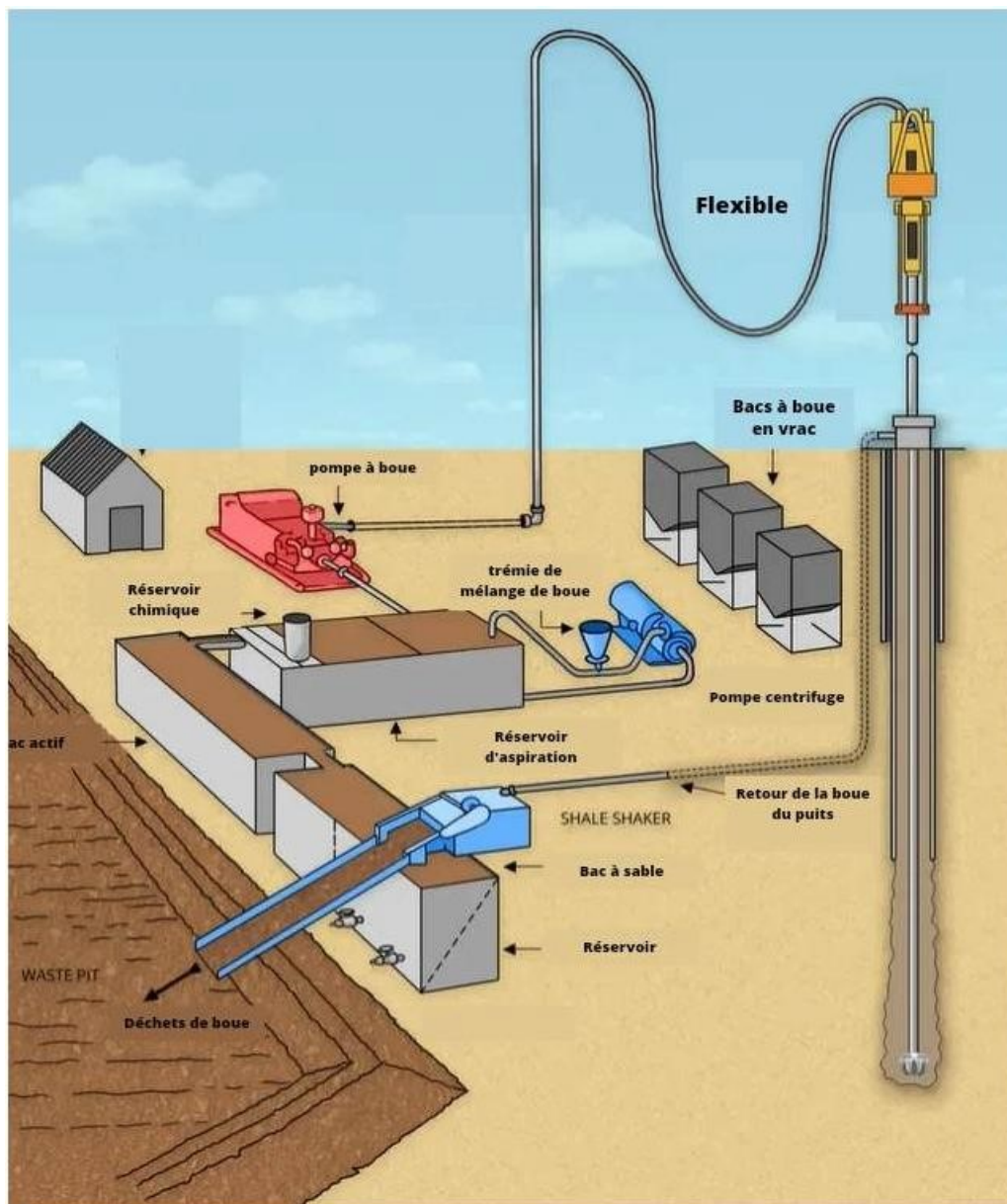


FIGURE 2.1 – Circuit de circulation du fluide de forage

2.3 Rôles des fluides de forage

Les fluides de forage ont plusieurs fonctions fondamentales :

Tout d'abord, leur fonction principale est d'assurer l'évacuation des déblais générés par l'action du trépan sur les formations géologiques. Grâce à leur viscosité, les fluides permettent le transport de ces particules solides vers la surface par circulation dans l'espace annulaire. En complément, leur capacité à maintenir les déblais en suspension lors des arrêts de circulation empêche leur sédimentation et prévient le coincement des outils, ceci étant rendu possible par les propriétés thixotropes du fluide.

Par ailleurs, les fluides de forage assurent la lubrification et le refroidissement des éléments métalliques en mouvement, notamment le trépan et la colonne de forage, limitant ainsi l'usure prématurée des équipements et les risques de surchauffe.

Une autre fonction essentielle est la stabilisation des parois du puits. La pression hydrostatique exercée par le fluide contribue à prévenir les effondrements et à maintenir l'intégrité structurale du trou de forage. En s'infiltrant dans les formations perméables, le fluide dépose un film appelé cake de filtration, qui réduit la perméabilité des parois et limite les échanges de fluides entre la formation et le puits. Ce film doit cependant rester mince afin d'éviter une réduction du diamètre effectif du puits.

Les fluides de forage jouent également un rôle crucial dans le contrôle de la pression au sein du puits. En ajustant leur densité, il est possible d'équilibrer les pressions des formations traversées, évitant ainsi les risques de venues incontrôlées de fluides (kicks), voire d'éruptions.

En outre, ces fluides participent à la préservation de la qualité des formations traversées. Ils forment une barrière empêchant la contamination croisée avec d'autres fluides souterrains, comme l'eau douce ou les hydrocarbures, contribuant ainsi à la protection des ressources naturelles.

Enfin, les fluides de forage facilitent l'échantillonnage géologique, par le transport de cuttings ou la collecte de carottes, permettant une meilleure compréhension des formations rencontrées. Ils soutiennent également l'outillage, stabilisant la colonne de forage et assurant une meilleure maîtrise de la trajectoire du puits.

De plus, la formulation des boues doit être rigoureusement contrôlée afin d'éviter toute action corrosive ou abrasive sur les équipements, ainsi que tout danger toxique ou inflammable pour le personnel sur site. [9] [10] [11]

2.4 Les types de fluide de forage

Il existe trois familles de fluides de forage :

2.4.1 Fluides de forage à base d'eau (WBM)

Les boues à base d'eau constituent la catégorie la plus couramment utilisée, notamment dans les premières sections d'un forage. Elles sont constituées d'eau douce ou salée comme phase continue, à laquelle sont ajoutés divers solides et additifs chimiques pour conférer au fluide les propriétés nécessaires à l'exécution optimale des opérations. Ce type de fluide présente l'avantage d'être relativement économique, facile à préparer et à traiter en surface.

Cependant, au cours du forage, la composition des boues à base d'eau peut être modifiée par l'interaction avec les matériaux des formations traversées, ce qui peut altérer leurs propriétés physico-chimiques. Les principaux composants d'une boue WBM incluent :

- **Eau** (douce ou saumâtre),
- **Traitement de la dureté** (ex. : carbonate de sodium – soda ash),
- **Régulateur de pH** (ex. : hydroxyde de sodium – NaOH),
- **Sels** (ex. : chlorure de sodium) pour contrôler l'activité de l'eau,
- **Inhibiteurs de schiste** et **encapsulateurs** pour limiter la dispersion des argiles,
- **Réducteurs de filtrat** pour limiter la perte de fluide dans la formation,
- **Viscosifiants** (ex. : bentonite) pour contrôler la rhéologie,
- **Agents alourdisants** (ex. : barytine),

- **Lubrifiants** et **stabilisateurs** pour réduire les frottements et améliorer la stabilité du puits.[9] [12]

2.4.2 Fluides de forage à base d'huile (OBM)

Les boues à base d'huile sont principalement utilisées dans les sections profondes ou complexes des forages, où les conditions de température et de pression sont élevées, ou lorsque les formations sont particulièrement réactives à l'eau. Elles sont constituées d'une phase continue huileuse (souvent du gasoil), avec une phase aqueuse dispersée sous forme d'émulsion inverse. La proportion d'eau dans ces fluides varie généralement de 5% à 30%, mais peut être ajustée selon les caractéristiques lithologiques et les objectifs techniques.

Les composants typiques d'une boue OBM comprennent :

- **Gasoil** (ou autre huile minérale) comme phase continue,
- **Viscosifiants** pour ajuster la rhéologie
- **Sources d'alcalinité** pour stabiliser le pH,
- **Émulsifiants** pour stabiliser l'émulsion eau-dans-huile,
- **Agents mouillants** pour favoriser l'adhésion de l'huile sur les surfaces solides,
- **Saumure (brine)** comme phase aqueuse (mélange d'eau et de NaCl),
- **Réducteurs de filtrat**,
- **Agents alourdissants** (ex. : barytine ou hématite),

Bien que plus coûteux et nécessitant un traitement spécifique pour minimiser leur impact environnemental, les fluides OBM offrent une excellente stabilité thermique et chimique, ainsi qu'un pouvoir lubrifiant supérieur aux WBM.[12]

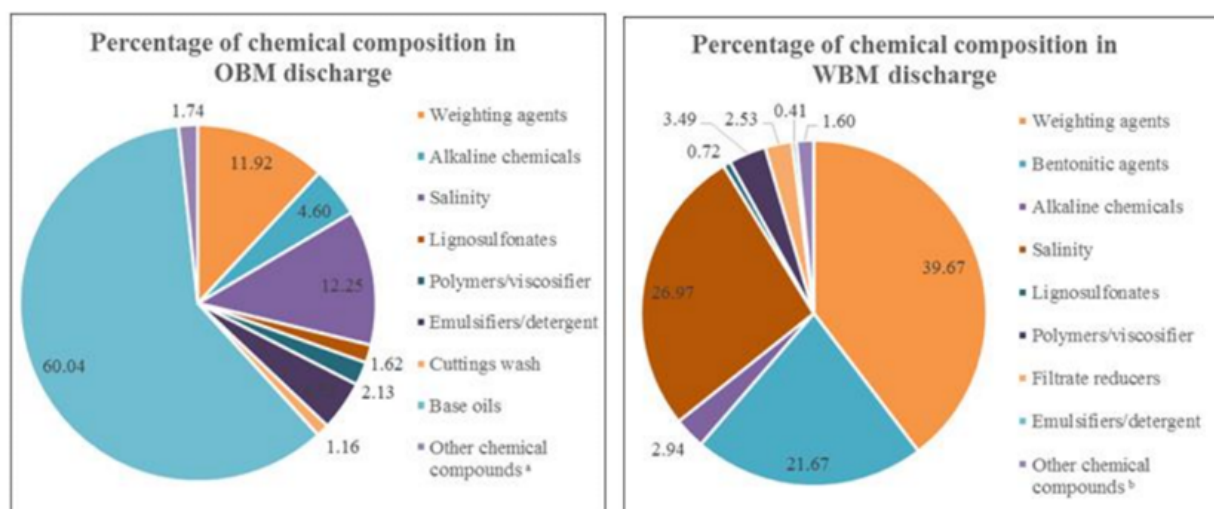


FIGURE 2.2 – Les différents composants des fluides de forage (les boues à base d'huile et les boues à base d'eau)

2.4.3 Fluides de forage à base de gaz

Moins courants, les fluides à base de gaz sont utilisés dans des contextes bien spécifiques, notamment pour des forages sous pression contrôlée ou en milieux peu perméables. Leur phase continue est constituée de gaz (air, azote ou gaz naturel), parfois mélangé avec de l'eau ou des agents moussants afin de former des mousses ou des brouillards. Ce type de fluide permet une

réduction significative de la pression hydrostatique exercée sur la formation, minimisant ainsi les risques de perte de fluide et de dommages à la formation.

Les gaz peuvent être introduits intentionnellement ou provenir partiellement des formations traversées. L'utilisation de ces fluides nécessite un contrôle rigoureux des équipements de surface et une gestion fine des paramètres de forage, en raison de leur faible densité et de leur comportement instable dans certaines conditions. [9]

2.5 Les propriétés des fluides de forage

La caractérisation et le contrôle des boues de forage sont des étapes fondamentales dans la gestion des opérations de forage. Ces opérations sont encadrées par des normes rigoureuses, principalement celles établies par l'American Petroleum Institute (API), qui définissent les procédures d'évaluation des propriétés physico-chimiques des fluides de forage. L'étude des caractéristiques des boues repose classiquement sur quatre paramètres principaux : la densité, la viscosité, le volume de filtrat, et la réactivité chimique du fluide vis-à-vis des formations traversées. Ces propriétés sont évaluées à l'aide de tests standardisés, réalisés à intervalles réguliers au cours du forage.[13]

Certaines mesures sont réalisées systématiquement sur tous les forages (viscosité, densité, filtration) et d'autres en fonction des besoins (taux de gaz, alcalinité). A partir des mesures réalisées et des connaissances acquises, on ajuste si nécessaire la composition de la boue en "temps réel" en ajoutant certains produits ou en reformulant la boue.

Historiquement, face aux problèmes majeurs rencontrés dans les premiers forages, la densité est le premier paramètre fondamental pour la caractérisation des fluides de forage. [9]

2.6 Le choix des boues de forage

La sélection du fluide de forage constitue une étape stratégique dans la planification des opérations, car elle dépend étroitement des conditions géologiques locales, des contraintes techniques rencontrées ainsi que des objectifs de chaque phase de forage. À chaque section du puits correspond un fluide spécifique, dont les caractéristiques physico-chimiques sont adaptées aux propriétés mécaniques des formations traversées, à la profondeur, à la température et à la pression en présence. De manière générale, les boues à base d'eau, notamment les boues bentonitiques, sont privilégiées pour les premières phases du forage, notamment en raison de leur simplicité de mise en œuvre, de leur coût relativement faible et de leur compatibilité avec les formations superficielles peu complexes. Dans les phases plus profondes, où les conditions deviennent plus contraignantes (hautes pressions, températures élevées, formations instables ou sensibles à l'eau), l'utilisation de boues à base d'huile devient plus appropriée. Ces dernières offrent de meilleures performances en termes de stabilité thermique, de lubrification et de contrôle de pression. Il convient toutefois d'ajuster le rapport eau/huile selon les exigences de chaque section forée, afin d'optimiser les propriétés du fluide tout en minimisant les impacts environnementaux et les risques opérationnels.[14]

2.7 Conclusion

En somme, les fluides de forage sont bien plus que de simples liquides utilisés dans les forages pétroliers et gaziers : ils sont au cœur de la réussite technique et sécuritaire des opérations. Leur formulation, soigneusement adaptée aux conditions géologiques et aux objectifs de chaque phase de forage, permet de répondre à des enjeux multiples allant du transport des déblais à la protection des équipements et des formations traversées. Maîtriser les propriétés, les types et les fonctions de ces fluides est donc indispensable pour optimiser la performance des forages tout en limitant les risques environnementaux et opérationnels.

Chapitre 3

Les déblais de forage (cuttings)

3.1 Introduction

Les opérations de forage pétrolier génèrent divers types de déchets, dont les plus courants sont les déblais de forage (cuttings) associés à des fluides de forage. Ces déchets contiennent généralement une grande proportion de fluide de base, qu'il soit aqueux ou non aqueux (huile minérale, oléfine, ester, etc.). La nature de ces déchets, ainsi que leur impact environnemental potentiel, dépendent fortement du type de fluide utilisé, de la méthode d'élimination et du milieu récepteur (eau, sol, air).

Les cuttings rejetés en mer, par exemple, peuvent provoquer des effets sur les fonds marins, la colonne d'eau et l'interface air/eau. Les fluides de forage à base de gasoil sont aujourd'hui interdits en rejet direct, en raison de leur toxicité, de leur pouvoir de dispersion et de leur persistance. Même les fluides à base d'eau, bien que généralement considérés comme moins nocifs, posent certains problèmes tels que l'étouffement de la faune benthique, la présence d'huile résiduelle, la toxicité, ou encore la présence de métaux lourds issus d'additifs comme la barite.

Ce chapitre vise à présenter les caractéristiques des déblais de forage, leurs sources de contamination, leurs impacts sur l'environnement et la santé humaine, ainsi que les normes réglementaires encadrant leur gestion. Une attention particulière est accordée à leur nature géochimique, leur transport, leur interaction avec les fluides de forage, et les risques liés aux métaux lourds, aux hydrocarbures et aux sels solubles.[\[15\]](#)

3.2 Définition et origine des déblais de forage

Les déblais de forage représentent les fragments solides issus de la roche, extraits d'un trou de forage réalisé par des méthodes rotatives ou par percussion. Ce type de forage est couramment utilisé dans les puits pétroliers et/ou gaziers, que ce soit en phase d'exploration ou de développement. Les déblais sont générés par l'action de l'outil de forage qui désagrège la roche, et sont ensuite transportés vers la surface par le fluide de forage circulant à partir du trépan. Une fois à la surface, les déblais peuvent être séparés du fluide par différents équipements tels que les tamis vibrants (shale shakers), les centrifugeuses ou encore les séparateurs cycloniques.[\[16\]](#)



FIGURE 3.1 – Zone de stockage des déblais de forage (bourbier)

3.3 Types de déblais de forage (cuttings)

Les déblais de forage, générés par l'action de l'outil de forage sur la roche, varient en nature selon le type de fluide de forage utilisé pour remonter ces fragments à la surface. On distingue généralement trois types principaux :

- Les déblais issus de boues à base d'eau, contenant des additifs tels que le chlorure de potassium (KCl), des polymères ou encore du glycol
- Les déblais issus de boues semi-huileuses (pseudo-oil based), souvent formulées à base d'oléfines ou d'esters biodégradables
- Les déblais issus de boues à base d'huile, comprenant des huiles minérales propres, et historiquement, du diesel

Les boues à base de diesel sont reconnues pour leur impact environnemental élevé, notamment en raison de leur toxicité et de leur persistance dans les milieux naturels. C'est pourquoi leur utilisation a été progressivement réduite au profit de formulations plus respectueuses de l'environnement, en particulier dans les sites de forage terrestre à proximité de zones sensibles ou agricoles. Jusqu'aux années 1990, les boues à base d'huile étaient couramment utilisées, surtout dans les sections profondes et étroites des puits, ce qui limitait la quantité de déblais produits par mètre foré. Cependant, face aux exigences croissantes en matière de protection environnementale, des solutions alternatives ont été développées. Ces boues modernes sont conçues pour être moins toxiques, plus facilement biodégradables, et moins polluantes.^[17]

3.4 Déblais de Forage : Propriétés, Usages et Enjeux Opérationnels

Les déblais sont des petits fragments de roches asymétriques, de diamètre allant de quelques dizaines de microns jusqu'à 1–2 cm, générés par l'action du trépan lors du forage des formations rocheuses pour atteindre les réservoirs d'huile ou de gaz. Ils diffèrent dans leurs tailles et textures, et leurs types varient de celui des grains de sables jusqu'au gravier selon la formation rocheuse atteinte et le type de boue utilisée. Pour éviter l'entassement des *cuttings* au fond du puits foré, ils sont remontés en surface par la boue de forage [17].

3.4.1 Propriétés Physiques et Chimiques

Les déblais de forage sont influencés par la formation rocheuse hôte, les sources de calcium et les boues de forage résiduelles. Par exemple, la barytine dans les boues résiduelles peut enrober les déblais et affecter la mobilité des métaux traces comme le strontium [18]. Cela peut entraîner une redistribution locale des éléments dans l'environnement du puits, avec des implications sur la contamination potentielle des eaux souterraines ou de surface.

Les déblais peuvent contenir des métaux traces principalement associés à la pyrite et au carbone organique total (TOC) dans le schiste [18]. La présence de pyrite favorise également des processus d'oxydation qui peuvent modifier le pH local et faciliter la libération d'autres éléments géochimiques.

3.4.2 Utilisation pour la caractérisation des réservoirs

Les déblais de forage peuvent être utilisés pour caractériser les propriétés minéralogiques et géochimiques des roches réservoirs. Des analyses comme la diffraction des rayons X (XRD) et l'analyse pyrolytique peuvent révéler des informations sur la composition minérale et le potentiel de génération d'hydrocarbures [19]. Ces méthodes permettent une évaluation rapide et non destructive de la formation sans perturber le processus de forage.

Ils offrent une alternative rentable à l'extraction de carottes pour l'évaluation des propriétés pétrophysiques des réservoirs [20], en couvrant une plus grande section du puits à moindre coût, particulièrement dans des environnements complexes.

3.4.3 Interaction avec les fluides de forage

Les propriétés mécaniques des déblais, telles que la résistance à la compression et le module de Young, peuvent influencer les propriétés rhéologiques des fluides de forage, surtout lorsque la teneur en sable dépasse 10 % en poids [21]. Ces interactions peuvent entraîner des fluctuations de pression, modifier le comportement d'écoulement et impacter significativement la stabilité du puits.

3.4.4 Transport et Dépôt des Déblais

Le transport des déblais est affecté par la rhéologie des fluides, la taille des particules et l'inclinaison du puits. Une mauvaise évacuation peut entraîner la formation de lits de déblais, augmentant le temps non productif [22, 23], causant des blocages, endommageant les équipements et réduisant l'efficacité du forage.

Les modèles numériques et expérimentaux aident à optimiser les paramètres de forage pour améliorer l'efficacité du transport des déblais [24, 25]. Ils permettent de simuler différents scénarios de forage et de sélectionner les meilleures conditions opérationnelles afin d'éviter les incidents liés à l'accumulation des déblais.

3.4.5 Méthodes de surveillance et de caractérisation en temps réel

Des méthodes automatiques utilisant des capteurs laser permettent une caractérisation en temps réel des déblais [26], ce qui est crucial pour évaluer la stabilité du puits et l'efficacité du nettoyage. Ces outils réduisent la dépendance à l'échantillonnage manuel tout en fournissant des données instantanées pour une prise de décision rapide.

3.5 Sources de contamination des déblais de forage

La toxicité des déblais de forage peut s'introduire par trois voies principales :

- Par la composition chimique des boues de forage
- Par leur mode de stockage
- Et par les matériaux géologiques rencontrés au cours du forage

3.5.1 Contamination d'origine géologique (formations rocheuses)

Il n'existe pas d'informations précises sur la toxicité des matériaux solides issus des formations traversées, en raison de l'hétérogénéité de leur distribution dans le sous-sol et de la variabilité de leur contenu en substances toxiques. Toutefois, ces formations présentent un potentiel de contamination notable par :

- Les métaux lourds
- Les hydrocarbures (dans certains cas),
- Les sels contenus notamment dans les dômes salins [27]

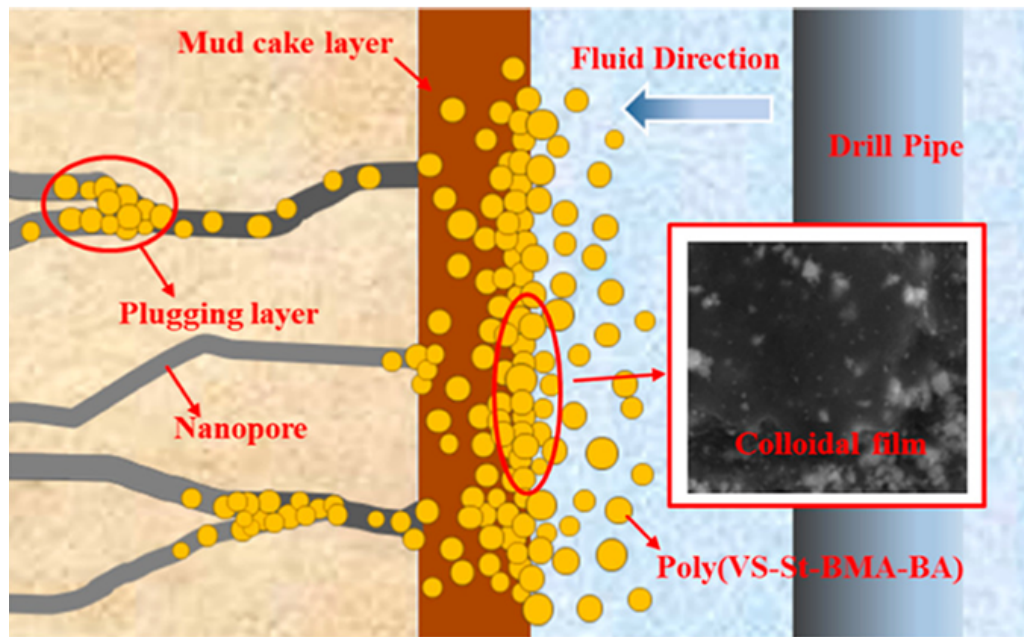


FIGURE 3.2 – Interaction entre les fluides de forage et la formation géologique

3.5.2 Boues de forage et additifs chimiques

La composition de la boue représente la source de contamination la plus importante, en raison de l'ajout volontaire d'agents chimiques pour assurer la performance des fluides de forage [28] [27]. Ces produits comprennent :

- Le baryum
- Le chrome (additifs : réducteurs de viscosité, inhibiteurs de corrosion)
- Le cadmium et le mercure (impuretés de la barytine)
- Le plomb, l'arsenic, le diesel (dans les boues à base d'huile)[28]

La boue d'huile est particulièrement préoccupante, car les déblais forés absorbent les hydrocarbures par fixation des matières hydrophobes, à cause des agents mouillants à base d'huile.[28]

Les fluides utilisés dans le champ de Hassi Messaoud contiennent des composants chimiques dont Voici quelques exemples présentés dans le tableau 3.1 [29] :

TABLE 3.1 – Composition chimique, quantité et toxicité des fluides de forage utilisés à Hassi Messaoud

Nom commercial	Composition chimique	Quantité (Tonnes)	Fonction	Information toxicologique
Baryte	Sulfate de baryum	1180,84	Agent alourdissant	Impact réduit
Bentonite	Argile montmorillonite	117,36	Viscosifiant et contrôleur de filtrat	Impact réduit
Soude caustique	Hydroxyde de sodium	1,956	Modificateur de pH	LD50 = 500 mg/kg
Polysal	Dazomet de Thiadiazine	6,52	Réducteur de perte de fluide	Non disponible
Versatrol	Gilsonite	6,84	Réducteur de perte de fluide	Impact réduit
Sel	Chlorure de sodium	95,766	Sel pour saumure d'accomplissement	LD50 = 3000 mg/kg
Lime	Hydroxyde de calcium	14,3	Source de calcium, modificateur de pH	LD50 = 7340 mg/kg
VG-69	Quartz et silice cristalline	24,145	Inverseur d'émulsion, viscosifiant	LD50 > 20000 mg/kg
VERSAMUL	Huile minérale, alcool méthylique, résidus pétroliers	4,56	Émulsifiant	LD50 > 5000 mg/kg ; LC50 = 1.6 mg/l (4h)
VERSACOAT	Huile minérale, alcool méthylique, résidus pétroliers	20,974	Émulsifiant	LD50 > 5000 mg/kg ; LC50 = 1.6 mg/l (4h)
VERSAWET	Non disponible	13,5605	Agent mouillant	Impact réduit

3.5.3 Contamination par les hydrocarbures

Les hydrocarbures présents dans les déblais peuvent provenir de plusieurs sources :

- La boue de forage, en particulier si elle est à base d'huile
- Les hydrocarbures piégés dans les réservoirs géologiques
- Ou les huiles industrielles utilisées durant les opérations de forage [28]

Ces hydrocarbures sont souvent adsorbés ou encapsulés dans les déblais, un phénomène mesuré par le taux d'huile sur déblais (OOC%). À Hassi Messaoud, un seul bourbier peut contenir jusqu'à 600 m³ de gasoil encapsulé. de plus, les eaux de rejet associées peuvent atteindre 12000 m³, intensifiant le phénomène de lessivage des contaminants.[30]

3.5.4 Contamination par les métaux lourds

Les métaux lourds représentent une menace environnementale majeure, surtout lorsqu'ils deviennent mobiles et solubles dans l'eau. Ils peuvent avoir deux origines :

- Naturelle, via les roches traversées
- Ou anthropique, par l'intermédiaire des additifs contenus dans les boues de forage, comme :
 - o La baryte (contenant parfois du cadmium, du mercure ou du manganèse)
 - o Ou des réducteurs chimiques à base de chrome [28, 27]

Certains composés organiques présents dans le sol, tels que les acides carboxyliques ou les phénols, favorisent la formation de complexes métalliques très solubles, augmentant ainsi la dispersion de ces métaux dans l'environnement.[28]

3.5.5 Présence de sels solubles

Les sels solubles, tels que le chlorure de sodium (NaCl) ou le chlorure de potassium (KCl), sont couramment détectés dans les déblais de forage. Leur origine peut être :

- Géologique, lors de la traversée de formations salifères ou de dômes de sel
- Ou technique, suite à l'utilisation de fluides de forage traités avec des saumures spécifiques, selon les besoins opérationnels [28].

3.6 Impacts des déblais de forage sur l'environnement

3.6.1 Effets sur l'environnement physique

L'un des impacts majeurs des déblais de forage est leur toxicité pour les organismes vivants. Cette toxicité peut provenir de la composition chimique des boues de forage, de leur stockage et des déblais eux-mêmes[31]. Parmi les substances problématiques figurent des métaux lourds comme le baryum, le chrome, le mercure, le cadmium, le plomb, ainsi que des hydrocarbures (diesel) et des éléments toxiques comme l'arsenic.

Les produits chimiques introduits intentionnellement dans les fluides de forage, pour en modifier les propriétés, constituent une source directe de danger environnemental. Le sodium, fréquemment utilisé dans les boues, représente un fort potentiel de pollution lorsqu'il est déversé dans les bourniers. Quant aux métaux lourds, leur mobilité dans le sol est limitée, mais leur présence reste préoccupante. [32]

Ces contaminations peuvent entraîner des modifications importantes des propriétés physico-chimiques des sols. Les zones désertiques étant particulièrement vulnérables – en raison de la faible couverture végétale, des températures élevées et d'un pouvoir d'auto-régénération réduit – les risques sont amplifiés. De plus, les nappes phréatiques situées à faible profondeur présentent un risque élevé de pollution par infiltration.

3.6.2 Risques pour la santé humaine

Les hydrocarbures présents dans les déblais de forage peuvent contaminer les sols et représenter un danger pour la santé humaine. Cela peut survenir par contact direct avec le sol contaminé ou par leur introduction dans la chaîne alimentaire via le phénomène de bioaccumulation.[33]

Les fluides de forage peuvent provoquer des irritations de la peau, des voies respiratoires, et même des effets cancérogènes, notamment à cause des hydrocarbures aromatiques polycycliques (HAP).[34] Ces composés sont également associés à des effets mutagènes et posent des problèmes de persistance dans l'environnement en raison de leur bioaccumulation, de leur faible dégradabilité et de leur adsorption sur les particules du sol.[35]

3.6.3 Impacts sur la faune et la flore

Dans les zones désertiques, les plantes développent des systèmes racinaires très étendus pour capter l'eau, mais leur croissance est lente. Toute atteinte à leur intégrité (notamment par

contamination du sol) est difficilement réversible.

La faune désertique dépend directement de la végétation pour l'alimentation, l'ombrage et la protection. Ainsi, la dégradation des écosystèmes végétaux entraîne automatiquement une perturbation de la faune locale. Insectes, reptiles, rongeurs et autres petits mammifères utilisent les plantes comme abri ou source d'eau.

Un autre danger vient des bourbiers laissés sur place après la fin des opérations de forage. Ces réservoirs non traités peuvent piéger les animaux et représenter un risque direct pour la faune.[33] Une illustration de ce phénomène est donnée par l'exemple d'un chameau sauvé après s'être enlisé dans un bourbier.



FIGURE 3.3 – Sauvetage d'un chameau piège dans un bourbier

3.6.4 Dégradation de la qualité de l'air

Les impacts sur l'air proviennent principalement des émissions de poussières et de gaz de combustion. La préparation des sites de forage – notamment les travaux de terrassement, la construction de pistes et le transport de matériaux – génère une importante quantité de particules en suspension.

Les moteurs des engins de chantier et les générateurs utilisés sur les plateformes émettent du monoxyde de carbone (CO), des oxydes d'azote (NOx) et du dioxyde de soufre (SO₂). Pendant les travaux civils, la consommation de diesel peut atteindre 3 m³/jour, et jusqu'à 5 m³/jour lors des phases de forage.

Même si les émissions sont généralement temporaires et limitées dans le temps, elles peuvent provoquer une dégradation locale de la qualité de l'air. Toutefois, ces émissions tendent à disparaître rapidement après la fin des activités.[35]

La figure 3.4 illustre la préparation du terrain sur un site de forage. [35]



FIGURE 3.4 – Préparation du terrain

3.7 Législation et normes

En Algérie les grands principes qui régissent la réglementation en matière de protection du sol et du sous-sol sont formulés par les textes suivants :

- La loi n°83-03 du 05/02/1993 prend en considération la pollution du sol. Le législateur définit les déchets et pose les principes et les règles générales de leur gestion. Concernant cette dernière, la loi fait obligation à tout détenteur ou producteur de déchets d'assurer ou de faire assurer leur élimination en évitant seulement, et non en supprimant, les effets néfastes sur le sol, la faune, la flore, etc. . .
- Le décret exécutif n°93-160 du 10/07/1993 régit le déversement des huiles et lubrifiants aussi bien dans le milieu naturel que dans le réseau d'assainissement.
- Le décret exécutif n°93-162 du 10/07/1993 fixe les modalités de récupération et de traitement des huiles usagées.
- Le décret exécutif n°94-43 du 30/01/1994 fixe les règles de conservation des gisements d'hydrocarbures et de protection des aquifères.
- Le Décret du 18/12/1992, agissant en tant que norme européenne, publié au Journal officiel de la République Française le 30/03/1993, est la norme suivie par SONATRACH pour le traitement des déblais de forage. Ce décret définit les paramètres qui devraient être analysés et leurs limites maximales à ne pas excéder pour que les déblais puissent être rejetés dans la nature. Ces valeurs limites sont données dans le tableau suivant, en cas de non-respect de ces valeurs, il convient de traiter ces déblais avant leur rejet.

TABLE 3.2 – Valeurs limites pour la mise en décharge des déblais de forage selon le standard européen

Éléments	Abréviation	Standard Européen
Hydrocarbures totaux	Hc	< 5% (massique)
Chrome hexavalent	Cr ⁶⁺	< 10 mg/kg ou 1 mg/l
Chrome total	Cr	< 100 mg/kg ou 10 mg/l
Plomb	Pb	< 100 mg/kg ou 10 mg/l
Zinc	Zn	< 500 mg/kg ou 50 mg/l
Cadmium	Cd	50 mg/kg ou 5 mg/l
Cyanure	CN	10 mg/kg ou 1 mg/l
Nickel	Ni	< 100 mg/kg ou 10 mg/l
Arsenic	As	< 10 mg/kg ou 1 mg/l
Mercure	Hg	< 10 mg/kg ou 1 mg/l

Chapitre 4

La gestion des déblais de forage

4.1 Introduction

Lors du processus de forage, l'action du trépan sur la formation géologique engendre une quantité importante de déblais. Ces derniers sont transportés en surface par la boue de forage, souvent à base d'huile, utilisée pour stabiliser le puits et lubrifier les outils. Le traitement des déblais issus de ces opérations se fait en deux étapes principales : un traitement primaire, qui vise essentiellement à séparer la phase solide de la boue afin de récupérer cette dernière pour un usage ultérieur, et un traitement secondaire, destiné à réduire la teneur en huile des déblais solides à des niveaux conformes aux normes environnementales.

Après le traitement primaire, les déblais conservent encore une teneur en huile relativement élevée, généralement comprise entre 15 % et 20 %. Cependant, dans la plupart des régions, les réglementations environnementales imposent une teneur résiduelle en huile inférieure à 1 %. Il est donc nécessaire de mettre en œuvre des méthodes de traitement secondaire, qui ont pour objectif à la fois de rendre ces déchets inoffensifs pour l'environnement et de permettre la récupération du fluide de base.[36]

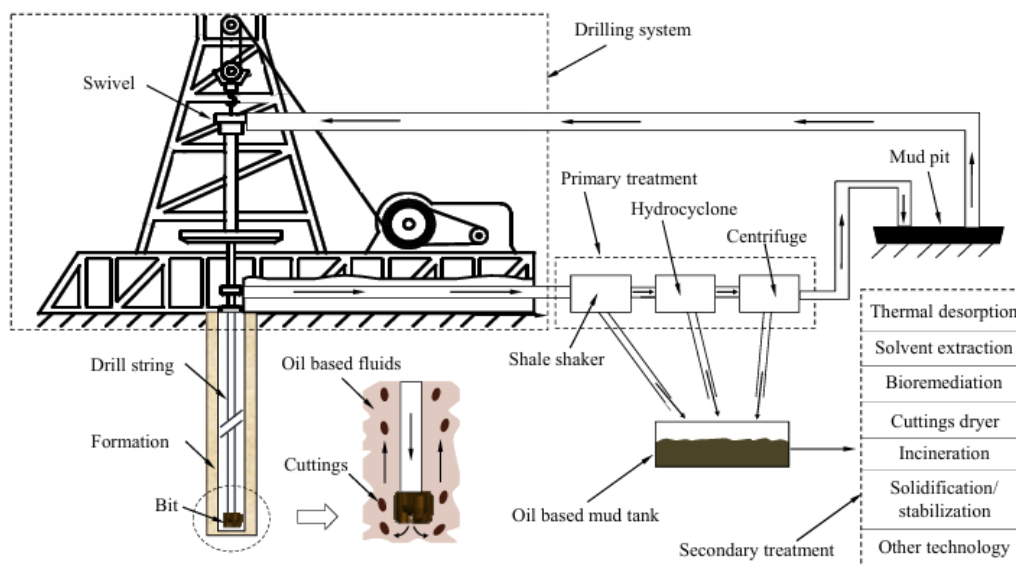


FIGURE 4.1 – Schéma du traitement des déblais de forage à base d'huile

Ce chapitre présente et analyse les principales méthodes de traitement appliquées aux déblais de

forage à base d'huile. Il explore successivement les approches physiques, chimiques, biologiques et thermiques, en mettant en évidence leurs mécanismes, leurs avantages et leurs limites. En complément, une dernière section est consacrée au bournier, structure essentielle pour la gestion temporaire de ces déchets sur site.

4.2 Les méthodes de traitement physique des cuttings

Les traitements physiques consistent à séparer mécaniquement les particules solides présentes dans la boue de forage sans en altérer la composition chimique. Ces procédés reposent essentiellement sur des principes physiques tels que la gravité, la vibration ou la force centrifuge. Dans ce contexte, le système de contrôle des solides constitue le cœur de ce traitement.

4.2.1 Séparation par criblage (Screen Separation)

Le shale shaker est l'équipement de criblage le plus couramment utilisé, comme illustré dans la Figure 4.2. Il est composé d'un ou plusieurs tamis vibrants à travers lesquels la boue chargée en solides est filtrée. Les particules liquides et fines traversent les tamis tandis que les déblais sont recueillis au fond de l'appareil. Lorsqu'il est bien sélectionné et utilisé efficacement, le shale shaker, associé aux tamis grossiers (scalper screens ou Gumbo shakers), peut éliminer jusqu'à 80 % des solides présents dans la boue de forage.

Les shale shakers se divisent en deux types selon leur mouvement :

- **Mouvement circulaire ou elliptique** (rumba shakers)
- **Mouvement linéaire** (va-et-vient)

L'expérience sur le terrain montre que les modèles elliptiques sont plus adaptés aux boues à base d'eau, tandis que les modèles à mouvement linéaire conviennent mieux aux boues à base d'huile. Il est recommandé d'utiliser au minimum trois shale shakers pour assurer une séparation efficace des solides.[37]

La figure 4.2 montre un exemple typique de vibreur de schiste utilisé sur site [37].



FIGURE 4.2 – Vibreur de schiste typique

4.2.2 Séparation par sédimentation (Settling Separation)

La séparation par sédimentation est une méthode de contrôle des solides peu coûteuse, consistant à laisser reposer la boue de forage pour permettre aux solides de se déposer. Dans ce processus, la boue contaminée circule à travers des fosses de décantation fonctionnant selon un principe de débordement.

Le piège à sable (sand trap) est le premier compartiment de cette chaîne, alimenté par la boue déjà filtrée par les shale shakers. Aucune agitation ne doit avoir lieu dans ces fosses pour ne pas empêcher la décantation.

Les solides les plus gros et les plus lourds se déposent rapidement, tandis que les particules plus fines, notamment celles de taille limoneuse, peuvent prendre beaucoup plus de temps – parfois plusieurs jours – pour se déposer. En général, la décantation est facilitée lorsque : a) les particules sont grandes et denses. b) la boue est légère et peu visqueuse. c) la gravité est renforcée par des moyens mécaniques.

4.2.3 Dégazage (Gas Removal)

Les gaz emprisonnés dans la boue doivent être éliminés afin de maintenir une densité conforme aux exigences de contrôle de pression en fond de puits. Un dégazeur typique est illustré dans la Figure 4.3.

Ce dispositif, généralement composé d'une pompe à vide et d'un système à flotteur, fonctionne en créant une basse pression interne qui aspire la boue chargée en gaz dans une cuve. La boue s'écoule ensuite en fine couche sur une plaque interne, ce qui permet aux bulles de gaz de se dilater et de remonter à la surface.

Le gaz ainsi séparé est extrait par la pompe à vide, puis évacué soit vers l'atmosphère, soit

brûlé (flaring), selon les procédures de sécurité en vigueur.[37]



FIGURE 4.3 – Un dégazeur à vide typique

4.2.4 Séparation centrifuge (Forced Settling)

Les désableurs (desanders) et **désileurs** (desilters) sont des dispositifs appelés **hydrocyclones**. Ils fonctionnent en créant une force centrifuge qui sépare les solides de la boue. La boue est injectée de manière tangentielle dans l'hydrocyclone, provoquant le déplacement des solides vers les parois. Ceux-ci sont ensuite évacués par la base avec une petite quantité de boue, tandis que la partie liquide est récupérée en haut comme débordement.[37]

Le principe de fonctionnement de ces hydrocyclones est illustré dans la figure 4.4 [37].

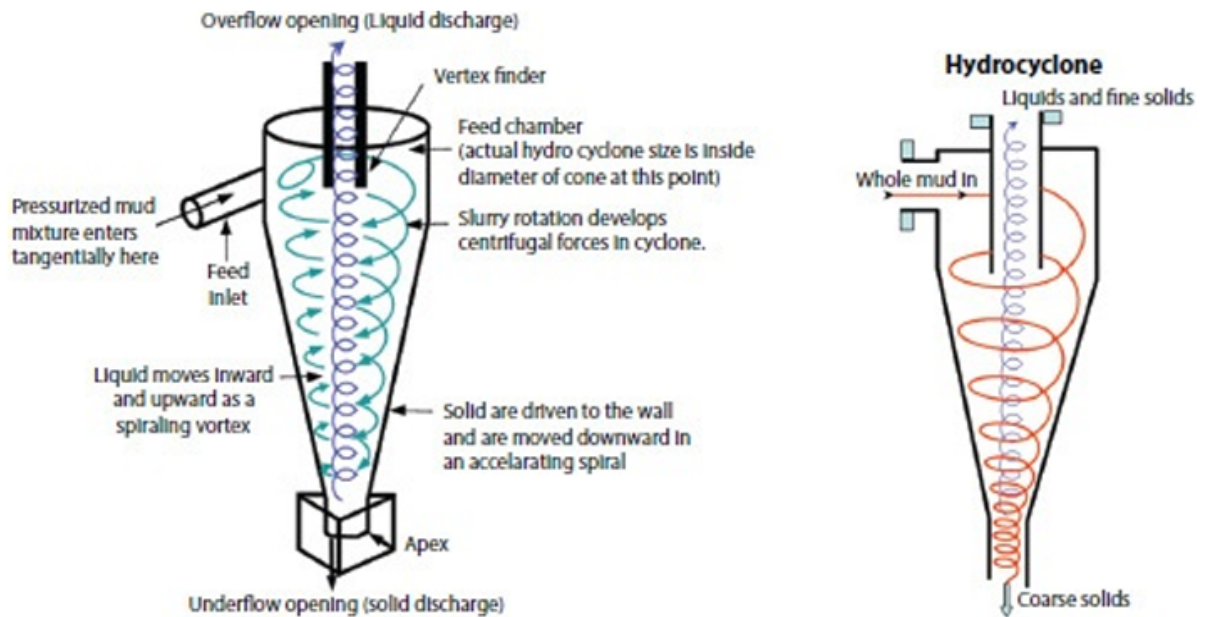


FIGURE 4.4 – Principe de fonctionnement des hydrocyclones

Les hydrocyclones existent en différentes tailles, adaptées à des plages spécifiques de particules. On distingue plusieurs types :

4.2.4.1 Désableurs

Ce sont des hydrocyclones de 6 pouces de diamètre interne ou plus (Figure 4.5), utilisés surtout en phase de forage superficielle avec des boues à base d'eau pour limiter le poids de la boue. Ils préviennent la surcharge des désileurs et protègent les pompes contre l'abrasion. Leur seuil d'efficacité se situe pour des particules entre 40 et 45 microns, mais ils sont inadaptés aux boues à base d'huile.[37]



FIGURE 4.5 – Un désableur typique

4.2.4.2 Désileurs

Composés de nombreux petits cônes (< 6 pouces), ils sont conçus pour éliminer les particules fines de type limon (Figure 4.6). Leur utilisation est recommandée en combinaison avec les désableurs pour traiter des boues légères lors des sections de forage supérieur. En cas d'augmentation du poids de la boue, celle-ci doit être assurée par l'ajout de barite, et non par l'accumulation de solides à faible densité.[37]



FIGURE 4.6 – Un désiltateur typique

4.2.4.3 Nettoyeurs de boue (Mud Cleaners)

Les mud cleaners combinent un shale shaker à tamis très fin (environ 320 mailles) situé sous un désileur, placé au-dessus d'un écran vibrant à haute énergie (Figure 4.7). Ils sont principalement

utilisés avec des boues pondérées, car la barite peut être entraînée avec les particules de limon. Ces équipements permettent donc de récupérer et de réutiliser la barite.[37]

Ils deviennent nécessaires lorsque les shale shakers seuls ne suffisent plus à maintenir un faible poids de boue. Toutefois, il est plus efficace de traiter le flux du désileur avec une centrifugeuse que d'utiliser uniquement les tamis d'un mud cleaner. Leur usage avec des boues à base d'huile doit être limité, car des pertes de boue de 3 à 5 barils par heure sont fréquemment observées.[37]

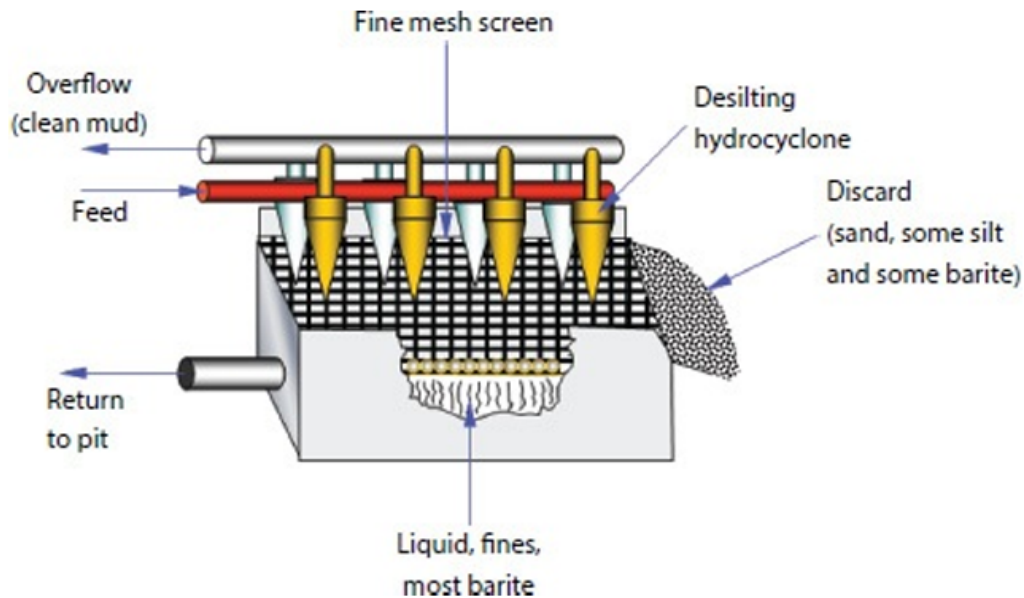


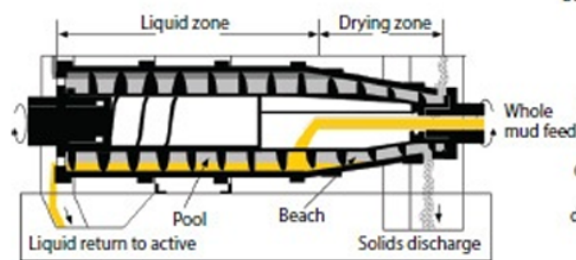
FIGURE 4.7 – Un nettoyeur de boue

4.2.4.4 Centrifugeuses (Centrifuges)

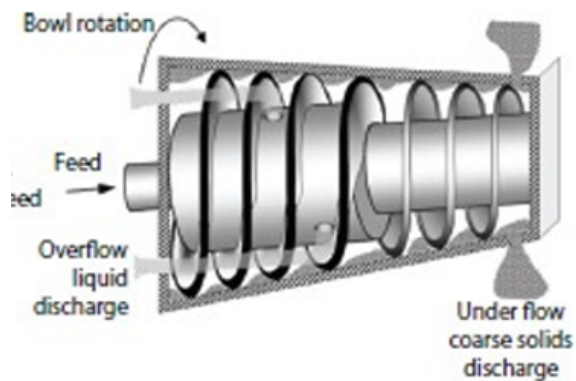
Les centrifugeuses utilisent la force centrifuge pour séparer les solides lourds des composants liquides et légers de la boue. La Figure 4.8b montre une centrifugeuse décanteuse, composée d'une cuve conique en acier tournant à haute vitesse, avec un convoyeur à double vis tournant légèrement plus lentement dans le même sens.[37]

La Figure 4.8a illustre les formes cylindriques ou coniques typiques des bols utilisés. La boue est projetée contre les parois, où les particules de limon et de sable se déposent et sont déplacées vers l'extrémité étroite par les pales du convoyeur, pour être récupérées sous forme de particules humides. Le liquide et les particules d'argile sont quant à eux évacués en débordement par l'extrémité large.[37]

L'efficacité de la centrifugeuse dépend principalement du débit d'alimentation, mais aussi de paramètres tels que la vitesse du bol, la différence de vitesse du convoyeur et la profondeur du bassin.[37]



(a) Centrifugeuse de décantation avec un bol conique en acier, horizontal et élargi



(b) Centrifugeuse de décantation avec un cylindre

FIGURE 4.8 – Une centrifugeuse de décantation typique

4.3 Les méthodes de traitement chimiques des déblais de forage

Le traitement chimique est une méthode qui repose sur l'utilisation de réactions chimiques ou de solvants spécifiques pour modifier, extraire ou stabiliser les substances présentes dans un déchet. Il permet notamment de transformer les composants polluants en formes moins nuisibles ou de les séparer pour faciliter leur réutilisation ou leur élimination.

4.3.1 Extraction par solvant

En général, des fluides supercritiques comme le CO_2 ou le propane sont utilisés pour extraire l'huile des déblais de forage à base d'huile. Ces fluides, à mi-chemin entre un gaz et un liquide, combinent une forte pénétration avec une bonne solubilité.

La méthode la plus utilisée repose sur l'extraction au CO_2 supercritique : les déblais sont mélangés au fluide supercritique, qui capte l'huile. Après décompression, l'huile et le solvant peuvent être récupérés et réutilisés. Cette technique nécessite toutefois une pression élevée (14,5 MPa) et une température d'environ 40 °C.

Pour simplifier le procédé, une alternative appelée extraction par gaz liquéfié utilise du propane ou du butane. À 500 psi et température ambiante, le butane a permis de réduire la teneur en huile de 21 % à 0,24 %, sans altérer la qualité du fluide de base récupéré.[36]

4.3.2 Solidification/Stabilisation

L'Agence de protection de l'environnement des États-Unis (USEPA) définit la solidification/stabilisation (S/S) de la manière suivante :

- **La solidification** désigne les techniques consistant à encapsuler les déchets dans une matrice solide monolithique dotée d'une haute intégrité structurale. L'encapsulation peut concerner de fines particules de déchets ou un grand volume de déchets (micro-encapsulation)

- **La stabilisation** correspond aux techniques visant à réduire le potentiel de dangerosité des déchets en transformant les contaminants dans leur forme la moins soluble, la moins mobile ou la moins toxique

Dans la gestion des déblais de forage à base d'huile, il est généralement nécessaire d'assurer leur confinement grâce à la combinaison des techniques de solidification et de stabilisation. Cette approche permet de réduire considérablement l'impact des ions nocifs et des matières organiques sur le sol. Le produit final issu de ce traitement peut être utilisé pour le remblayage, le pavage ou même la construction.

Cependant, cette technologie présente certains inconvénients : un coût élevé, l'utilisation d'agents chimiques coûteux, des besoins en espace importants, ainsi que des risques potentiels pour la sécurité.[36]

4.4 Traitement thermique

4.4.1 Incinération

La technologie d'incinération repose sur l'oxydation thermique des composants organiques présents dans les déblais de forage à base d'huile. Ce procédé est réalisé à des températures très élevées, généralement comprises entre 1200 et 1500°C, ce qui permet de transformer les contaminants en résidus inertes, réduisant ainsi significativement leur nocivité environnementale. L'incinération est reconnue pour sa capacité à traiter une grande diversité de déchets, y compris des volumes importants de déblais, ce qui en fait une méthode adaptée aux opérations de forage générant des quantités massives de résidus [36].

L'utilisation de ce procédé présente plusieurs avantages :

- Une réduction significative du volume des déchets, facilitant leur gestion ultérieure ;
- Une dégradation quasi complète des composés organiques polluants ;
- Une valorisation énergétique possible, par la récupération de chaleur produite lors de la combustion des fumées, pouvant être réutilisée dans d'autres processus industriels [38].

Cependant, malgré son efficacité, l'incinération présente certaines limites notables :

- Elle est inefficace pour le traitement des composants inorganiques, notamment les métaux lourds contenus dans les déblais de forage à base d'huile ;
- Elle requiert l'installation de systèmes de dépoussiérage sophistiqués afin de capter les particules métalliques présentes dans les émissions gazeuses ;
- Elle engendre une consommation énergétique particulièrement élevée, ce qui peut compromettre sa viabilité économique et environnementale ;
- Elle émet d'importantes quantités de gaz à effet de serre, principalement du dioxyde de carbone (CO_2) et des oxydes d'azote (NO_x), contribuant ainsi au changement climatique et à la pollution atmosphérique [36].

4.4.2 Désorption thermique

La désorption thermique est une technologie de traitement des déblais de forage à base d'huile reposant sur le chauffage en conditions anaérobies. Ce procédé permet de séparer les phases liquides, notamment les hydrocarbures, par évaporation suivie d'une condensation. Lors du

processus, l'eau contenue dans les déblais s'évapore en premier, générant de la vapeur qui contribue à abaisser le point d'ébullition des huiles présentes. Par conséquent, l'évaporation des hydrocarbures intervient à une température inférieure à leur point d'ébullition théorique, ce qui optimise l'efficacité énergétique du traitement.

Différents systèmes de désorption thermique sont couramment utilisés :

- Les unités à tambour chauffé indirectement, où le chauffage est assuré par des sources thermiques externes appliquées à un tambour rotatif.
- Les unités à vis sans fin chauffée, qui utilisent un fluide caloporteur circulant dans une vis creuse pour transmettre la chaleur aux déblais.
- Les unités TCC (Thermal Cuttings Cleaner), qui combinent chauffage direct et mécanique par friction pour atteindre la température désirée tout en fragmentant les particules solides.

Ces différentes unités ont en commun la capacité de réduire la teneur en huile résiduelle des solides à moins de 1%, tout en permettant la récupération des hydrocarbures par condensation.

En fonction des températures de fonctionnement, ces systèmes peuvent être classés en deux grandes catégories :

- Les systèmes à basse température (typiquement entre 250°C et 350°C), comme les TCC, adaptés au traitement des déchets contenant des huiles légères.
- Les systèmes à haute température (pouvant atteindre jusqu'à 520°C), tels que les unités de désorption thermique indirecte, destinés aux déchets plus chargés en huiles lourdes.

Les systèmes à basse température présentent plusieurs avantages significatifs :

- Lorsqu'il est possible d'extraire les hydrocarbures à basse température, la teneur résiduelle en huile sur les déblais (ROC – Retention On Cuttings) peut être maintenue en dessous de 1%, tout en réduisant la consommation énergétique globale du processus.
- Par ailleurs, les propriétés physico-chimiques des hydrocarbures récupérés restent quasiment inchangées, ce qui facilite leur réutilisation ou leur valorisation.[36]

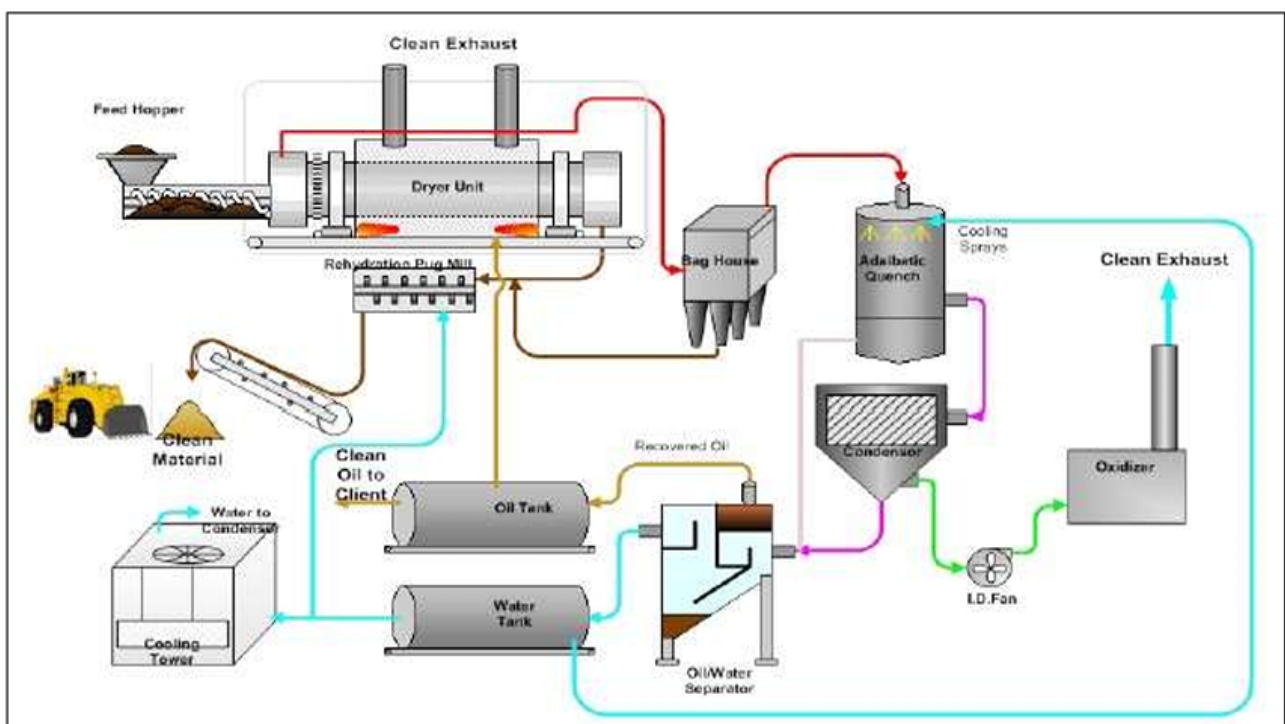


FIGURE 4.9 – Schéma simplifié de l'unité de traitement par désorption thermique.

Cependant, certains inconvénients subsistent selon le type de système utilisé. Par exemple, les systèmes de désorption thermique indirecte nécessitent une large surface d'échange thermique, ce qui les rend encombrants et très énergivores. À l'inverse, la technologie TCC parvient à surmonter ces limites : elle intègre un système de broyage mécanique par marteaux qui facilite la libération des hydrocarbures en brisant les particules solides, ce qui limite leur diffusion et améliore l'efficacité de récupération de l'huile.



FIGURE 4.10 – Thermal Cuttings Cleaner (TCC)

4.4.3 Séparation Thermique des Phases (TPS – Thermal Phase Separation)

La séparation thermique des phases (TPS) est une technologie avancée de traitement thermique qui vise à séparer les différents composants des déblais de forage contaminés, principalement les fractions aqueuses, huileuses et solides. Ce procédé repose sur l'application d'un chauffage contrôlé en environnement fermé, permettant l'évaporation sélective des fluides présents dans les déblais, suivie de leur condensation et de leur récupération sous forme de phases distinctes.

Le principe fondamental du système TPS est basé sur la différence de points d'ébullition entre l'eau, les hydrocarbures légers et les huiles lourdes. Sous l'effet de la chaleur, les phases volatiles (eau et huiles) sont vaporisées successivement, puis condensées dans des circuits de refroidissement distincts, tandis que les fractions solides sont récupérées sous forme de résidus secs. L'ensemble du procédé se déroule en conditions contrôlées, généralement sous pression réduite ou atmosphérique, afin de limiter la dégradation thermique des hydrocarbures et de maximiser la qualité des produits récupérés.

La technologie TPS offre plusieurs avantages opérationnels et environnementaux notables :

- **Séparation complète des trois phases** : le système permet de récupérer de manière distincte l'eau, les hydrocarbures et les solides, facilitant leur réutilisation ou leur élimination dans des filières adaptées.
- **Haute efficacité de récupération des hydrocarbures** : le rendement de récupération peut dépasser 95%, avec une teneur résiduelle en huile sur les solides inférieure à 0,5%, ce qui respecte largement les normes environnementales internationales.
- **Valorisation énergétique possible** : les hydrocarbures récupérés peuvent être utilisés comme combustibles, et l'énergie thermique du système peut être partiellement recyclée dans le processus, réduisant ainsi la consommation énergétique globale.

En termes de conditions opératoires, les systèmes TPS fonctionnent généralement à des températures allant de 300°C à 500°C, selon la nature des déblais et les objectifs de séparation. La

conception modulaire de certains équipements TPS permet également une installation mobile, facilitant leur déploiement sur les sites de forage isolés ou temporaires.

Toutefois, cette technologie présente aussi certaines contraintes :

- **Investissement initial élevé** : en raison de la complexité des équipements et du besoin de matériaux résistants à la chaleur et à la corrosion.
- **Maintenance technique rigoureuse** : notamment pour garantir l'efficacité des circuits de condensation et éviter les pertes de produit.
- **Production de gaz résiduels** : qui nécessitent un traitement approprié avant rejet, souvent par des systèmes complémentaires comme les torchères ou les unités de post-combustion.

En résumé, la séparation thermique des phases constitue une solution performante et polyvalente pour le traitement des déblais de forage à base d'huile. Elle s'inscrit dans une logique de valorisation des déchets et de réduction de l'impact environnemental, en permettant à la fois la récupération de ressources valorisables et le traitement efficace des résidus solides.[39]



FIGURE 4.11 – Thermal Phase Separation (TPS)

4.5 Traitement biologique

4.5.1 La biorémédiation

La biorémédiation est une technologie de traitement biologique qui exploite la capacité de certains micro-organismes (bactéries, champignons, ou plantes) à dégrader les hydrocarbures présents dans les déblais de forage à base d'huile. Elle repose sur des procédés naturels tels que le compostage, le landfarming, l'épandage contrôlé, ou encore les bioréacteurs. L'objectif est de réduire la toxicité des déchets pétroliers en les transformant en composés moins nocifs pour l'environnement.

Malgré son efficacité écologique, la biorémédiation présente plusieurs limites. Elle nécessite des temps de traitement longs, de plusieurs semaines à plusieurs années, selon la méthode utilisée — les bioréacteurs étant les plus rapides. La vitesse de dégradation dépend de plusieurs facteurs :

- les conditions environnementales (température, humidité, aération, etc.)
- la composition des déblais
- le type de procédé appliqué

Par ailleurs, cette technologie est sensible aux variations climatiques, ce qui réduit sa robustesse et limite son utilisation dans certaines zones. Elle requiert également de grandes surfaces de terrain, ce qui peut poser des contraintes logistiques. Enfin, elle ne permet ni la récupération d'énergie ni la valorisation des matériaux contenus dans les déblais, ce qui en limite le rendement global. [36]

4.5.2 Types de biorémédiation

4.5.2.0.1 Le compostage : Le compostage consiste à mélanger les sols excavés avec des amendements organiques afin de stimuler l'activité microbienne responsable de la dégradation des contaminants, notamment les hydrocarbures. C'est une méthode simple à mettre en œuvre, qui s'inscrit dans une logique de traitement naturel et durable.[40]

4.5.2.0.2 La biopile : La biopile est une version plus technique du compostage, adaptée aux sols fortement contaminés. Elle consiste à former des monticules de 2 à 4 mètres de hauteur sur des drains inclinés, recouverts d'une membrane imperméable. Les paramètres physico-chimiques (température, humidité, pH, oxygénation) sont rigoureusement contrôlés, et des nutriments ou micro-organismes peuvent être ajoutés pour accélérer la biodégradation. Le traitement dure généralement entre 8 et 56 jours. [41][42]

4.5.2.0.3 Le landfarming : Le landfarming est une méthode de traitement ex situ qui consiste à étendre les sols contaminés en fines couches sur une surface imperméable, généralement une géomembrane, afin d'éviter toute pollution secondaire. Le sol est ensuite régulièrement aéré et amendé pour favoriser l'activité biologique. Cette technique, peu coûteuse, est particulièrement efficace dans les environnements froids, comme les zones arctiques.[43]

Partie Pratique

Chapitre 5

Présentation du processus Verti-G et des données industrielles

5.1 Introduction générale à la partie pratique

5.1.1 Contexte industriel et environnemental

Les opérations de forage pétrolier engendrent d'importants volumes de déblais, appelés **cuttings**, imprégnés d'huile de forage. Ces cuttings posent un défi environnemental majeur car ils peuvent être sources de pollution des sols et des nappes phréatiques s'ils ne sont pas traités de manière adéquate.

Afin de limiter cet impact environnemental et de valoriser l'huile de forage récupérable, l'industrie pétrolière met en place des procédés de traitement spécifiques, dont le **Verti-G (Vertical Dryer)**, un équipement industriel destiné à séparer l'huile résiduelle des cuttings. Ce processus est essentiel pour permettre la réutilisation de l'huile dans les opérations de forage, tout en rejetant les cuttings dans le respect des normes environnementales en vigueur.

L'entreprise ADFC, filiale de Sonatrach, a fourni une base de données issue de ses opérations industrielles afin d'optimiser les performances de séparation réalisées par le Verti-G. L'objectif est d'explorer et de mettre en œuvre des techniques d'intelligence artificielle permettant d'identifier les paramètres optimaux de fonctionnement du Verti-G et de réduire efficacement la teneur en huile résiduelle dans les cuttings.

5.1.2 Objectifs du projet : séparation et valorisation de l'huile

L'objectif principal de ce projet est de développer et d'implémenter une approche d'intelligence artificielle pour optimiser le traitement des *cuttings* au niveau du Verti-G. Plus précisément, ce projet vise à :

- Prédire avec précision le pourcentage d'huile résiduelle dans les *cuttings* traités à partir des données industrielles fournies par ADFC.
- Identifier et sélectionner les variables influentes permettant d'améliorer les performances du traitement.
- Optimiser les modèles de Machine Learning afin d'obtenir un traitement prédictif robuste et fiable.

- Optimiser les paramètres de fonctionnement du Verti-G après la prédiction pour déterminer les meilleures combinaisons de réglages permettant de minimiser la teneur en huile résiduelle dans les *cuttings*.
- Valoriser l’huile récupérée pour une réutilisation efficace dans les opérations de forage futures, tout en assurant un rejet de *cuttings* conforme aux exigences environnementales.

La mise en œuvre de cette approche contribue non seulement à la rentabilité du processus, mais également à la préservation de l’environnement, en réduisant considérablement la pollution potentielle liée aux déchets pétroliers.

5.2 Présentation des données industrielles

5.2.1 Description du processus Verti-G

Le Verti-G, souvent désigné sous le nom de centrifugeuse verticale ou de *cuttings dryer*, constitue un dispositif industriel essentiel dans le traitement des déblais issus des opérations de forage. Cet appareil est spécifiquement conçu pour traiter les *cuttings* imprégnés d’hydrocarbures et de métaux lourds, dans le but de convertir les substances potentiellement polluantes en formes plus stables et moins nocives. Cette transformation permet de respecter les exigences environnementales et économiques de l’industrie pétrolière.

L’unité Verti-G est équipée d’un rotor conique muni d’un tamis rotatif et de systèmes de transport mécanique des déchets. Les déblais provenant du shale shaker sont convoyés par des vis sans fin jusqu’à la partie supérieure du cône de la centrifugeuse. À l’intérieur, ces déblais sont soumis à la force centrifuge générée par la rotation du cône, ce qui permet de séparer les phases liquide et solide. Le liquide est ainsi filtré à travers le tamis et collecté pour être réutilisé ou traité, tandis que les solides se concentrent à la base de l’équipement sous forme de poudre sèche et sont ensuite évacués vers le bournier (*dry pit*) au moyen d’une pelle mécanique.



FIGURE 5.1 – Vue extérieure du Verti-G

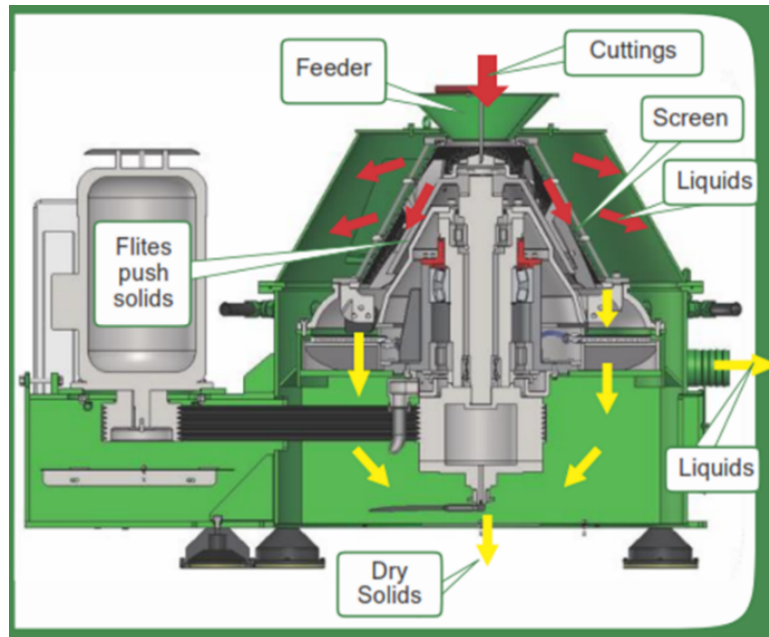


FIGURE 5.2 – Vue en coupe détaillée du Verti-G

Ce procédé joue un rôle déterminant dans la réduction de l’huile résiduelle contenue dans les cuttings et participe à la valorisation de l’huile récupérée, contribuant ainsi à une meilleure gestion environnementale et à la rentabilité du processus.[38]

5.2.2 Base de données fournie par ADFC

Dans le cadre de ce projet, l’entreprise ADFC, filiale de Sonatrach, a fourni une base de données industrielle issue des opérations réelles de traitement du Verti-G. Cette base comprend un ensemble de variables décrivant à la fois les propriétés des cuttings à traiter et les paramètres de fonctionnement du Verti-G.

Les principales colonnes présentes dans la base de données sont :

- **Density (SG)** : densité des cuttings (gravité spécifique)
- **Pourcentage huile** : proportion d’huile présente dans le fluide de forage
- **Concentration d’huile** et **Concentration d’eau** : mesures précises des concentrations respectives
- **Weighting Material, Salt, Lime, VersaCoat, VersaWet, VG 69** : additifs chimiques utilisés pour stabiliser le fluide de forage
- **Teneur en boue dans les cuttings (%)** : proportion de boue dans les *cuttings* avant traitement
- **Granulométrie des cuttings (μm)** : taille moyenne des particules des cuttings
- **Vitesse de rotation (rpm)** : vitesse de rotation du rotor conique
- **Débit d’alimentation (L/min)** : quantité de cuttings traités par minute
- **Température de fonctionnement ($^{\circ}\text{C}$)** : température interne du processus
- **Temps de séjour (s)** : durée de passage des cuttings dans le Verti-G
- **Type de tamis utilisé** : caractéristique du tamis conique

- **Fréquence de vibration (Hz)** : paramètre mécanique influant sur l’efficacité de la séparation
- **Teneur en huile résiduelle sur les cuttings secs (% OOC)** : indicateur de performance à la sortie du Verti-G

Ces variables constituent la base pour l’étape de modélisation prédictive et pour l’optimisation des paramètres du Verti-G dans le cadre de ce projet. Un prétraitement des données a été réalisé pour nettoyer et corriger les valeurs manquantes et vérifier la cohérence des mesures avant leur utilisation dans les modèles d’intelligence artificielle.

5.2.3 Illustration des données initiales

La figure ci-dessous présente un extrait représentatif de la base de données fournie par ADFC. On y retrouve notamment des variables telles que la densité (SG), le pourcentage d’huile, la concentration d’huile et d’eau, ainsi que les paramètres de fonctionnement du Verti-G (vitesse de rotation, débit d’alimentation, température, temps de séjour, type de tamis utilisé, etc.).

Note : Pour des raisons de confidentialité, seule une partie des données est présentée dans cette figure.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	
1	Density (SG)	Percentage	Concentration	Concentration	ghting Mat	Salt	Lime	VersaCoat	VersaWet	VG 69	ue dans le	trie des cu	de rotation	limentation	de fonction	de séjoi	de tamis	ce de vibra	Teneur en hui	
2	1.18	75	0.644201	0.214734	284.69	79.41788	28	11.40	2.85	11.40		336.4117	2173.813		49.01226	54.0374		37.53083	7.339614	
3	1.80	85	0.570443	0.100666	1113.605	37.23085	28	11.40	8.55	7.13	41.30868		2096.549	130.0903	37.49439	57.42632	Mesh 100		7.826584	
4	1.26	75	0.625815	0.208605	392.0589	77.15123	28		2.85	11.40	34.43844	335.435	2126.327	95.47167	48.58103	60.52494	Mesh 120	38.91096	6.909116	
5	2.02	85	0.518997	0.091588	1400.209	33.87313	20	12.83	8.55	8.55	54.01515	738.4382	1995.538	114.7996	45.3139	54.32878	Mesh 80	40.14201	6.894894	
6	1.58	90	0.590317	0.065591	425.7935	24.25832		11.40	8.55	8.33	36.52191	723.8462	1948.78	113.7353	38.00655	51.99852	Mesh 80	39.28708	7.739654	
7	1.50	85	0.587428	0.103664	372.8275	38.33938	28	11.40	8.55		34.84132	788.2761	2045.364	110.4264	44.60256	56.32749	Mesh 120		6.980495	
8	1.30	80	0.565411	0.141353	586.1854	52.27844	28	11.40	5.70		39.89606	590.0115		114.2286	45.51889	55.62527			6.928573	
9	1.20	90	0.684092	0.07601	422.2406	28.11189	40	12.00	2.00	10.00	33.83717	300.4102	2103.864	111.9093	37.09634			35.73971	8.098981	
10	1.40	90	0.592657	0.065851	65.02295	24.35448	28	12.00	4.00	20.00	31.65263	365.7314	2076.528		44.22438	66.82727	Mesh 80	37.19399	7.35482	
11	1.30	90	0.630958	0.070106	578.9604	25.92842		8.00	4.00	20.00	34.7128	392.2317	2171.363	113.9017	45.39676	60.02491	Mesh 120	37.69031	7.310627	

FIGURE 5.3 – Échantillon des données industrielles brutes

5.2.4 Analyse et interprétation

On constate que la base de données contient des informations riches et variées permettant d’étudier la relation entre les paramètres de fonctionnement du Verti-G et la teneur en huile résiduelle dans les cuttings. Plusieurs variables semblent cruciales pour l’analyse :

- Les paramètres opératoires comme la vitesse de rotation, le débit d’alimentation, la température de fonctionnement et le temps de séjour.
- Les caractéristiques physico-chimiques des cuttings (pourcentage d’huile, granulométrie).
- La variable cible, à savoir la teneur en huile résiduelle sur les cuttings secs (% OOC), qui sera au cœur de la modélisation prédictive.

On note également la présence potentielle de valeurs manquantes (NaN) qui devront être traitées lors de l’étape de nettoyage.

Afin de compléter cette première analyse descriptive, un résumé statistique des variables quantitatives a été généré. Ce tableau fournit des statistiques essentielles telles que la moyenne, l’écart-type, le minimum, le maximum et les quartiles, permettant ainsi de mieux cerner la distribution des valeurs et d’identifier d’éventuelles anomalies.

	Density (SG)	Pourcentage huile	Concentration d'huile	Concentration d'eau	Weighting Material	Salt	Lime	VersaCoat	VersaWet	VG 69	Teneur en boue dans les cuttings (%)	Granulor des cu
count	2000.000000	2000.000000	1840.000000	1840.000000	1840.000000	1840.000000	1840.000000	1840.000000	1833.000000	1839.000000	1837.000000	1831.0
mean	1.428935	85.875830	0.609886	0.105000	605.326348	38.871924	29.809239	11.671440	5.847316	12.313910	33.876086	564.0
std	0.335579	7.407277	0.092526	0.066187	481.136468	24.542423	5.169543	2.260543	2.291674	5.239705	9.277355	160.7
min	0.850000	68.690000	0.410000	0.020000	0.000000	9.000000	19.400000	7.670000	1.750000	3.300000	13.590000	213.1
25%	1.220000	84.115000	0.560000	0.060000	311.287500	21.732500	27.830000	11.227500	4.030000	8.500000	28.110000	459.8
50%	1.330000	89.520000	0.600000	0.080000	460.960000	29.180000	28.050000	11.900000	5.770000	11.690000	32.780000	564.0
75%	1.512500	90.200000	0.670000	0.140000	641.352500	52.882500	28.350000	12.317500	7.080000	15.480000	39.380000	654.6

FIGURE 5.4 – Extrait du tableau récapitulatif des statistiques descriptives des variables

Chapitre 6

Nettoyage, traitement et analyse exploratoire des données

6.1 traitement des valeurs manquantes (NaN)

6.1.1 Définition et objectif

Les données industrielles collectées peuvent contenir des valeurs manquantes, appelées NaN (Not a Number), souvent dues à des erreurs de mesure, des capteurs défectueux ou des oublis de saisie. Une gestion rigoureuse de ces valeurs manquantes est indispensable pour garantir la qualité des données utilisées lors des étapes de modélisation prédictive.[\[44\]](#)

6.1.2 Visualisation des valeurs manquantes (heatmap)

Pour mieux appréhender la répartition des valeurs manquantes dans l'ensemble des variables, une première visualisation a été réalisée à l'aide d'une heatmap. Cette représentation graphique permet de repérer rapidement les colonnes et les lignes affectées par les valeurs manquantes.



FIGURE 6.1 – Heatmap des valeurs manquantes dans la base de données

6.1.3 Analyse du pourcentage des valeurs manquantes par variable

Après la visualisation graphique, un calcul précis du pourcentage de valeurs manquantes a été effectué pour chaque variable. Les résultats sont classés par ordre croissant du pourcentage de valeurs manquantes.

TABLE 6.1 – Pourcentage de valeurs manquantes (NaN) par variable

Variable	Pourcentage de NaN
Density (SG)	0.00%
Teneur en huile résiduelle sur les cuttings secs (% OOC)	0.00%
Pourcentage huile	0.00%
Concentration d'huile	8.00%
Concentration d'eau	8.00%
Weighting Material	8.00%
Salt	8.00%
Lime	8.00%
VersaCoat	8.00%
VG 69	8.05%
Température de fonctionnement (°C)	8.10%
Vitesse de rotation (rpm)	8.15%
Débit d'alimentation (L/min)	8.15%
Temps de séjour (s)	8.15%
Type de tamis utilisé	8.15%
Fréquence de vibration (Hz)	8.15%

6.1.4 Interprétation des résultats

Les résultats de la heatmap et du calcul du pourcentage des valeurs manquantes montrent que la majorité des colonnes de la base de données contiennent des valeurs manquantes. Cependant, ces valeurs manquantes représentent un très faible pourcentage pour chaque colonne (inférieur à 9 %), ce qui laisse envisager des méthodes de traitement adaptées pour éviter de réduire significativement la taille de l'échantillon exploitable.

6.2 Méthodes de gestion des valeurs manquantes

Une fois les valeurs manquantes identifiées, plusieurs méthodes peuvent être envisagées pour les traiter. La solution la plus simple consiste à éliminer directement les lignes contenant des valeurs manquantes en utilisant la fonction `dropna()`. Cependant, bien que le pourcentage de valeurs manquantes par colonne soit modéré, cette approche risquerait de réduire fortement la taille de l'échantillon global et de laisser environ 500 lignes seulement, ce qui pourrait affecter la fiabilité, la stabilité du modèle prédictif et entraîner un phénomène de surapprentissage (overfitting).

Pour pallier ces limitations, d'autres techniques d'imputation ont été privilégiées, notamment le remplacement des valeurs manquantes par la moyenne ou la médiane selon la nature de la variable.^[45]

6.3 Mise en place du prétraitement

Le prétraitement a été mis en œuvre à l'aide de pipelines pour automatiser l'imputation et la transformation des données. Dans ce pipeline :

- Les variables numériques sont traitées en remplaçant les valeurs manquantes par la moyenne de la colonne, puis standardisées à l'aide du `StandardScaler`.
- La variable catégorielle est traitée en remplaçant les valeurs manquantes par la modalité la plus fréquente, puis transformée en variables indicatrices grâce à l'encodeur `OneHotEncoder`.

Cette stratégie garantit une préparation cohérente et homogène des données pour l'étape de modélisation prédictive.^{[44][45]}

6.4 Distribution de la cible : Teneur en huile résiduelle sur les cuttings secs (% OOC)

L'objectif de cette analyse est de visualiser la distribution de la variable cible, à savoir la teneur en huile résiduelle sur les cuttings secs (% OOC), afin de mieux comprendre comment les valeurs sont réparties dans l'ensemble des données. Cette étape permet également de détecter d'éventuelles anomalies dans la distribution (comme des outliers ou des biais).

6.4.1 Visualisation de la distribution de la cible

Afin d'obtenir une première vue d'ensemble de la distribution de cette variable, un histogramme représentant la fréquence des différentes valeurs de % OOC dans les cuttings traités par le Verti-G a été généré.

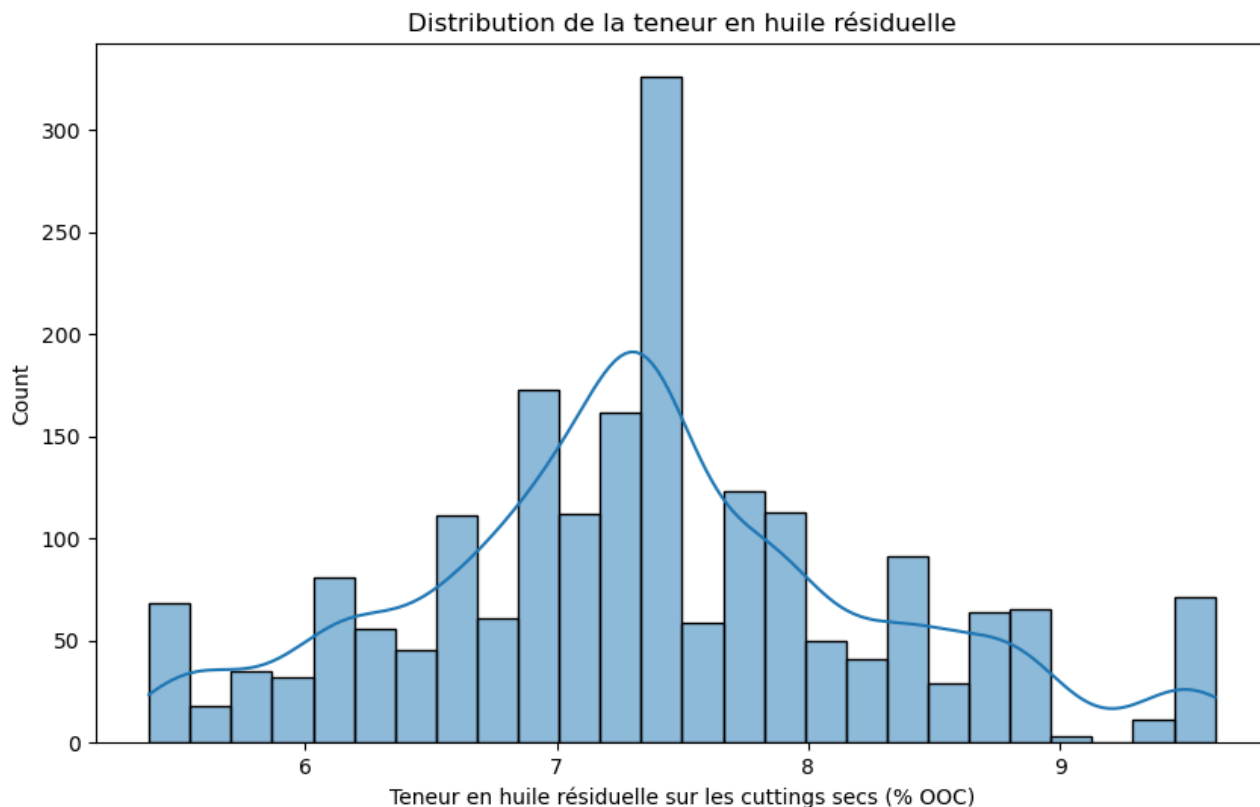


FIGURE 6.2 – Distribution de la teneur en huile résiduelle sur les cuttings secs (% OOC)

6.4.2 Analyse et interprétation

En observant le graphique de distribution, on peut noter que :

- La variable % OOC présente une distribution légèrement asymétrique avec une concentration de valeurs autour de 7.
- Les pics dans l'histogramme indiquent qu'une proportion importante des données a une teneur en huile résiduelle proche de certaines valeurs spécifiques, ce qui peut suggérer que le processus de traitement Verti-G génère des résultats plus homogènes dans certaines plages de la variable.
- Il existe également une légère tendance à une répartition plus faible des valeurs autour des extrêmes de l'intervalle, ce qui peut être un indicateur d'un traitement efficace mais pas toujours parfait pour certaines conditions de fonctionnement.

Cela nous aide à mieux comprendre la variabilité dans les performances de séparation de l'huile par le Verti-G, et peut également guider le choix des méthodes de modélisation adaptées à ces caractéristiques.

6.5 Heatmap de la matrice de corrélation

L'objectif de cette analyse est d'étudier les relations linéaires entre les différentes variables de la base de données. La matrice de corrélation permet de visualiser rapidement la force et la direction de ces relations, ce qui peut aider à mieux comprendre comment chaque variable influence la variable cible Teneur en huile résiduelle sur les cuttings secs (% OOC) ainsi que les autres caractéristiques des cuttings.

6.5.1 Visualisation de la matrice de corrélation

La figure ci-dessous montre la matrice de corrélation des variables de la base de données. Les couleurs plus proches de 1 ou -1 indiquent une forte corrélation positive ou négative, tandis que celles proches de 0 suggèrent une faible corrélation entre les variables.[46]

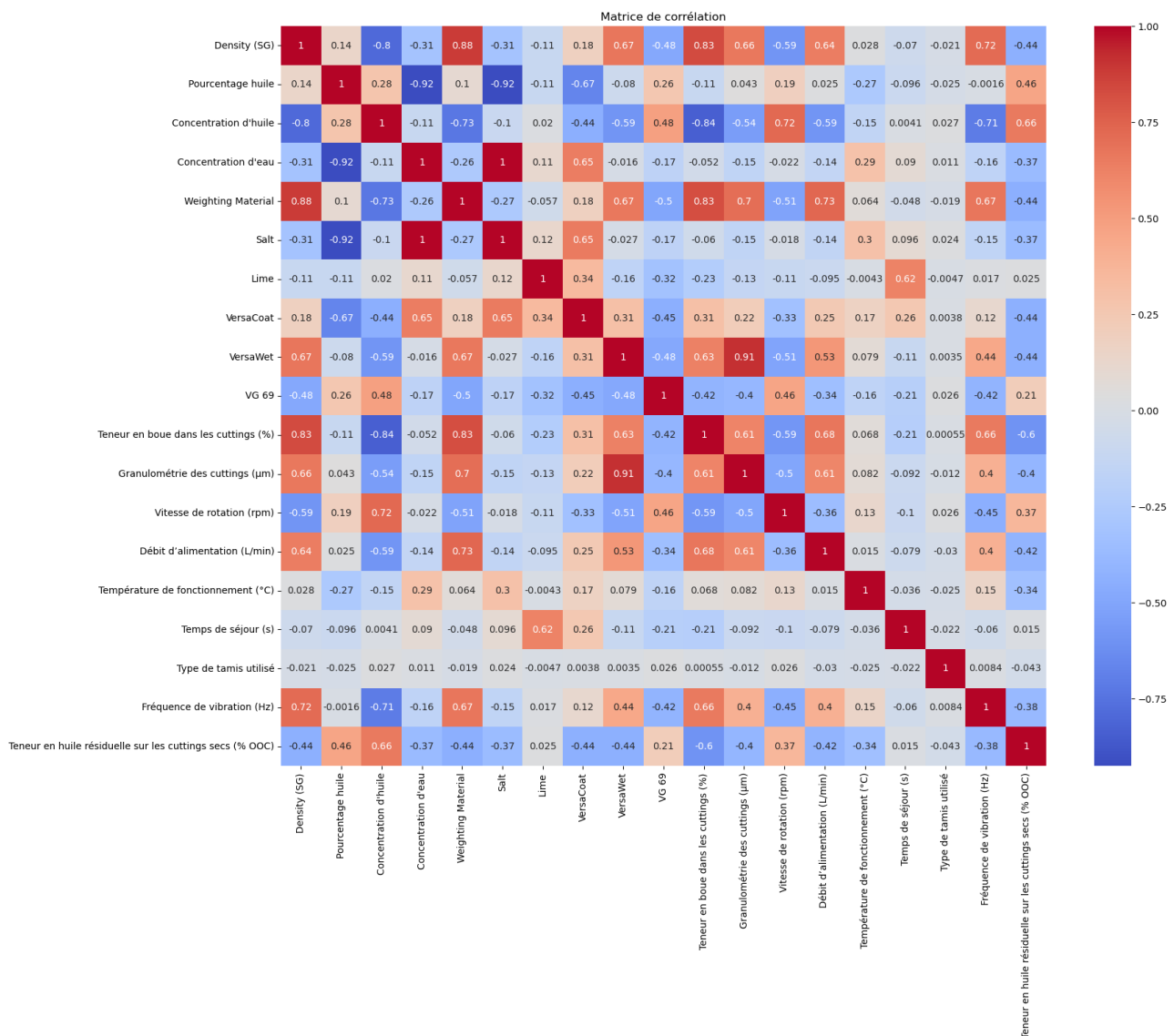


FIGURE 6.3 – Matrice de corrélation des variables

6.5.2 Analyse et interprétation

– Variables fortement corrélées :

- **Pourcentage huile** et **Concentration d’huile** présentent une corrélation très forte (0,92), ce qui indique qu’elles sont fortement liées et qu’elles capturent probablement des informations similaires sur la composition des *cuttings*.
- **Teneur en boue dans les cuttings (%)** est fortement corrélée avec plusieurs variables, notamment **Concentration d’huile** et **Concentration d’eau** (-0,83, -0,88). Cela peut indiquer que plus la boue est présente, plus la séparation de l’huile devient complexe.
- **Teneur en huile résiduelle sur les cuttings secs (% OOC)** montre une corrélation négative avec plusieurs variables comme **Density (SG)** (-0,44) et **Vitesse de rotation (rpm)** (-0,37), ce qui suggère que ces paramètres peuvent influencer la quantité d’huile résiduelle après traitement.

– Variables faiblement corrélées :

- Les variables comme **Type de tamis utilisé** et **Fréquence de vibration (Hz)** ont des corrélations proches de zéro avec la majorité des autres variables, ce qui pourrait suggérer qu’elles n’ont pas une influence directe sur le processus de séparation.

6.5.3 Importance pour le choix des modèles de Machine Learning

Le fait qu’il existe une certaine corrélation entre les différentes variables est un bon indicateur pour le choix des modèles de Machine Learning. Les modèles qui bénéficient de relations linéaires ou non linéaires entre les variables peuvent être particulièrement efficaces dans ce cas.

Ces modèles peuvent être optimisés en tenant compte des relations entre les variables afin de maximiser la précision des prédictions, tout en réduisant les risques d’instabilité grâce à des techniques de pénalisation.

6.6 Importance des variables

L’objectif de cette analyse est d’identifier les variables ayant le plus grand impact sur la prédiction de la teneur en huile résiduelle sur les cuttings secs (% OOC). Cette étape est essentielle pour sélectionner les variables pertinentes et optimiser la modélisation prédictive. Deux méthodes ont été utilisées pour mesurer l’importance des variables :

- L’importance des variables calculée par le modèle RandomForestRegressor
- Les valeurs SHAP (SHapley Additive exPlanations) qui expliquent l’impact des variables sur les prédictions du modèle.

6.6.1 Graphique de l’importance des variables (barplot)

Cette méthode fournit une estimation de l’importance de chaque variable en fonction de sa contribution à la réduction de l’erreur de prédiction. Plus la barre est longue, plus la variable a une influence sur la prédiction du modèle.

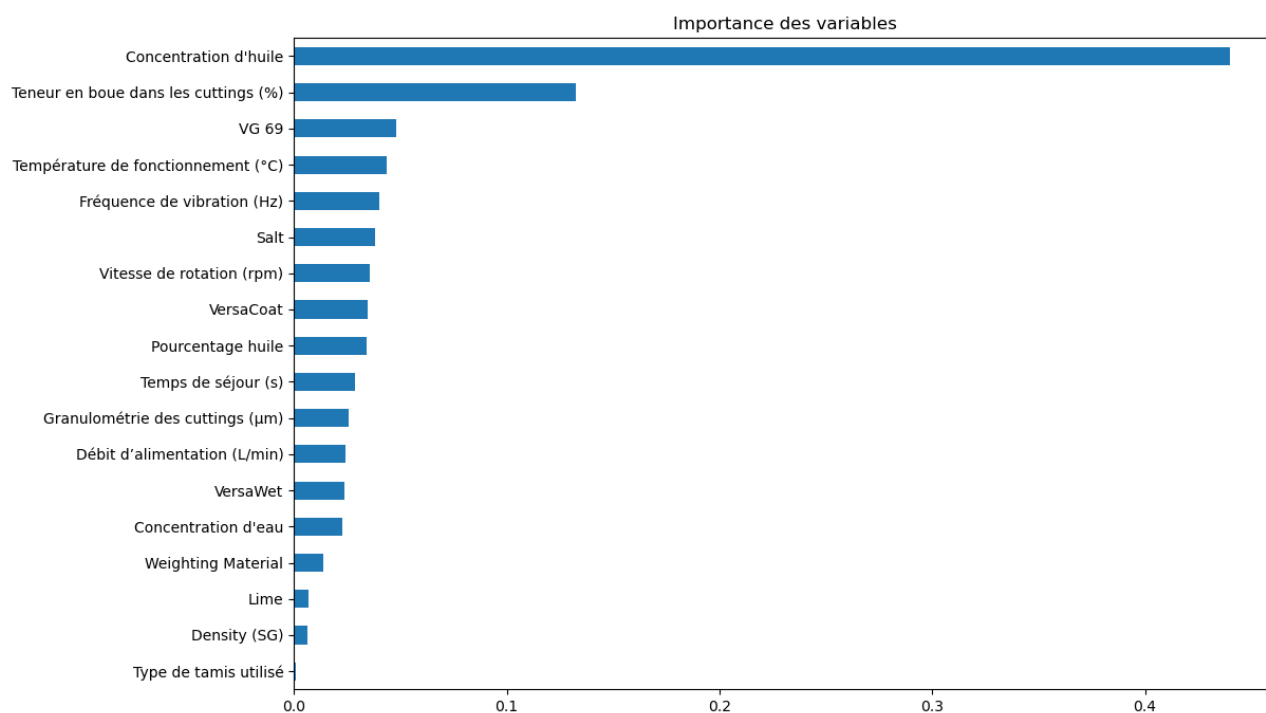


FIGURE 6.4 – Importance des variables (barplot)

6.6.2 Graphique SHAP

Les valeurs SHAP sont utilisées pour comprendre l'impact individuel de chaque variable sur les prédictions du modèle. Le graphique montre l'impact de chaque valeur de variable sur le résultat final, où les couleurs indiquent l'intensité de l'impact.[\[47\]](#)

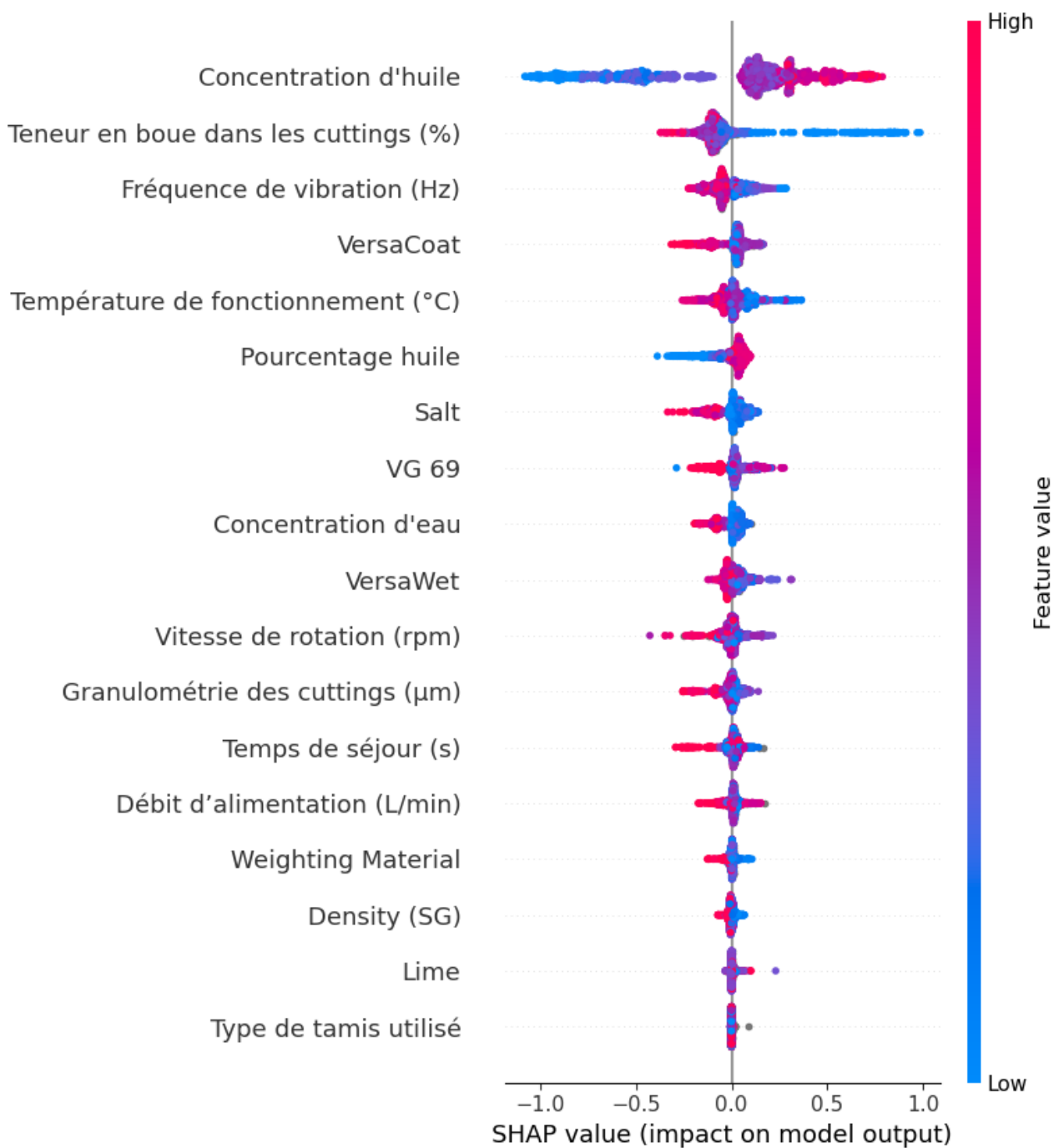


FIGURE 6.5 – Visualisation des valeurs SHAP pour l'importance des variables

6.6.3 Analyse et interprétation

- Graphique d'importance des variables (barplot) :
 - **Concentration d'huile** et **Teneur en boue dans les cuttings (%)** ressortent comme les variables les plus influentes dans la prédiction de la teneur en huile résiduelle, ce qui suggère qu'elles jouent un rôle central dans la performance du modèle.
 - Des variables comme **VG 69**, **Température de fonctionnement (°C)** et **Fréquence de vibration (Hz)** viennent ensuite avec des importances modérées, ce qui peut indiquer qu'elles influencent également la séparation de l'huile mais de manière secondaire.

- Les variables **Density (SG)** et **Type de tamis utilisé** ont une importance plus faible, ce qui pourrait signifier qu'elles ont un impact limité sur la prédiction du modèle.
- **Graphique SHAP :**
 - Le graphique SHAP montre que la **concentration d'huile** et la **teneur en boue dans les cuttings** sont les variables avec les plus grandes valeurs SHAP, ce qui confirme leur impact important sur la décision du modèle.
 - Les valeurs SHAP montrent également des comportements non linéaires pour certaines variables, ce qui suggère que les effets de ces variables varient en fonction de leurs valeurs spécifiques dans l'ensemble des données.

6.7 Conclusion

Dans cette section, nous avons procédé à une analyse exploratoire des données issues du traitement Verti-G afin de mieux comprendre leur structure. Après un échantillonnage et nettoyage, nous avons traité les valeurs manquantes, visualisé la distribution de la variable cible, analysé les corrélations entre variables et évalué leur importance. Ces étapes nous ont permis de préparer les données dans des conditions optimales pour la modélisation.

À l'issue de cette exploration, il devient évident que certaines variables influencent significativement la teneur en huile résiduelle. Cela ouvre la voie à l'application de techniques de Machine Learning supervisé, notamment des modèles capables de gérer des relations linéaires et non linéaires.

Le choix du modèle adéquat dépend de la nature des données, du nombre d'échantillons, et de la structure des relations entre les variables. Pour nous orienter dans cette décision, le schéma ci-dessous (fourni par scikit-learn) offre une vue d'ensemble des algorithmes applicables selon différents scénarios de régression.

Dans le chapitre suivant, nous allons donc sélectionner et comparer plusieurs modèles de régression tels que RidgeRegression, SVR, Random Forest, Lasso en analysant leurs performances selon des critères bien définis (R^2 , RMSE, MAE). Cette étape nous permettra de construire une base solide pour des optimisations avancées comme le tuning hyperparamétrique et l'assemblage de modèles.

La figure 6.6 présente un schéma de sélection d'un estimateur en apprentissage automatique [48].

Chapitre 7

Modélisation prédictive de la teneur en huile résiduelle

7.1 Sélection et comparaison des modèles de Machine Learning

7.1.1 Objectif de la modélisation

- L’objectif principal de la modélisation dans cette étude est de développer un outil prédictif capable d’estimer avec précision la teneur en huile résiduelle sur les *cuttings* après traitement par le Verti-G, à partir d’un ensemble de paramètres mesurés en amont. Ce modèle repose sur les données combinées issues de la composition du mélange boue-*cuttings* et des paramètres de fonctionnement de la centrifugeuse verticale (Verti-G).
- L’enjeu fondamental est double :
 - **D’un point de vue opérationnel** : il s’agit de fournir aux opérateurs un outil de prévision qui permette d’évaluer l’efficacité de séparation attendue en fonction de réglages choisis.
 - **D’un point de vue stratégique** : le modèle vise à être inversé pour proposer, à partir d’un mélange donné, les paramètres optimaux de fonctionnement de la machine afin de réduire au maximum la teneur en huile résiduelle dans les *cuttings* en sortie.
- Plus précisément, la modélisation doit répondre aux objectifs suivants :
 - Construire un modèle robuste capable de prendre en compte la diversité des mélanges (densité, taux d’huile, granulométrie, etc.) et les différents paramètres opérationnels (vitesse de rotation, débit d’alimentation, etc.).
 - Évaluer la performance de plusieurs algorithmes d’apprentissage supervisé en termes de capacité de prédiction, notamment en comparant les métriques d’erreur (*RMSE*, *MAE*) et de qualité d’ajustement (R^2).
 - Fournir une base solide pour une optimisation inverse des paramètres de réglage du Verti-G, en s’appuyant sur le modèle prédictif retenu.

7.1.2 Sélection initiale des modèles

Afin de déterminer les algorithmes les plus performants pour la prédiction de la teneur en huile résiduelle sur les cuttings, une première phase expérimentale a consisté à tester une large gamme de modèles de régression issus de différentes familles d'apprentissage automatique. L'objectif de cette étape préliminaire était d'avoir une vision comparative globale des performances de ces modèles sur notre jeu de données industriel.

Parmi les algorithmes évalués, on retrouve des approches linéaires (régression linéaire, Ridge, Lasso, ElasticNet), des méthodes à base d'arbres (Decision Tree, Random Forest, Gradient Boosting, AdaBoost), des modèles à noyaux (Support Vector Regressor – SVR), ainsi que des techniques de boosting plus récentes telles que CatBoost, LightGBM et XGBoost.

- Chaque modèle a été intégré dans une pipeline complète comprenant :
 - l'imputation des valeurs manquantes par la moyenne,
 - la normalisation des variables par standardisation,
 - et une évaluation rigoureuse via validation croisée à 5 plis (cross-validation $k = 5$).
- Les performances ont été comparées selon trois indicateurs standards :[49]
 - R^2 (coefficient de détermination) : mesure la proportion de la variance de la variable cible expliquée par le modèle. Il est défini par la formule :

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

- **RMSE** (Root Mean Squared Error) : indique l'erreur quadratique moyenne entre les valeurs prédites et observées :

$$RMSE = \sqrt{\frac{1}{n} \sum (y_i - \hat{y}_i)^2}$$

- **MAE** (Mean Absolute Error) : mesure l'erreur absolue moyenne des prédictions :

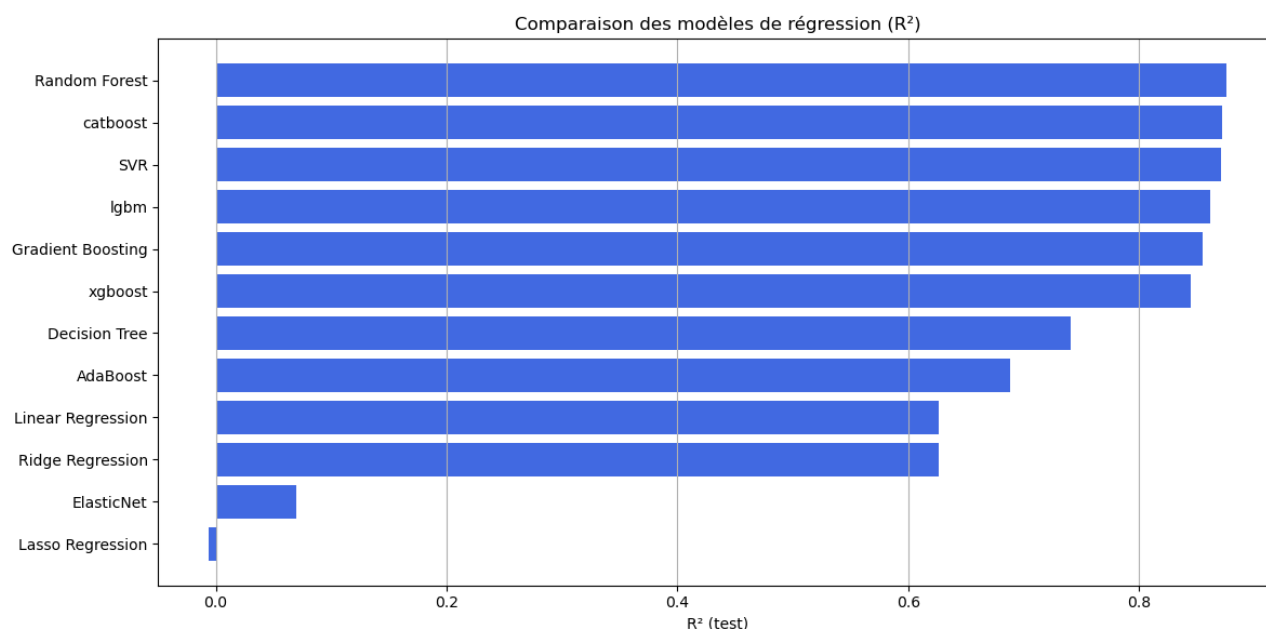
$$MAE = \frac{1}{n} \sum |y_i - \hat{y}_i|$$

Les résultats obtenus sont présentés dans le tableau suivant :

TABLE 7.1 – Résultats comparatifs des performances des modèles de régression

Modèle	R^2 (test)	RMSE	MAE	Score CV (5-fold)
Random Forest	0.8756	0.3107	0.1762	0.8822
CatBoost	0.8716	0.3157	0.1845	0.8836
SVR	0.8711	0.3162	0.1704	0.8864
LightGBM	0.8611	0.3282	0.1922	0.8722
Gradient Boosting	0.8549	0.3355	0.2130	0.8701
XGBoost	0.8444	0.3475	0.2028	0.8643
Decision Tree	0.7408	0.4485	0.2031	0.7479
AdaBoost	0.6879	0.4921	0.3987	0.7180
Linear Regression	0.6262	0.5385	0.4293	0.6190
Ridge Regression	0.6262	0.5385	0.4293	0.6191
ElasticNet	0.0696	0.8496	0.6417	0.0646
Lasso Regression	-0.0058	0.8833	0.6646	-0.0052

La figure suivante illustre visuellement la comparaison des performances des principaux modèles sur la base des scores R^2 :

FIGURE 7.1 – Comparaison des modèles de régression selon le score R^2

Les résultats montrent que les modèles **Random Forest**, **CatBoost**, **SVR** et **LightGBM** surpassent nettement les autres en termes de précision et de robustesse, avec des scores R^2 supérieurs à 0,86 et une bonne stabilité en validation croisée. Le **SVR** se distingue notamment par la plus faible MAE (0,1704), tandis que **CatBoost** affiche le meilleur score moyen en validation croisée (0,8836), indiquant une excellente généralisation.

Ces modèles peuvent être brièvement caractérisés comme suit :

- **Random Forest Regressor** : algorithme d'ensemble basé sur la moyenne des prédictions de multiples arbres de décision entraînés sur des sous-échantillons aléatoires. Il est robuste, efficace pour modéliser des relations non linéaires, et peu sensible aux surapprentissage.[50]
- **Support Vector Regressor (SVR)** : extension des SVM à la régression, cherchant à maintenir les erreurs dans une marge de tolérance. Il est efficace pour les données de petite à moyenne taille avec des relations complexes.[51]
- **CatBoost Regressor** : méthode de *gradient boosting* développée pour gérer nativement les variables catégorielles. Elle construit des arbres séquentiels en corrigeant les erreurs du modèle précédent, avec une excellente stabilité.[52]
- **LightGBM Regressor** : algorithme de *boosting* rapide et efficace, conçu pour les grands jeux de données. Il utilise une stratégie de croissance “leaf-wise” avec histogramme, permettant une meilleure optimisation du gain.[52]

Ces observations ont conduit à la sélection définitive de ces quatre modèles pour la suite du travail, qui portera sur l'optimisation de leurs hyperparamètres, puis sur l'évaluation d'une approche combinée par empilement (**stacking**).

7.2 Optimisation des hyperparamètres avec Optuna

L'ajustement des hyperparamètres est une étape déterminante dans le processus de construction de modèles de régression performants. Ces paramètres, propres à chaque algorithme, influencent directement la qualité de l'approximation, la stabilité des prédictions et la capacité de généralisation du modèle. Afin d'optimiser les performances des modèles sélectionnés, différentes méthodes d'optimisation ont été explorées (Grid Search, Random Search, Optuna).

Dans un premier temps, deux approches classiques ont été mises en œuvre : **Grid Search** et **Random Search**.

- **Grid Search** consiste à explorer exhaustivement toutes les combinaisons possibles d'un ensemble de valeurs prédéfinies pour chaque hyperparamètre. Bien que méthodique, cette approche devient rapidement coûteuse en temps de calcul dès que l'espace de recherche s'élargit.[52]
- **Random Search**, quant à elle, génère aléatoirement des combinaisons d'hyperparamètres dans un espace défini, avec un gain d'efficacité par rapport à Grid Search, mais au prix d'une couverture incomplète et parfois sous-optimale de l'espace de recherche.[53]

Pour surmonter ces limitations, la méthode **Optuna** a été testée. Celle-ci repose sur une approche d'optimisation bayésienne séquentielle, où chaque itération est guidée par les performances précédentes, permettant une exploration intelligente, rapide et ciblée de l'espace de recherche.

7.2.1 Justification du choix d'Optuna

La comparaison des performances entre Grid Search et Optuna sur les quatre modèles retenus (Random Forest, SVR, CatBoost, LGBM) a démontré la supériorité d'Optuna en termes de précision et de minimisation des erreurs.

Le tableau suivant résume les performances observées :

TABLE 7.2 – Comparaison des performances (R^2 , RMSE) des modèles optimisés par Grid Search et Optuna

Modèle	Méthode	R^2	RMSE
LGBM	Grid Search	0.861	0.329
	Optuna	0.872	0.315
Random Forest	Grid Search	0.879	0.306
	Optuna	0.880	0.305
SVR	Grid Search	0.881	0.304
	Optuna	0.884	0.300
CatBoost	Grid Search	0.875	0.311
	Optuna	0.877	0.309

Dans tous les cas, Optuna a permis d’affiner les hyperparamètres au-delà de ce que permettaient les méthodes conventionnelles, et a donc été retenue comme méthode d’optimisation principale pour la suite du travail.

7.2.2 Présentation d’Optuna

Optuna est une bibliothèque d’optimisation automatique d’hyperparamètres, open source et en Python, conçue pour les tâches de Machine Learning. Elle repose sur l’algorithme *Tree-structured Parzen Estimator* (TPE), une méthode bayésienne qui estime la distribution des performances passées afin de proposer, à chaque itération, les combinaisons les plus prometteuses.^[54]

Ses principaux avantages sont :

- Une exploration adaptative de l’espace de recherche, plus efficace que la grille fixe ou l’aléatoire
- Une capacité à gérer des espaces complexes et conditionnels (paramètres discrets, continus, dépendants)

- Une intégration facile avec les bibliothèques populaires comme `scikit-learn`, `XGBoost`, `CatBoost` et `LightGBM`
- Un temps de calcul optimisé, en réduisant le nombre d’itérations nécessaires pour converger vers des solutions performantes.[55]

7.2.3 Application d’Optuna aux modèles retenus

Chaque modèle a été soumis à un processus d’optimisation par Optuna avec des espaces de recherche adaptés, en maximisant le score R^2 et en validant par validation croisée à 5 plis. Voici les paramètres optimaux et les performances obtenues :

1. LightGBM Regressor

- Paramètres optimaux : `n_estimators=368`, `learning_rate=0.0127`, `max_depth=10`, `num_leaves=34`, `min_child_samples=24`, `subsample=0.832`, `colsample_bytree=0.523`
- Résultats : $R^2 = 0.872$, RMSE = 0.315

2. Random Forest Regressor

- Paramètres optimaux : `n_estimators=187`, `max_depth=16`, `min_samples_split=9`, `min_samples_leaf=4`, `max_features='log2'`, `bootstrap=False`
- Résultats : $R^2 = 0.880$, RMSE = 0.305

3. Support Vector Regressor (SVR)

- Paramètres optimaux : `C=0.484`, `gamma=0.110`, `epsilon=0.176`, `kernel='rbf'`
- Résultats : $R^2 = 0.884$, RMSE = 0.300

4. CatBoost Regressor

- Paramètres optimaux : `iterations=500`, `learning_rate=0.0491`, `depth=6`, `l2_leaf_reg=6.9`, `random_strength=0.262`, `bootstrap_type='Bernoulli'`
- Résultats : $R^2 = 0.877$, RMSE = 0.309

Ces résultats démontrent l’efficacité d’Optuna pour exploiter pleinement le potentiel de chaque modèle, et justifient son utilisation comme outil central d’optimisation dans cette étude. L’ensemble de ces modèles optimisés sera par la suite intégré dans une approche d’assemblage par stacking, afin d’en évaluer la complémentarité.

7.3 Assemblage de modèles (stacking)

7.3.1 Principe du stacking

Le stacking, ou empilement de modèles, est une technique d’assemblage qui combine plusieurs modèles de Machine Learning afin d’améliorer la performance de prédiction. Plutôt que de se fier à un seul modèle, l’idée est d’exploiter les forces de différents modèles pour obtenir une prédiction plus robuste et précise. Cette méthode est particulièrement utile dans des situations où des modèles individuels peuvent être bons mais présentent des lacunes dans certaines parties du jeu de données ou dans la gestion des différents types de relations entre les variables.[56]

7.3.2 Pourquoi combiner les modèles ?

Combiner plusieurs modèles permet de tirer parti des complémentarités entre eux. Chaque modèle peut être efficace pour une partie spécifique des données ou de la structure du problème, mais aucun modèle n'est parfait pour toutes les situations. Par exemple :

- **Les modèles linéaires** (comme la régression linéaire ou SVR) peuvent bien modéliser les relations linéaires mais échouer sur les relations non linéaires complexes.
- **Les arbres de décision** (comme Random Forest ou XGBoost) peuvent gérer des relations non linéaires, mais ils sont susceptibles de surapprendre ou de sous-apprendre selon la configuration des données.

Le stacking combine ainsi ces modèles en une seule méthode d'assemblage pour augmenter les précisions globales et réduire le biais et la variance associés aux prédictions. Le modèle combiné est souvent plus stable et plus fiable, car il exploite les points forts de chaque modèle tout en compensant leurs faiblesses respectives.

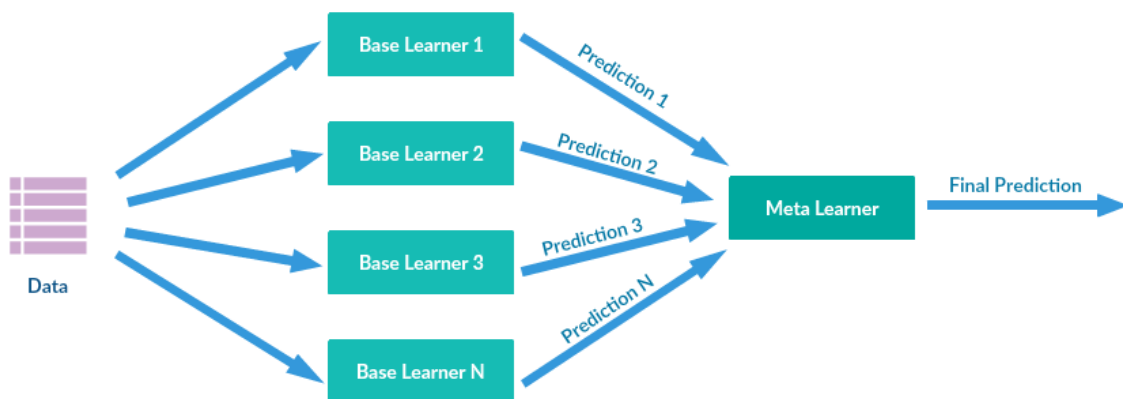


FIGURE 7.2 – Schéma de fonctionnement d'un modèle de stacking

7.3.3 Avantages du stacking :

- **Réduction du biais** : le stacking permet de réduire le biais en combinant les prédictions de différents modèles, ce qui permet d'obtenir une estimation plus fiable.
- **Amélioration de la précision** : en intégrant plusieurs modèles, on peut améliorer la précision globale de la prédiction, car chaque modèle pourrait capturer différentes dynamiques dans les données.
- **Flexibilité** : le stacking peut être utilisé avec différents types de modèles (linéaires, arbres de décision, réseaux de neurones, etc.), ce qui le rend très flexible.

En résumé, le stacking offre une approche plus robuste pour prédire des résultats, en s'appuyant sur la diversité des modèles pour mieux capturer les variations et nuances des données.

7.4 Construction de l'assembleur

Le processus de stacking repose sur la construction de deux composants clés : les modèles de base et le méta-modèle. Les modèles de base sont des modèles simples ou complexes qui génèrent des prédictions indépendantes sur les données d'entrée, tandis que le méta-modèle combine ces prédictions pour fournir une prédiction finale. L'objectif principal est de maximiser la performance en intégrant différentes perspectives de plusieurs modèles afin de réduire l'erreur de prédiction par rapport aux modèles individuels.[56, 57]

7.4.1 Modèles de base

Les modèles de base sont les modèles qui vont fournir les prédictions sur lesquelles le méta-modèle va travailler. Dans ce projet, nous avons choisi plusieurs modèles déjà optimisés via Optuna. Ces modèles sont :

- **Random Forest Regressor**
- **SVR** (Support Vector Regressor)
- **CatBoost Regressor**
- **LGBM Regressor**

Chaque modèle de base est appliqué indépendamment sur les données d'entraînement, puis ses prédictions sont utilisées comme nouvelles caractéristiques (features) pour l'entraînement du méta-modèle. Ces modèles sont choisis pour leurs capacités à gérer différents types de relations (linéaires, non linéaires, etc.) et leurs performances respectives.[56, 57]

7.4.2 Méta-modèle

Le méta-modèle est le modèle final qui va combiner les sorties des modèles de base. Typiquement, le méta-modèle est un modèle simple (par exemple une régression linéaire ou Logistic Regression), mais il peut aussi s'agir de modèles plus complexes comme des arbres de décision ou des modèles d'ensemble. Le méta-modèle apprend à combiner les prédictions des modèles de base de manière optimale.

L'objectif du méta-modèle est d'apprendre quelle pondération donner à chaque modèle de base pour maximiser la précision de la prédiction finale. Il est donc formé en utilisant les prédictions des modèles de base comme données d'entrée.[56, 57]

En résumé :

- **Modèles de base** : fournissent des prédictions sur les données d'entrée.
- **Méta-modèle** : combine ces prédictions pour produire une estimation plus précise et plus robuste.

7.4.3 Pourquoi utiliser un méta-modèle ?

Le méta-modèle permet de corriger les erreurs de prédiction des modèles de base en apprenant comment chaque modèle se comporte face à différents types de données et comment il contribue

à l'erreur globale. En combinant les points forts de chaque modèle, le méta-modèle crée une prédiction plus fiable et robuste, en particulier dans des cas où un modèle seul ne suffirait pas.

Explication pratique :

- **Modèles de base** : chaque modèle (`RandomForest`, `SVR`, `CatBoost`, `LGBM`) est d'abord entraîné séparément sur les données d'entraînement.
- **Méta-modèle (`RidgeCV`)** : après l'entraînement des modèles de base, leurs prédictions sont utilisées comme nouvelles variables pour entraîner un modèle `RidgeCV`. Celui-ci apprend à combiner les prédictions des modèles de base afin de produire une estimation finale plus précise.
- **Validation croisée** : le stacking utilise une validation croisée à 5 plis (`cv=5`) pour évaluer la performance sur différentes portions des données d'entraînement et optimiser les paramètres.

En résumé, cette approche de *stacking* permet de combiner plusieurs modèles puissants et de tirer parti de leurs forces respectives, tout en améliorant la robustesse et la précision des prédictions.

7.5 Évaluation des performances du stacking

7.5.1 Comparaison avec les modèles `SVR` et `Random Forest`

Dans cette sous-section, nous évaluons l'intérêt de combiner deux modèles optimisés, `SVR` et `RandomForestRegressor`, dans une stratégie de stacking. L'objectif est de comparer les performances individuelles des deux modèles avec celles de leur assemblage via un stacking pour vérifier si cette combinaison améliore la qualité des prédictions

7.5.2 Résultats obtenus

TABLE 7.3 – Comparaison des performances des modèles `SVR`, `Random Forest` et `Stacking` (`RF + SVR`)

Modèle	RMSE	MAE	R^2	R^2 moyen (CV)	Écart-type CV
SVR	0.312	0.1912	0.875	0.8852	0.0137
Random Forest	0.305	0.1740	0.880	0.8980	0.0095
Stacking (RF + SVR)	0.298	0.1562	0.886	0.9009	0.0098

7.5.3 Analyse et interprétation

Les résultats montrent que la combinaison des modèles **RandomForest** et **SVR** via le *stacking* permet d'obtenir une amélioration mesurable sur tous les indicateurs :

- Le **RMSE** baisse à 0,298, contre 0,305 pour **RandomForest** seul et 0,312 pour **SVR**, ce qui indique une diminution de l'erreur quadratique moyenne.
- Le **MAE** atteint 0,1562, ce qui représente une amélioration nette par rapport aux performances des deux modèles séparés.
- Le R^2 passe à 0,886, un score légèrement supérieur à celui de **RandomForest** seul, mais surtout accompagné d'un R^2 moyen de 0,9009 en validation croisée, preuve d'une excellente capacité de généralisation.
- L'**écart-type du R^2** (0,0098) reste très stable, ce qui confirme la robustesse du modèle combiné.

Ces résultats indiquent que le *stacking* exploite efficacement les complémentarités entre **SVR** (qui modélise bien les relations lisses et continues) et **RandomForest** (efficace sur les variables fortement non linéaires). Le méta-modèle apprend à pondérer intelligemment les sorties des deux pour améliorer la prédiction finale.

7.6 Sélection du meilleur modèle et discussion des résultats

Après avoir exploré, nettoyé et préparé la base de données expérimentale, plusieurs modèles de régression ont été testés pour prédire avec précision la teneur en huile résiduelle sur les cuttings secs (% OOC). Ces modèles incluent à la fois des algorithmes individuels optimisés à l'aide de la bibliothèque Optuna, ainsi que des modèles d'assemblage (*stacking*) combinant plusieurs prédicteurs de base.

L'objectif de cette section est d'évaluer, comparer et sélectionner le modèle le plus performant à l'aide de métriques standards de régression. Ces indicateurs permettront de juger la précision, la robustesse et la capacité de généralisation de chaque modèle.

7.6.1 Méthodologie d'évaluation

Chaque modèle a été évalué à l'aide d'un jeu de test indépendant (issu de la division *train/test*) et validé par une validation croisée à 5 plis (5-fold cross-validation).

Les métriques utilisées sont :

- **RMSE** (*Root Mean Squared Error*)
- **MAE** (*Mean Absolute Error*)
- R^2 (coefficient de détermination)
- R^2 moyen en validation croisée
- **Écart-type du R^2** en validation croisée

7.6.2 Résultats obtenus

Le tableau suivant résume les performances des modèles testés :

TABLE 7.4 – Performances comparées des modèles individuels et empilés (stacking) selon cinq métriques d'évaluation

Modèle	RMSE	MAE	R^2	R^2 moyen (CV)	Écart-type CV
SVR	0.312	0.1912	0.875	0.8852	0.0137
Random Forest	0.305	0.1740	0.880	0.8980	0.0095
LGBM	0.314	0.1836	0.873	0.8881	0.0085
CatBoost	0.316	0.1884	0.872	0.8862	0.0100
RandomF + SVR	0.298	0.1562	0.886	0.9009	0.0098
SVR + LGBM	0.300	0.1619	0.884	0.8955	0.0086
RandomF + CatB	0.302	0.1607	0.882	0.9007	0.0105
RandomF + LGBM	0.303	0.1604	0.882	0.9012	0.0107
CatB + SVR	0.299	0.1667	0.885	0.8939	0.0105
CatB + SVR + LGBM	0.299	0.1616	0.885	0.8957	0.0092
RandomF + CatB + SVR	0.299	0.1559	0.885	0.9012	0.0098
RandomF + SVR + LGBM	0.294	0.1514	0.889	0.9023	0.0124
RandomF + CatB + LGBM	0.301	0.1606	0.883	0.9011	0.0108
RandomF + CatB + SVR + LGBM	0.299	0.1556	0.885	0.9017	0.0102

7.6.3 Analyse des performances

De manière générale, les modèles individuels tels que Random Forest, SVR, LGBM ou CatBoost offrent des performances satisfaisantes, avec des scores R^2 compris entre 0.872 et 0.880, et des erreurs (RMSE) autour de 0.31 à 0.305. Cependant, les résultats montrent que les modèles de stacking surpassent clairement les prédicteurs simples. Cela est cohérent avec la logique de l'assemblage, qui vise à combiner les forces de plusieurs modèles pour réduire les erreurs et améliorer la précision.

Le modèle RandomF_SVR_LGBM s'est démarqué :

- Il enregistre le **RMSE** le plus bas (**0,294**) et le **MAE** le plus faible (**0,1514**).
- Il atteint un R^2 de **0,889**, le plus élevé parmi toutes les configurations testées.
- Il offre également le meilleur score moyen en validation croisée (**0,9023**), confirmant sa capacité à bien généraliser sur des données non vues.

7.6.4 Sélection du meilleur modèle

En tenant compte de l'ensemble des critères de performance (précision sur le test, régularité en validation croisée, stabilité), le modèle de *stacking* RandomF_SVR_LGBM est retenu comme le meilleur modèle prédictif de ce travail.

Sa capacité à combiner des modèles complémentaires lui permet de tirer parti à la fois :

- des caractéristiques linéaires, captées par SVR,
- et des relations non linéaires, apprises par Random Forest et LGBM,

tout en réduisant les erreurs globales et en assurant une excellente généralisation.

7.7 Validation du modèle sélectionné : RandomF_SVR_LGBM

Pour confirmer la fiabilité et la généralisation du modèle sélectionné (stacking entre Random-Forest, SVR et LGBM), deux analyses graphiques supplémentaires ont été réalisées :

- **une courbe d'apprentissage**
- **une distribution des erreurs (résidus)**

7.7.1 Courbe d'apprentissage

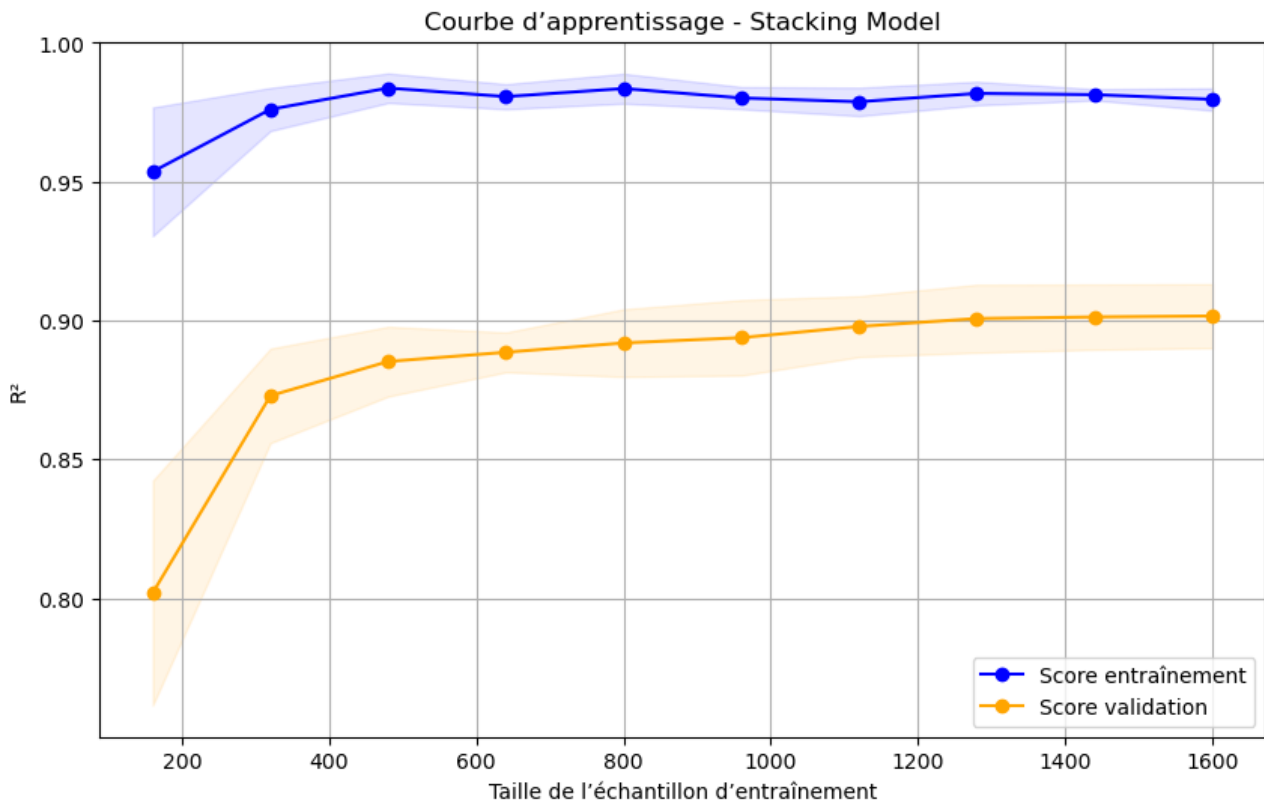


FIGURE 7.3 – Courbe d'apprentissage du modèle `RandomF_SVR_LGBM`

Cette courbe montre l'évolution du score R^2 en fonction de la taille de l'échantillon d'apprentissage. On y distingue deux courbes :

- En **bleu** : score sur l'ensemble d'apprentissage,
- En **orange** : score en validation croisée.

Interprétation :

- Le score d'apprentissage reste très élevé et stable (environ $R^2 \approx 0,98$), ce qui confirme que le modèle apprend efficacement sans difficulté majeure.
- Le score de validation augmente rapidement avec la taille de l'échantillon et se stabilise autour de $R^2 \approx 0,90$, ce qui montre que le modèle généralise bien à de nouvelles données.
- L'écart entre les deux courbes est raisonnablement faible et constant, ce qui traduit une bonne capacité de généralisation sans surapprentissage (*overfitting*).

7.7.2 Distribution des erreurs (résidus)

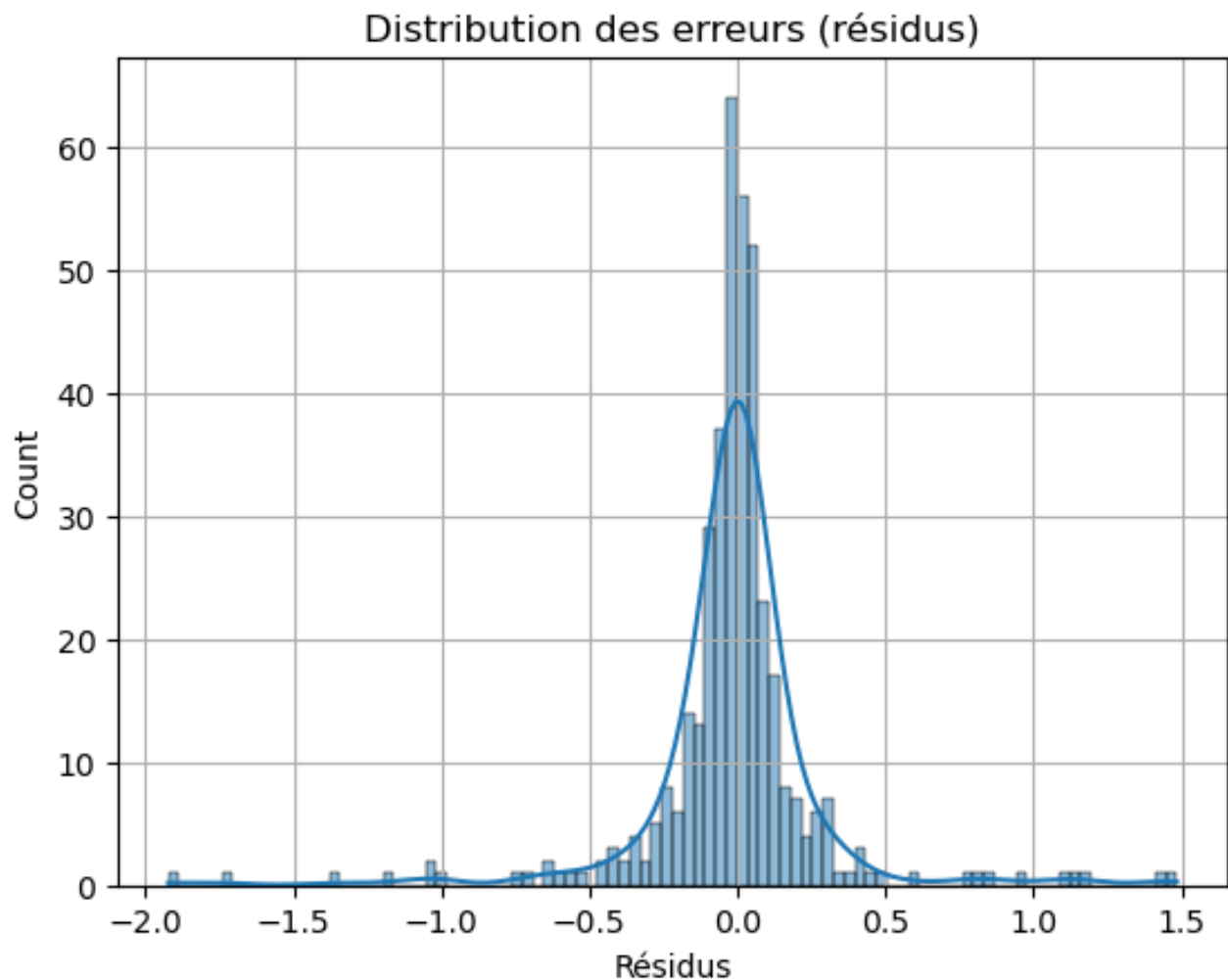


FIGURE 7.4 – Distribution des résidus du modèle `RandomF_SVR_LGBM`

Ce graphique présente la distribution des erreurs de prédiction (résidus), c'est-à-dire la différence entre les valeurs réelles et les valeurs prédites.

Interprétation :

- La courbe suit une distribution quasi-gaussienne centrée autour de zéro, ce qui est un excellent indicateur de normalité des erreurs.
- Il y a peu de valeurs extrêmes (*outliers*), ce qui signifie que le modèle fait peu d'erreurs importantes.
- Cette forme de distribution confirme que les hypothèses de stabilité et de symétrie des erreurs sont respectées.

7.7.3 Conclusion

Ces deux analyses graphiques viennent renforcer la conclusion précédente : le modèle `RandomF_SVR_LGBM` est le plus performant et le plus fiable dans le cadre de ce projet.

Il présente non seulement d'excellents résultats numériques ($RMSE$, R^2 , MAE), mais également un comportement cohérent, stable et généralisable, tel que démontré par la courbe d'apprentissage et la distribution des résidus.

Ce modèle sera donc retenu comme solution finale recommandée pour la prédiction de la teneur en huile résiduelle des *cuttings* dans un contexte industriel.

Chapitre 8

Optimisation inverse des réglages Verti-G

8.1 Formulation du problème inverse

L'objectif principal de cette dernière étape est d'exploiter le modèle prédictif entraîné afin de déterminer, de manière inverse, les réglages optimaux de l'unité **Verti-G** permettant de minimiser la teneur en huile résiduelle (*OOC* – *Oil on Cuttings*) pour un mélange donné de boues et de *cuttings*.

Contrairement à l'usage classique d'un modèle de régression, où l'on prédit une sortie à partir d'entrées connues, il s'agit ici de trouver les meilleures entrées (réglages machine) pour atteindre une sortie cible minimale. Cela constitue un **problème d'optimisation inverse**, dans lequel l'objectif est d'inverser le modèle prédictif pour guider les décisions opérationnelles.

Ce cadre inverse est particulièrement pertinent dans un contexte industriel, où les paramètres des boues sont imposés par les conditions de forage, et où les seuls leviers d'action opérationnels résident dans le réglage des équipements de traitement. La problématique se formule alors comme suit :

- **Objectif** : minimiser la teneur en huile résiduelle (% OOC) prédite par le modèle de stacking basé sur RandomForest, SVR et LightGBM.
- **Paramètres fixés** : caractéristiques physico-chimiques du mélange de boues et de cuttings (densité, teneur en huile, concentration en eau, granulométrie, additifs chimiques, etc.). Ces paramètres sont considérés comme non modifiables car dépendants des conditions de forage.
- **Paramètres à optimiser** : réglages opérationnels de la machine Verti-G :
 - Vitesse de rotation (rpm)
 - Débit d'alimentation (L/min)
 - Température de fonctionnement (°C)
 - Temps de séjour (s)
 - Type de tamis utilisé (catégoriel)
 - Fréquence de vibration (Hz)

L'ensemble de ces variables constitue l'espace de recherche de l'optimisation. Le problème peut donc être exprimé comme suit :

$$\min_{\theta \in \Theta} \hat{f}(x_{\text{fixe}}, \theta)$$

où :

- $\hat{f}$ est la fonction prédictive apprise par le modèle de stacking (RandomForest + SVR + LGBM) ;
- x_{fixe} est le vecteur des caractéristiques fixes du mélange (boues + *cuttings*) ;
- θ représente les réglages de la machine Verti-G à optimiser.

Dans la suite, ce problème inverse sera résolu à l’aide de la méthode **Optuna**, déjà utilisée pour l’optimisation des modèles, afin de tirer profit de son efficacité dans l’exploration d’espaces de recherche multidimensionnels et mixtes (variables numériques et catégorielles).

8.1.1 Méthodologie d’optimisation

L’approche retenue pour résoudre le problème inverse formulé précédemment repose sur l’utilisation du modèle de stacking préalablement entraîné, combinant les prédictions de trois régressions optimisées (RandomForest, SVR, LGBM), intégré dans un pipeline de traitement des données numériques et catégorielles. Ce modèle sert de simulateur numérique capable de prédire la teneur en huile résiduelle (OOC) à partir d’une combinaison donnée de paramètres de boues et de réglages machine.

Pour identifier la configuration optimale des réglages Verti-G, une fonction objectif a été définie. Celle-ci prend en entrée un vecteur de réglages candidats (valeurs continues et catégoriques), auquel s’ajoutent les caractéristiques fixes d’un mélange réel de boues et de cuttings. Le modèle prédictif évalue ensuite la teneur en huile correspondante. Cette valeur, que l’on cherche à minimiser, constitue la fonction à optimiser.

8.1.1.1 Fonction objectif

La fonction objectif suit la structure suivante :

1. Fusion des variables fixes (issues d’un mélange réel) avec les paramètres variables à tester.
2. Préparation des données sous forme d’un **DataFrame** adapté au pipeline du modèle.
3. Prédiction du *OOC* à l’aide du modèle empilé (*stacking*).
4. Renvoi de la prédiction comme valeur de la fonction à minimiser.

Les paramètres optimisés sont les suivants :

- **Vitesse de rotation (rpm)** : variable continue entre 800 et 2500
- **Débit d’alimentation (L/min)** : de 100 à 500
- **Température de fonctionnement (°C)** : de 30 à 90
- **Temps de séjour (s)** : de 10 à 120
- **Type de tamis utilisé** : variable catégorielle parmi {Mesh 80, Mesh 100, Mesh 120}
- **Fréquence de vibration (Hz)** : de 10 à 80

8.1.1.2 Optimisation avec Optuna

La bibliothèque Optuna a été utilisée pour explorer l'espace des réglages à l'aide de son algorithme TPE (*Tree-structured Parzen Estimator*). L'optimisation a été conduite sur 700 itérations, chaque essai consistant à :

- Générer aléatoirement une combinaison plausible de réglages selon les distributions définies
- L'évaluer avec la fonction objectif (i.e., prédiction du modèle *stacking*)
- Retenir les combinaisons minimisant la valeur de sortie (OOC prédite)

Cette approche permet de gérer efficacement :

- Des paramètres continus et discrets
- Une fonction objectif non analytique (boîte noire)
- Et des contraintes industrielles réalistes (plages opérationnelles)

Le résultat final de cette procédure est la configuration optimale des réglages **Verti-G** pour un mélange donné, c'est-à-dire celle qui permet, selon le modèle, de réduire au minimum la teneur en huile résiduelle sur les *cuttings*. Cette configuration est présentée et analysée dans la section suivante.

8.2 Cas pratique : exemple sur un mélange réel

Afin d'illustrer concrètement l'approche d'optimisation inverse développée, un cas pratique a été réalisé à partir d'un mélange réel de boues et de cuttings issu de la base de données industrielle (ligne 59). Les caractéristiques physico-chimiques de ce mélange ont été fixées (densité, pourcentage d'huile, composition en eau et en additifs, granulométrie, etc.), représentant ainsi les conditions initiales imposées en contexte de forage.

Ces caractéristiques ont ensuite été introduites dans la fonction objectif d'Optuna, et l'optimisation a été lancée pour déterminer la configuration optimale des réglages Verti-G permettant de minimiser la teneur en huile résiduelle (OOC) à la sortie.

8.2.1 Caractéristiques du mélange testé

Le mélange sélectionné (ligne 59 de la base de données) présente les propriétés suivantes :

- **Densité (SG)** : 1,25
- **Pourcentage d'huile** : 79,75 %
- **Concentration d'huile** : 0,5959
- **Concentration d'eau** : 0,1476
- **Weighting Material** : 509,88
- **Sel** : 54,20
- **Chaux** : 28,00
- **VersaCoat** : 11,35

- **VersaWet** : 5,77
- **VG 69** : 11,83
- **Teneur en boue dans les cuttings (%)** : 26,15
- **Granulométrie des cuttings (μm)** : 577,06

Ces valeurs ont été maintenues constantes pendant toute la procédure d'optimisation.

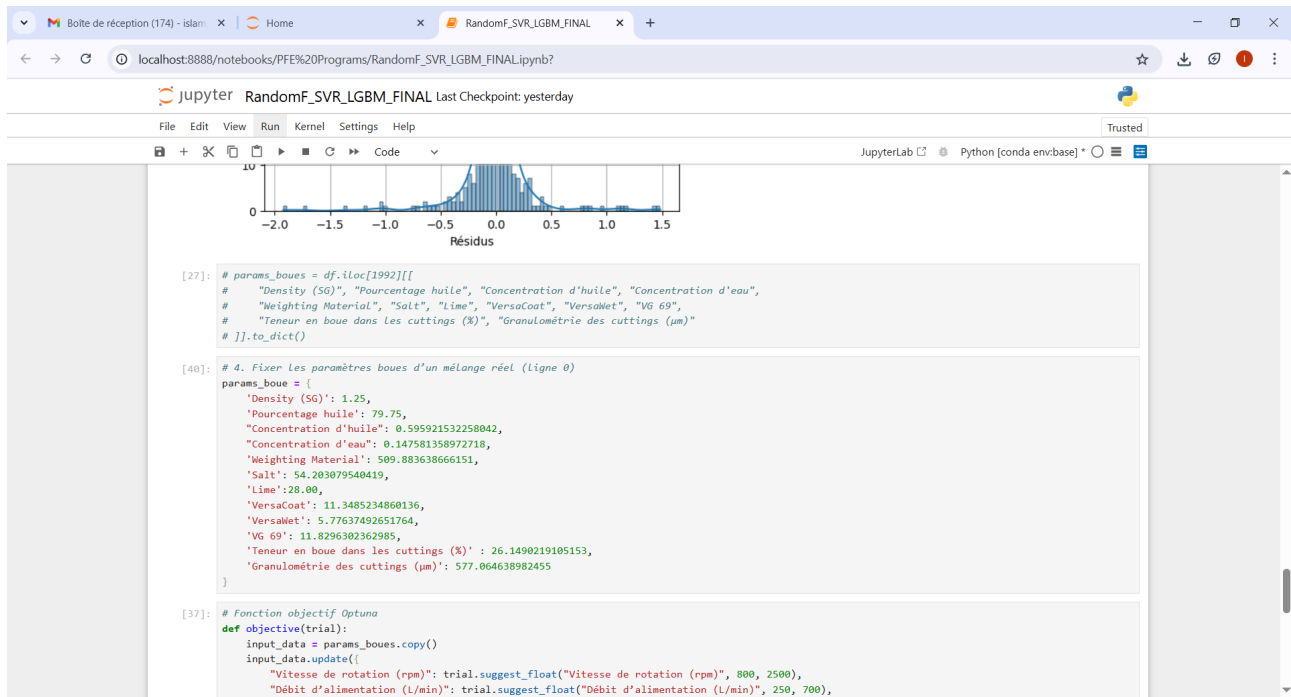


FIGURE 8.1 – Définition des paramètres d'entrée du mélange réel (ligne 59 de la base de données)

8.2.2 Réglages Verti-G optimaux proposés

À l'issue de l'optimisation, les meilleurs réglages **Verti-G** recommandés par **Optuna** sont les suivants :

- **Vitesse de rotation** : 2498,12 rpm
- **Débit d'alimentation** : 699,77 L/min
- **Température de fonctionnement** : 51,14 °C
- **Temps de séjour** : 283,87 s
- **Type de tamis utilisé** : Mesh 80
- **Fréquence de vibration** : 36,69 Hz

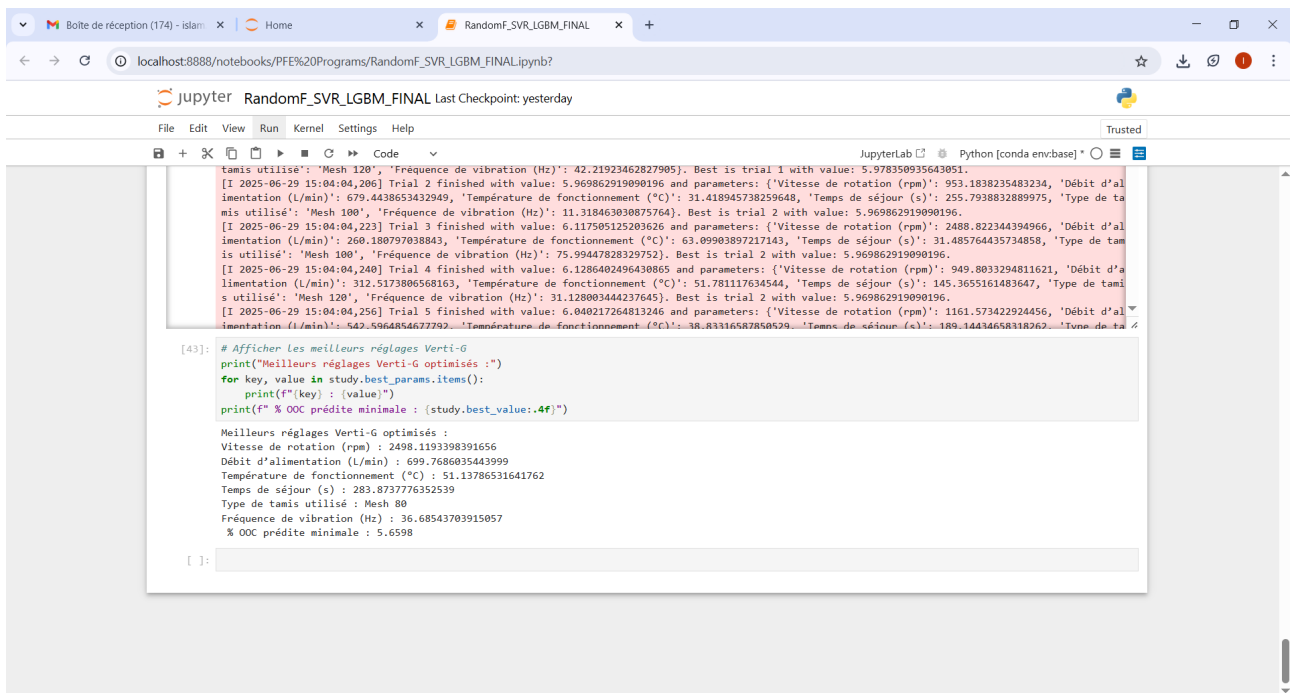


FIGURE 8.2 – Réglages Verti-G optimaux proposés par Optuna et prédiction de la teneur en huile

8.2.3 Résultat obtenu

Avec cette configuration, la teneur en huile résiduelle prédite par notre modèle est de :

5,6598 % OOC

Ce résultat a été obtenu à partir des caractéristiques du mélange réel (ligne 59), pour lequel la teneur en huile mesurée expérimentalement était de 7,41 %.

On observe que la teneur prédite par le modèle est inférieure à la valeur réelle mesurée. Cela indique que l'optimisation automatique a permis d'identifier une configuration de paramètres plus performante, capable de réduire davantage la teneur en huile résiduelle.

Ainsi, le modèle ne se contente pas de reproduire les réglages existants, mais propose des réglages optimaux conduisant à une meilleure efficacité de séparation. Cette capacité à recommander des conditions améliorées constitue un atout majeur pour l'optimisation du procédé de traitement des boues.

8.3 Discussion sur la faisabilité opérationnelle des réglages

Les résultats obtenus à l'issue de l'optimisation inverse démontrent le potentiel concret de l'intelligence artificielle dans l'assistance au réglage des équipements industriels, en l'occurrence la centrifugeuse Verti-G. Cependant, la pertinence d'une recommandation issue d'un modèle prédictif ne repose pas uniquement sur sa précision statistique. Elle doit également être évaluée

à l'aune de sa faisabilité technique, de sa robustesse en conditions réelles et de son acceptabilité par les opérateurs de terrain.

8.3.1 Compatibilité technique

Les réglages proposés par Optuna se situent tous dans des plages de fonctionnement réalistes pour les équipements Verti-G :

- La **vitesse de rotation** (2498 rpm) et le **débit d'alimentation** (700 L/min) sont cohérents avec les capacités industrielles usuelles.
- Le **temps de séjour** relativement long (284 s) correspond à une stratégie de séparation plus lente et plus efficace.
- L'utilisation d'un **tamis de type Mesh 80**, moins fin que Mesh 120, assure un bon compromis entre performance de séparation et limitation du risque de colmatage.

8.3.2 Contraintes opérationnelles

Certains réglages, bien qu'optimaux selon le modèle, peuvent entraîner des contraintes techniques ou économiques :

- Un **débit élevé** (près de 700 L/min) peut solliciter davantage les pompes et engendrer une consommation énergétique accrue.
- La **fréquence de vibration spécifique** (36,69 Hz) requiert une surveillance attentive des composants mécaniques et de la calibration.
- Le **temps de séjour prolongé** peut ralentir le débit global de traitement, ce qui impose un équilibre entre productivité et qualité de séparation.

8.3.3 Vers une intégration terrain assistée

L'implémentation industrielle de cette approche ne suppose pas un remplacement des décisions humaines, mais plutôt leur assistance. Un tel système pourrait être intégré dans une interface logicielle embarquée, capable de :

- Recommander en temps réel les réglages optimaux en fonction des boues détectées
- Alerter l'opérateur si des valeurs sortent des tolérances mécaniques
- Apprendre de l'historique des performances et s'ajuster dynamiquement

Cette perspective ouvre la voie à une optimisation continue du traitement, avec un double bénéfice :

- **Réduction de la teneur en huile** rejetée dans l'environnement
- **Amélioration de la performance énergétique et économique** du traitement

Conclusion générale

À l'issue de ce travail de fin d'études, il apparaît que la gestion des déblais de forage imprégnés de fluides huileux constitue un enjeu critique pour l'industrie pétrolière, tant du point de vue opérationnel qu'environnemental. L'étude menée a permis de mieux comprendre les interactions entre les paramètres du procédé Verti-G et la performance de traitement mesurée par la teneur en huile résiduelle sur les cuttings. En analysant un ensemble de données industrielles fournies par ADFC, il a été possible d'identifier les variables les plus influentes, de les exploiter à travers des algorithmes d'apprentissage automatique et de proposer un modèle prédictif robuste basé sur une approche d'assemblage de modèles.

Les résultats obtenus montrent qu'il est possible d'estimer avec une grande précision la performance du système Verti-G à partir d'un nombre restreint de variables opérationnelles. De plus, la phase d'optimisation inverse du modèle, rendue possible grâce à des techniques de recherche automatique, a permis de proposer des combinaisons de réglages permettant de réduire efficacement la teneur en huile résiduelle. Cette approche ouvre ainsi la voie à un pilotage plus intelligent et plus efficace du procédé, contribuant à une meilleure valorisation des déchets de forage et à la réduction de leur impact environnemental.

Ce projet constitue une illustration concrète de l'apport de l'intelligence artificielle dans le domaine du génie des procédés et du traitement des déchets industriels. Il démontre qu'une approche fondée sur les données permet non seulement de modéliser des systèmes complexes, mais également de les optimiser dans une perspective de durabilité. Il serait pertinent d'envisager à l'avenir l'intégration de capteurs en temps réel pour automatiser l'ajustement des paramètres du Verti-G, ainsi que l'extension de cette méthodologie à d'autres unités de traitement, comme le TCC ou le TPS. La création d'interfaces interactives à destination des opérateurs de terrain pourrait également renforcer l'implémentation industrielle de ces solutions.

Au-delà des résultats techniques, ce travail a représenté une expérience formatrice, en mobilisant des compétences pluridisciplinaires allant de l'analyse environnementale à la modélisation mathématique, en passant par la science des données. Il s'inscrit dans une dynamique de progrès technologique responsable et constitue une contribution concrète à la modernisation durable du secteur pétrolier.

Bibliographie

- [1] R. Akkal. *Optimisation des fluides de forage et endommagement de la formation*. Thèse de doctorat, École Nationale Polytechnique, Département de Génie Minier, 2014.
- [2] National Research Council. *Drilling and Excavation Technologies for the Future*. National Academies Press, 1994.
- [3] N. J. Hyne. *Nontechnical Guide to Petroleum Geology, Exploration, Drilling, and Production*. PennWell Books, Tulsa, 3rd edition, 2012.
- [4] Somaa Ismail Hafsi Takieddine. Appareil de forage, etude et dimensionnement cas : Champs hassi messaoud, 2011. 2011/2012.
- [5] Dyhia Oumessaoud. Élaboration et caractérisation d'un matériau composite en carbure de tungstène-alliage utilisé dans la fabrication des outils de forage. Master's thesis, Université Mouloud Mammeri, 2017. Année universitaire 2017/2018.
- [6] Oussama Hebbal, Azzedine Touaghzeit, and Med Abd Harkati. Étude et maintenance de treuil de forage oil well 840e. Master's thesis, Université Hassiba Benbouali de Chlef, 2012.
- [7] M. R. Alexandre. *Études rhéophysiques de systèmes d'émulsions inverses / argile organophile. Applications aux boues de forage pétroliers*. PhD thesis, Université Pierre et Marie Curie, 2007. Soutenue le 17 septembre 2007.
- [8] A.S.S.A.S.S. Committee et al. *Drilling Fluids Processing Handbook*. Elsevier, 2011.
- [9] M. Khodja, M. Khodja-Saber, J. P. Canselier, N. Cohaut, and F. Bergaya. Drilling fluid technology : Performances and environmental considerations. In Igor Fuerstner, editor, *Products and Services, from R&D to Final Solutions*. InTech, 2010.
- [10] R. Caenn and G. V. Chillingar. Drilling fluids : State of the art. *Journal of Petroleum Science and Engineering*, 14(3-4) :221–230, 1996.
- [11] A. S. Apaleke, A. Al-Majed, and M. E. Hossain. Drilling fluid : State of the art and future trend. In *SPE North Africa Technical Conference and Exhibition*, page SPE-149555, 2012.
- [12] John C. Reis. *Environmental Control in Petroleum Engineering*. Elsevier, 1996.
- [13] C. G. Ryen and V. Chillingar. Drilling fluids : State of the art. *Journal of Petroleum Science and Engineering*, 14 :221–230, 1996.
- [14] P. Reid, R. Minton, and A. Twynam. Field evaluation of a novel inhibitive water-based drilling fluid for tertiary shales. In *SPE EUROPEC Featured at EAGE Conference and Exhibition*, pages SPE-24979, 1992.
- [15] ASME Shale Shaker Committee. *Drilling Fluids Processing Handbook*. Elsevier Science & Technology, Amsterdam, 1 edition, 2004.
- [16] R. Ali and I. Ghiniwa. Drill cuttings as an economic tool to understand the petroleum system for exploratory and development wells. In *International Symposium of the Society of Core Analysts*, page 2, Snowmass, Colorado, USA. Présenté du 21 au 26 août 2016.
- [17] B. Niall, Simon J. Cripps, J. Troels, K. Grethe, and Gordon B. Picken. Review of drill cuttings piles in the north sea. final report, 1998. Report for the Offshore Decommissioning Communications Project, pp. 4–5.

- [18] M. Stuckman, C. Lopano, S. Berry, and J. Hakala. Geochemical solid characterization of drill cuttings, core and drilling mud from marcellus shale energy development. *Journal of Natural Gas Science and Engineering*, 2019.
- [19] I. Yousef and V. Morozov. Exploring the potential of drill cuttings for reservoir characterization : A case study from the volga-ural basin, russia. *Petroleum Research*, 2024.
- [20] B. Bageri, M. Benaafi, M. Mahmoud, A. Mohamed, S. Patil, and S. Elkatatny. Effect of formation cutting’s mechanical properties on drilling fluid properties during drilling operations. *Arabian Journal for Science and Engineering*, 2020.
- [21] R. Kadyrov, D. Nurgaliev, E. Saenger, M. Balcewicz, R. Minebaev, E. Statsenko, M. Glukhov, A. Nizamova, and B. Galiullin. Digital rock physics : Defining the reservoir properties on drill cuttings. *Journal of Petroleum Science and Engineering*, 2021.
- [22] M. Huque, M. Rahman, S. Zendehboudi, S. Butt, and S. Imtiaz. Experimental and numerical study of cuttings transport in inclined drilling operations. *Journal of Petroleum Science and Engineering*, 208, 2022.
- [23] B. Pang, S. Wang, Q. Wang, K. Yang, H. Lu, M. Hassan, and X. Jiang. Numerical prediction of cuttings transport behavior in well drilling using kinetic theory of granular flow. *Journal of Petroleum Science and Engineering*, 161, 2018.
- [24] M. Storchak, T. Stehlé, and H. Möhring. Numerical modeling of cutting characteristics during short hole drilling : Modeling of kinetic characteristics. *Journal of Manufacturing and Materials Processing*, 2023.
- [25] A. Montes, S. Callerio, Ç. Turhan, A. Safarov, P. Ashok, and E. Van Oort. Automatic determination of cuttings and cavings properties for hole cleaning and wellbore stability assessment using a laser-based sensor. *SPE Journal*, 2024.
- [26] E. Epelle and D. Gerogiorgis. Drill cuttings transport and deposition in complex annular geometries of deviated oil and gas wells : A multiphase flow analysis of positional variability. *Chemical Engineering Research and Design*, 2019.
- [27] UKOOA Drill Cuttings Initiative. Research and development programme project 1.2 : Cuttings pile toxicity. Technical report, Environment & Resource Technology Ltd, RF – Rogaland Research, 2000. Projet technique.
- [28] Y. Yaiche. Environmental impact assessment of the drilling activities in the hassi messaoud field. Master’s thesis, Faculty of Design and Technology, School of Engineering, The Robert Gordon University, Aberdeen, 2006.
- [29] Dossier technique et réglementaire sur le traitement des bourbiers de forage et de workover. Technical report, Non spécifié, 2011. Rapport technique.
- [30] J. Reis. *Environmental Control in Petroleum Industry*. Gulf Publishing Company, Houston, Texas, 1996.
- [31] W. W. Crawley and R. T. Branch. Characterization of treatment zone soil conditions at a commercial non-hazardous oilfield waste land treatment unit. In *Premier colloque international sur l’exploration et la production de l’hydrocarbure et du gaz*, New Orleans, 1990. Environmental Protection Agency.
- [32] J. C. Means, S. G. Ward, J. J. Hassett, and W. L. Banwart. Sorption of polynuclear aromatic hydrocarbons by sediment and soil. *Science and Technology*, 1980.
- [33] P. Moszkowicz. Solidification et stabilisation des déchets. Technical report, Laboratoire d’analyse environnementale des procédés et des systèmes industriels, Institut National des Sciences Appliquées de Lyon.
- [34] R. Scriban. *Biotechnologie*. Tec. et Doc, Lavoisier, Paris, France, 5 edition, 1999.
- [35] Étude d’impact environnemental et social des activités de forage du azs-2 gassi chergui, algérie. Technical report, Non spécifié, 2007.

- [36] Z. Huang, Z. Xu, Y. Quan, H. Jia, J. Li, Q. Li, Z. Chen, and K. Pu. A review of treatment methods for oil-based drill cuttings. In *2nd International Symposium on Resource Exploration and Environmental Science*, volume 170, page 022074. IOP Publishing, 2018.
- [37] M. E. Hossain and A. A. Al-Majed. *Fundamentals of Sustainable Drilling Engineering*. John Wiley & Sons, Hoboken, 2015.
- [38] Chevron Texaco and Marathon. Drilling waste management technology descriptions. Technical report, Argonne National Laboratory, Washington, USA, September 2008.
- [39] F. Damou. The adaptability study for the use of thermal desorption and solidification/stabilization processes for the treatment of drill cuttings in hassi messaoud field. Master’s thesis, Faculty of Design and Technology, School of Engineering, The Robert Gordon University, Aberdeen, 2007.
- [40] Manan Sharma and Russell Reynnells. Importance of soil amendments : survival of bacterial pathogens in manure and compost used as organic fertilizers. In *Preharvest Food Safety*, pages 159–175. 2018.
- [41] Romina L. Dias, David Juck, Lyle G. Whyte, and Charles W. Greer. Hydrocarbon removal and bacterial community structure in on-site biostimulated biopile systems designed for bioremediation of diesel contaminated antarctic soil. *Polar Biology*, 38 :677–687, 2015.
- [42] Ajay Singh, Brajesh Singh, and Owen Ward. Potential applications of bioprocess technology in petroleum industry. *Biodegradation*, 23 :865–880, 2012.
- [43] Borislava Lukić, Milica Stojanović, Jelena Radulović, Zorica Lopičić, Obrad Petrović, and Lidija Đokić. A review on the efficiency of landfarming integrated with composting as a soil remediation treatment. *Environmental Technology Reviews*, 6(1) :94–116, 2017.
- [44] Ashish Kumar. *Learning Predictive Analytics with Python*. Packt Publishing, Birmingham, 1 edition, 2016.
- [45] Youssef Benzaki. Données manquantes en data science : méthodes et stratégies. <https://mrmint.fr/donnees-manquantes-data-science>, 2018. MRMint.fr, 24 octobre 2018.
- [46] GeeksforGeeks. How to create a seaborn correlation heatmap in python, March 2025. GeeksforGeeks.org, 15 mars 2025.
- [47] A. A. Awan. Introduction to shap values for machine learning interpretability, October 2024. DataCamp.com, 4 octobre 2024.
- [48] scikit-learn. Machine learning algorithm cheat-sheet, n.d. Consulté en juin 2025.
- [49] Davide Chicco, Matthijs J. Warrens, and Giuseppe Jurman. The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation. *PeerJ Computer Science*, 7 :e623, 2021.
- [50] Michael Bowles. *Machine Learning in Python : Essential Techniques for Predictive Analysis*. John Wiley & Sons, Incorporated, Somerset, 1 edition, 2015.
- [51] D. Chopra and R. Khurana. *Introduction to Machine Learning with Python*. Bentham Science Publishers, 1 edition, 2023. Disponible sur : ProQuest Ebook Central.
- [52] Alberto Boschetti and Luca Massaron. *Python Data Science Essentials : A Practitioner’s Guide Covering Essential Data Science Principles, Tools, and Techniques*. Packt Publishing, Limited, Birmingham, 3 edition, 2018.
- [53] S. Nayak, editor. *Fundamentals of Optimization Techniques with Algorithms*. Academic Press, 2020.
- [54] J. Mao, J. Chen, and Y. Deng. Optimization strategy for ensemble learning models based on fusing resampling, adaptive dimensionality reduction, and optuna in intelligent flight technology evaluation. *Aerospace Science and Technology*, 162 :110251, 2025. Article n° 110251.
- [55] C. Pinichka, S. Chotpantarat, K. H. Cho, and W. Siri Wong. Comparative analysis of swat and swat coupled with xgboost model using optuna hyperparameter optimization

- for nutrient simulation : A case study in the upper nan river basin, thailand. *Journal of Environmental Management*, 388 :126053, 2025. Article n° 126053.
- [56] George Kyriakides and K. G. Margaritis. *Hands-On Ensemble Learning with Python : Build Highly Optimized Ensemble Machine Learning Models Using Scikit-Learn and Keras*. Packt Publishing, Birmingham, 1 edition, 2019.
- [57] Pierre-Henri Juillard. Théorie : apprentissages d'ensemble – bagging, boosting et stacking, March 2022. Medium.com, 31 mars 2022.