

REPUBLIC ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR ET RECHERCHE SCIENTIFIQUE
UNIVERSITE DE M'HAMED BOUGARA-BOUMERDES



**Mémoire de fin d'études
En vue de l'obtention du diplôme :**

MASTER

Présenté par :

MEGHERBI Zakaria

LAOUFI Hichem

Filière : Hydrocarbures

Option : Génie pétrolier : Production des hydrocarbures

Thème :

**Développement d'un outil d'aide à la prise de décision lors de
la complétion des puits d'hydrocarbures**

Devant le jury :

- | | | | |
|----------------------|------------|------|-----------|
| - BENYOUNES Khaled | Professeur | UMBB | Président |
| - ZERAIBI Noureddine | Professeur | UMBB | Encadrant |
| - GARECHE Mourad | Professeur | UMBB | Examineur |
| - SAIFI Redha | MAA | UMBB | Examineur |

Année universitaire : 2021/2022

Dédicace

“

A mes chers parents

Boudor Bouamama Fafa et Megherbi Abdelsalem

Pour tous leurs sacrifices, leur amour, leur tendresse, leur soutien et leurs prières tout au long de mes études

*A mon **grand frère** Abdenour pour son encouragement permanent, et son soutien moral*

*A ma **chère sœur** Farah pour ses prières*

*A mon binôme **Laoufi Hichem***

A toute ma famille pour leur soutien tout au long de mon parcours universitaire, Que ce travail soit l'accomplissement de vos vœux tant allégués, et le fruit de votre soutien infaillible

Merci d'être toujours là pour moi

”

-Zakaria-

Dédicace

“

*A ma **mère**, ma réussite, mon amour, mon soutien,
A mon **père**, mon exemple de la vie et ma source d'inspiration,
A mes adorables **sœurs**, mon exemple de persévérance, de courage et de générosité,
Mon cher **frère**, pour son encouragement indéfectible
Et a mes chers **amis**, qui j'ai partagé les meilleurs moments de ma vie d'étude
A mon binôme **Megherbi Zakaria***

Merci, tous ceux qui m'aime et que et que j'aime et je respecte

”

-Hichem -

Remerciements

Nous tenons tout d'abord à remercier Dieu le tout puissant et miséricordieux, qui nous a donné la force et la patience d'accomplir ce modeste travail.

Nous tenons à remercier en premier degré notre encadrant **Professeur ZERAIBI Nouredine**, pour l'orientation, la confiance, la patience qui ont constitué un apport considérable sans lequel ce travail n'aurait pas pu être mené au bon port. Qu'il trouve dans ce travail un hommage vivant à sa haute personnalité.

Nous tenons à remercier également **M. MERZOUG Ahmed** pour son immense aide, la qualité de son suivie ainsi que pour tous les conseils et les informations qu'il nous a prodigués avec un degré de patience. Nous voudrions lui témoigner notre gratitude pour son soutien qui nous a été précieux afin de mener notre travail à bon port.

Nous tenons aussi à adresser mes plus sincères remerciements à l'équipe **SMARTEST** incluant **M. AOUNAOUK Lotfi** pour l'aide compétente qu'il nous a apportée, pour sa patience et son encouragement. Son regard critique nous a été très précieux pour structurer le travail et pour améliorer la qualité des différentes sections.

Que les membres de jury trouvent, ici, l'expression de nos sincères remerciements pour l'honneur qu'ils nous font en prenant le temps de lire et d'évaluer ce travail. Nos remerciements s'étendent également à tous nos enseignants durant les années d'études.

Pour finir, nous tenons à remercier tous ceux qui, de près ou de loin, ont contribué à la réalisation de ce travail.

Nomenclature

| | |
|--------------------------------|--|
| DD : | <i>Drawdown (Différence de pression entre la pression du réservoir et le fond du puit)</i> |
| WC : | <i>Water Cut (Saturation d'eau)</i> |
| σ_v : | <i>Contrainte verticale</i> |
| σ_H : | <i>Contrainte horizontale maximale</i> |
| σ_h : | <i>Contrainte horizontale minimale</i> |
| σ_r : | <i>Contrainte radiale</i> |
| σ_t : | <i>Contrainte tangentielle</i> |
| σ_a : | <i>Contrainte axiale</i> |
| Pp : | <i>Pression des pores</i> |
| UCS : | <i>Résistance à la compression uniaxiale</i> |
| τ : | <i>Résistance au cisaillement</i> |
| σ_n : | <i>Contrainte normale au plan de rupture</i> |
| C : | <i>Force de cohésion</i> |
| Θ : | <i>L'angle de friction</i> |
| SC : | <i>Sorting coefficient (coefficient de tri)</i> |
| UC : | <i>Uniformity coefficient (coefficient d'uniformité)</i> |
| GP : | <i>Gravel Pack</i> |
| SAS : | <i>Standalone Screens</i> |
| ESS : | <i>Expandable Screens</i> |
| WWS : | <i>Wire-Wrapped Screens</i> |
| CHGP : | <i>Cased Hole Gravel Pack</i> |
| OHGP : | <i>Open Hole Gravel Pack</i> |
| FP : | <i>Frac Pack</i> |
| HRWP : | <i>High Rate Water Pack</i> |

Nomenclature

- PSD :** *Particle size distribution*
- SRT :** *Sand retention testing (Essai de rétention de sable)*
- CDP :** *Pression de drawdown critique*
- MDP :** *Pression de drawdown minimale*
- CEDP :** *Drawdown critique d'érosion*
- NPV :** *Net present value (Valeur actuelle nette)*
- NTG :** *Net to gross ratio (Proportion du gré dans le réservoir)*
- BHT :** *Température de fond du puit*
- BHP :** *Pression de fond du puit*
- Kh :** *Perméabilité *Épaisseur de zone productrice*
- PI :** *Indice de Productivité*

Résumé

La production de sable est l'un des principaux problèmes de l'industrie pétrolière et gazière, plus précisément l'aspect amont. Elle affecte gravement les aspects techniques et économiques en diminuant la récupération des hydrocarbures et en augmentant les coûts opérationnels. Heureusement, il existe différentes techniques pour traiter ce phénomène mais le défi reste tout le temps d'actualité et ce pour comprendre les critères de sélection de la méthode de contrôle du sable appropriée afin de remédier à ce problème.

L'objectif de cette thèse est de développer une approche de classification pour prédire la méthode optimale de contrôle du sable en utilisant des algorithmes d'apprentissage automatique. Quatre modèles différents ont été utilisés, à savoir :

1. K-Nearest Neighbors.
2. Support Vector Machine.
3. Random Forest.
4. Arbre de Decision.

Cette étude a été menée sur une base de données appartenant au Dataset internationale regroupant les données de plus de 1900 puits, utilisée à l'origine pour comparer la fiabilité de différentes formes de contrôle du sable.

Le long de ce mémoire, nous allons proposer une démarche fiable en utilisant les techniques sus citées, l'analyse et le traitement des données ont montré que jusqu'à un certain niveau, ces méthodes peuvent nous fournir des solutions optimales assurant une complétion adéquate pour le puit candidat. De ce fait, cette approche s'avère une approche potentielle qui s'élève au niveau d'un outil d'aide à la prise de décision. Et par conséquent ces techniques développées peuvent venir en aide auprès des ingénieurs et les superviseurs en charge de l'étape de choix des complétions.

Mots clés : Production du sable, Méthodes de contrôle du sable, L'intelligence artificielle, Apprentissage automatique, Conception de complétion

Abstract

Sand production is one of the main problems of the oil and gas industry, more precisely the upstream aspect. It severely affects the technical and economic aspects by decreasing hydrocarbon recovery and increasing operational costs. Fortunately, there are different techniques to deal with this phenomenon but the challenge remains to understand the criteria for selecting the appropriate sand control method to remedy this problem.

The objective of this thesis is to develop a classification approach to predict the optimal sand control method using machine learning algorithms. Four different models were used, namely :

1. K-Nearest Neighbors.
2. Support Vector Machine.
3. Random Forest.
4. Decision Tree.

This study was conducted on a database belonging to the International Dataset gathering data from more than 1900 wells, originally used to compare the reliability of different forms of sand control.

Along this thesis, we will propose a reliable approach using the above mentioned techniques, the analysis and processing of the data showed that up to a certain level, these methods can provide us with optimal solutions ensuring an adequate completion for the candidate well. Thus, this approach is a potential approach that rises to the level of a decision support tool. Therefore, these techniques developed can help engineers and supervisors in charge of the completion selection stage.

Keywords: Sand production, Sand control methods, Artificial intelligence, Machine learning, Completion design

ملخص

إن إنتاج الرمال يشكل واحدة من المشاكل الرئيسية التي تواجه صناعة النفط والغاز، وعلى وجه التحديد مشكلة المنبع ويؤثر بشدة على الجوانب التقنية والاقتصادية من خلال تقليل إستعادة المحروقات وزيادة التكاليف التشغيلية. لحسن الحظ، هناك تقنيات مختلفة للتعامل مع هذه الظاهرة، ولكن يبقى التحدي هو فهم معايير إختيار التقنية الأنسب للتحكم في الرمال من أجل علاج هذه المشكلة.

الهدف من هذه المذكرة هو وضع نهج تصنيف للتنبؤ بالطريقة المثلى للتحكم في الرمال باستخدام خوارزميات التعلم الآلي. استخدمت أربعة نماذج مختلفة، وهي:

K-Nearest Neighbors 1

Support Vector Machine 2

Decision Tree 3

Random Forest 4

وأجريت هذه الدراسة على قاعدة بيانات تابعة لمجموعة البيانات الدولية التي تجمع بيانات من أكثر من 1900 بئر، استخدمت أصلاً لمقارنة موثوقية مختلف أشكال مراقبة الرمال.

وفي هذا الإطار، سنقترح نهجا موثوقا به يستخدم التقنيات المذكورة أعلاه، وقد أظهر تحليل البيانات ومعالجتها أن هذه الأساليب يمكن أن توفر لنا، إلى مستوى معين، حلاً مثلى تضمن اكتمال كاف للبئر المرشح. وبالتالي، فإن هذا النهج نهج محتمل يرقى إلى مستوى أداة دعم القرارات. وبالتالي، فإن هذه التقنيات المطورة يمكن أن تساعد المهندسين والمشرفين المسؤولين عن مرحلة إختيار الإنجاز.

الكلمات المفتاحية: إنتاج الرمال ، طرق التحكم في الرمال ، الذكاء الاصطناعي ، التعلم الآلي ، تصميم البئر

Table des matières

| | |
|---|-----------|
| Dédicace..... | I |
| Dédicace..... | II |
| Remerciements | III |
| Nomenclature..... | IV |
| Résumé | VI |
| Abstract..... | VII |
| ملخص..... | VIII |
| Liste des figures | XII |
| Liste des tableaux | XV |
| Introduction générale..... | 1 |
| Chapitre 1 : Phénomène de production du sable..... | 4 |
| 1.1) Nature, causes et conséquences de la production du sable | 4 |
| 1.1.1) La nature de production de sable | 4 |
| 1.1.2) Effets de la production de sable | 5 |
| 1.1.3) Causes de la production de sable | 6 |
| 1.2) Stabilité du puits et critères de rupture du sable | 10 |
| 1.2.1) Stabilité du puits | 10 |
| 1.2.2) Critères de rupture du sable | 11 |
| Chapitre 2 : Les différentes méthodes de contrôle du sable | 14 |
| 2.1) Méthode de contrôle du débit | 14 |
| 2.2) Perforation | 15 |
| 2.2.1) Perforation sélective | 16 |
| 2.2.2) Perforation orientée | 17 |
| 2.3) Screens | 17 |
| 2.3.1) Slotted Liner | 18 |
| 2.3.2) Wire-Wrapped Screens | 19 |
| 2.3.3) Pre-packed Screens | 20 |
| 2.3.4) Premium Screens | 20 |
| 2.3.5) Expandable Screens | 21 |
| 2.4) Gravel Pack | 21 |
| 2.4.1) Open hole gravel pack | 22 |
| 2.4.2) Cased hole gravel pack | 22 |
| 2.4.3) Sélection du gravier | 23 |
| 2.5) Consolidation Chimique | 24 |
| 2.5.1) Consolidation in-situ | 24 |
| 2.5.2) Gravier enrobé de résine | 25 |
| 2.6) Frac Pack | 26 |
| 2.6.1) Critères de sélection du Frac Pack | 26 |

| | |
|--|-----------|
| 2.7) High Rate Water Pack | 26 |
| 2.8) Comparaison entre HRWP and Frac Pack | 27 |
| Chapitre 3 : Sélection des méthodes de contrôle du sable | 29 |
| 3.1) Les méthodologies utilisées dans le processus de sélection | 29 |
| 3.1.1) Approche de Tiffin, 1998 | 30 |
| 3.1.2) Approche de Price-Smith, 2003 | 31 |
| 3.1.3) Approche de Mathisen, 2007 | 31 |
| 3.1.4) Approche de Farrow, 2004 | 33 |
| 3.1.5) Approche de Michael Bryne, 2010 | 35 |
| 3.1.6) Approche de Keng Seng Chan, 2013 | 39 |
| 3.1.7) Approche d'Ehsan Khamsehchi, 2015 | 41 |
| 3.1.8) Approche de Raymond Rayes, 2018 | 43 |
| 3.1.9) Approche de Shaofeng Hu 2020 | 45 |
| 3.2) Nouvelle approche de sélection des méthodes de contrôle du sable | 48 |
| Chapitre 4 : L'intelligence artificielle..... | 49 |
| 4.1) Notion de base sur AI, DM, ML | 49 |
| 4.1.1) L'intelligence artificielle | 49 |
| 4.1.2) Data Mining (Exploration des données) | 50 |
| 4.1.3) Apprentissage automatique (ML) | 50 |
| 4.2) Étapes de modélisation de l'apprentissage automatique | 50 |
| 4.3) Prétraitement des données | 51 |
| 4.3.1) Nettoyage des données | 51 |
| 4.3.2) Imputation des valeurs manquantes | 51 |
| 4.3.3) Détection des valeurs aberrantes | 52 |
| 4.3.4) Visualisation des données | 52 |
| 4.3.5) Classement et sélection des données | 53 |
| 4.3.6) Mise à l'échelle (Normalisation ou Standardisation) | 53 |
| 4.4) Types d'apprentissage automatique | 54 |
| 4.4.1) Apprentissage supervisé | 54 |
| 4.4.2) Apprentissage non supervisé | 54 |
| 4.5) Classification et régression | 54 |
| 4.6) Concept de biais et de variance | 55 |
| 4.7) Types de modèle Machine Learning | 57 |
| 4.7.1) K-nearest neighbors | 57 |
| 4.7.2) Machine à vecteurs de support (SVM) | 58 |
| 4.7.3) Arbre de décision | 60 |
| 4.7.4) Forêt aléatoire (Random Forest) | 61 |
| 4.8) Métriques d'évaluation des modèles | 62 |
| Chapitre 5 : Prétraitement des données et développement des modèles | 64 |
| 5.1) Description de l'ensemble des données | 64 |

Table des matières

| | |
|---|-----------|
| 5.2) Prétraitement des données | 65 |
| 5.2.1) Analyse statistique | 65 |
| 5.2.2) Traitement des valeurs manquantes | 66 |
| 5.2.3) La sélection des principaux paramètres | 67 |
| 5.2.4) Traitement des valeurs aberrantes | 71 |
| 5.2.5) Normalisation des données | 73 |
| 5.3) Développement des modèles | 73 |
| 5.3.1) La création du modèle de sélection des méthodes de contrôle du sable | 74 |
| Conclusion..... | 78 |
| Bibliographie..... | 79 |
| Annexes..... | 81 |
| 1) Prédiction de production du sable | 81 |
| 1.1) Description des données | 81 |
| 1.2) Sélection des paramètres principaux | 81 |
| 1.3) Développement du modèle | 84 |

Liste des figures

| | |
|--|----|
| FIGURE 1:GEOMETRIE D'UN ARC STABLE ENTOURANT UNE PERFORATION | 5 |
| FIGURE 2: ÉROSION DE L'EQUIPEMENT DE PRODUCTION DUE A LA PRODUCTION DE SABLE :(A) EROSION DU SCREEN, (B) EROSION DE LA DUSE DE SURFACE | 6 |
| FIGURE 3: ESTIMATION DU TAUX D'ÉROSION POUR UNE TETE DE Puits, LES CARRES SONT DANS DES CONDITIONS DE GAZ SEC, ET LES DIAMANTS SONT DANS DES CONDITIONS DE GAZ/LIQUIDE | 6 |
| FIGURE 4: FENETRE DU DRAWDOWN SECURISE TRACE A UNE SEULE PROFONDEUR..... | 8 |
| FIGURE 5: CORRELATION DE TERRAIN ENTRE LE DEBIT DE PRODUCTION ET LE VOLUME DE SABLE PRODUIT . | 8 |
| FIGURE 6:PRODUCTION DE SABLE EN FONCTION DU DRAWDOWN ET DE LA SATURATION D'EAU..... | 9 |
| FIGURE 7:PRODUCTION DE SABLE CALCULEE EN FONCTION DE DIFFERENTS DROWDOWN POUR UNE SATURATION EN EAU FIXE | 9 |
| FIGURE 8:PRODUCTION DE SABLE CALCULEE EN FONCTION DE DIFFERENTES SATURATIONS POUR UN DRAWDOWN FIXE | 10 |
| FIGURE 9:ÉCHANTILLONS TYPQUES POUR LES ESSAIS UNIAXIAUX OU TRIAXIAUX | 10 |
| FIGURE 10:VARIATION DES CONTRAINTES A L'ECART DU Puits DE FORAGE..... | 11 |
| FIGURE 11:CRITERE DE DEFAILLANCE DE MOHR-COULOMB..... | 12 |
| FIGURE 12:RUPTURE PAR TRACTION ET CISAILLEMENT..... | 13 |
| FIGURE 13: FORMATION DES ARCS NATURELS (A) DETERMINATION DE LA PRODUCTION DE SABLE EN FONCTION DU DEBIT (B)..... | 15 |
| FIGURE 14:SONIC LOG POUR LA PERFORATION SELECTIVE | 16 |
| FIGURE 15:EFFET DES PERFORATIONS ORIENTEES SUR LE DRAWDOWN CRITIQUE | 17 |
| FIGURE 16:SCREEN COMPLETION | 18 |
| FIGURE 17:LES GEOMETRIES DE SLOTTED-LINERS..... | 19 |
| FIGURE 18:COMPARAISON DES SURFACES EFFECTIVES D'ENTREE (20 GAUGE)..... | 19 |
| FIGURE 19:SCHEMA DE CONSTRUCTION DE WIRE WRAPPED SCREEN..... | 20 |
| FIGURE 20:(A) SINGLE-SCREEN PRE-PACK (B) SLIM-PAK (C) DUAL-SCREEN..... | 20 |
| FIGURE 21:EXEMPLE DU PREMIUM SCREEN..... | 21 |
| FIGURE 22:LES METHODES D'EXPANSIONS D'EXPANDABLE SCREEN..... | 21 |
| FIGURE 23:TAMIS ET GRAVIER DANS TROU OUVERT /CASING PERFORE..... | 23 |
| FIGURE 24:ANALYSE GRANULOMETRIQUE DU SABLE DE FORMATION | 24 |
| FIGURE 25:LES CHANGEMENTS DE SATURATION PENDANT LA CONSOLIDATION DE LA RESINE | 25 |
| FIGURE 26:CRITERES DE SELECTION | 30 |
| FIGURE 27:RESULTATS DES ESSAIS DE RETENTION DE SABLE DE DIFFERENTS TAMIS A SABLE | 32 |
| FIGURE 28:RESULTATS DES ESSAIS DE RETENTION DE SABLE DE DIFFERENTS TAMIS A SABLE EN TERMES DE | |

| | |
|--|----|
| PERMEABILITE | 32 |
| FIGURE 29:L'ORGANIGRAMME DE SELECTION INITIAL | 34 |
| FIGURE 30:TEST DE MIGRATION DES FINES | 35 |
| FIGURE 31:REPRESENTATION DE LA MANIERE DONT LES PARTICULES FINES PEUVENT ETRE DEFINIES SUR LA BASE DE LEUR COMPORTEMENT ET DE LEUR IMPACT | 37 |
| FIGURE 32:PROCESSUS DE SELECTION POUR LE CONTROLE DU SABLE | 38 |
| FIGURE 33:SKIN ET PI DETERMINES PAR DES TESTS DE PUIITS POUR DES PUIITS AVEC DES COMPLETIONS CHFP ET OHGP DANS UN RESERVOIR MALAISIE EN EAU PROFONDE..... | 40 |
| FIGURE 34:TRACE DE LA NPV POUR UN RESERVOIR A FAIBLE IP | 41 |
| FIGURE 35:TRACE DE LA NPV POUR UN RESERVOIR A IP ELEVE..... | 42 |
| FIGURE 36:TRACE DE LA NPV POUR UN BARIL DE PETROLE A 20 \$..... | 42 |
| FIGURE 37:TRACE DE LA NPV POUR UN BARIL DE PETROLE A 100 \$..... | 42 |
| FIGURE 38:DIAGRAMME DE CONCEPTION DU CONTROLE OPTIMALE DU SABLE | 44 |
| FIGURE 39:DIAGRAMME SCHEMATIQUE DE LA SIMULATION ISMR | 45 |
| FIGURE 40:SCHEMA DE LA ROTATION DYNAMIQUE DU PUIITS DE FORAGE..... | 46 |
| FIGURE 41:L'EFFET DE DIFFERENTES VALEURS DE NTG SUR LA PERMEABILITE DU SCREEN | 46 |
| FIGURE 42:L'EFFET DES DEFERENTES INCLINAISONS DE PUIITS SUR LA PERMEABILITE DU SCREEN | 47 |
| FIGURE 43:METHODE DE CONTROLE DU SABLE POUR L'ISMR | 48 |
| FIGURE 44:LES ETAPES DE MODELISATION ML | 50 |
| FIGURE 45:BOX PLOTS DES CARACTERISTIQUES DANS L'ENSEMBLE DE DONNEES POUR IDENTIFIER LA PRESENCE DES VALEURS ABERRANTES | 52 |
| FIGURE 46:LA DIFFERENCE ENTRE LES METHODES DE CLASSIFICATION ET DE REGRESSIO | 54 |
| FIGURE 47:LES CONCEPTS DE BIAIS ET DE VARIANCE SONT ILLUSTRÉS PAR LE LANCEMENT DE FLECHETTES SUR UNE CIBLE | 56 |
| FIGURE 48:ILLUSTRATION DE L'ALGORITHME KNN..... | 57 |
| FIGURE 49:MACHINE A VECTEURS DE SUPPORT POUR LA CLASSIFICATIO | 58 |
| FIGURE 50:L'IMPACT DE L'AUGMENTATION DE C SUR LA MARGE DU SVM..... | 59 |
| FIGURE 51:L'IMPACT DU GAMMA SUR LA FRONTIERE DES SVM..... | 59 |
| FIGURE 52:ILLUSTRATION DE L'ARBRE DE DECISION..... | 60 |
| FIGURE 53:ARBRE DE DECISION VERSUS FORET ALEATOIRE..... | 61 |
| FIGURE 54:L'ENSEMBLE DES DONNEES APRES TRAITEMENT DES VALEURS MANQUANTES..... | 66 |
| FIGURE 55:LA CARTE THERMIQUE DE CORRELATION POUR NOTRE ENSEMBLE DE DONNEES | 69 |
| FIGURE 56:TRACE DE COURBE TOP PERF DEPTH_TV D VS BHP | 70 |
| FIGURE 57:LA REPRESENTATION GRAPHIQUE DE LA RELATION PAR PAIRE DE L'ENSEMBLE DES DONNEES.... | 71 |
| FIGURE 58:DIAGRAMME EN BOITE DE L'ENSEMBLE DES DONNEES..... | 72 |
| FIGURE 59:DIAGRAMMES A BARRES DE F1_SCORE, PRECISION, RECALL DES QUATRES MODELES..... | 74 |

| | |
|--|----|
| FIGURE 60:DIAGRAMME A BARRES DE F1_MACRO, PRECISION_MACRO, RAPPEL_MACRO DU MODELE SELECTIONNE..... | 77 |
| FIGURE AN 1 : LA CARTE THERMIQUE DE CORRELATION POUR L'ENSEMBLE DES DONNEES DE PRODUCTION DU SABLE..... | 82 |
| FIGURE AN 2 : TRACE DE COURBE COH vs TVD..... | 83 |
| FIGURE AN 3 : TRACE DE COURBE TT vs COH..... | 83 |
| FIGURE AN 4 : TRACE DE COURBE COH vs EOVS | 84 |

Liste des tableaux

| | |
|--|----|
| TABLEAU 1: AVANTAGES DU FRAC PACK ET DU HRWP | 27 |
| TABLEAU 2: LES RISQUES DU FRAC PACK ET DU HRWP | 28 |
| TABLEAU 3: LISTE DES CRITERES DE CONCEPTION DU CONTROLE DU SABLE..... | 29 |
| TABLEAU 4: CONTROLE DU SABLE RECOMMANDE SELON LA METHODOLOGIE DE PRICE-SMITH..... | 31 |
| TABLEAU 5: LA SELECTION DES METHODES DE CONTROL DU SABLE..... | 40 |
| TABLEAU 6: RESUME DES DONNEES CALCULEES POUR TOUTES LES METHODES DE METHODES DE CONTROLE DU SABLE | 41 |
| TABLEAU 7:LES PRINCIPAUX PARAMETRES DANS L'ENSEMBLE DES DONNEES | 64 |
| TABLEAU 8: NOMBRE TOTALE DES PUIITS POUR CHAQUE TYPE DE COMPLETION..... | 65 |
| TABLEAU 9: DESCRIPTION STATISTIQUE DE L'ENSEMBLE DES DONNEES..... | 65 |
| TABLEAU 10:LES VARIANCES DES PARAMETRES NUMERIQUES | 67 |
| TABLEAU 11: LES SCORES DE CHI SQUARED POUR LES VARIABLES CATEGORIELLES | 68 |
| TABLEAU 12: DESCRIPTION STATISTIQUE DES DONNEES APRES LA NORMALISATION | 73 |
| TABLEAU 13: ANALYSE DES PERFORMANCES DES MODELES INTELLIGENTS IMPLEMENTEES POUR PREDIRE LA METHODE APPROPRIEE DE CONTROLE DU SABLE..... | 74 |
| TABLEAU 14: SELECTION DES HYPERPARAMETRES POUR NOTRE MODELE RANDOM FOREST..... | 75 |
| TABLEAU 15: ANALYSE DE LA PERFORMANCE DU RANDOM FOREST APRES L'OPTIMISATION..... | 75 |
| TABLEAU 16: LE RAPPORT DE CLASSIFICATION DU RANDOM FOREST | 76 |
| TABLEAU 17: ANALYSE DE LA PERFORMANCE DE RANDOM FOREST LORS DE LA VALIDATION..... | 76 |
| TABLEAU AN 1 : LES PRINCIPAUX PARAMETRE DANS LE JEU DES DONNEES | 81 |
| TABLEAU AN 2 : LA MATRICE DE CONFUSION DU RANDOM FOREST | 84 |

Introduction générale

Souvent la production de pétrole et de gaz est toujours suivie d'une certaine production de particules solides provoquant la réduction de la performance des puits. Les occurrences sont très limitées dans les schistes, les roches volcaniques et les roches basiques fracturées (basalte). Si l'on compare l'importance des réservoirs de grès et de carbonates, on constate que les grès sont plus abondants, mais que les calcaires sont plus importants en tant que réservoirs pour les hydrocarbures.

Pendant et après la production de sable, les puits peuvent s'ensabler et la productivité peut alors diminuer. L'élimination du sable produit est également un coût important associé à la production de ces formations. Les procédures correctives nécessitent des heures de travail non souhaitées sur les plates-formes. Des cas extrêmes avec des défaillances catastrophiques ont montré un sablage massif et l'abandon du puits. Dans les puits à débit élevé, le sable peut être transporté à la surface et provoquer l'érosion des conduites, des joints, des duses et des vannes. Cela pose de sérieux risques pour la sécurité des travailleurs. Lorsque la production de sable est identifiée, les opérateurs ont besoin de dispositifs de surveillance et de systèmes d'élimination en plus du traitement correctif ou des recomplétions.

L'industrie pétrolière et gazière n'a cessé de développer et d'utiliser des complétions de contrôle du sable dans les réservoirs afin de maîtriser ce problème. Cette technologie a joué un rôle essentiel, car les exigences deviennent plus difficiles et que les normes de performance augmentent. Avec des coûts d'exploitation et d'intervention élevés, l'impact de la production de sable ne peut être ignoré. L'effet du sable de formation dans un puits peut entraîner une perte d'intégrité et, par conséquent, l'effondrement du puits de forage. Il est absolument crucial pour l'industrie de gérer activement le sable.

L'objectif simple du contrôle du sable est de retenir tous les grains de sable, ou la majorité d'entre eux, tout en permettant l'écoulement des fluides de production ou d'injection. Un objectif commun est souvent de permettre à de petites particules du réservoir de passer à travers le dispositif de contrôle du sable.

Il existe plusieurs méthodes de contrôle du sable, les méthodes mécaniques et chimiques. Les méthodes mécaniques impliquent l'utilisation d'écrans (Screens) pour retenir le sable de la formation (avec ou sans gravier) ou l'utilisation de gravier pour retenir le sable (avec ou sans screen pour retenir le gravier), y compris le massif filtrant (Gravel Pack), Screen Pre-Packed, Slotted Liner et Wire-Wrapped. Les méthodes chimiques utilisent une résine liquide qui est injectée à partir d'un puits de forage dans la roche non consolidée entourant le puits. Les méthodes chimiques font appel à des techniques de consolidation du sable in situ et au le massif filtrant enrobé de résine. Autres méthodes nommées le Frac pack qui comprend la fracturation hydraulique simultanée du réservoir et la mise en place d'un massif filtrant, les packs d'eau à haut débit (HRWP) s'installent de la même manière que le frac pack mais permettent de réduire le coût des fluides de fracturation.

La conception d'un contrôle optimal du sable est un processus compliqué car le choix d'une méthode optimale de contrôle dépend de différents facteurs efficaces. Ces facteurs comprennent la taille du gravier du massif filtrant, la largeur des fentes et la longueur de la crépine, la réduction de l'IP et les coûts d'exploitation. La sélection optimale est encore plus complexe en raison de l'incertitude associée aux variables influençant les méthodes de contrôle du sable. De nombreux auteurs ont étudié diverses méthodes de complétion de puits dans différentes conditions.

Certains d'entre eux ont discuté des conséquences de la production de sable, tandis que peu de spécialistes ont travaillé sur la sélection des méthodes de contrôle du sable.

Un examen des études antérieures montre que toutes les méthodes de sélection du contrôle du sable ne prennent en compte qu'un nombre limité de critères pour déterminer la meilleure méthode. Néanmoins, de nombreux facteurs influencent simultanément la sélection en raison de la complexité et des incertitudes rencontrées dans le domaine de l'ingénierie pétrolière.

Chaque industrie est confrontée à des défis opérationnels au quotidien, que ce soit en raison de l'arrêt des machines ou de la défaillance des équipements. Toutefois, les dernières avancées technologiques, telles que l'intelligence artificielle, aident les industries à relever efficacement ces défis. Après en avoir été témoin, l'industrie pétrolière et gazière a finalement commencé l'intégration de ces technologies dans ses opérations. Parmi celles-ci, l'apprentissage automatique (Machine Learning) et la science des données (Data Science). L'apprentissage automatique contribue à transformer la découverte des gisements pétroliers en permettant aux entreprises de collecter des informations massives en temps réel et afin de retrouver des intuitions utiles à partir de grands ensembles de données.

L'apprentissage automatique est aujourd'hui employé dans le secteur pétrolier pour l'analyse et la modélisation, le forage et la caractérisation du sous-sol, l'anticipation des besoins de maintenance, l'optimisation des réseaux d'approvisionnement et l'optimisation des ressources financières, autant d'éléments qui améliorent l'efficacité. Lorsque la mise en œuvre de l'IA dans l'industrie pétrolière et gazière donne des résultats significatifs. Cela permet de débloquer des efficacités jamais imaginées auparavant en proposant des méthodes innovantes d'exploration, de développement, de production, etc. Comme l'IA a un immense potentiel d'augmentation des processus, elle permet de travailler avec une main-d'œuvre moins nombreuse et de la libérer pour des activités plus ingénieuses.

Cette thèse a été élaborée dans le but d'apporter une valeur ajoutée et une contribution constructive à ce domaine. Notre étude consiste à implémenter les différentes techniques d'apprentissage automatique supervisé pour sélectionner efficacement la méthode appropriée de contrôle du sable afin de réduire l'incertitude de la sélection

Cette mémoire est organisée en cinq chapitres :

Dans le premier chapitre « **Phénomène de production du sable** » revue de la littérature qui couvre le problème de production du sable, sa nature, ses effets, les causes et les conséquences de ce dernier. Le deuxième chapitre « **les méthodes de contrôle du sable** » résume les techniques les plus courantes de contrôle du sable utilisées aujourd'hui, y compris les caractéristiques de chaque technique. Le troisième chapitre « **Sélection des méthodes de contrôle du sable** » passe en revue les travaux antérieurs et les différentes approches de la sélection du contrôle du sable. En couvrant dans le quatrième chapitre les techniques d'intelligence artificielle et d'apprentissage automatique et illustrant les mathématiques qui les composent ainsi que les différentes techniques d'exploration de données avant d'entrer dans le dernier chapitre qui consiste à préparer le jeu de données, à créer des modèles ML pour la sélection des méthodes de contrôle du sable et à les évaluer. La dernière partie est consacrée à la conclusion et aux perspectives d'avenir.

Théorie et Littérature

Chapitre 1 : Phénomène de production du sable

1.1) Nature, causes et conséquences de la production du sable :

1.1.1) La nature de production de sable :

Les conditions qui peuvent provoquer la production de sable peuvent être déterminées par les facteurs qui affectent le début de ce phénomène. Ces facteurs doivent décrire à la fois la nature du matériau de la formation et les forces qui provoquent la rupture de la roche.

La résistance d'un grès est contrôlée par :

- La quantité et le type de matériau cimentaire qui maintient les grains individuels ensemble.
- Les forces de frottement entre les grains
- La pression du fluide dans les pores de la roche
- Les forces de pression capillaire

Le type de défaillance susceptible de se produire dans le grès a été étudié par plusieurs chercheurs. Les travaux menés chez Exxon indiquent que la nature d'un tunnel de perforation défaillant est révélatrice d'une rupture par cisaillement qui se produira lorsque la résistance à la compression de la roche sera dépassée. En outre, les travaux d'Exxon indiquent que dans les grès faiblement consolidés, la production de sable pourrait entraîner des vides derrière le tubage. Ils ont conclu que la résistance à la compression de la roche devrait être un bon indicateur du potentiel de production de sable.

En général, la résistance à la compression d'une roche est principalement contrôlée par les forces de frottement intergranulaires. Par conséquent, l'augmentation de la résistance de la roche augmente avec l'augmentation de la contrainte du confinement. Dans la situation de rupture de la matrice rocheuse entourant un tunnel de perforation, la roche sera dans un état de contrainte non confinée, la production de sable doit donc être liée à la résistance à la compression non confinée de la roche. Les contraintes qui provoquent la rupture de la roche dans cette situation comprennent la contrainte mécanique résultant du matériau de recouvrement et les forces de traînée associées à l'écoulement de fluides visqueux dans la matrice rocheuse. La contrainte verticale est partiellement supportée par la pression interstitielle à l'intérieur de la roche ; ainsi, la contrainte qui provoque réellement la rupture de la roche (c'est-à-dire la contrainte effective) est la différence entre la contrainte verticale et la pression interstitielle.

Après avoir produit un peu de sable autour d'un tunnel de perforation, il se forme un arc qui a une résistance suffisante pour supporter le poids du matériau environnant. Dans certaines conditions, la production d'une quantité limitée de sable de formation peut être tolérée pour permettre à un arc de se développer, après quoi la production de sable de formation cesse. La figure 1 illustre le concept d'un arc stable autour d'une perforation ; cependant, la stabilité de l'arc est compliquée par le fait que l'état de contrainte entourant la perforation change constamment en raison des éléments suivants : la différence de pression, le débit d'eau etc. [1]

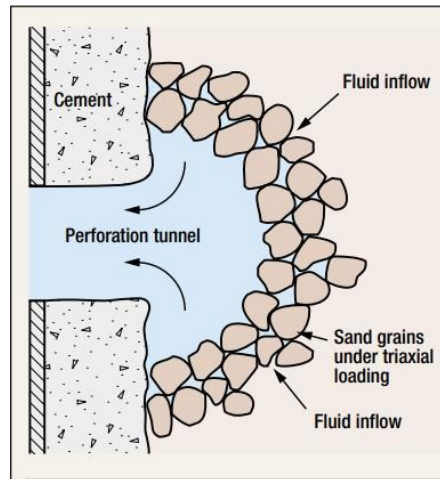


Figure 1: Géométrie d'un arc stable entourant une perforation [2]

1.1.2) Effets de la production de sable :

Les effets de la production de sable sont presque toujours préjudiciables à la productivité à court et/ou à long terme du puits. Bien que certains puits connaissent régulièrement une production de sable "gérable", ces puits sont l'exception et non la règle. Dans la plupart des cas, tenter de gérer les effets d'une production importante de sable pendant la durée de vie du puits n'est pas une alternative d'exploitation prudente ou économiquement intéressante.

- **Accumulation dans les équipements de surface.** Si la vitesse de production est suffisamment élevée pour entraîner le sable dans les flowlines, celui-ci peut se retrouver piégé dans le séparateur, le dispositif de traitement thermique ou la conduite de production. Si un volume suffisamment important de sable est piégé dans l'une de ces zones, un nettoyage sera nécessaire pour permettre une production efficace du puits. Pour rétablir la production, il faut fermer le puits, ouvrir l'équipement de surface et enlever le sable manuellement. En plus du coût du nettoyage, le coût de la production différée doit être pris en compte. Si un séparateur est partiellement rempli de sable, la capacité du séparateur à traiter le pétrole, le gaz et l'eau est réduite. Par exemple, un pied cube de sable dans un séparateur huile/eau avec un temps de séjour de 2 minutes fera que le séparateur traitera 128 barils de liquide de moins par jour. Si le rapport entre le pétrole et l'eau entrant dans le séparateur est d'un pour un (c'est-à-dire 50 % water cut), le séparateur fournira 64 barils de pétrole commercialisable de moins par jour. À 114,52 \$ le baril, cela représente 48 153 368,60 \$ perdus par an.
- **Accumulation en fond de puits.** Si la vitesse de production n'est pas suffisante pour transporter le sable peut commencer à remplir l'intérieur du tubing. Finalement, l'intervalle de production peut être complètement recouvert de sable. Dans ce cas, le taux de production diminue jusqu'à ce que le puits soit "ensablé" et que la production cesse. Dans de telles situations, des opérations correctives sont nécessaires pour nettoyer le puits et rétablir la production. Si la production de sable est continue, les opérations de nettoyage peuvent être nécessaires sur une base régulière, aussi souvent qu'une fois par mois ou même par semaine.
- **Érosion des équipements de fond et de surface.** Dans les puits très productifs, les fluides circulant à grande vitesse et transportant du sable peuvent produire une érosion excessive de l'équipement de fond et de surface, entraînant une maintenance fréquente pour remplacer l'équipement endommagé. La figure 2 montre une photographie d'une section de screen exposée à une perforation qui produisait du sable. Ainsi qu'une duse qui s'est rompue en raison d'une érosion excessive.

Si cette dernière est grave ou se produit sur une période suffisamment longue, une défaillance complète de l'équipement de surface et/ou de fond peut se produire, entraînant des problèmes critiques de sécurité et d'environnement ainsi qu'une production différée. Pour certaines défaillances d'équipement, un workover assisté par un appareil de forage peut être nécessaire pour réparer les dommages. [1]

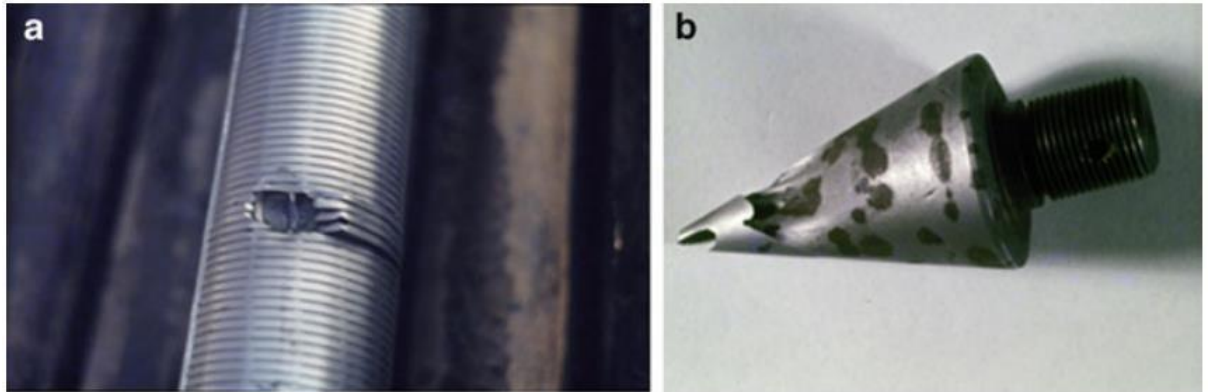


Figure 2: Érosion de l'équipement de production due à la production de sable : (a) érosion du screen, (b) érosion de la duse de surface [6]

La figure 3 décrit le taux d'érosion estimé pour une tête de puits. Elle indique que si l'érosion de la tête ne constitue pas une préoccupation majeure avant 2017, maintenant c'est le bon moment pour contrôler la production de sable il faut soit reconstituer le puits avec un dispositif de contrôle du sable en fond de trou, soit utiliser un trou d'écoulement à diamètre intérieur élargi sur l'arbre pour réduire la vitesse d'écoulement.

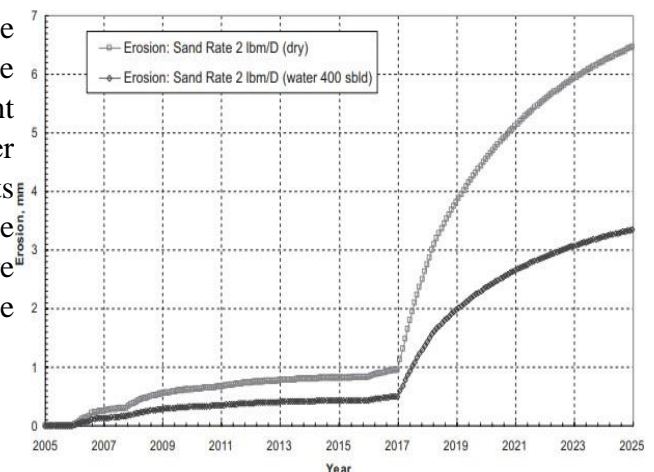


Figure 3: Estimation du taux d'érosion pour une tête de puits, les carrés sont dans des conditions de gaz sec, et les diamants sont dans des conditions de gaz/liquide [3]

- **Effondrement de la formation** : De grands volumes de sable peuvent être transportés hors de la formation avec le fluide produit. Si le taux de production de sable est suffisamment élevé et se poursuit pendant une période de temps suffisante, une zone vide ou un vide se développera derrière le tubage et continuera à s'agrandir au fur et à mesure de la production de sable. Lorsque le vide devient suffisamment grand il y a un risque d'effondrement.

1.1.3) Causes de la production de sable :

Le facteur critique pour évaluer le risque de production de sable à partir d'un puits particulier est de savoir si oui ou non la production des particules solides peut être maintenue en dessous d'un niveau acceptable aux débits prévus et aux conditions économiques du puits. La liste suivante résume un grand nombre des facteurs qui influencent la tendance d'un puits à produire du sable :

- Le degré de consolidation
- Réduction de la pression interstitielle pendant la durée de vie d'un puits
- Taux de production
- La viscosité du fluide du réservoir
- Augmentation de la production d'eau pendant la durée de vie du puits.

Ces facteurs peuvent être classés en deux catégories : les effets sur la résistance de la roche et les effets sur l'écoulement des fluides. Chacun de ces facteurs et leur rôle dans le déclenchement de la production de sable sont abordés dans la suite :

- **Degré de consolidation** : La cimentation d'un grès est généralement un processus géologique secondaire et, en règle générale la résistance à la compression est une caractéristique mécanique de la roche qui est liée au degré de consolidation. Les formations de grès mal consolidées ont généralement une résistance à la compression faible. De plus, même les formations de grès bien consolidées peuvent être modifiées par la dégradation du matériau de la matrice, ce qui permettrait la production de sable.
- **Réduction de la pression interstitielle** : La pression dans le réservoir supporte une partie du poids de la roche sus-jacente. Au fur et à mesure que la pression du réservoir diminue pendant la durée de production d'un puits, une partie du soutien de la roche sus-jacente est supprimée. L'abaissement de la pression du réservoir crée une contrainte croissante sur le sable de formation lui-même. À un moment donné, les grains de sable de formation peuvent se détacher de la matrice ou être écrasés, créant des fines qui sont produites avec les fluides du puits. La compaction de la roche réservoir due à une réduction de la pression interstitielle peut entraîner un affaissement de la surface. [1]
- **Taux de production** : La production des fluides du réservoir crée un différentiel de pression et des forces de frottement qui peuvent se combiner pour dépasser la résistance à la compression de la formation. Cela indique qu'il existe un débit critique pour la plupart des puits en dessous duquel les forces de pression différentielle et de frottement ne sont pas suffisamment importantes pour dépasser la résistance à la compression et provoquer la production de sable. Le débit critique d'un puits peut être déterminé en augmentant lentement le taux de production jusqu'à ce que la production de sable soit détectée. Une technique utilisée pour minimiser la production de sable consiste à étrangler le débit jusqu'au débit critique où la production de sable ne se produit pas ou à un niveau acceptable. Dans de nombreux cas, ce débit est nettement inférieur au taux de production acceptable pour le puits. [1]

Le diagramme de la figure 4 montre un "drawdown sécurisé" pour des combinaisons particulières de complétion, une certaine résistance de la roche et des conditions initiales de pression et de contrainte du réservoir. Le tracé de la pression d'écoulement du fond du trou en fonction de la pression du réservoir est divisé en deux zones:

- Zone de non production au-dessus de la ligne diagonale où la pression d'écoulement du fond du trou est supérieure à la pression du réservoir.
- Zone de production du puits (la pression d'alimentation du fond est inférieure à la pression du réservoir) qui est subdivisée par la ligne de drawdown sécurisé en:
 1. Zone de production sans sable (zone verte) et.
 2. Zone de production de sable (plus pétrole) (zone rouge).

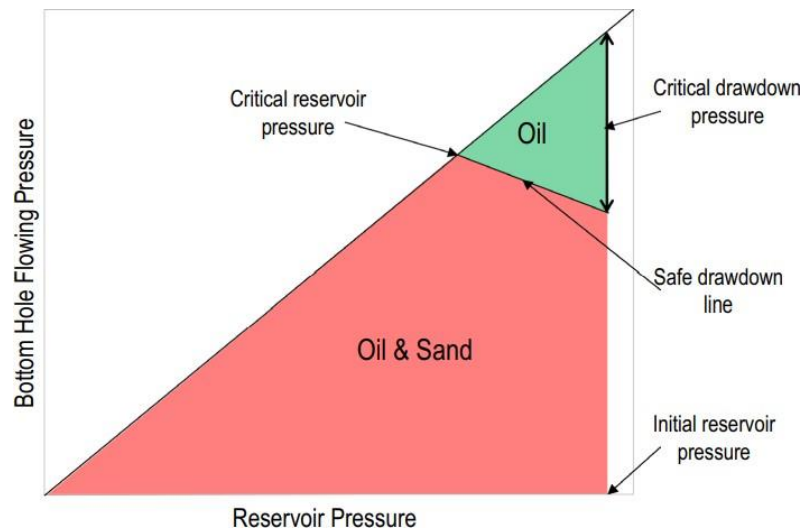


Figure 4: Fenêtre du drawdown sécurisé tracé à une seule profondeur [4]

L'expérience a montré qu'il existe un taux de production critique en dessous duquel la production de sable diminue de façon gérable. Le taux de production critique peut éventuellement se situer à un niveau de production non économique. Il est donc utile de déterminer ce taux avant de concevoir une stratégie de complétion pour une formation donnée. La figure 5 présente une corrélation de terrain entre le débit d'un puits et la quantité de sable produite dans l'un des réservoirs de grès d'Arabie Saoudite.

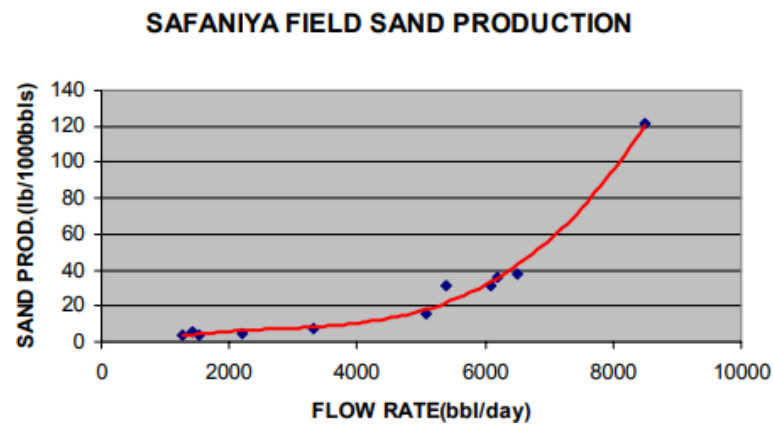


Figure 5: Corrélation de terrain entre le débit de production et le volume de sable produit [5]

- **Viscosité du fluide de réservoir :** La force de frottement exercée sur les grains de sable de la formation est créée par l'écoulement du fluide du réservoir. Cette force de frottement est directement liée à la vitesse d'écoulement du fluide et sa viscosité. Une viscosité élevée du fluide appliquera une grande force de frottement sur les grains de sable qu'un fluide à faible viscosité. L'influence de la résistance visqueuse fait que du sable est produit à partir de réservoirs de pétrole lourd qui contiennent des huiles de faible gravité et de haute viscosité, même à des faibles vitesses d'écoulement.
- **Augmentation de la production d'eau :** La production de sable peut augmenter ou commencer lorsque l'eau commence à être produite ou lorsque la coupe d'eau augmente. Deux possibilités peuvent expliquer bon nombre de ces occurrences. Tout d'abord, pour une formation typique de grès mouillé par l'eau, une certaine cohésion entre les grains est assurée par la tension superficielle de l'eau connatée entourant chaque grain de sable.

Au début de la production d'eau, l'eau connatée a tendance à s'unir à l'eau produite, ce qui entraîne une réduction des forces de tension superficielle et une réduction subséquente de la cohésion entre les grains. Il a été démontré que la production d'eau limite fortement la stabilité de l'arc de sable autour d'une perforation. Ce qui entraîne le début de la production de sable.

Un deuxième mécanisme par lequel la production d'eau affecte la production de sable est lié aux effets de la perméabilité relative. Lorsque la quantité d'eau augmente, la perméabilité relative du pétrole diminue. Il en résulte une augmentation de la pression différentielle nécessaire pour produire du pétrole au même rythme. Une augmentation de la pression différentielle près du puits de forage crée une force de cisaillement plus importante sur les grains de sable de la formation. Une fois encore, les contraintes plus élevées peuvent entraîner une instabilité de l'arc de sable autour de chaque perforation et la production de sable qui s'ensuit. [1]

Sachant que la saturation en eau et le drawdown appliqué ont un effet sur la quantité de sable produite, les deux effets ont été combinés et corrélés comme le montre la figure 6 ci-dessous :

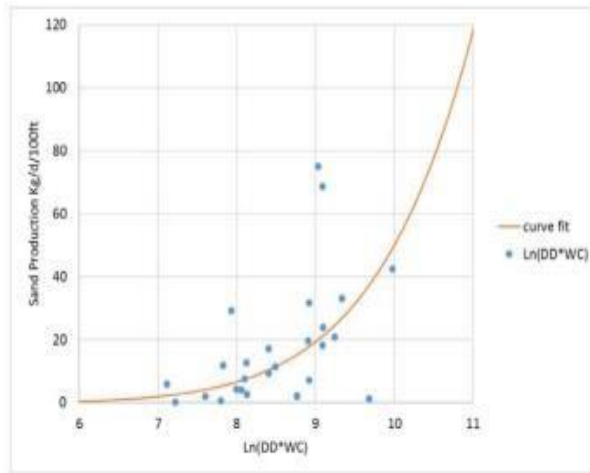


Figure 6: Production de sable en fonction du drawdown et de la saturation d'eau [7]

Pour voir comment la corrélation se comportera en fonction de différentes saturations d'eau, la saturation d'eau a été maintenue constante tandis que le drawdown de l'eau a été varié comme le montre la figure 7 :

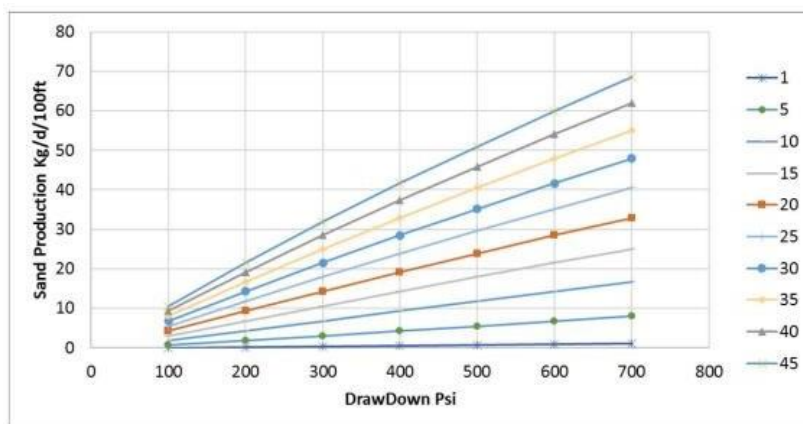


Figure 7: Production de sable calculée en fonction de différents drawdown pour une saturation en eau fixe [7]

Le même effet a été tracé dans la figure 8, mais cette fois en gardant le drawdown constant et en changeant la saturation d'eau.

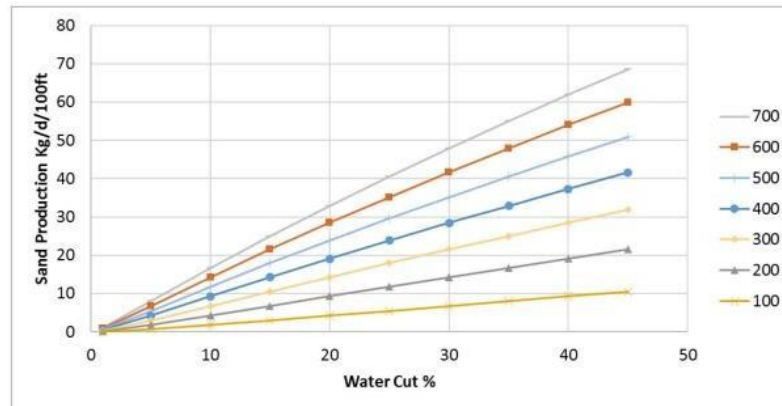


Figure 8: Production de sable calculée en fonction de différentes saturations pour un drawdown fixe [7]

L'expérience a montré que l'augmentation du drawdown et de la saturation en eau agissait sur l'augmentation de la production de sable.

1.2) Stabilité du puits et critères de rupture du sable :

1.2.1) Stabilité du puits :

La résistance de la roche est sa capacité à résister à une contrainte sans céder ou se fracturer. La figure 9 illustre un essai typique, un cylindre dont le rapport longueur/diamètre est de 2 : 1. Sur les faces d'extrémité du cylindre une contrainte est exercée tandis qu'un bain d'huile de confinement exerce une contrainte d'une ampleur éventuellement différente sur la circonférence (Fjær et al. 2008). Il s'agit d'un essai uniaxial si la contrainte de confinement est nulle et d'un essai triaxial s'il est réalisé avec une pression de confinement non nulle. Deux techniques de laboratoires sont généralement utilisées pour estimer La résistance des roches : les essais de résistance à la compression uniaxiale et les essais de résistance à la compression triaxiale ou confinée. [17]

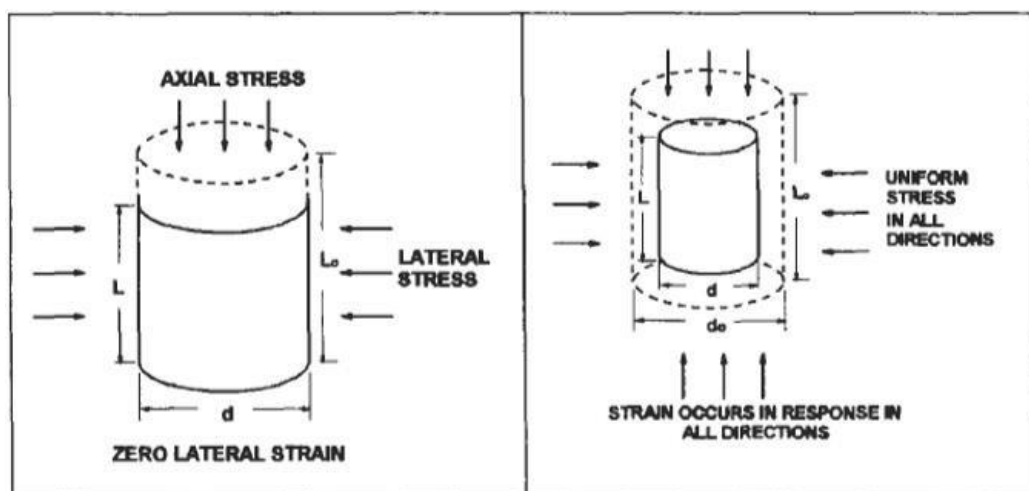


Figure 9: Échantillons typiques pour les essais uniaxiaux ou triaxiaux [17]

Avant le forage d'un puits, la roche est dans un état d'équilibre. Les contraintes dans la roche dans ces conditions sont connues sous le nom des contraintes du champ lointain σ_v , σ_H , σ_h ou contraintes in-situ (Gaurina-Međimurec 1994).

Lors de forage du puits le support initialement offert par la roche est remplacé par la pression hydrostatique de la boue ce qui induit à la redistribution des contraintes à proximité du puits.

Ces dernières peuvent être résolues en une contrainte verticale (overburden) σ_v , et deux contraintes horizontales σ_H (la contrainte horizontale maximale in situ) et σ_h (la contrainte horizontale minimale in situ) qui sont généralement inégales (McLean et Addis 1990). Si l'état des contraintes redistribuées dépasse la résistance de la roche, que ce soit en tension ou en compression, il peut en résulter une instabilité. La figure 10 montre les contraintes du puits après le forage. On constate la contrainte radiale σ_r , la contrainte tangentielle (contrainte circonférentielle ou cerclée) σ_t , et la contrainte axiale σ_a . La contrainte radiale agit dans toutes les directions perpendiculaires à la paroi du puits, la contrainte tangentielle agit aux alentours du puits et la contrainte axiale agit parallèlement à l'axe du puits, tenant en compte que les contraintes du puits changent rapidement lorsqu'on s'éloigne et se convertissent en contraintes du champ lointain. Cela s'explique par le fait que la roche est dans un état original et non perturbé. [6]

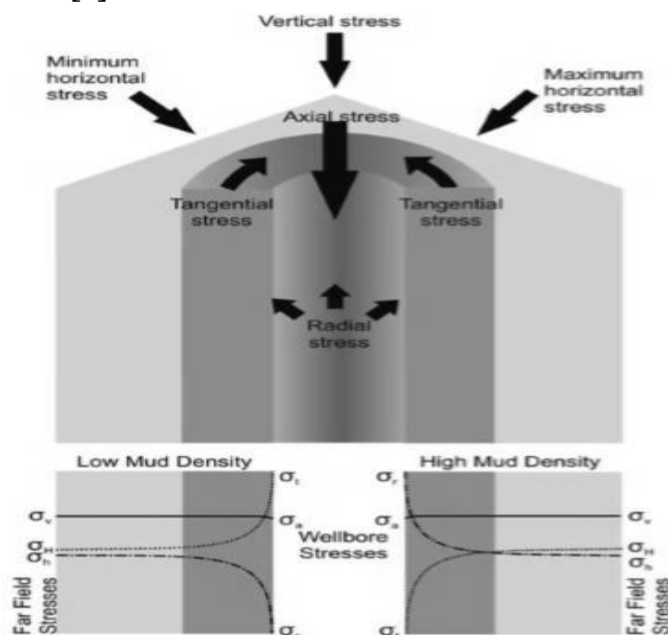


Figure 10: Variation des contraintes à l'écart du puits de forage [6]

Pendant la production, le drawdown et/ou l'épuisement provoquent une augmentation des contraintes effectives ($\sigma_{\text{effective}} = \sigma_{\text{total}} - \alpha \cdot P_p$) et conduisent à une rupture de la roche c'est-à-dire un volume désastreux de production du sable.

En d'autres termes, des contraintes effectives plus élevées dans la circonférence du puits ou autour des tunnels de perforation se manifesteront à partir d'un drawdown important et si ces contraintes dépassent la résistance de la roche gréseuse intacte (UCS) une défaillance se produira. Aussi, la diminution de la pression des pores pendant la production aggravera la situation car elle peut modifier les contraintes in situ agissant sur la roche. Par conséquent, ce phénomène est associé avec des contraintes plus élevées d'où le risque de la production du sable sera élevé. [12]

1.2.2) Critères de rupture du sable :

Comme la rupture de la roche est la cause principale de toute production de sable il est essentiel de comprendre les différents mécanismes de défaillance. Les formations sableuses peuvent subir des défaillances en compression, en tension et en cohésion qui déclenchent la production de sable. Une rupture est assez fréquente pendant le forage est la rupture par compression soit à cause de la détérioration des matériaux de cimentation par le filtrat de la boue ou bien lorsque la roche ne peut pas résister au nouveau champ des contraintes.

Pendant la production, un rabattement de pression est établi pour un taux de production donné. Cette baisse de pression peut provoquer la rupture de la roche en tension ou en cohésion (érosion). Tenant compte des trois formes de rupture les propriétés mécaniques associées à la production de sable sont les suivantes : La résistance à la compression uniaxiale (UCS), la force de cohésion et la résistance à la traction. [15]

1.2.2.1) Rupture par cisaillement :

La résistance au cisaillement se compose de deux éléments : la cohésion ou les liens physiques entre les particules de sable adjacents et la friction. La détérioration de la formation associée avec de petits et gros grains de sable aura lieu au niveau du plan de rupture une fois que la rupture par cisaillement se produit. Divers critères de rupture peuvent être utilisés pour prédire le mécanisme de rupture par cisaillement d'une roche.

Parmi eux, on trouve Von Mises, Drucker-Prager, Mohr-Coulomb, Hoek-Brown, Lade modifié et Weibols & Cook modifié.

Le critère de Mohr-Coulomb est le plus largement utilisé pour la prédiction de la rupture par cisaillement. Il stipule que la rupture se produit lorsque la contrainte de cisaillement sur un plan donné à l'intérieur de la roche atteint une valeur critique donnée par :

$$\tau = C + \sigma_n \tan \theta \quad (E)$$

Où : τ est la résistance au cisaillement, psi

σ_n : est la contrainte normale au plan de rupture, psi C : est la force cohésive, psi

θ : est l'angle de friction interne, degré

L'équation (E) se compose de deux éléments : la cohésion (C) et la friction ($\sigma_n \tan \theta$). Il faut noter que la rupture par cohésion produit des particules de sable tandis que la rupture par cisaillement brise la roche le long du plan de cisaillement. [13]

Plusieurs facteurs déterminent si une roche (réservoir) va se rompre mécaniquement. Parmi eux, on trouve la résistance de la roche (UCS), la contrainte moyenne effective agissant sur la roche et la distribution des contraintes à proximité du puits et le drawdown dû à la production.

La valeur de la résistance de la roche (UCS) peut être obtenue en laboratoire en effectuant des essais uniaxiaux et triaxiaux appropriés. Un diagramme circulaire de Mohr sert à générer une enveloppe de défaillance en représentant les contraintes agissant sur les échantillons de roche. Le cercle de Mohr (rouge) représente l'état de contrainte à toute orientation dans un corps de matériau, allant de la plus petite contrainte effective principale σ_3 , à la plus grande σ_1 . La rupture en cisaillement se produit si le cercle de Mohr intersecte la condition de rupture (**bleu**). Le cercle de Mohr présenté dans la figure 11 fournit également la contrainte normale (σ_n) et la contrainte de cisaillement (τ) à travers le plan de rupture, ainsi que l'angle de rupture β . L'angle β connu sous le nom d'angle de rupture est mesuré à partir de la direction perpendiculaire à la contrainte principale maximale. [12]

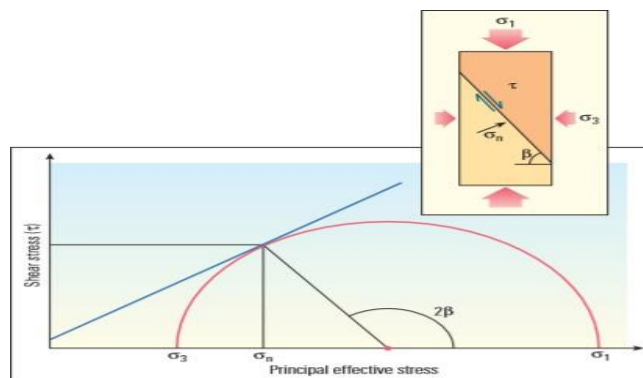


Figure 11: Critère de défaillance de Mohr-Coulomb [14]

1.2.2.2) Rupture par traction :

La rupture par traction fait référence à une contrainte radiale de traction dépassant l'enveloppe de rupture par traction et est déclenchée exclusivement par la pression de rabattement. Ce mécanisme est assez fréquent autour d'un tunnel de perforation généralement lorsqu'un changement brusque de la pression différentielle entre le puits et le réservoir induit la contrainte radiale à dépasser la résistance à la traction de la formation, provoquant ainsi la production de sable et l'élargissement du tunnel de perforation. [15]

La résistance à la traction est sensible à la présence de fissures dans le matériau. Il faut savoir qu'au niveau du laboratoire un échantillon qui subit une rupture en traction se fend généralement le long des plans de fracture, comme l'illustre la figure 12. Les fissures préexistantes orientées plus ou moins à la direction de la contrainte de traction souvent forment des plans de fracture. Probablement il y aura un endommagement de la roche là où elle situe la plus grande de ces fissures, par conséquent, la ou les plus grandes fissures se développeront de plus en plus rapidement et fendront rapidement l'échantillon. [17]

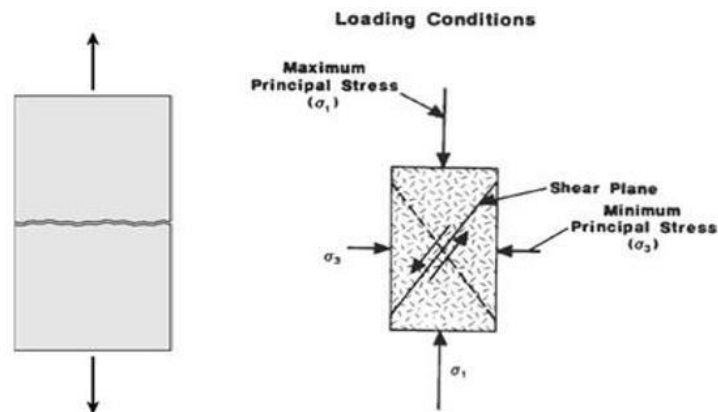


Figure 12: Rupture par traction et cisaillement [17]

1.2.2.3) Rupture par cohésion :

La rupture par l'érosion (ou la cohésion) fait référence à l'enlèvement ou à la production progressive des particules de sable individuelles de la surface de la cavité (tunnel de perforation). [13] Ce mécanisme de défaillance cohésive est particulièrement important dans le cas des sables mal consolidés. L'érosion est contrôlée par la force de cohésion, elle aura lieu si la force de traînée exercée sur une particule de surface dépasse la cohésion (apparente) entre les particules de surface. La vitesse d'écoulement du fluide est un paramètre important puisqu'elle est liée directement à la résistance par frottement. La cohésion est due à deux facteurs : (1) le matériau de cimentation et (2) les forces capillaires [15]. Une remarque très importante et confirmée par l'expérience du terrain est que dans les formations faiblement consolidées, la production de sable à partir des trous ouverts tend à être inférieure à celle des perforations, cela à cause que la vitesse du fluide à la surface de perforation est supérieure à celle du trou ouvert. [13]

Chapitre 2 : Les différentes méthodes de contrôle du sable

Le concept de contrôle du sable est basé sur la tolérance zéro pour la production de sable à la surface. Les problèmes liés à ce phénomène ont justifié la mise en place des dispositifs de contrôle du sable en fond du puit. Plusieurs méthodes ont été utilisées dans un réservoir et le choix dépend des pratiques d'exploitation, des conditions du réservoir (caractéristiques du sable de formation), des expériences réussies sur le terrain et des considérations économiques. Traditionnellement, les principales techniques sont soit de type mécanique ou chimique [13].

Les techniques de contrôle du sable disponibles dans l'industrie comprennent :

1. Contrôle du débit
2. Perforations
3. Screens (sans gravel packs)
 - Slotted liners
 - Wire-wrapped screens
 - Premium screens
 - Expandable screens
 - Pre-packed screens
4. Gravel pack
 - Open hole gravel pack
 - Cased hole gravel pack
5. Frac pack
6. HRWP (Pack d'eau à haut débit)
7. Consolidation chimique
 - Consolidation in situ de formation
 - Gravier enrobé de résine

2.1) Méthode de contrôle du débit :

Cette méthode est utilisée par les opérateurs pour contrôler la production de sable. L'expérience en laboratoire et sur le terrain a montré que la réduction de taux de production et la baisse de la différence de pression réduit le flux de sable. La procédure consiste à réduire ou à augmenter le taux/le drawdown jusqu'à ce qu'un niveau acceptable de production de sable soit atteint ou bien en identifiant un seuil de taux/drawdown, où du sable peut être produit. [13]

L'objectif de cette technique est d'obtenir un débit maximal qui permettra la formation d'un arc stable au niveau du puits de forage. La recherche d'un débit maximal doit être répétée plusieurs fois, car les conditions du réservoir changent (par exemple, les pressions, la saturation du fluide). La méthode de contrôle du débit est généralement plus efficace lorsque peu de sable a déjà été produit. [6]

Le contrôle des débits est utilisé temporairement jusqu'à ce qu'une méthode de contrôle du sable plus efficace soit installée (Mécanique ou chimique). L'utilisation permanente de cette méthode est généralement lorsque le coût d'une autre méthode de contrôle ne peut pas être justifié ou n'est pas réalisable sur le plan opérationnel. [13]

La méthode est basée sur le fait qu'en régime permanent, des arcs naturels s'établissent pour contrôler la production de sable (Figure 13.a). En augmentant le débit par paliers, la concentration de sable fait un bond (Figure 13.b).

Mais lorsque la vitesse critique du fluide est atteinte, les arcs ne se forment pas et la production de sable continue. Le taux de production doit donc être réduit et maintenu en dessous de la plage critique. [6]

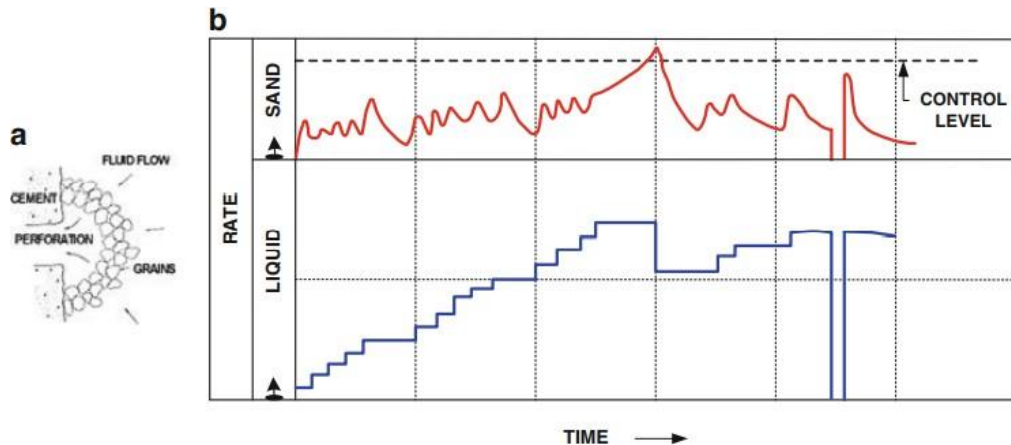


Figure 13: Formation des arcs naturels (a) Détermination de la production de sable en fonction du débit (b)

2.2) Perforation :

Les complétions tubées et perforées sont la base de nombreux champs. La perforation d'un puits a pour but d'établir une bonne communication de flux entre le puits de forage et le réservoir. Ceci est plus courant dans les champs onshore mais existe également dans les zones offshore. La productivité d'un puits dépend largement de la conception de la perforation. Les paramètres de perforation qui influent sur la productivité du puits comprennent les dimensions de la perforation (diamètre et pénétration), le phasage, la densité des tirs et le type de charge. La dimension de la perforation est en fonction de la conception et de la qualité de la charge/du canon de perforation, de la position du canon/des charges dans le puits de forage lorsqu'ils sont tirés, et des conditions du puits telles que la température, la pression, les fluides de forage, la taille et la métallurgie du tubage, le ciment et les propriétés de la formation (Navjeet, 2004). L'objectif de la plupart des complétions de trous tubés est de générer des charges pénétrantes profondes d'une longueur maximale de perforation. L'énergie explosive d'une perforation crée un trou par pression extérieure.

Cette pression écrase le ciment et la roche. Ils ne sont pas détruits dans le processus, mais ils se retrouvent à l'intérieur de la perforation, ainsi que des parties de l'assemblage de ce dernier. Ils doivent être retirés pour que la perforation soit productive. Si elles ne sont pas enlevées, il en résulte une chute de pression plus importante au niveau des perforations qui peut contribuer à la rupture par traction.

La plupart de ces débris seront des roches écrasées/fracturées, avec des quantités mineures de débris de charge dont il existe plusieurs façons d'éliminer ces endommagements. L'écoulement du puits après la perforation créera un drawdown sur toutes les perforations. Cela entraînera une partie des débris de certaines perforations (Bellarby, 2009). Il est important de minimiser l'altération des perforations afin de réduire les contraintes de production qui tendent à provoquer la production de sable. Pour contrôler le sable, il est impératif de procéder à un nettoyage plus approfondi avant le gravel pack, le frac pack ou l'injection de produits chimiques de consolidation (Completion Tech.).

L'objectif du gravel pack est d'avoir des perforations de grand diamètre qui ont été bien nettoyées, laissant une cavité ouverte où le gravier peut être placé.

Le nettoyage des perforations est nécessaire pour éliminer les endommagements causés par l'écrasement et la compaction de la formation, les boues de forage, le ciment, les fluides de complétion et les débris des tirs de perforation. Le nettoyage des perforations peut être réalisé par une perforation en underbalance et un lavage des perforations. La perforation en underbalance permet la production du sable pendant les étapes initiales et évite ainsi d'avoir à gérer la production transitoire de sable pendant les étapes ultérieures de la production du puits (Venkitaraman et al., 2000). Le lavage des perforations est une méthode largement utilisée pour éliminer les dommages. Elle consiste à pomper à travers chaque perforation avec une pression et un débit suffisant car il est essentiel de maintenir la circulation pour éliminer les débris du puits. Un système de fluide propre ou des fluides avec des systèmes de filtration de surface ou une combinaison des deux peuvent être utilisés pour le lavage. [13]

2.2.1) Perforation sélective :

Dans les formations hétérogènes, la résistance de la roche varie considérablement avec la profondeur. En évitant de perforer les sections les plus faibles de la formation, on peut maintenir un taux de production sans sable pendant toute la durée de vie du puits. Un drawdown critique plus élevé peut être obtenu en perforant les zones fortes (ayant un degré de cimentation plus élevé). Malheureusement, ces sections plus faibles sont souvent les principales zones productives que les intervalles plus forts ayant une perméabilité plus faible. Pour permettre le drainage du réservoir, la formation doit avoir une bonne perméabilité verticale de sorte que les fluides des sections les plus faibles puissent s'écouler vers les sections les plus fortes. Avant de prendre cette décision, il est nécessaire de procéder à une analyse de la productivité à l'aide de l'analyse nodale et l'analyse de la résistance (à l'aide de carottes, de diagraphies, etc.). [13]

L'intervalle B de la figure 14 est un réservoir gréseux sans production de sable, tandis que l'intervalle A présente une production de sable. La vitesse sonique dans le sable A est de 70-91 msec/ft alors qu'elle est d'environ 60 msec/ft pour le sable B. [15]

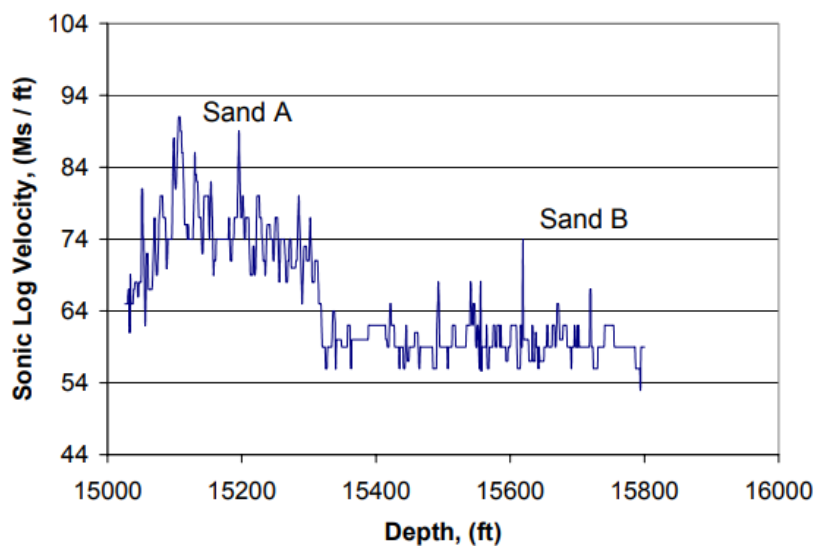


Figure 14: Sonic Log pour la perforation sélective [15]

2.2.2) Perforation orientée :

Les tests de laboratoire effectués dans le passé ont indiqué que la stabilité mécanique des cavités de perforation dépend de la direction de la perforation par rapport aux contraintes in situ. Cela a confirmé l'idée d'une perforation orientée pour minimiser les contraintes de cisaillement agissant sur la paroi des cavités de perforation. [1] Dans les régions où il existe un contraste important entre les contraintes in situ, la perforation doit être orientée dans la direction de la stabilité maximale.

Les tunnels de perforation orientés peuvent être utilisés pour ne perforer que dans une seule direction, retardant ou évitant la production de sable. Des perforations orientées dans la direction de la contrainte horizontale maximale dans un puits vertical augmentent la probabilité des tunnels de perforation plus stables. Ainsi, le contrôle conventionnel du sable n'est plus nécessaire (Najveet, 2009). [13]

Cette méthode peut être utilisée avec succès dans des champs soumis à des considérations économiques. La perforation orientée utilise un tir en phase à 180°. L'utilisation de ce tir est censé réduire le risque de production de sable en raison d'une réduction de probabilité de toucher la direction de perforation la plus défavorable.

La figure 15 montre le gain important en termes d'augmentation du drawdown critique d'une orientation correcte des perforations et la sensibilité de l'orientation du canon des perforations. Même avec de grandes incertitudes de l'ordre de 20°, les bénéfices sont toujours prononcés. [2]

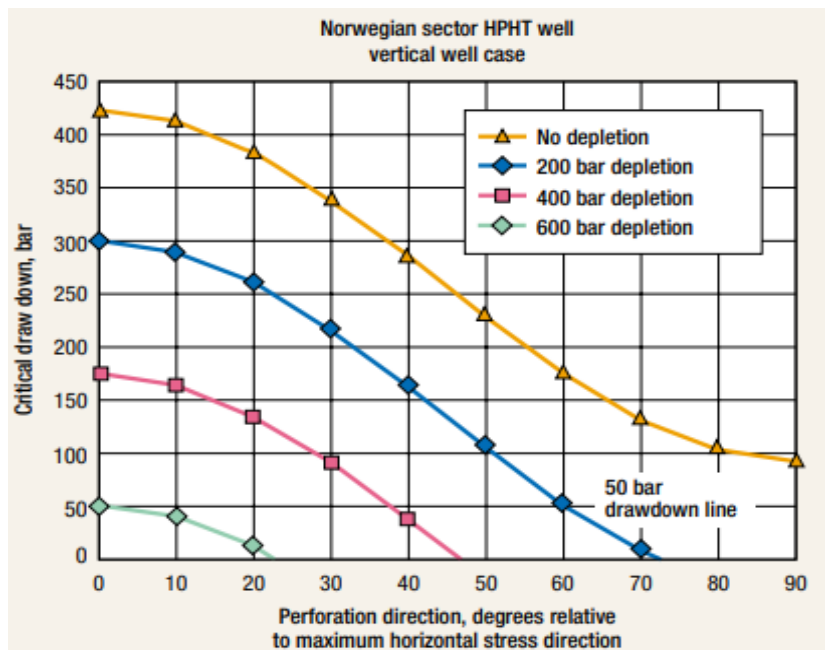


Figure 15: Effet des perforations orientées sur le drawdown critique [2]

2.3) Screens :

Toutes les formes de screen peuvent être utilisées dans un puits tubé ou ouvert, avec ou sans gravel pack, bien que chacune ait son environnement optimal. Les crépines peuvent également être utilisées dans un trou ouvert avec un liner préinstallé et pré-perforé pour fournir une protection supplémentaire à l'installation (Bellarby, 2009). La méthode de contrôle du sable la plus simple, la plus ancienne et relativement peu coûteuse qui utilise uniquement un screen pour limiter la production de sable (Navjeet, 2004).

Lorsqu'ils sont utilisés seuls comme dispositifs de contrôle du sable, les slotted liners ou les screens sont placés à travers l'intervalle productif et le sable de la formation est piégé mécaniquement sur une fente (slot) ou une ouverture. [1] Il existe plusieurs règles pour le dimensionnement des screens. Normalement, la largeur de la fente doit être dimensionnée pour être égale à la taille de grain du sable de formation au niveau le plus élevé de 10 %. Coberly a présenté ce critère de dimensionnement (Navjeet, 2004) dont la désignation D10 indique la taille du screen qui retiendrait 10 % des grains de sable.

Puisqu'il n'y a pas de zone de gravel pack autour du screen (Figure 16), tous les grains de sable produits se rassemblent autour de lui, bouchant les fentes du screen, ce qui entraîne une rupture de ce dernier. C'est pourquoi les considérations relatives à la qualité du sable sont extrêmement importantes lors de la conception d'une installation de standalone screen. En tant que dispositifs délicats, les screens sont des outils principalement fabriqués en acier inoxydable, spécialement conçus pour empêcher le sable de pénétrer dans les installations de complétion et de surface et de les éroder. [13]

Lors du choix d'un type de screen approprié et de ses dimensions pour la complétion les considérations suivantes doivent être prises en compte : la résistance aux endommagements, la taille de l'ouverture, la résistance au bouchage et à l'érosion, les tests en laboratoire avec des échantillons de sable de formation, l'expérience des travaux précédents. [6]

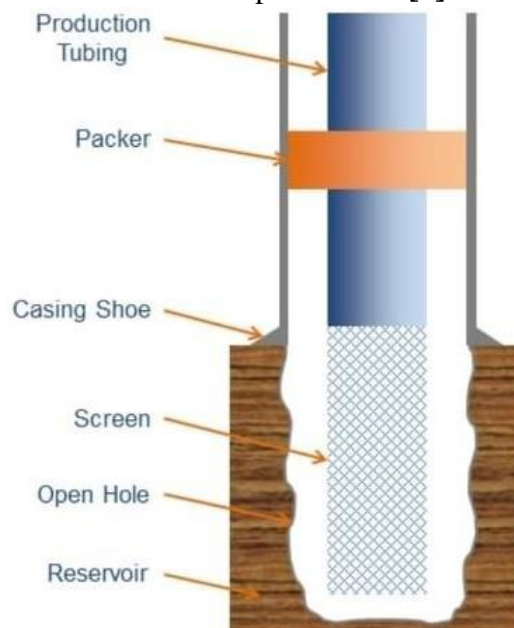


Figure 16:Screen complétion [21]

Il existe de nombreux types de screen dans le marché, mais les principales catégories sont résumées ci-dessous.

2.3.1) Slotted Liner :

Slotted liners sont fabriquées selon divers modèles, comme l'illustre la figure 17. La largeur de la fente varie généralement de 0,012 in à 0,250 in. Bien que slotted liners soient généralement moins coûteuses que wire-wrapped screens, elles ont une surface d'entrée plus petite et subissent des chutes de pression plus importantes pendant la production. Slotted liners se bouchent également plus facilement que les screens et sont utilisées lorsque la productivité du puits est faible.

Le modèle single-slotted staggered-row est généralement préféré car une plus grande partie de la résistance originale du tuyau est préservée. Le motif décalé donne également une répartition plus uniforme des fentes sur la surface du tuyau. Ce type de slotted liner est rainuré à la machine avec un nombre pair de rangées autour du tuyau. Il y a généralement un espace longitudinal de 15,2 cm (6 in. (15,2 cm) entre les rangées de fentes. [2]

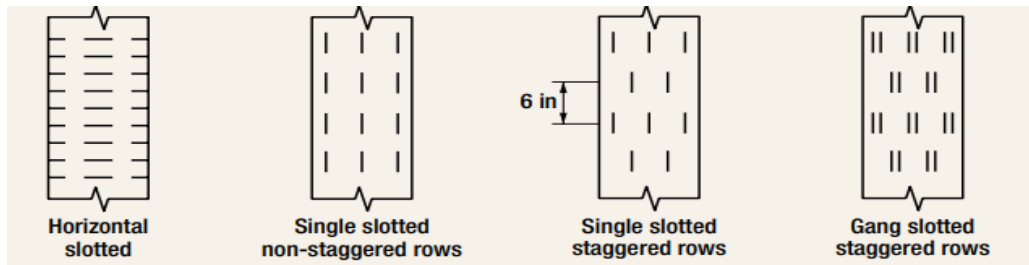


Figure 17: les géométries des Slotted-liners [2]

2.3.2) Wire-Wrapped Screens :

Wire-wrapped screens offrent une alternative pour retenir le gravier dans un anneau entre le screen et la formation. Ils sont souvent utilisés dans les complétions de gravel pack et les complétions standalone screens. L'avantage d'un wire-wrapped screen par rapport à un slotted liner est que la surface d'entrée est nettement plus grande (Figure 18).

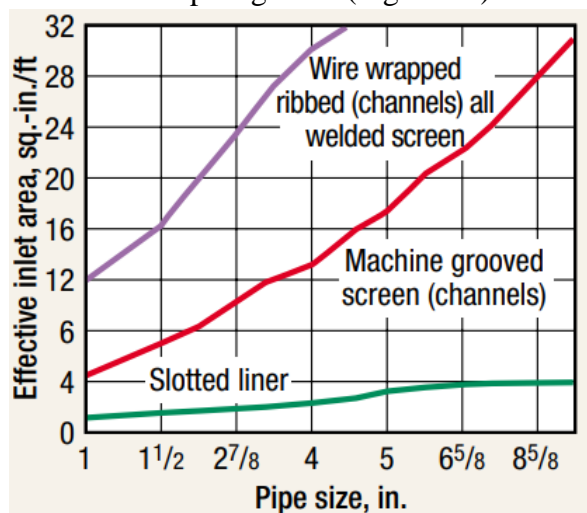


Figure 18: Comparaison des surfaces effectives d'entrée (20 gauge) [2]

Le terme "gauge" peut prêter à confusion. La gauge fait référence à la largeur de la fente ou à l'espacement des fils (slots) mesuré en millièmes de pouce. Par exemple, un screen de gauge 20 a un espace entre les fils de 0,020 in. (0,051 cm), et un slotted liner de gauge 20 a un slot de 0,020 in.

Le screen est constitué d'une enveloppe extérieure, qui est fabriquée sur des machines spéciales (wrapping) qui ressemblent à un tour. L'enveloppe est simultanément enroulée et soudée à des tiges longitudinales pour former une seule fente hélicoïdale.

La gaine est ensuite placée et soudée à chaque extrémité à une base de tuyau (pipe base) de support (contenant des trous percés) pour fournir un support structurel. Cette conception standard est générique et fabriquée par plusieurs sociétés. Un schéma de la construction de screen est illustré à la figure 18.

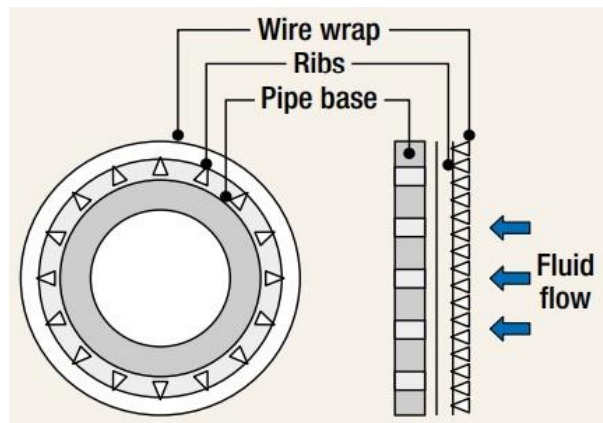


Figure 19: Schéma de construction de Wire Wrapped Screen [2]

2.3.3) Pre-packed Screens:

Leur construction est similaire à celle des wire-wrapped screens, mais avec deux screens. Pre-packed screens sont constitués d'un screen intérieur entouré d'une couche de gravier enrobé de résine constituant l'anneau et d'un tamis extérieur. La taille des fentes des screens est telle qu'elle empêche le gravier emballé entre les screens de s'échapper. Les Pre-packed screens offrent un certain degré de filtration en profondeur, et la porosité relativement élevée de plus de 30 % combinée à leur très grande perméabilité permettent des pertes de charge minimales (Bellarby, 2009).

Cependant, ils peuvent avoir tendance à se boucher en raison des sables plus fins qui s'incrustent entre les gorges des pores. Ils se bouchent plus facilement qu'un wire-wrapped screen (William et Joe, 2003), des exemples de différents prepacked screens sont présentés dans la figure 20. [13]



A

B

C

Figure 20: (A) Single-Screen Pre-pack (B) Slim-Pak (C) Dual-Screen [2]

2.3.4) Premium Screens :

Il s'agit du screen constitué d'une maille tissée (woven mesh) et d'une sorte d'enveloppe de protection. Il existe une variété de modèles différents chez les vendeurs. Les screens Premium sont constitués de plusieurs couches tissées, ce qui signifie qu'ils ont des ouvertures non uniformes. Ils sont plus fins que les pre-packed screens, mais légèrement plus épais que wire-wrapped screens. En raison de leur construction plus robuste, elles sont utilisées dans des environnements difficiles - puits horizontaux et à trou ouvert (Bellarby, 2009). La figure 21 présente un exemple du screen premium.



Figure 21: Exemple du Premium Screen

2.3.5) Expandable Screens :

Un expandable screen est composé de trois couches : (1) un tube de base à fentes (slotted base pipe), (2) un milieu filtrant et (3) une enveloppe protectrice extérieure. La conception de ce type de screen offre de nombreux avantages uniques. Il offre un diamètre interne plus grand que tout autre type de screens ; l'ESS élimine l'espace annulaire entre le screen et la surface durissable. Cela permet de stabiliser la formation et de minimiser le mouvement du sable.

La productivité des complétions des screens expansibles est généralement supérieure à d'autres méthodes de contrôle du sable, en particulier lorsqu'ils sont déployés dans un trou ouvert et offrent l'avantage d'un accès proche du puits pour les opérations d'intervention si cela s'avère nécessaire au cours de la vie du puits (Ayoola, 2009). Les ESS sont largement utilisés dans les complétions horizontales à trou ouvert. Leur plus grand inconvénient est leur susceptibilité à s'effondrer dans les formations comprimées. [13]

L'expansion du screen est réalisée en forçant un dispositif d'élargissement vers le bas à travers la corde ("string"). Les méthodes les plus courantes pour ce faire sont illustrées à la Figure 22 : l'expansion par cône à diamètre fixe, l'expansion rotative fixe avec des rouleaux et la méthode d'expansion hydraulique. [6]

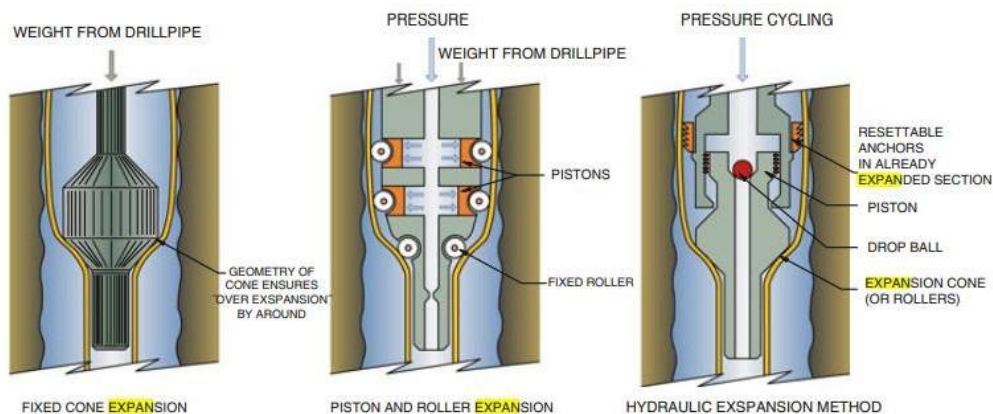


Figure 22: Les méthodes d'expansions d'expandable screen [6]

2.4) Gravel Pack :

Gravel Pack est la méthode la plus largement utilisée aujourd'hui pour contrôler la production de sable. Cette méthode est plus adaptée aux complétions initiales mais même les vieux puits qui produisent du sable peuvent également être des bons candidats pour ce type de traitement.

Elle peut être réalisée en tant que complétion à trou ouvert ou en tant que traitement de complétion à trou tubé. Ce système filtrant consiste à placer un pack de sable calibré afin de stabiliser la formation. Le gravillonnage comprend l'installation des équipements spéciaux conçus pour différents environnements capables de résister aux pressions et aux températures les plus élevées, et le traitement de la ou des zones avec une boue de gravier en suspension pompée simultanément dans un tubing pour combler l'annulaire de la formation/screen. De cette manière les grains de sable seront piégés à l'intérieur de la zone de massif filtrant ou derrière l'interface de la crépine ce qui permet d'obtenir une production des fluides sans sable. Une procédure typique de mise en place de gravillonnage, le massif filtrant se forme dans l'espace traité tandis que le liquide pompé pénètre dans le liner, traverse le wash pipe et remonte vers la surface. [6]

De manière générale, un gravier de taille optimale est placé contre la formation et maintenu en place par un tamis (screen). Il peut être placé soit dans un trou ouvert (massif filtrant externe) soit dans un casing perforé (massif filtrant interne). [18]

2.4.1) Open hole gravel pack :

Contrairement aux formations non consolidées, qui sont dans la plupart des cas complétées avec un casing perforé, le contrôle du sable en trou ouvert peut être appliqué dans le cas de formations consolidées susceptibles de produire du sable [6]

Le gravier à trou ouvert était utilisé à l'origine dans les puits déviés ou verticaux depuis les années 1990 et il a devenu après une forme courante de contrôle du sable spécialement pour sa forte productivité dans les puits horizontaux.

S'il est installé avec succès, il réduit la production des fines et empêche la formation de s'effondrer également. [19]

2.4.2) Cased hole gravel pack :

Les packs de gravier avec un casing et en particulier leur extension aux techniques de Frac Pack sont largement utilisés dans le golfe du Mexique et occasionnellement ailleurs. Dans autres environnements, comme la mer du Nord, ils ont rarement utilisé. D'après King et al 2003, ils sont les plus fiables pour le contrôle du sable surtout dans les régions où d'autres techniques sont difficiles (formation à faible perméabilité et haute teneur en fines) [19].

Parmi leurs avantages on cite :

- 1) La possibilité d'une isolation zonale en utilisant les packs empilés.
- 2) Le massif filtrant permet de traiter plusieurs zones l'une après l'autre en commençant par la zone la plus basse.
- 3) Ils sont plus adaptés aux formations moyennement à extrêmement non consolidées où l'existence de casing facilite la stabilisation de l'intervalle produit et l'obtention d'une zone propre. [6]

Tenant compte de leur installation, l'inconvénient est la complexité opérationnelle, la logistique et le temps considérables. Pour les longues sections du réservoir le coût et la complexité les rendent considérablement moins intéressants. [19]

En gros, il est primordial que l'espace annulaire et les perforations soient remplis du gravier afin d'obtenir un massif filtrant efficace. Lorsque ces espaces sont complètement comblés par le gravier, la perte de pression sera minimale. Egalement pour le tamis (screen), il doit être complètement recouvert par le gravier pour éviter la production de sable et l'érosion. [6]

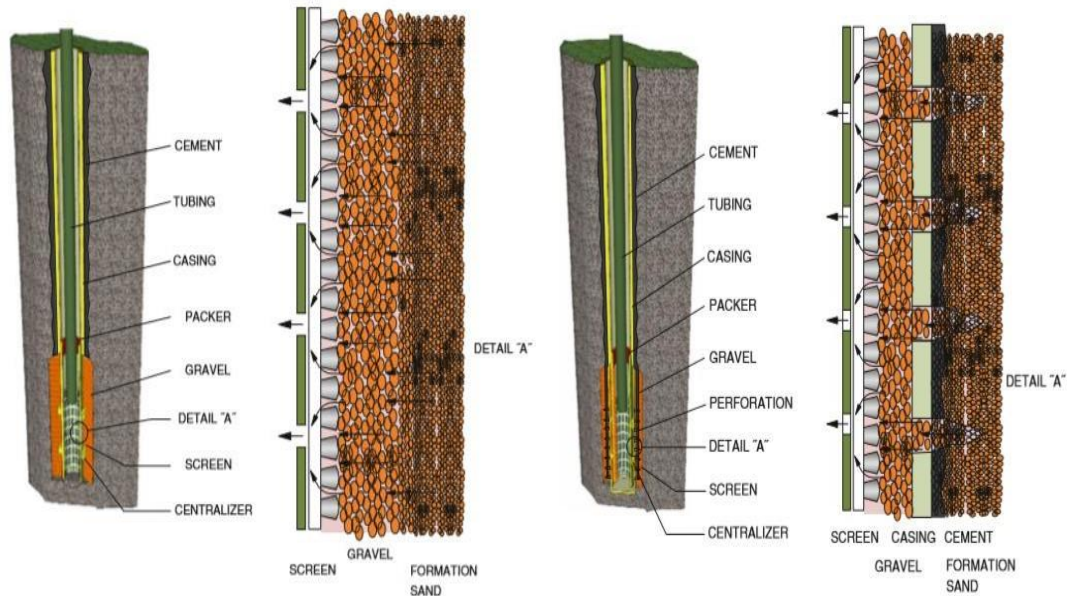


Figure 23: Tamis et gravier dans trou ouvert / Casing perforé [6]

2.4.3) Sélection du gravier :

Quelle que soit la méthode, ce qui est important est de dimensionner correctement le gravier, d'utiliser un fluide de complétion propre, des équipements de fond du puits optimaux notamment les tamis (screen). [6] Il faut se rendre compte que la taille du gravier doit être choisie de manière à empêcher le sable de formation sans réduire le flux vers le puits.

D'après Coberly et Wagner 1938 ; Tausch et Corley 1958, trois choses doivent être faites : (1) la sélection d'un gravier de taille appropriée pour maintenir le sable en place, (2) la mise en place correcte du gravier, et (3) la sélection des ouvertures du tamis (screen) pour contenir le gravier sélectionné. [6]

Les premiers travaux sur la taille du gravier ont été réalisés par Coberly et Wagner (1938), qui ont suggéré d'utiliser une taille de gravier égale à 10 fois le D10 du sable de formation. Hill (1941) a proposé de réduire cette taille à 8 fois le D10 mais des échecs ont encore eu lieu. Certains auteurs ont souligné qu'il faut tenir compte de l'équilibre entre l'arrêt de sable et le colmatage du gravier, parmi eux Saucier (1974) qui a proposé que la taille du gravier doit être entre 5 et 7 fois la taille médiane (D50) des particules et ça basant sur des expériences en laboratoire. Le critère de Saucier de six fois la D50 a été largement utilisé par la suite et a résisté à l'épreuve du temps. Par comparaison, le dimensionnement de tamis (screen) est plus facile et l'ouverture du tamis doit être taillée correctement de manière à empêcher tout le gravier de passer, généralement pas plus de 75% du plus petit diamètre du gravier. [19]

Par rapport au dimensionnement il faut mentionner aussi la distribution de la taille des grains qui est caractérisée par deux paramètres :

1. La médiane, le diamètre au point de 50% de la courbe, d50.
2. Le coefficient d'uniformité (Uniformity Coefficient) exprime l'uniformité de la distribution par le rapport entre le diamètre du point à 40% et le diamètre du point à 90% : $UC = D_{40}/D_{90}$.

Tiffin et al 1998 ont introduit deux nouveaux paramètres liés à la distribution des particules de sable basant sur l'expérience de champ et les expériences menées sur des carottes de différentes régions. On constate le paramètre de tri (d_{10}/d_{95}) et la fraction massique des fines (particules inférieures à $44 \mu\text{m}$). [1] La procédure à suivre pour caractériser efficacement une formation susceptible d'avoir un gravel pack comme méthode de prévention du sable se base sur la distribution granulométrique (Figure 24). La distribution granulométrique est déterminée en tamisant les grains de sable sur une série de tamis à maillage décroissant. Les résultats de l'analyse granulométrique sont représentés sous la forme d'un pourcentage pondéral cumulé retenu par le tamis en fonction de l'ouverture de maille (taille du grain) sur une échelle semi-logarithmique. Certains paramètres seront obtenus à partir du graphe : taille médiane D_{50} : C'est la taille des particules correspondant au 50ème percentile de l'analyse granulométrique ; coefficient de tri du sable ; coefficient d'uniformité. [18]

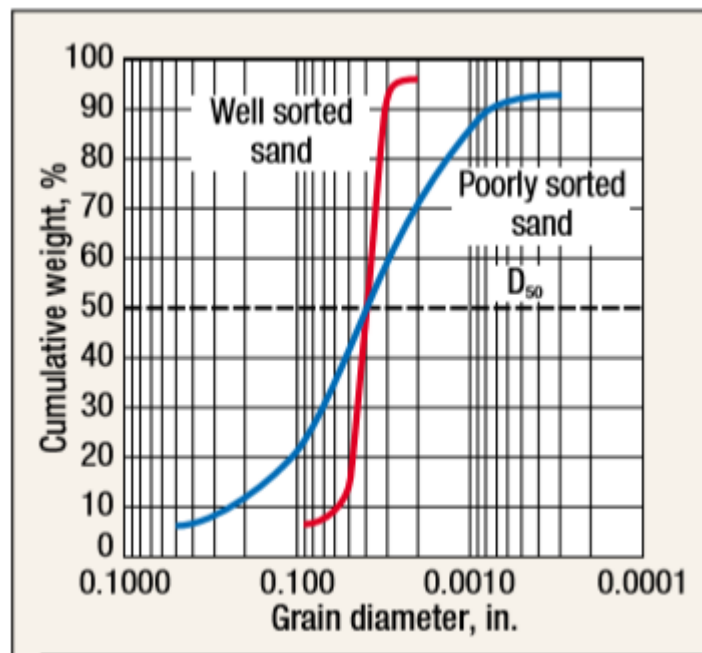


Figure 24: Analyse granulométrique du sable de formation [2]

2.5) Consolidation Chimique :

La consolidation chimique fait référence à une méthode qui utilise une résine liquide injectée dans la roche non-consolidée. Une masse poreuse et perméable se forme à partir de la polymérisation de résine par le biais d'un catalyseur. Avec la plupart des résines, il est primordial d'injecter un fluide de déplacement, c'est-à-dire un fluide qui achemine la résine en assurant qu'il n'est pas miscible avec elle (Talaghat et al. 2009). La consolidation chimique peut être subdivisée en deux catégories : (1) la consolidation in-situ, (2) le gravier enrobé de résine. [13]

2.5.1) Consolidation in-situ : La formation non-consolidée est traitée avec de la résine pour cimenter les grains de sable ensemble à leurs points de contact. Pour cela on injecte la résine sous forme liquide à travers les perforations, qui sont ensuite rincées par un catalyseur. Trois types de résine sont utilisés dans l'industrie : époxy, furannes et phénoliques [13]. Un traitement de consolidation réussi doit fournir la résistance supplémentaire requise tout en maintenant autant que possible la perméabilité de la formation.

Également, il faut assurer une couverture complète de toutes les perforations de la zone utile et pour cela ce genre de traitement est effectué sur des intervalles courts de 15 pieds ou moins. [5] Les traitements de consolidation in situ sont de deux types : les systèmes de séparation de phase et les systèmes de débordement. Toujours les systèmes de débordement contiennent un pourcentage élevé de résine par rapport aux systèmes de séparation de phase qui sont connues par une solution inerte avec de 15 à 25 % de résine active. Souvent, les stabilisateurs d'argile sont utilisés comme pré-flush puisque la concentration d'argile nuit à l'efficacité du processus de consolidation. En raison des difficultés à réaliser un placement efficace, du coût élevé des résines, ainsi que des problèmes de compatibilité et de contamination, l'utilisation de cette méthode de contrôle du sable est limitée. [13]

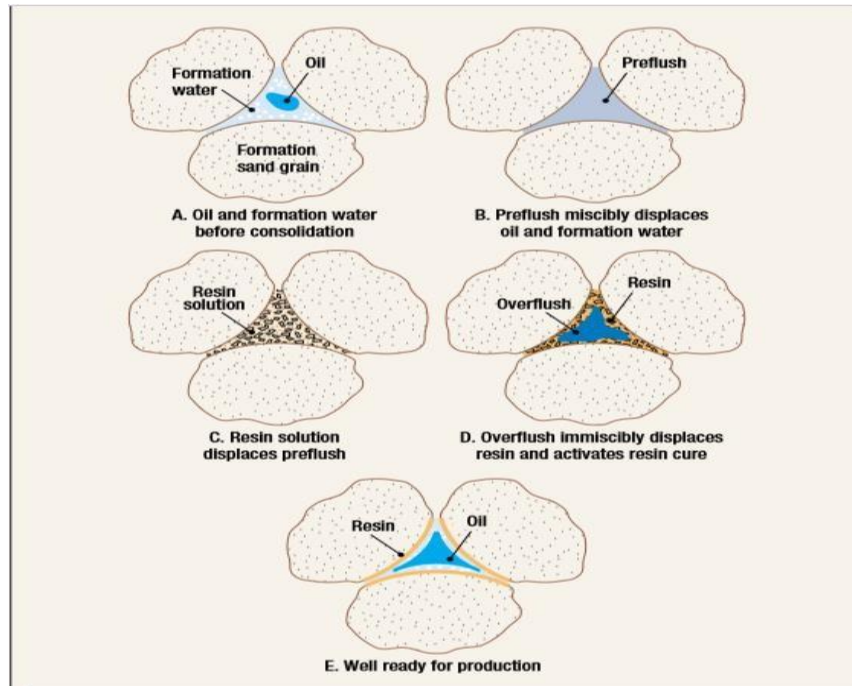


Figure 25: Les changements de saturation pendant la consolidation de la résine [2]

2.5.2) Gravier enrobé de résine :

Le gravier enrobé de résine est un sable de gravier à haute perméabilité recouvert d'une fine couche de résine. Une masse de sable consolidée se forme lorsque cette résine est exposée à la chaleur.

L'utilisation de gravier enrobé de résine comme technique de contrôle du sable implique le pompage du gravier dans le puits pour remplir complètement les perforations et le casing.

Le durcissement de résine en masse consolidée est dû à la température de fond du puits ou l'injection du vapeur. Le gravier consolidé va rester juste dans les perforations après l'avoir extrait du casing afin de laisser le puits non obstrué, alors le filtre perméable inclut dans les perforations va empêcher la production du sable.

Bien que simple dans son concept, l'utilisation de gravier enrobé de résine peut être une opération complexe. Tout d'abord, pour assurer la réussite de l'opération il faut que toutes les perforations soient complètement remplies de gravier enrobé de résine et que le gravier durcisse. Ensuite, le gravier enrobé doit durcir suffisamment en assurant une bonne résistance à la compression tenant compte que cette dernière dépend de la température et du temps. Pour obtenir une résistance à la compression élevée, des températures supérieures à 300°F sont nécessaires pendant plusieurs heures et c'est le cas juste des puits associés aux techniques de récupération thermique.

Malheureusement, il n'y a pas assez d'informations dans la littérature sur l'utilisation, le succès ou l'échec du gravier enrobé avec de la résine comme une méthode de prévention du sable. [1]

2.6) Frac Pack :

Un Frac Pack est essentiellement une forme de massif filtrant qui est pompé à un débit au-dessus de la pression de fracture afin de créer des fractures dans la formation. Similaire aux puits perforés avec du gravier (Cased Hole Gravel Pack), le gravier pompé sera acheminé vers ces fractures pour les compacter. Le gravier empêchera également les fractures de se refermer après la fin de l'opération du pompage. Comme les fractures permettent de contourner l'endommagement subis par la formation à proximité du puits et d'augmenter l'exposition au réservoir, cette technique de contrôle du sable sert à augmenter la production et fournit également des valeurs de skin négatives. Dans des régions comme le Brésil et le golfe du Mexique - où les Frac Packs sont devenus la méthode préférée de contrôle du sable dans les puits tubés et perforés- il est courant d'avoir des navires de stimulation dédiés aux frac packs. [20]

Le Frac Packing est une alternative qui devrait être envisagée lors de la planification du développement des champs qui produisent du sable. Cette technique augmente de plus en plus, on observe qu'au début des années 1990 c'était environ 100 jobs par an et maintenant dix fois plus chaque année. En Afrique de l'Ouest, environ 5 % des traitements de contrôle du sable sont des frac packs et au moins 3 % des puits en Amérique latine.

La différence entre cette méthode et le gravel pack et la fracturation hydraulique classique ça semble plus claire due au progrès continu en matière de conception de la stimulation, d'équipement de complétion des puits, des fluides de traitement et d'agents de soutènement. Les opérateurs américains appliquent désormais cette méthode de contrôle du sable pour compléter plus de 60% des puits offshore. [2]

2.6.1) Critères de sélection du Frac Pack :

Plusieurs critères de candidatures doivent être remplis afin d'appliquer des complétions de type frac and pack à un puits. Les puits présentant les caractéristiques suivantes pourraient être les meilleurs candidats (Meese et al. 1994) :

- a) Réservoirs avec des endommagements significatifs.
- b) Formations mal consolidées ayant des problèmes de fines et de migration du sable.
- c) Séquences laminées de sable et de schiste
- d) Réservoirs à faible perméabilité qui doivent être stimulés pour améliorer la productivité globale de la zone.

2.7) High Rate Water Pack :

Dans cette méthode, un fluide porteur de type saumure à faible viscosité est utilisé pour créer une fracture de longueur entre 5 et 15 pieds. Généralement, Les débits de pompage sont plus élevés que ceux des opérations classiques de gravillonnage, mais généralement inférieurs à ceux du fracpack. On constate un taux de pompage typique de l'ordre de 8 à 12 bpm.

Pour le but d'améliorer la capacité à traiter efficacement plusieurs sous-intervalles de sable en un seul traitement l'opération de HRWP peut être réalisée en multi-étages, et cela pour assurer une bonne couverture de gravier dans toutes les perforations. Le pompage du proppant au-dessus de la pression de fracturation est fait de sorte qu'une courte fracture (<15 pieds) est placée à travers la zone endommagée autour du puits. Il a été reporté que plus de 50 % des HRWP peuvent avoir des valeurs de skin inférieures à cinq, contre 85 % des frac packs. [2]

2.8) Comparaison entre HRWP and Frac Pack :

La principale différence entre ces deux techniques est la méthode de placement :

- Mise en place de Frac Pack : pompage d'un fluide viscosifiant avec une concentration de gravier progressive (généralement 1/2 à 15 ppa) et un débit d'injection 5 à 40 bpm.
- Mise en place de HRWP : pompage d'un fluide de complétion à faible viscosité et à faible concentration en gravier (généralement 1 à 2 ppa) avec un débit d'injection 5 à 15 bpm.

La méthode de HRWP présente un avantage par rapport aux packs de gravier mais ne permet pas d'améliorer la productivité comme les packs de fracturation. Une autre étude de comparaison confirme que les deux techniques permettent l'obtention d'un bon contournement des endommagements malgré la différence dans la géométrie des fractures créées sachant que le HRWP génère des fractures plus courtes que celles du Frac Pack. Même si dans la majorité des cas ces deux techniques sont interchangeables, en raison des différences dans les propriétés des réservoirs, les propriétés des fluides et les pratiques de pompage, certaines applications spécifiques sont mieux adaptées à l'une ou l'autre. La figure suivante énumère les avantages et les risques associés au frac pack et au HRWP. [2]

Tableau 1: Avantages du Frac Pack et du HRWP [2]

| Frac pack | HRWP |
|--|---|
| Avantages | |
| Des longueurs de fracture beaucoup plus importantes <ul style="list-style-type: none"> • Important pour les formations à faible perméabilité | Amélioration de la qualité du gravel pack <ul style="list-style-type: none"> • Meilleur remplissage des perforations • Meilleur remplissage annulaire |
| Une croissance verticale accrue des fractures <ul style="list-style-type: none"> • Important pour les longs intervalles de formations faiblement stratifiées | Mécanisme de transport du gravier <ul style="list-style-type: none"> • Le gravier concentré près du puits de forage • Ne pas se déplacer pendant les opérations multi-étage |
| Une concentration plus élevée de proppant à proximité du puits <ul style="list-style-type: none"> • Permet une meilleure pénétration de gravier dans la formation | Fluide moins efficace <ul style="list-style-type: none"> • Moins de risque de développement de fractures hors zone |

Tableau 2: Les risques du Frac Pack et du HRWP [2]

| Frac pack | HRWP |
|--|---|
| Risques | |
| Impossibilité d'obtenir un TSO <ul style="list-style-type: none"> • Nécessité d'une analyse mini-frac et d'une redéfinition des tâches sur place | Incapacité à se fracturer <ul style="list-style-type: none"> • Formations à haut kh modérément endommagées avec des fluides de réservoir à viscosité relativement faible • Traitement de formations à haute température avec une saumure de faible densité • Puissance hydraulique limitée |
| La rupture du système de gel <ul style="list-style-type: none"> • Combinaison gel/casseur appropriée pour la température de la formation | Croissance en hauteur insuffisante <ul style="list-style-type: none"> • Sable stratifié de plus de 50 pieds d'épaisseur |
| Croissance défavorable de la fracture <ul style="list-style-type: none"> • Ne pas essayer de traiter des sables à intervalles multiples • Dangereux à proximité d'un contact huile/eau ou gaz/pétrole | Longueur de fracture insuffisante <ul style="list-style-type: none"> • Formations à faible perméabilité • Dommages "profonds" avérés |
| Faible croissance annulaire <ul style="list-style-type: none"> • Ne pas traiter les intervalles avec des déviations >60° • Peut être traité en réduisant la viscosité du fluide à la fin du travail | Réduction de la conductivité des fractures <ul style="list-style-type: none"> • Possibilité d'un skin plus élevée dans les complétions à haut débit |

Chapitre 3 : Sélection des méthodes de contrôle du sable

La consultation d'un grand nombre de publications a donné lieu à de nombreuses études de cas publiées dans le monde entier sur le contrôle du sable dans les puits pétroliers. Ces études de cas ont permis de lister les méthodologies individuelles utilisées pour le choix d'une méthode de contrôle spécifique pour un puits particulier. Ces études tiennent compte des différents facteurs lors du choix de la technique appropriée pour leur puits, et puisque la recherche continue pour améliorer les méthodes existantes une variété de facteurs remarquables devrait être considérée lors du choix d'une technique donnée.

Suivant ces études de cas en tant que des modèles pour les appliquer dans d'autres réservoirs tenant compte de leur variété peut prendre beaucoup de temps lors de l'application dans la pratique. D'après ça Il semble qu'il n'existe pas de guide pratique simplifié indiquant les paramètres clés fondamentaux à prendre en compte lors du choix d'une technique de contrôle du sable. Puisqu'il a été constaté que les critères et les paramètres clés pris en compte dans le choix d'une technique sont les mêmes, il est nécessaire d'avoir un outil universel simplifié que les opérateurs et les sociétés de services puissent utiliser dans la pratique lorsqu'ils décident d'implémenter une méthode de prévention du sable. [22]

La sélection d'une solution de complétion pour un puits donné est un processus complexe qui peut donner lieu à plusieurs solutions acceptables ou, à l'inverse, à des alternatives parmi les quelles il faut choisir la solution optimale. De nombreux défis doivent être pris en compte lors de la sélection à cause de la diversité des caractéristiques du réservoir ainsi que la géomécanique locale et la façon dont elles interagissent. Bien que les caractéristiques du réservoir soient des considérations importantes, il faut tenir compte aussi des performances historiques de ces options de complétion. Une évaluation de la fiabilité est essentielle dans le processus de sélection des systèmes de contrôle du sable. [23]

D'après plusieurs études de cas, des données pertinentes qui devraient être prédéterminées avant la mise en œuvre d'une technique ont été identifiées et se sont résumées dans le tableau ci-dessous : [22]

Tableau 3: Liste des critères de conception du contrôle du sable [22]

| Critères de conception du contrôle du sable à prendre en compte avant la mise en œuvre | |
|---|---------------------------------------|
| No | Critères |
| 1 | Les conditions du réservoir |
| 2 | Les changements lithologiques |
| 3 | Les données du puits |
| 4 | Caractérisation du sable de formation |
| 5 | Evaluation économique |

3.1) Les méthodologies utilisées dans le processus de sélection :

Nous avons déjà discuté du fait qu'il existe plusieurs études de cas concernant la sélection du contrôle du sable, donc dans cette section nous allons présenter les articles importants qui ont travaillé sur le processus de sélection en suivant un ordre chronologique :

3.1.1) Approche de Tiffin, 1998 :

Les auteurs de cet article ont proposé un critère qui se concentre principalement sur la distribution de la taille des grains de sable. La méthodologie utilisée se base sur la distribution des tailles des particules en termes de sorting coefficient, un coefficient qui représente une très large gamme de tailles entre la taille minimale et maximale des grains. L'étude évalue également les formations à très forte teneur en fines. En terme pratique, ils utilisent l'expérience de terrain et les tests laboratoires sur des carottes depuis différents réservoirs du monde. Cette méthode développe le critère de Saucier couramment utilisé pour la sélection de la taille des tamis et du gravier en utilisant trois paramètres clés D50, D40/D90, D10/D95 et la mobilité des particules fines (moins de $44\mu m$) pour la sélection d'une méthode de contrôle du sable. Le D10/D95 est un nouveau critère et est utilisé pour distinguer la variation entre les plages de taille du sable de formation. En gros, l'article présente une méthode solide pour la sélection des techniques de contrôle du sable en se concentrant sur quatre paramètres :

- D50 : la taille médiane des particules correspond au 50% dans la courbe de l'analyse granulométrique.
- D40/D90 : coefficient d'uniformité
- D10/D95 : coefficient de tri (Sorting coefficient)
- Teneur en fines moins de $44\mu m$

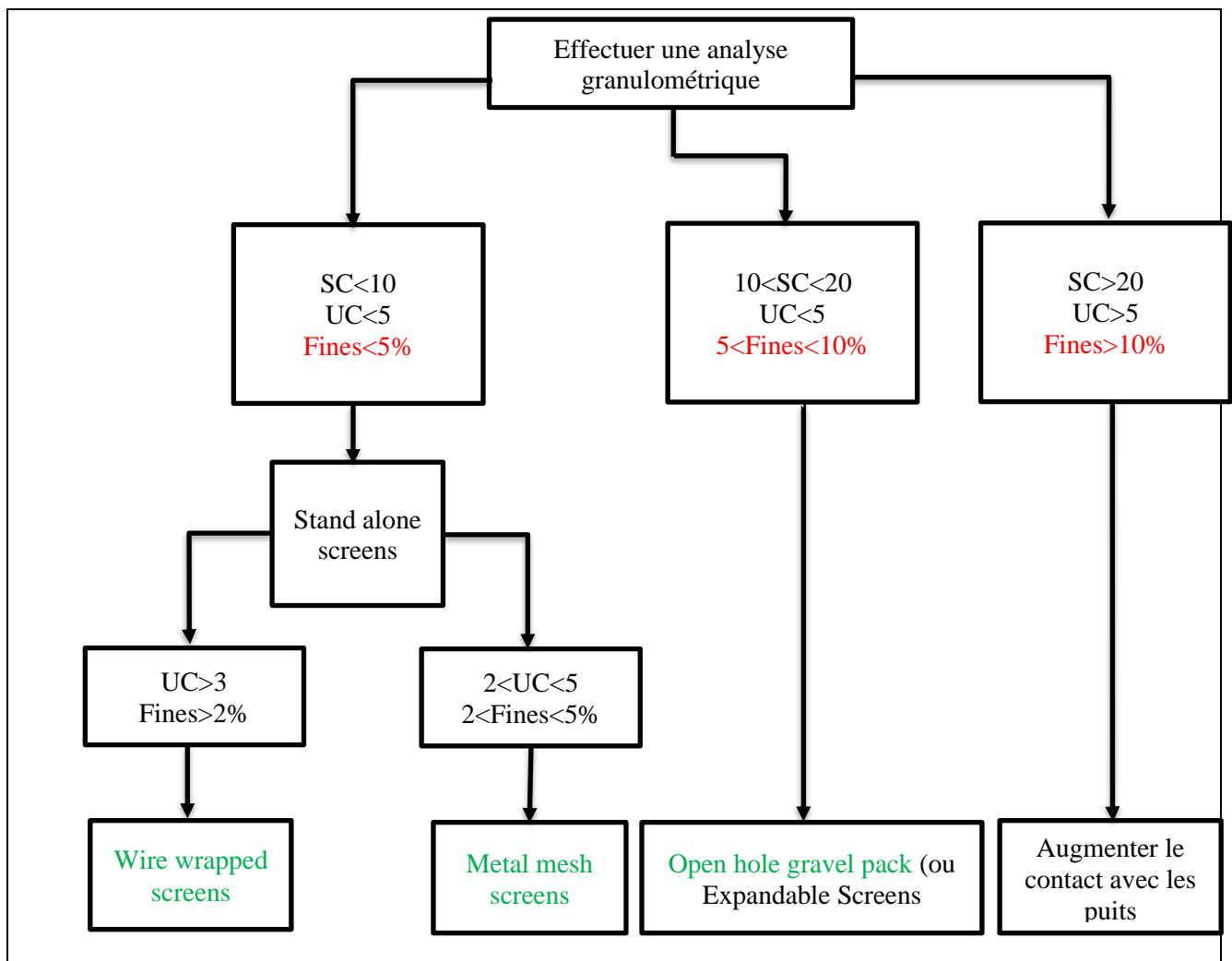


Figure 26: Critères de sélection [24]

3.1.2) Approche de Price-Smith, 2003 :

L'objectif de cet article est de proposer une méthode de conception pour la sélection d'une complétion horizontale à trou ouvert pour le contrôle du sable. Ce travail fournit un ensemble unifié de directives basées sur l'expérience et l'expertise technique de 5 opérateurs et d'un fournisseur de services de complétion.

Les auteurs abordent une méthodologie généralisée et unifiée pour traiter trois stades importants : quand installer le contrôle du sable, ce qu'il faut installer comme méthode de prévention et comment l'installer dans les complétions horizontales à trou ouvert. Ils considèrent certains facteurs critiques comme l'architecture du puits, la lithologie du réservoir, les propriétés pétrophysiques et la fiabilité de l'équipement de contrôle du sable. En détail, leur travail consiste à choisir entre la méthode de gravel packing (GP) et standalone screens (SAS) en se focalisant sur l'expérience de terrain et les données expérimentales. Cette méthodologie est validée par dix cas d'études détaillées portant sur différents types de complétions horizontales à trou ouvert. [25]

En conclusion, les directives doivent être utilisées avec prudence jusqu'à ce qu'elles soient validées et/ou affinées par de nombreuses autres études de cas. Sur la base de la méthodologie validée par les études de cas, les résultats sont résumés ci-dessous dans le tableau 4 :

Tableau 4: Contrôle du sable recommandé selon la méthodologie de Price-Smith [26]

| Median Grain Size (D50) | Uniformity (D40/D90) | Teneur en fines | Débit de production | Méthode de contrôle du sable recommandée |
|--|----------------------|-----------------|---------------------|--|
| >75 μm | Elevé | Faible | Faible | Wire Wrapped Screen |
| >75 μm | Modéré | Modéré | Faible | Prepacked Screen or Premium Screen |
| >75 μm | Elevé | Faible | Elevé | Prepacked Screen or Premium Screen |
| Utilisez le pack de gravier lorsque $D50 < 75\mu\text{m}$. <ul style="list-style-type: none"> ▪ GP-Shunt Pack est recommandé pour les environnements à faible NTG, schisteréactif, OBM et faible gradient de fracture. ▪ GP-Water Pack est destiné aux environnements à NTG élevé, à l'absence de schisteréactif, à l'OBM et à un gradient de fracture élevé. | | | | |

3.1.3) Approche de Mathisen, 2007 :

L'objectif de cet article est de recommander une pratique pour la sélection des screens utilisés dans les complétions Stand alone screens (SAS). L'étude se base sur l'essai et le classement de différentes conceptions de screen et un processus de qualification des fluides en s'assurant que le fluide de complétion est compatible pendant l'installation du screen.

La méthodologie proposée se subdivise en trois étapes :

- 1) Les essais de rétention du sable et de colmatage ont été utilisés pour classer les modèles de tamis. Les données pour les essais ont été obtenues à partir de l'analyse de la distribution des particules de sable (PSD) en prenant juste les portions de la formation avec des faibles coefficients d'uniformité et de sorting. Des tamis Wire- Wrapped, Premium and Expandable ont été utilisés pour les tests.
- 2) L'évaluation géologique de l'hétérogénéité du réservoir a été utilisée pour identifier les sections de schiste entre la carotte et la trajectoire prévue du puits.
- 3) L'utilisation de technologie de contrôle de débit (ICD) pour minimiser le débit annulaire

En conclusion, Les études sur la rétention du sable et le colmatage montrent que les screens Premium présentent la plus faible chute de pression, associée à une rétention du sable et une perméabilité élevée. Les techniques de contrôle du débit (ICD) sont recommandées pour les formations comportant de longues sections de schiste et/ou une forte teneur de fines. Cet article a souligné que la méthode de complétion par SAS peut avoir des applications plus larges que celles reconnues jusqu'à présent. [26]

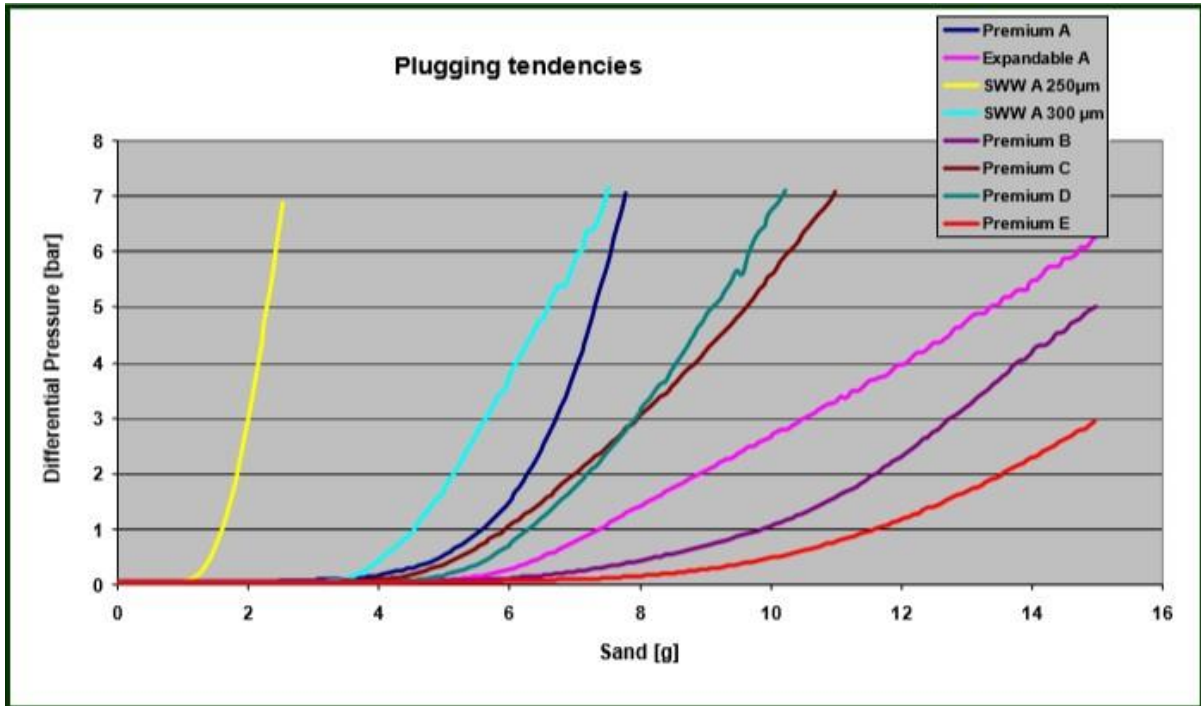


Figure 27: Résultats des essais de rétention de sable de différents tamis à sable [28]

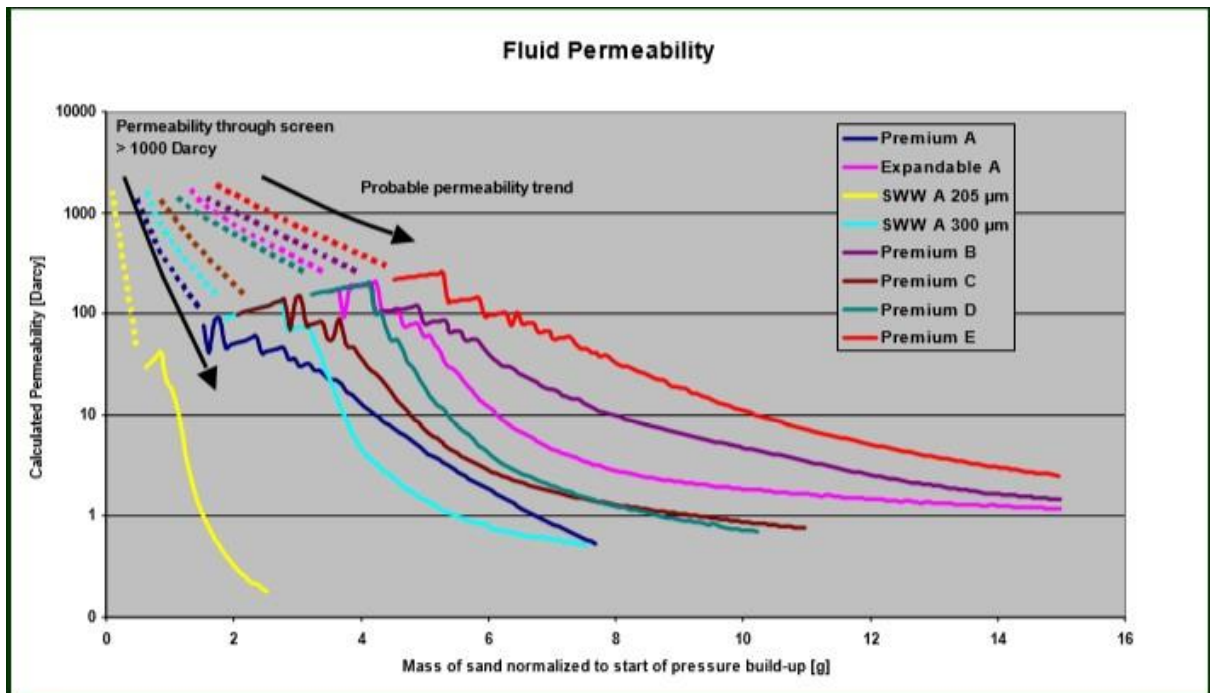


Figure 28: Résultats des essais de rétention de sable de différents tamis à sable en termes de perméabilité [28]

3.1.4) Approche de Farrow, 2004 :

Les auteurs ont développé une méthodologie de sélection destinée à évaluer et classer les techniques de contrôle du sable disponibles pour les champs du North West Shelf (NWS). La méthode est une combinaison d'organigramme et de matrice de sélection des méthodes de prévention du sable.

L'organigramme est utilisé initialement pour déterminer le type du puits (existant ou nouveau), l'emplacement (un puits sous-marin ou avec plateforme) et pour traiter aussi les déviations dans la section du réservoir. Ensuite, la sélection finale sera par le biais de système matriciel, cette matrice de sélection utilise un système de feux de circulation pour représenter les résultats en se basant sur un système coloré : vert = pas inquiétude, jaune = une certaine inquiétude, orange = une inquiétude majeure, rouge = exclu. Cette matrice de couleur est sélectionnée en fonction d'un classement probabilité*conséquence.

Les critères de sélection utilisés sont classés sous forme des catégories :

- Caractéristiques de la rupture du sable
- Distribution de la taille des particules
- Conditions du puits et schistes
- Caractéristiques du réservoir et du fluide du réservoir
- Performances de production
- Exigences en matière d'intervention et de gestion des réservoirs
- Fiabilité
- Coût et risque d'installation

En conclusion, les résultats des 3 études de cas mises en œuvre ont montré que :

- Cas 1 et 2 : puits de plate-forme existant– **le gravel pack est la technique préférée.**
- Cas 3 : nouveau puits de développement sous-marin – **les standalone screens sont l'option préférée.**

Enfin, nous pouvons dire que cette méthodologie de sélection a permis une évaluation et un classement transparents des options de contrôle du sable. [27]

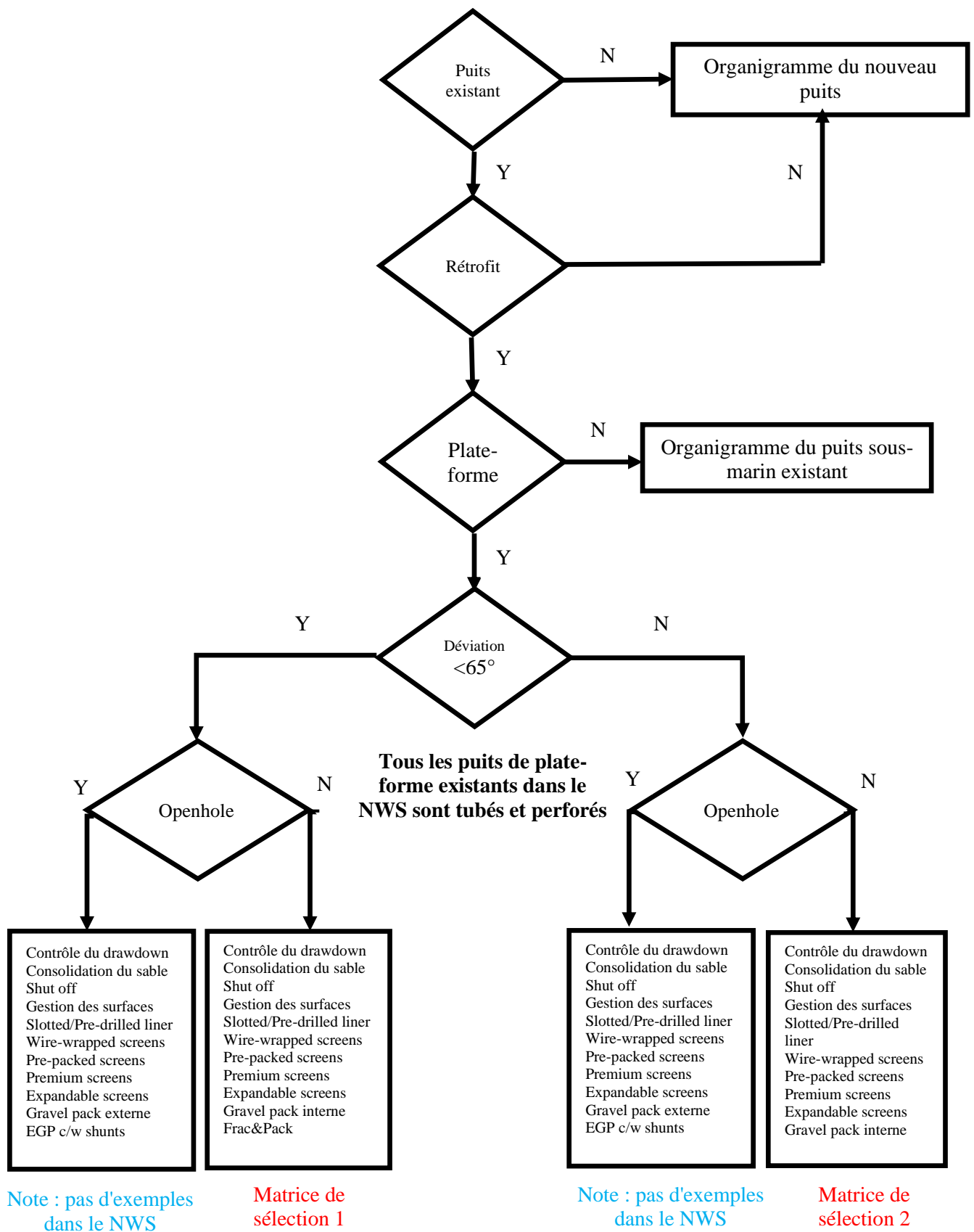


Figure 29: L'organigramme de sélection initial [27]

3.1.5) Approche de Michael Bryne, 2010 :

Les auteurs mettent en évidence la nouvelle définition des fines. Ceci afin de garantir la sélection d'un contrôle du sable optimal et adapté. Un processus de sélection du contrôle du sable modifié pour tenir compte des nouvelles définitions est proposé dans cet article.

Ils passent également en revue les publications antérieures sur les fines et les directives existantes en matière de sélection de prévention du sable.

Une revue initiale montre les guides conventionnels utilisés dans la détermination des systèmes de contrôle du sable : Schwartz (1969) puis Tiffin et al. (1998). Les deux méthodologies ainsi que quelques directives supplémentaires ont été consolidées et intégrées.

Les définitions précédentes des fines ne sont pas adaptées à la sélection du contrôle du sable. Elles ne sont pas suffisamment spécifiques et ne peuvent s'appliquer qu'à certains réservoirs. Dans un réservoir à grain très fin, par exemple, certains grains de la roche peuvent être proches ou inférieurs à 45 microns. Ils pourraient être classés comme fines, mais n'en sont pas vraiment.

Compte tenu de la façon dont les fines peuvent se déplacer et boucher les pores d'une roche et être produites au-delà de la roche et boucher l'annulaire/le gravier/les screens, il est proposé que les fines soient considérées comme "la partie de la roche qui peut se déplacer à travers les pores de la roche intacte et non perturbée".

Leur potentiel de déplacement dépend de nombreux facteurs : le type de particule et sa mouillabilité, l'adhérence à la surface de la roche, le type de fluide mobile, la vitesse du fluide et la taille de la gorge des pores. Sachant que ces particules peuvent ou non sortir de la roche pour pénétrer dans le puits et potentiellement dans l'équipement de surface, leur comportement est évalué de manière plus efficace par des tests de migration des fines.

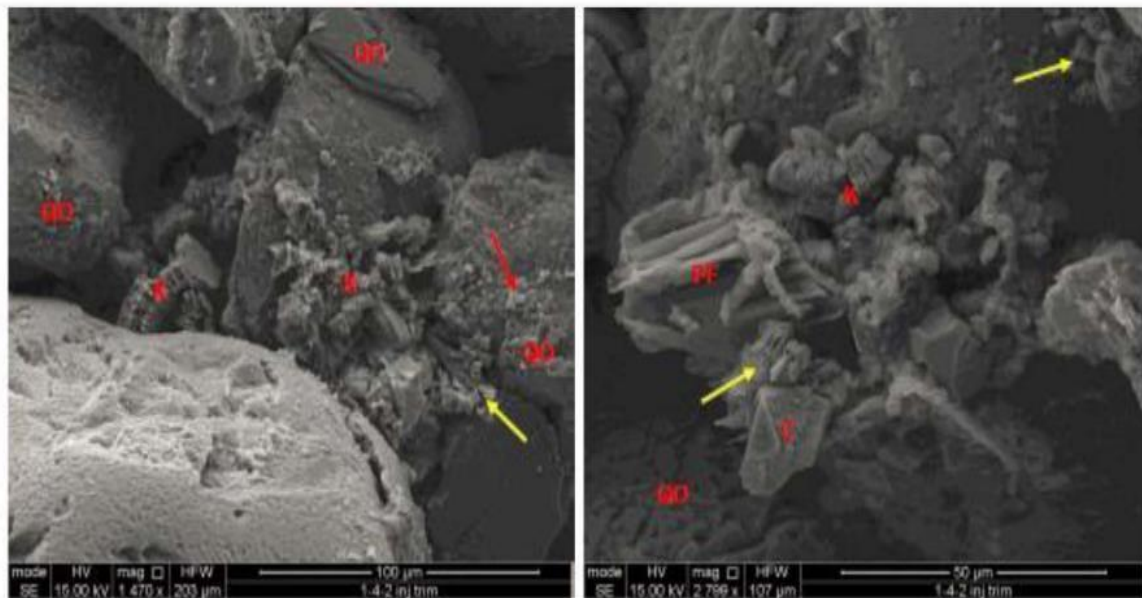


Figure 30: Test de migration des fines [29]

La figure 30 présente EFM (Extended Fine Migration) et les Micrographies SEM (Scanning Electron Microscopy) de la carotte. La micrographie SEM de gauche montre de la kaolinite (K), la présence de deux tailles de cristaux différentes dans une zone (flèche jaune) suggère une possible migration.

La micrographie de droite montre de la kaolinite qui a probablement migré dans des positions non naturelles (flèches jaunes).

L'impact potentiel d'une mauvaise uniformité de la formation peut se réorganiser dans un espace confiné (tel que l'espace annulaire entre le trou ouvert et un screen de sable, ou dans le tunnel de perforation d'un massif filtrant interne). Il est probable que la perméabilité de cette zone réarrangée soit inférieure à celle de la matrice non perturbée. Cela peut potentiellement nuire à la productivité et à l'intégrité du puits. Le nom de ce phénomène est parfois appelé "autopack".

En appliquant la définition proposée, ils permettent de prendre en compte tous les réservoirs et toutes les fines. Dans une roche à grain très fin, l'utilisation de la définition conventionnelle "moins de 45 microns" pourrait conduire à une sélection inappropriée du contrôle du sable. Une roche à grain fin peut être bien triée (well sorted) et présenter un potentiel minimal de colmatage des tamis par la partie de la roche qui peut se déplacer à travers les pores, mais peut présenter une fraction "inférieure à 45 microns" très élevée. De même, une roche à très gros grains peut avoir une teneur en fines mobiles importante, mais une grande partie de ces fines peut avoir une taille supérieure à 45 microns. Notez cependant qu'il peut ne pas être possible de dimensionner un screen autonome (SAS) pour les sables très fins et qu'il peut être nécessaire de sélectionner un gravel pack sur cette seule base, comme décrit dans (Price- Smith, C. et al, 2003) où des sables avec un $d_{50} < 75$ microns ont été identifiés pour le GP.

L'importance d'une meilleure classification est de mieux identifier la source des problèmes potentiels, alors les auteurs précisent non seulement le comportement des fines (c'est-à-dire qu'elles se déplacent) mais aussi sur leur impact (c'est-à-dire qu'elles bouchent la formation, les screen ou les installations). Cette clarification est illustrée dans la page suivante (Figure 31).

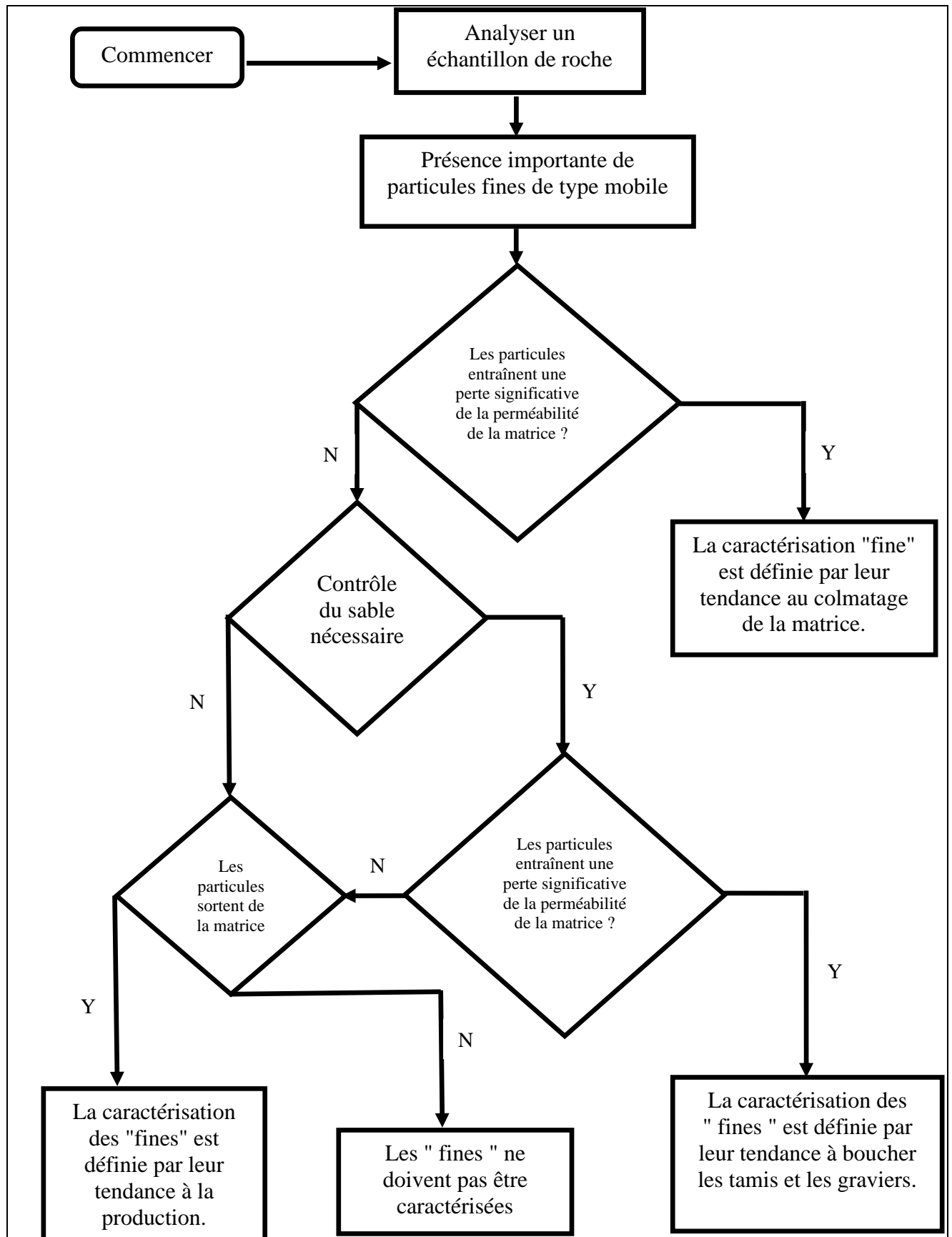


Figure 31:Représentation de la manière dont les particules fines peuvent être définies sur la base de leur comportement et de leur impact [29]

Suite à la définition modifiée des fines et à la nouvelle approche du sable décrite ci-dessus,

un processus de sélection amélioré est proposé et représenté dans la Figure 32 ci-dessous :

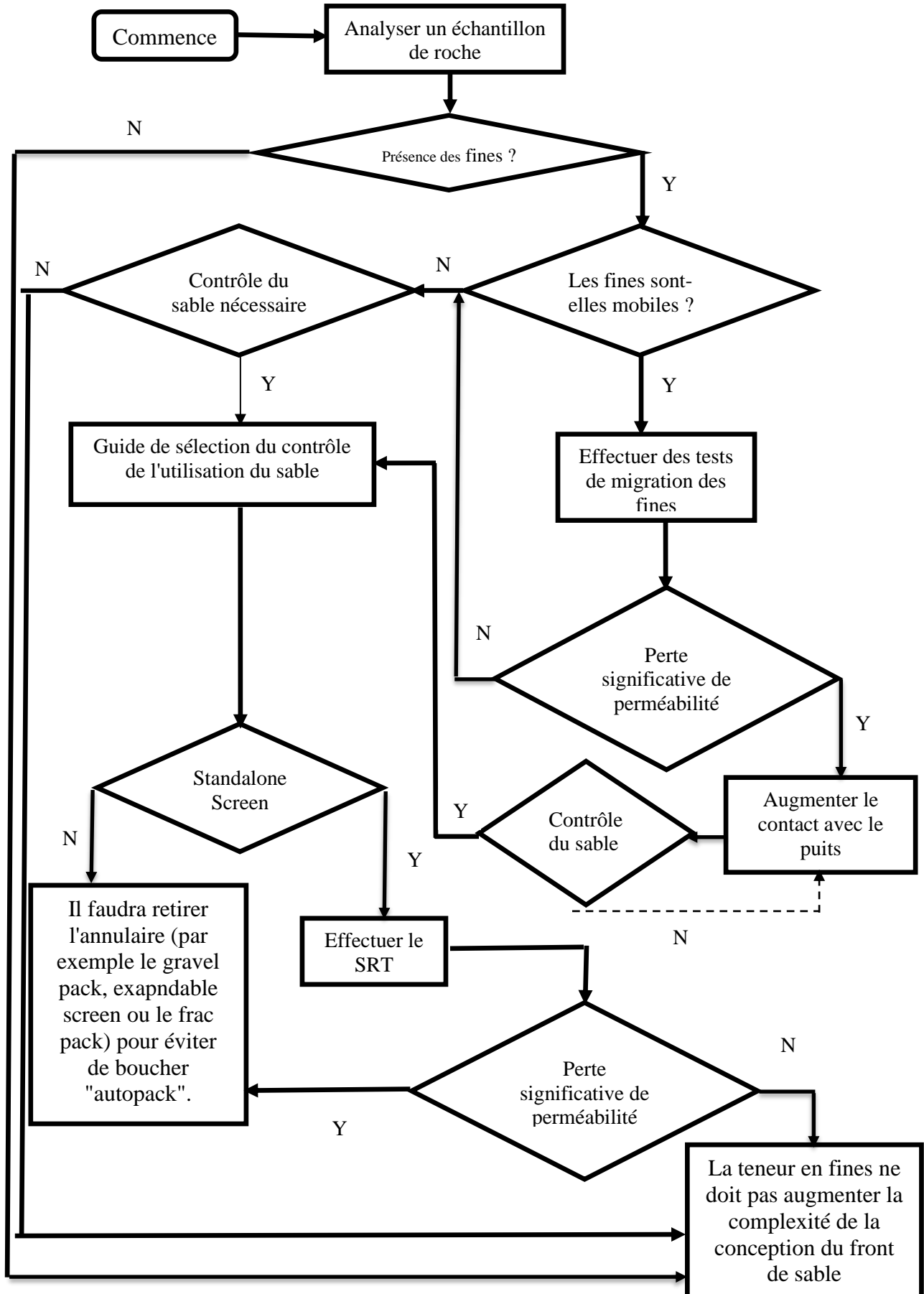


Figure 32:Processus de sélection pour le contrôle du sable [29]

Étant donné l'incertitude de la mesure de la taille des particules et de la définition des fines, l'essai de rétention de sable (SRT) est apparu comme une technique utile pour tester le sable de formation contre le milieu de contrôle du sable afin de déterminer l'efficacité de différentes solutions. Il s'agit de tester l'efficacité de la rétention du sable tout en n'entraînant pas de chute de pression importante.

L'augmentation du contact entre le puits et la formation diminue le débit des fluides produits pour un taux de production donné. L'avantage de ce phénomène est double : il réduit la vitesse d'écoulement à travers la matrice de la formation et diminue le risque ou l'étendue du mouvement des fines. La réduction du taux de flux signifie également que toute réduction de la perméabilité (due au déplacement des fines) aura moins d'impact sur la productivité.

L'augmentation du contact avec le puits est obtenue par des techniques telles que la fracturation, l'augmentation de l'angle du puits dans le réservoir. La plupart de ces technologies sont coûteuses, d'où l'importance de prendre en compte le comportement et l'impact attendus des "fines".

3.1.6) Approche de Keng Seng Chan, 2013 :

Ces auteurs présentent une approche technique pour déterminer où et quand le contrôle du sable sera nécessaire dans l'offshore malaisien sur la base d'une évaluation géomécanique de la pression de drawdown critique (CDP) pour le type de puits, la configuration et la complétion sélectionnés. La décision de mettre en œuvre une méthode appropriée de contrôle du sable peut être prise en comparant la CDP avec la pression de drawdown minimale (MDP) requise pour atteindre le taux de production cible. Le choix de la méthode de contrôle du sable était basé non seulement sur la distribution de la taille des particules, la durée de vie du puits et le mode de production du puits, mais aussi sur la maximisation de la récupération du pétrole et du gaz par puits.

Les objectifs de la mise en place de contrôle du sable pour le champ malaisien étaient les suivants :

- Réduire les coûts d'achèvement du contrôle du sable de plus de 50%
- Favoriser un changement dans la conception des puits afin de développer des solutions rentables pour les réservoirs de grès.
- Déterminer sélectivement où, quand et comment le contrôle du sable est nécessaire dans les réservoirs en développement.
- Permettre le placement des puits et la conception des trajectoires de pénétration dans les réservoirs afin de maximiser la pression d'abaissement critique sans sable (CDP).
- Optimiser l'utilisation des méthodes coûteuses de contrôle du sable.
- Pour estimer le rabattement de la pression d'érosion critique du sable.

La comparaison du profil MDP et du profil CDP devrait aider à déterminer la nécessité d'une complétion de prévention du sable par un critère évident de $MDP > CDP$. Avec tous les puits prévus évalués par cette méthode, il est possible de déterminer techniquement où et quand le contrôle du sable est nécessaire pour les réservoirs en développement.

Par la suite, la nature et les propriétés du sable doivent être réexaminées et les distributions granulométriques dans les intervalles de production doivent être mesurées. En fait, lorsque les sables sont bien triés et uniformes, un écran autonome (standalone screen) dont la conception a été testée en laboratoire peut être utilisé. Lorsque le tri du sable est bon et que la teneur en fines est relativement faible, un écran à mailles tissées (woven mesh) plus élaboré peut être utilisé. Lorsque les sables sont mal triés et non uniformes avec une quantité importante de fines, le choix de contrôle du sable doit se porter sur les complétions avec gravel pack ou avec frac pack.

Tableau 5: La sélection des méthodes de contrôle du sable [30]

| D10/D95 | D40/D95 | Sub 325 Mesh | Recommandation d'une méthode de contrôle du sable |
|---------------|-------------------|--------------|---|
| Triage | Uniformité | Fines | |
| <10 | <3 | <2 | Bare Screens |
| <10 | <5 | <5 | Bare Woven Mesh Screens (WWS) |
| <20 | <5 | <5 | Gravel pack avec gros gravier |
| <20 | <5 | <10 | Gravel pack avec gros gravier et WWS |
| >20 | >5 | >10 | Frac Pac, ou HZW, ou multilatérales |

- Ouverture du tamis D10 en pouces ou mm ou microns au-dessus de laquelle 10% du sable est retenu
- D10/D95 rapport de tri entre les tailles de particules minimale et maximale
- Rapport des dommages causés par les screens D40/D95
- Maille inférieure à 325 quantités de particules inférieures à 44 microns (limons, fines et argiles)

Dans le cas du développement d'un champ en eaux profondes en Malaisie, selon les lignes directrices suggérées pour le choix d'une méthode de contrôle du sable dans le Tableau 5, le choix devrait se porter sur le CHFP ou l'OHGP. La sélection des méthodes de gravel pack telles que CHFP, CHGP et OHGP doit être basée sur la maximisation de l'IP et la minimisation de skin qui dépendent certainement de l'angle de déviation du puits et du coût équilibré.

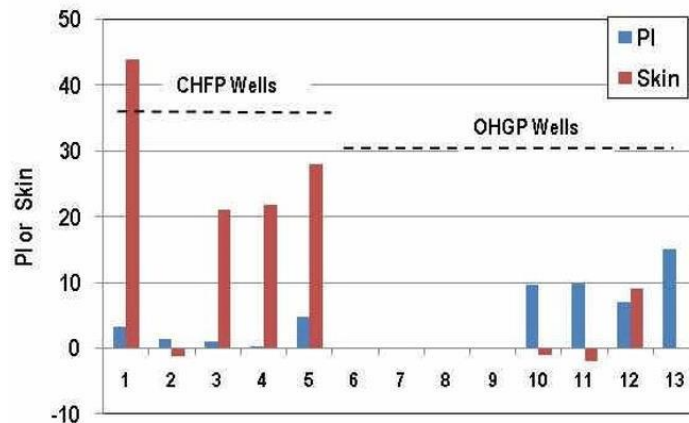


Figure 33: Skin et PI déterminés par des tests de puits pour des puits avec des complétions CHFP et OHGP dans un réservoir malaisien en eau profonde [30]

La figure 33 montre la comparaison de l'IP et le skin initiale pour les puits de complétion CHFP et OHGP. Pour les puits à angle de déviation élevé, il semble que la complétion OHGP soit plus appropriée pour atteindre l'objectif d'avoir un IP élevé, et un coût minimal.

Il ne suffit pas d'avoir une bonne complétion de contrôle du sable. Il faut s'efforcer d'évaluer et de développer un plan de gestion des drawdown sur le cycle de vie afin de réduire le potentiel de défaillance de la complétion. Les puits peuvent toujours produire du sable et des fines, et le skin peut augmenter avec le temps.

3.1.7) Approche d'Ehsan Khamehchi, 2015 :

Les principaux paramètres étudiés dans ce document sont le prix mondial du pétrole, la durée du projet de contrôle du sable et le coût des équipements nécessaires pour chaque méthode (conditions économiques et politiques), ainsi que l'indice de productivité des puits (paramètre du réservoir).

Toutes les données importantes calculées pour chaque méthode sont disponibles dans le tableau 6 et une comparaison générale est possible. Ce tableau indique que la meilleure méthode est la réalisation du gravel pack qui se base sur un point de fonctionnement de l'analyse nodal (IPR et VLP) élevé et le facteur du skin faible. Alors que ce n'est pas le cas.

Tableau 6: Résumé des données calculées pour toutes les méthodes de méthodes de contrôle du sable [31]

| Type de contrôle du sable | Taux de production au point de fonctionnement | Facteur du skin total |
|---------------------------|---|-----------------------|
| Gravel pack | 1410 | 8.7 |
| Slotted liner | 1078 | 14.8 |
| Wire Wrapped screen | 1310 | 10.2 |
| Pre-packed screen | 1370 | 9.5 |

Comme nous l'avons dit précédemment, la valeur du skin et le point de fonctionnement ne sont pas les seuls paramètres principaux pour choisir le meilleur projet. L'évaluation économique est l'étape la plus importante de la prise de décision. Toutes les méthodes ont été comparées économiquement en utilisant le coût d'exploitation approximatif de chaque méthode. Enfin, le meilleur projet sera trouvé.

La valeur actuelle nette (NPV) est une mesure de la rentabilité d'un investissement calculée comme la somme des flux de trésorerie engendrés par cette opération, chacun étant actualisé de façon à réduire son importance dans cette somme à mesure de son éloignement dans le temps.

La figure 34 représente le tracé de la NPV pour un réservoir dont la valeur de l'indice de productivité est faible. Les résultats montrent que le wire wrapped screen est la meilleure méthode de contrôle du sable.

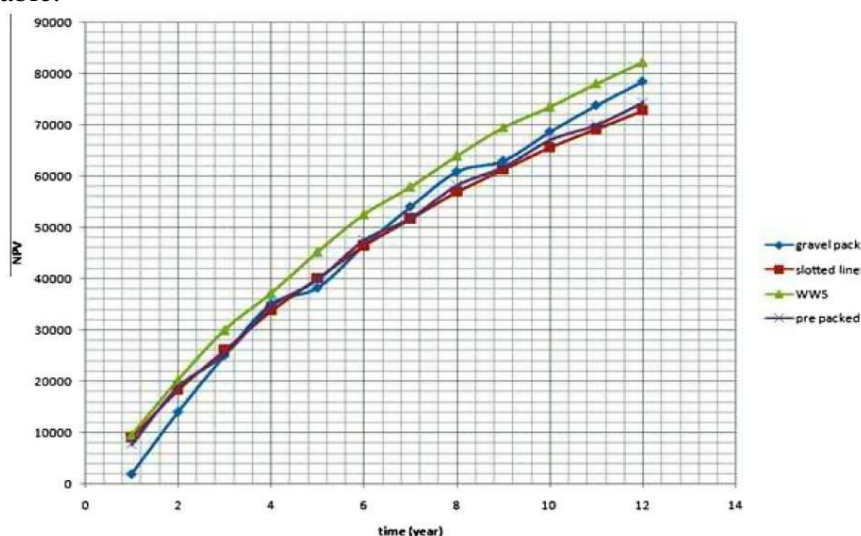


Figure 34:Tracé de la NPV pour un réservoir à faible IP [31]

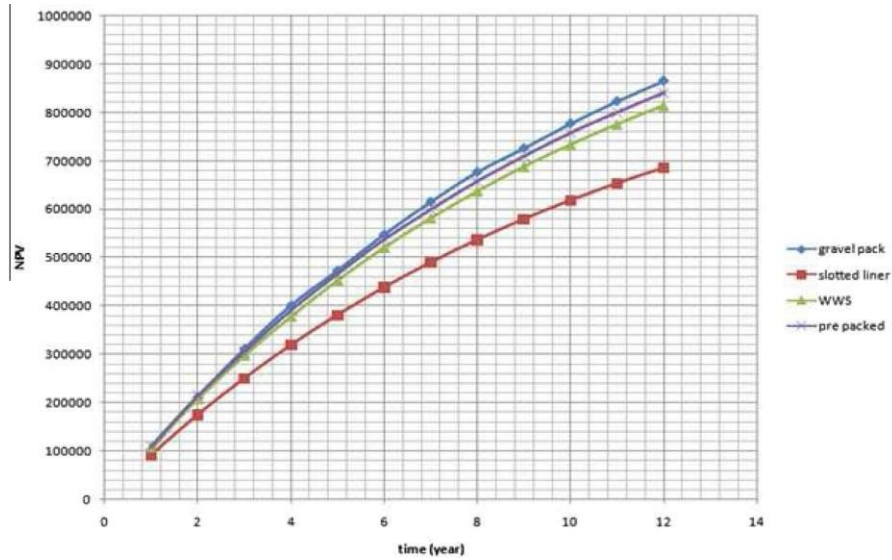


Figure 35:Tracé de la NPV pour un réservoir à IP élevé [31]

La figure 35 représente le tracé de la NPV pour un réservoir à valeur PI élevée. Comme on peut le voir, le gravel pack est la meilleure option.

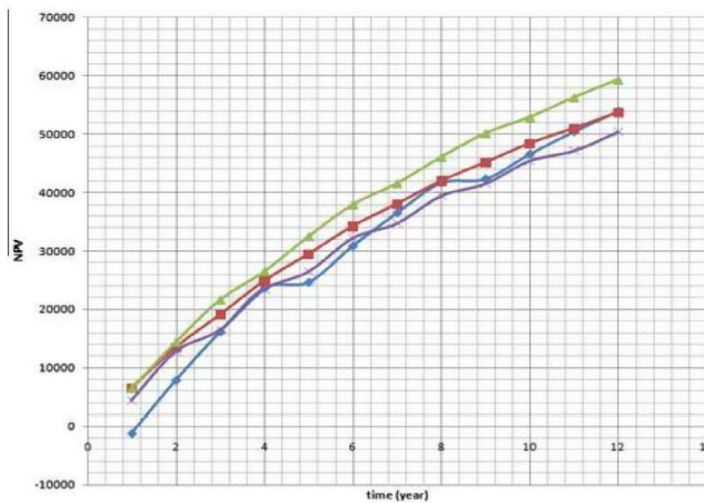


Figure 36:Tracé de la NPV pour un baril de pétrole à 20 \$ [31]

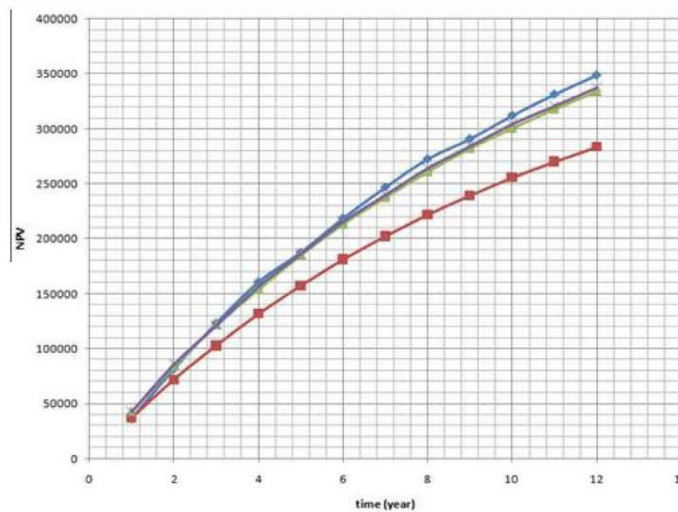


Figure 37:Tracé de la NPV pour un baril de pétrole à 100 \$ [31]

Figure 36 et 37 montrent l'effet du prix du pétrole sur le choix de la meilleure méthode de contrôle du sable. Lorsque les bénéfices sont faibles, le WWS est le meilleur, tandis que lorsque les bénéfices sont élevés, le gravel pack sera choisi. Ces analyses montrent que le prix global du pétrole, le temps de retour du capital et l'indice de productivité du réservoir sont plus importants que l'effet du skin de contrôle du sable, en particulier pour un faible taux de production de pétrole.

En conclusion :

- Tout équipement de contrôle du sable possède certaines caractéristiques dont les variations affecteront à la fois leur prix et l'ampleur du skin qu'ils provoquent.
- À faible taux de production et pour les réservoirs à faible IP, le choix de la méthode appropriée est influencé par le prix mondial du pétrole et le taux d'intérêt des investissements. En d'autres termes, le choix n'est pas directement affecté par l'effet de skin du puits.
- Pour les réservoirs à IP élevé et les projets à long terme, la meilleure méthode est celle qui permet une production de pétrole plus élevée. Ainsi, la valeur du skin a un effet intense sur le choix de la méthode de contrôle.
- Selon les conditions économiques, la meilleure méthode de contrôle est la méthode WWS (Wire Wrapped Screen).

3.1.8) Approche de Raymond Rayes, 2018 :

L'objectif principal de cette étude était de concevoir un guide sous la forme d'un organigramme qui peut potentiellement être utilisé dans la pratique lorsque des méthodes optimales de contrôle du sable et la mise en œuvre de techniques de complétion sont envisagées dans les puits producteurs de sable.

Des données distinctes sur les réservoirs et des paramètres clés ont été intégrés dans la conception du diagramme. Cet article se concentre principalement sur deux des techniques les plus courantes utilisées actuellement dans la pratique : les screens de sable et le gravel pack. Un travail d'analyse rétrospective a été réalisé. Une recherche systématique a été entreprise pour analyser les méthodologies de contrôle du sable utilisées dans les études récentes afin de déterminer les principales considérations prises en compte lors du choix du gravel pack et/ou du screens de sable. Avant de mettre en œuvre une technique de contrôle du sable dans un puits de pétrole individuel, plusieurs considérations doivent être prises en compte. Diverses mesures, évaluations et analyses doivent être effectuées afin de mettre en œuvre une technique spécifique.

L'organigramme présenté de la figure 38, a été conçu en quatre étapes distinctes. Première étape : analyse de la prédiction du sable, deuxième étape : analyse du sable, troisième étape : sélection de la technique de contrôle du sable et quatrième étape : sélection de la méthode de complétion. Une étude importante qui n'a cessé d'émerger est celle de Tiffin et al. 1998 ; cette étude a été la première étude publiée qui a été utilisée comme ligne directrice pour le choix de la technique de contrôle du sable ; le présent document a été développé en conséquence sur la base de cette étude ainsi que de diverses autres études largement liées à ce document. Il existe également des technologies plus récentes dans les méthodes de complétion qui peuvent s'avérer plus rentables.

Ces méthodes devraient faire l'objet d'une étude approfondie et être éventuellement incorporées dans le diagramme conçu dans le présent document avant que son utilisation ne soit appliquée dans la pratique à titre d'orientation une fois que les avantages, les inconvénients et la rentabilité auront été étudiés plus en détail.

L'objectif de cette étude a été atteint avec succès, bien qu'avec plusieurs limitations identifiées telles que le manque de temps, le manque d'avis d'experts de l'industrie et le manque de données de terrain sur site à utiliser activement dans l'évaluation de l'efficacité de l'organigramme.

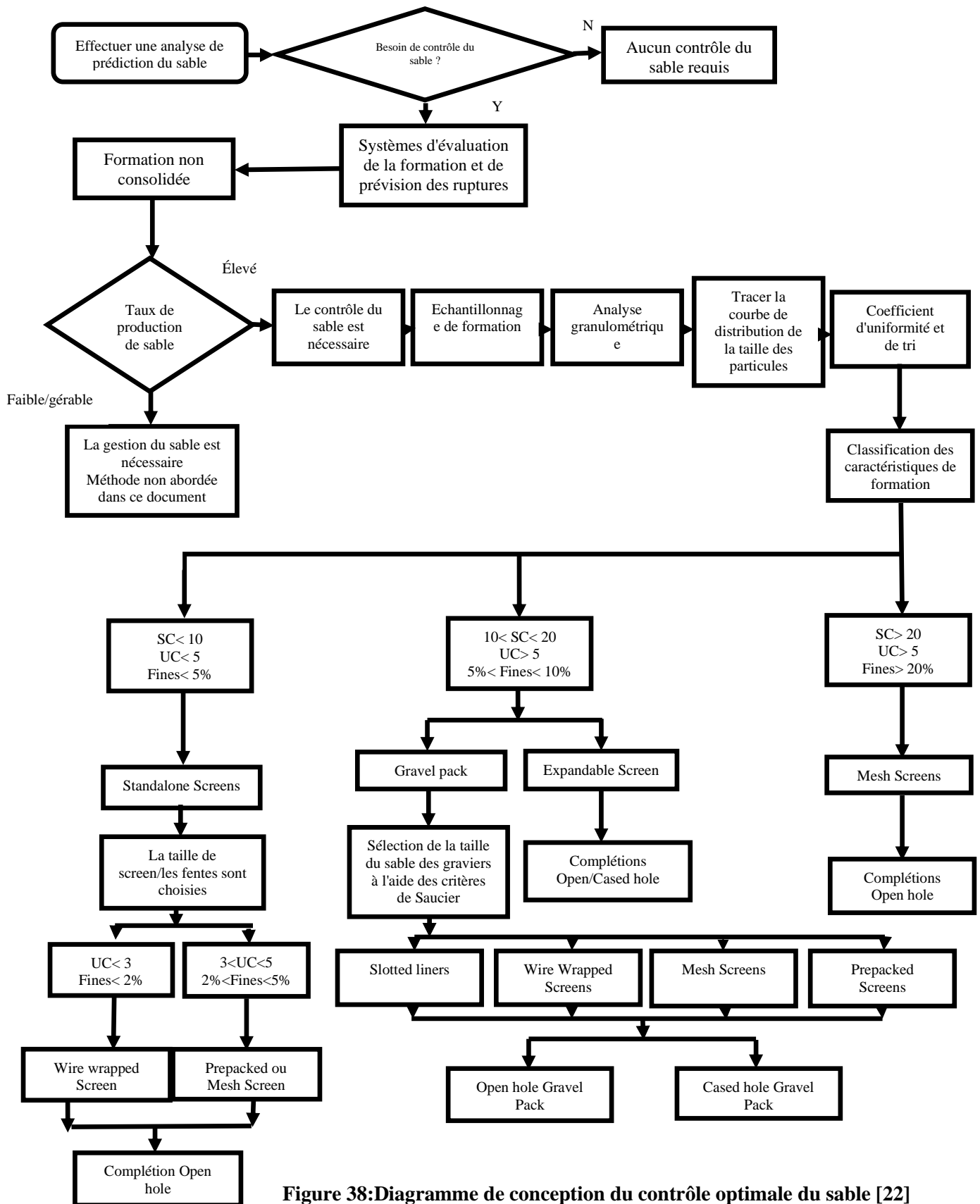


Figure 38: Diagramme de conception du contrôle optimale du sable [22]

3.1.9) Approche de Shaofeng Hu 2020 :

Il est essentiel de sélectionner une méthode de contrôle du sable optimisée pour un réservoir de grès et de mudstone interstratifiés (ISMR) (Figure 39) en raison des graves risques de production de sable. Cependant, la plupart des méthodes générales de contrôle du sable ne peuvent pas répondre aux exigences des réservoirs de grès et de mudstone interstratifiés (par exemple, les champs de pétrole et de gaz de la baie de Bohai en Chine). Assurer une efficacité de contrôle du sable et augmenter le taux de production de pétrole et de gaz dans ce grès et ce mudstone interstratifiés deviennent de plus en plus importants.

Dans cet article, un "dispositif expérimental rotatif multicouche de contrôle du sable" pour le réservoir de grès et de mudstone interstratifiés a été développé. Une série d'études expérimentales sur le contrôle du sable a été réalisée à l'aide du dispositif proposé.

Le rapport (NTG) et les inclinaisons des puits sont deux facteurs majeurs pris en compte dans l'analyse expérimentale. En outre, une analyse de sensibilité concernant la distribution granulométrique de la formation (PSD), la teneur en argile et les compositions minérales est réalisée dans l'expérience sous un mode de contrôle modéré du sable. Les résultats des tests expérimentaux systématiques de ce travail, combinés aux nombreux modèles de contrôle du sable existants, ont permis de développer un ensemble de conception optimale et le modèle d'optimisation associé pour l'ISMR, qui ont été appliqués avec succès dans la baie de Bohai. Les résultats de l'application sur le terrain montrent que le NTG et l'inclinaison du puits sont deux paramètres critiques dans la conception du contrôle du sable dans l'ISMR.

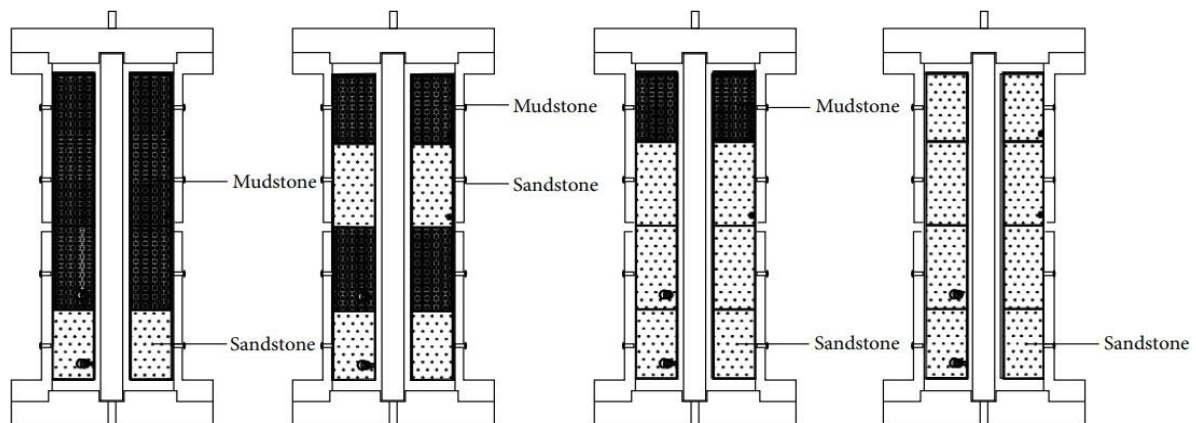


Figure 39: Diagramme schématique de la simulation ISMR [32]

Les principaux problèmes de la production de pétrole et de gaz dans l'ISMR sont les suivants :
(1) L'extrême hétérogénéité des grès et des mudstones entrelacés rend la conception du contrôle du sable plus difficile
(2) le mécanisme de bouchage des particules de sable et des minéraux argileux sur la surface extérieure du screen est différent du mécanisme de bouchage sur le milieu de blocage à l'intérieur de screen.

Cet article a construit un dispositif expérimental réaliste pour le contrôle du sable dans l'ISMR en se concentrant sur deux facteurs majeurs : (1) le rapport (NTG), qui est défini comme la proportion du contenu en grès dans l'ensemble du réservoir et reflète les caractéristiques de distribution du grès et du mudstone intercalés, et (2) les inclinaisons des puits, qui reflètent les trajectoires des puits de production.

Le réglage dynamique des inclinaisons des puits dans le dispositif de contrôle du sable (Figure40) peut simuler l'effet des inclinaisons arbitraires des puits sur la rupture de formation et donc le fluid flow et le colmatage des particules dans le contrôle du sable.

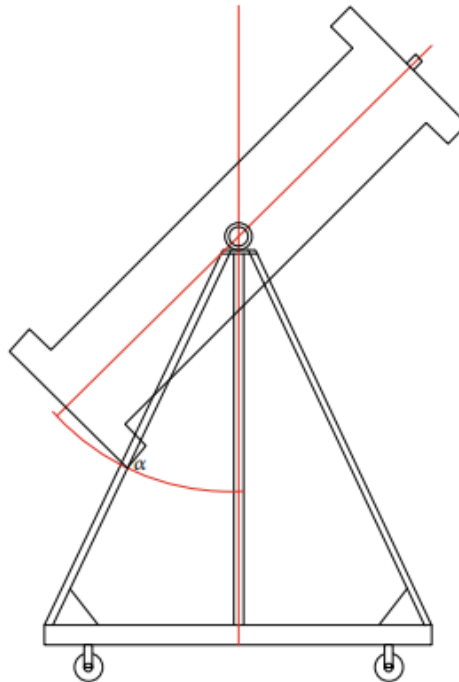


Figure 40:Schéma de la rotation dynamique du puits de forage [32]

Afin d'analyser quantitativement l'effet du NTG sur le choix des méthodes de contrôle du sable, ils prennent les réservoirs de Minghuazhen dans la baie de Bohai comme référence, des paramètres tels que la PSD (particle size distribution) et la teneur en argile sont simulés.

Compte tenu de ces caractéristiques, l'expérience de simulation de la production de sable sous différents NTG (0/0,25/0,50/0,75/1) a été réalisée.

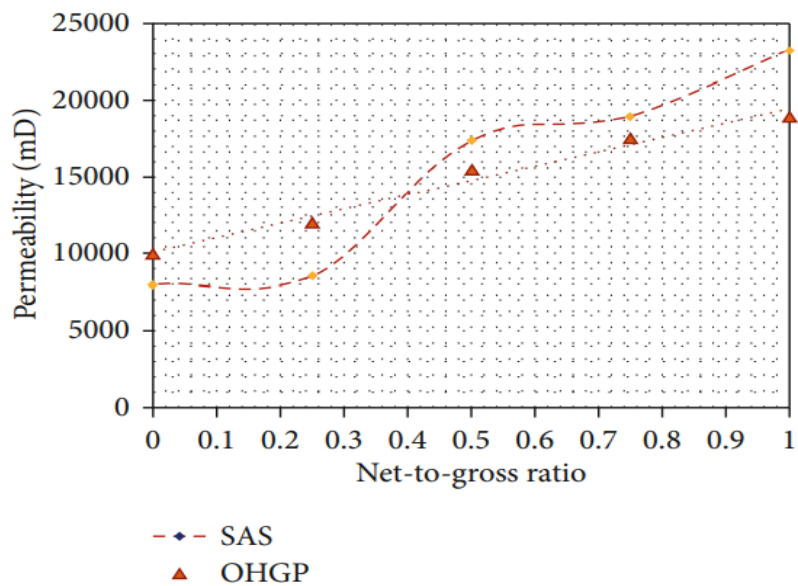


Figure 41:L'effet de différentes valeurs de NTG sur la perméabilité du screen [32]

La figure 41 présente l'effet des différentes valeurs NTG sur la perméabilité du screen. Les résultats expérimentaux montrent que le NTG de 0,4 est le point d'intersection du SAS et de l'OHGP. Lorsque le NTG est supérieur à 0,4, le colmatage du screen causé par l'argile accumulée pour le SAS est inférieur au colmatage de l'OHGP dans les mêmes conditions.

La figure 42 présente les résultats expérimentaux de différentes inclinaisons, lorsque le NTG est constant, la variation des inclinaisons du puits a un faible effet sur le blocage de l'OHGP. Pour le SAS, l'effondrement du puits s'intensifie aux inclinaisons comprises entre 30° et 60°. Le degré de blocage du SAS est plus élevé que celui de l'OHGP lorsque l'inclinaison du puits dépasse 45°. Par conséquent, l'inclinaison du puits doit être prise en compte dans la conception d'un mode de contrôle du sable optimisé dans les champs ISMR.

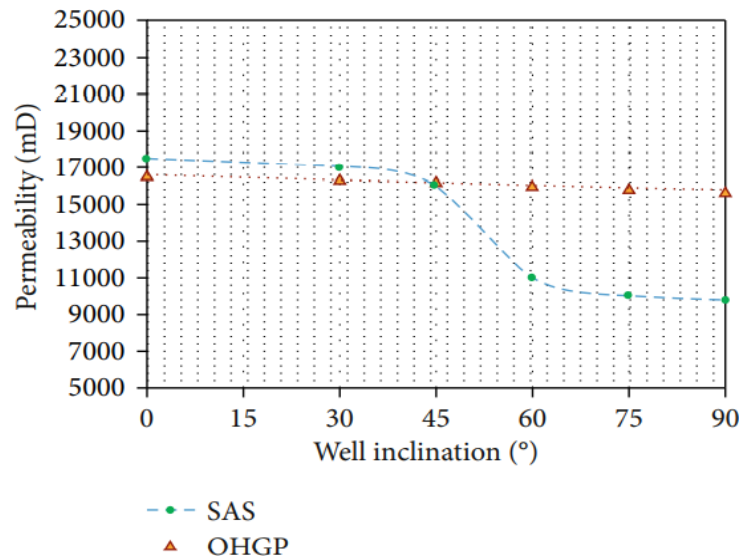


Figure 42: L'effet des différentes inclinaisons de puits sur la perméabilité du screen [32]

Sur la base de nombreuses méthodes de contrôle du sable existantes et d'une série de résultats de tests dans ce document, un ensemble de conception optimale de contrôle et le modèle d'optimisation associé pour l'ISMR ont été développés comme le montre la figure 43 :

1. $d_{50} < 50 \mu\text{m}$ (OHGP)
 - $50 \mu\text{m} < d_{50} < 250 \mu\text{m}$:
 - $\text{NTG} < 0.4$ (OHGP)
 - $\text{NTG} > 0.4$ and well inclination $< 45^\circ$ (SAS)
 - $\text{NTG} > 0.4$ and well inclination $> 45^\circ$ (OHGP)
2. $d_{50} > 250 \mu\text{m}$ (SAS):
 - $\text{UC} > 5$ (metal cotton premium screen)
 - $\text{UC} < 5$ (metal mesh premium screen)

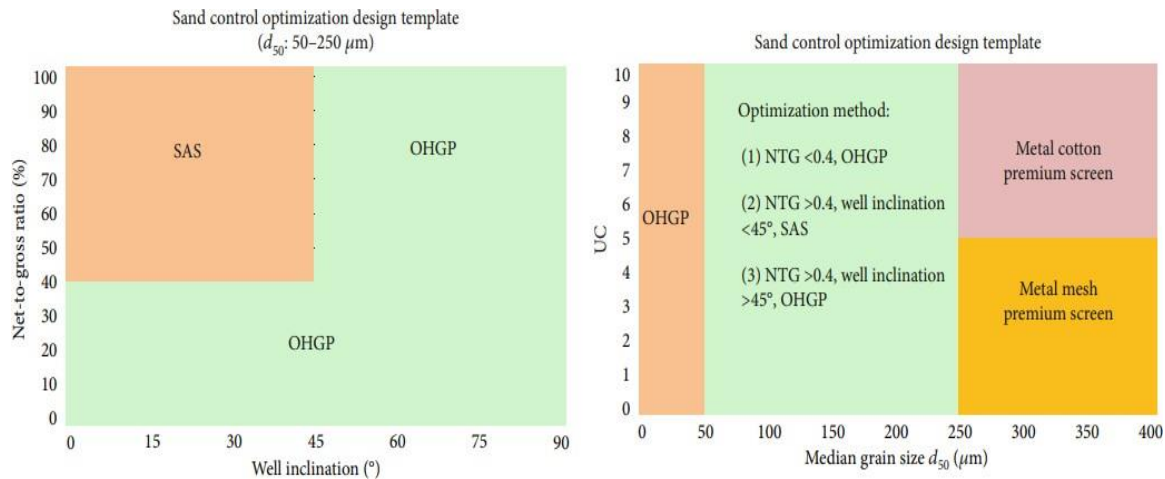


Figure 43: Méthode de contrôle du sable pour l'ISMR [32]

3.2) Nouvelle approche de sélection des méthodes de contrôle du sable :

Notre nouvelle approche de sélection vise à développer un modèle de classification qui permet de prédire la méthode optimale de contrôle du sable en utilisant des algorithmes d'apprentissage automatique (Machine Learning).

Après avoir recherché profondément sur les méthodes précédentes de sélection, on a remarqué qu'il y a pas une méthodologie basant sur l'analyse des données ou les nouvelles technologies d'AI et ça qui nous a conduit à réfléchir d'implémenter un outil puissant comme les algorithmes de machine learning pour solutionner ce problème.

On va utiliser dans notre étude une base de données internationale de plus de 1900 puits, utilisée à l'origine pour comparer la fiabilité de différentes formes de contrôle du sable. Les paramètres utilisés doivent être prédéterminés avant de faire la conception de la complétion du puits puisque on veut sélectionner la méthode convenable de contrôle du sable pour un champ donné en se basant sur les caractéristiques du réservoir, les paramètres du puits et les propriétés de production de sable.

L'apprentissage automatique et l'intelligence artificielle peuvent fournir des solutions axées sur les données qui peuvent améliorer considérablement notre capacité à optimiser les opérations de complétion. La méthodologie proposée dans ce travail réduit le temps et les incertitudes dans la sélection et la gestion des techniques de contrôle du sable.

Chapitre 4 : L'intelligence artificielle

L'intelligence artificielle (IA) et l'apprentissage automatique (ML) ont gagné en popularité dans divers secteurs. Les entreprises, les universités, les gouvernements et les groupes de recherche ont remarqué le véritable potentiel de diverses applications de l'IA et de l'apprentissage automatique pour automatiser divers processus tout en augmentant les capacités de prévision. Les progrès technologiques de l'IA dans le domaine des voitures à conduite autonome, de la détection des fraudes, de la reconnaissance vocale, du filtrage des spams, des recommandations de produits et de contenus d'Amazon et de Facebook, ont généré des quantités massives de valeur d'actif net pour diverses sociétés. Le secteur de l'énergie est dans la phase initiale de l'application de l'IA à différentes applications. La hausse de popularité dans le secteur de l'énergie est due aux nouvelles technologies telles que les capteurs et les services de calcul haute performance (par exemple, Apache Hadoop, NoSQL, etc.) qui permettent l'acquisition et le stockage de données volumineuses dans différents domaines d'étude. Le terme "big data" désigne une quantité de données trop importante pour être traitée (c'est-à-dire collectée, stockée et analysée) à l'aide d'outils et de techniques courants, par exemple des téraoctets de données. Le nombre de publications dans ce domaine a augmenté de façon exponentielle au cours des dernières années. Une recherche rapide sur le nombre de publications dans l'industrie pétrolière et gazière avec OnePetro de la Society of Petroleum Engineer ou l'American Association of Petroleum Geologists (AAPG) ces dernières années en attestent. Plus les entreprises seront nombreuses à réaliser la valeur ajoutée de l'intégration de l'IA dans leurs opérations quotidiennes, plus les idées créatives se multiplieront.

Aujourd'hui, les systèmes intelligents sont utilisés dans notre industrie dans de nombreux domaines. Ils couvrent des questions et des analyses de haut niveau, allant de la prévision de la production de gaz naturel aux États-Unis pour les 15 prochaines années et de la prise de décision au niveau de la direction, tout en traitant des données incomplètes, à des questions techniques plus difficiles qui concernent les géoscientifiques et les ingénieurs, comme le forage, la caractérisation des réservoirs, les questions d'ingénierie de la production, le traitement des puits et la gestion de la surface. Comme toute autre technique d'analyse, les systèmes intelligents ont des limites, il est important de comprendre les limites de ces techniques pour augmenter la probabilité de leur succès et leur efficacité. À titre d'exemple, considérons le groupe de techniques des systèmes intelligents qui sont développées sur la base de données, comme les réseaux neurones. Ces systèmes sont vulnérables à l'insuffisance de données. En d'autres termes, la principale limite de ces techniques est qu'elles ne peuvent être développées efficacement dans les cas où les données sont rares. [33]

4.1) Notion de base sur AI, DM, ML :

4.1.1) L'intelligence artificielle :

L'IA est tout simplement l'utilisation de l'intelligence de la machine ou de l'ordinateur plutôt que de l'intelligence humaine. Il s'agit d'une branche de l'informatique qui étudie la simulation par les ordinateurs des processus d'intelligence humaine tels que l'apprentissage, le raisonnement, la résolution de problèmes et l'autocorrection. L'objectif principal de l'IA est de créer des machines intelligentes qui fonctionnent, réagissent et imitent les fonctions cognitives des humains. Parmi les exemples d'IA, citons la classification des courriels (catégorisation), les assistants personnels intelligents tels que Siri, Alexa et Google, les répondants automatisés, l'automatisation des processus, la surveillance de la sécurité, la détection et la prévention des fraudes, la reconnaissance des formes et des images, la recommandation de produits et la prédiction d'achat, les recherches intelligentes, les ventes, les volumes et les prévisions commerciales, le ciblage des publicités,

la personnalisation des flux d'informations, la détection des activités terroristes, les voitures à conduite autonome, les diagnostics de santé, la prédiction des défauts de paiement des prêts hypothécaires, la prédiction des prix des maisons, les robo-advisors (gestionnaires de portefeuille automatisés) et l'assistant de voyage virtuel. [33]

4.1.2) Data Mining (Exploration des données) :

L'exploration de données est une terminologie utilisée en informatique et se définit comme le processus d'extraction d'informations spécifiques d'une base de données qui étaient cachées et non explicitement disponibles pour l'utilisateur, en utilisant un ensemble de techniques différentes telles que le ML. Elle est également appelée découverte de connaissances dans les bases de données. [35]

L'exploration de données est utilisée par des algorithmes ML pour trouver des liens entre diverses relations linéaires et non linéaires. Elle est souvent utilisée pour aider à collecter des données sur divers aspects de l'entreprise tels que le temps non productif, la tendance des ventes, les indicateurs clés de performance de la production, les données de forage, les données de complétion, les indicateurs et informations clés du marché boursier, etc. [33]

4.1.3) Apprentissage automatique (ML) :

L'apprentissage automatique est le domaine dans lequel les ordinateurs sont programmés de manière algorithmique pour apprendre et s'adapter à partir de l'expérience afin d'améliorer une tâche lorsqu'elle est évaluée à l'aide d'une métrique, plutôt que d'être explicitement informés du résultat. Cette définition est expliquée à l'aide de l'exemple simple d'une tâche dont l'objectif est de reconnaître des chiffres manuscrits, et la mesure est le pourcentage de chiffres correctement classés à l'aide de la base de données d'images de chiffres manuscrits étiquetés par l'homme servant d'expérience. La partie "apprentissage" utilise de grandes quantités de données pour permettre à l'ordinateur de comprendre les modèles sous-jacents. [33]

4.2) Étapes de modélisation de l'apprentissage automatique :

Les étapes de modélisation de l'apprentissage peuvent être décomposées en 7 étapes majeures :



Figure 44: Les étapes de modélisation ML [36]

- **Collecter des données :** Les machines apprennent initialement à partir des données qu'on leur fournit. Il est de la plus haute importance de collecter des données fiables afin que le modèle d'apprentissage automatique puisse trouver les bons modèles. La qualité des données fournies à la machine déterminera la précision du modèle. Si on dispose de données incorrectes, on obtiendra des résultats erronés ou des prédictions qui ne sont pas pertinentes.
- **Préparer les données :** Il existe plusieurs procédures de préparation des données lesquelles :
 - Rassembler toutes les données collectées et en les rendant aléatoires. Cela permet de s'assurer que les variables sont réparties de manière homogène et que l'ordre n'affecte pas le processus d'apprentissage.
 - Nettoyer les données pour supprimer les variables indésirables, les valeurs manquantes, les valeurs en double, la conversion des types de données, etc.
 - Visualisez les données pour comprendre comment elles sont structurées et comprendre la relation entre les différentes variables et classes présentes.

- Divisez les données nettoyées en deux ensembles : un ensemble d'apprentissage et un ensemble de test. L'ensemble d'apprentissage est l'ensemble à partir duquel le modèle apprend. Un ensemble de test est utilisé pour vérifier la précision du modèle.
- **Choix d'un modèle** : Un modèle d'apprentissage automatique détermine la sortie (output) obtenue après avoir exécuté un algorithme donné sur les données collectées. Il est important de choisir un modèle qui soit pertinent pour la tâche à accomplir. Au fil des ans, les scientifiques et les ingénieurs ont développé divers modèles adaptés à différentes tâches telles que la reconnaissance vocale, la reconnaissance d'images, la prédiction, etc. En outre, on doit également déterminer si notre modèle est adapté aux données numériques ou catégorielles.
- **Entraînement (training) du modèle** : L'entraînement est l'étape la plus importante de l'apprentissage automatique. Lors de l'entraînement, on transmette les données préparées à notre modèle d'apprentissage automatique pour qu'il trouve des schémas et fasse des prédictions. Le modèle apprend ainsi à partir des données afin d'accomplir la tâche qui lui a été confiée. Au fil du temps, avec la formation, le modèle s'améliore dans ses prédictions.
- **Évaluation du modèle** : Après avoir formé le modèle, il faut vérifier ses performances. Pour ce faire, il faut tester les performances du modèle sur des données inédites. Les données inédites utilisées sont l'ensemble de test dans lequel nous avons divisé les données précédemment. Si le test est effectué sur les mêmes données que celles utilisées pour la formation, on n'obtiendra pas une mesure précise, car le modèle est déjà habitué aux données et y trouve les mêmes schémas que précédemment. On obtiendra ainsi une précision disproportionnée. Lorsqu'il est utilisé sur des données de test, on obtient une mesure précise de la performance du modèle et de sa vitesse.
- **Réglage des paramètres** : Une fois qu'on a créé et évalué le modèle, il faut voir si sa précision peut être améliorée d'une manière quelconque. Cela se fait en ajustant les paramètres présents dans le modèle (hyperparamètres)
Les paramètres sont les variables du modèle que le programmeur décide généralement. À une valeur particulière de notre paramètre, la précision sera maximale. Le réglage des paramètres consiste à trouver ces valeurs.
- **Faire des prédictions** : Au final, on peut utiliser le modèle sur des données non vues pour faire des prédictions futures. [36]

4.3) Prétraitement des données :

4.3.1) Nettoyage des données :

Après la collecte et l'intégration des données, l'étape cruciale suivante est le nettoyage des données. De nombreux projets de ML échouent parce qu'ils ne consacrent pas suffisamment de temps à l'interprétation et au nettoyage des données avant d'appliquer les algorithmes de ML. Il n'existe pas une analyse rapide, car l'application correcte du ML à tout projet nécessite une interprétation et un nettoyage minutieux des données. Dans le cadre du nettoyage, il y a quelques étapes principales qui sont les suivantes :

4.3.2) Imputation des valeurs manquantes :

Les différentes approches des données manquantes sont les suivantes :

- Il suffit de supprimer toute ligne de données contenant des valeurs N/A : Il s'agit de la méthode la plus simple pour traiter les données manquantes, et l'inconvénient de cette approche est que le nombre d'échantillons pourrait être considérablement réduit en raison de la suppression d'un grand nombre des échantillons avec des informations manquantes.

- Remplacer les valeurs manquantes par la moyenne et la médiane des valeurs existantes. Bien que cette approche soit très simpliste et facile à réaliser, nous ne la recommandons pas pour combler les données manquantes. À moins qu'il n'y ait des données prouvées derrière cette approche, il est préférable de simplement supprimer les lignes de données manquantes plutôt que d'introduire des données erronées. [33]

4.3.3) Détection des valeurs aberrantes :

Comme mentionné précédemment, la visualisation des données permet de trouver les points aberrants. En outre, l'examen des paramètres à l'aide de la distribution, des diagrammes de dispersion, des diagrammes croisés, des cartes thermiques et des diagrammes en boîte permet de visualiser facilement les paramètres dont les plages sont raisonnables par rapport à ceux qui sont complètement à côté de la plaque. En outre, les statistiques de base des données, notamment la fréquence, le minimum, le maximum, la moyenne, la médiane et l'écart type de chaque paramètre, doivent être obtenues avant et après le nettoyage des données et la détection des valeurs aberrantes, afin de s'assurer que la gamme finale de paramètres est satisfaisante.[33]

La figure 45 montre un exemple de diagramme en boîte pour un exemple d'ensemble de données. En examinant la caractéristique VSH (V/V), nous pouvons voir que quelques points se situent beaucoup plus loin de la plage du quartile supérieur. Dans cette section, l'analyse des valeurs aberrantes n'est pas abordée. Cependant, nous pouvons procéder à une étape de filtrage supplémentaire pour éliminer toute valeur aberrante existante dans les données.

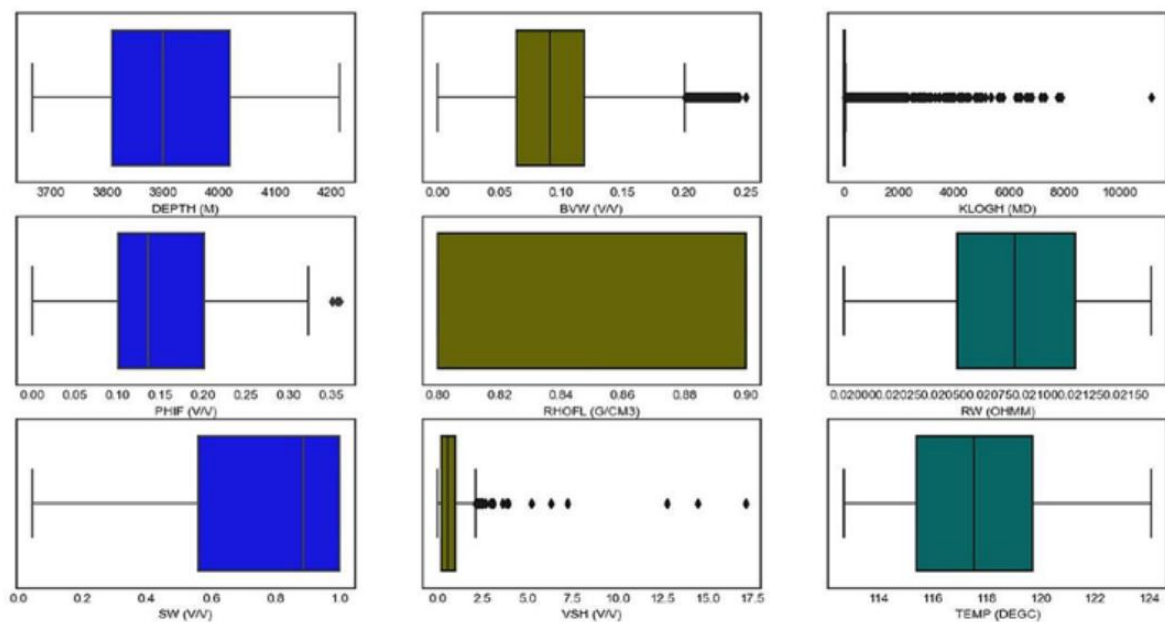


Figure 45:Box plots des caractéristiques dans l'ensemble de données pour identifier la présence des valeurs aberrantes [33]

4.3.4) Visualisation des données :

La deuxième étape du nettoyage des données est la visualisation des variables. La création de graphiques et de visualisations statiques ou interactives est l'une des tâches les plus importantes de l'analyse des variables. La vérification de la qualité et le contrôle des données sont peut-être les éléments les plus importants de toute analyse ML. De mauvaises données peuvent facilement briser un modèle et produire des résultats erronés inattendus. [37].

Les bibliothèques de visualisation telles que Matplotlib, Seaborn et Plotly en Python sont idéales pour la visualisation qui permet de trouver les points aberrants. Il est également crucial de comprendre la distribution de chaque caractéristique. Les graphiques recommandés à évaluer sont la distribution, box et scatter. [33]

4.3.5) Classement et sélection des données :

La prochaine étape importante de toute analyse ML est le classement des données, suivi de leur sélection. C'est à ce moment-là que l'expertise du domaine joue un rôle important dans la réussite d'un projet de ML. Si un scientifique des données n'a pas d'expérience pratique dans un secteur, il est extrêmement difficile de comprendre la relation entre les variables.

Dans certaines occasions, l'ingénieur et le data scientist ne connaissent pas nécessairement l'impact de chaque paramètre et ne sont pas sûrs de certaines des paramètres qui doivent être sélectionnés dans un projet ML. Par conséquent, il est important d'effectuer un classement des caractéristiques afin d'observer l'impact de chaque caractéristique sur le résultat souhaité. Il existe de nombreux algorithmes tels que la forêt aléatoire (RF), la forêt aléatoire extrême ou les arbres supplémentaires (XRF), le gradient boost (GB), ainsi que de nombreux autres algorithmes disponibles dans la littérature pour effectuer le classement des caractéristiques. [33]

4.3.6) Mise à l'échelle (Normalisation ou Standardisation) :

Pour s'assurer que l'algorithme d'apprentissage n'est pas biaisé par l'ampleur des données, les données (caractéristiques d'entrée et de sortie) doivent être mises à l'échelle. Cela peut également accélérer les algorithmes d'optimisation, tels que la descente de gradient, qui seront utilisés dans le développement du modèle en ayant chacune des valeurs d'entrée dans approximativement la même gamme. Les deux principaux types d'échelonnement (scaling) sont les suivants :

- **Normalisation des données :** La normalisation des données garantit que chaque paramètre sera mis à l'échelle de l'intervalle [0,1]. Il s'agit d'une étape critique pour la simple raison qu'un modèle sera biaisé lorsque les caractéristiques ont des magnitudes différentes. L'un des principaux inconvénients de la normalisation des données est que l'on obtient un écart-type plus faible lorsque la limite est comprise entre 0 et 1, ce qui peut avoir pour effet de supprimer les effets des valeurs aberrantes (le cas échéant). L'équation pour la normalisation des paramètres est :

$$\text{Normalisations des donnees}(X') = \frac{X - \min(X)}{\max(X) - \min(X)} \quad (4.1)$$

Où X' est le point de données normalisé, $\min(X)$ est le minimum de l'ensemble de données, et $\max(X)$ est le maximum de l'ensemble de données.

- **Standardisation des données :** La standardisation transforme chaque paramètre avec une distribution gaussienne en une distribution gaussienne avec une moyenne de 0 et un écart-type de 1. Notez que la standardisation ne transforme pas la structure de distribution sous-jacente des données. Certains algorithmes d'apprentissage, comme le SVM, supposent que les données sont distribuées autour de 0 avec une variance du même ordre. Si cette condition n'est pas remplie, l'algorithme sera biaisé vers les caractéristiques ayant une plus grande variance.

La standarization peut être effectuée en utilisant l'équation (4.2) :

$$\text{Standardisation des donnees}(X') = \frac{X - \mu}{\sigma} \quad (4.2)$$

Où X' est le point de données standardisé, μ est la moyenne de l'ensemble de données, et σ est l'écart standard de l'ensemble de données. [33]

4.4) Types d'apprentissage automatique :

4.4.1) Apprentissage supervisé :

Dans l'apprentissage supervisé, la variable de sortie ou de réponse souhaitée est connue, et l'algorithme d'apprentissage automatique fournit une correspondance entre les caractéristiques d'entrée et la variable de sortie.

Les deux principales sous-catégories de l'apprentissage supervisé sont les problèmes de régression et de classification, qui sont dictés par le type de variable de sortie. Lorsque la variable de sortie est continue, elle entre dans la catégorie de la régression. En revanche, pour les problèmes de classification, la variable de sortie contient plusieurs classes ou étiquettes. Dans l'apprentissage supervisé, le processus de formation du modèle se poursuit avec l'évaluation de l'erreur et l'apport d'améliorations jusqu'à ce qu'un niveau de précision souhaité soit atteint. [37]

4.4.2) Apprentissage non supervisé :

Dans l'apprentissage non supervisé, il n'y a pas de variable de sortie explicite, et les relations sont générées sur la base des données fournies à l'algorithme. Certains des algorithmes appartenant à cette catégorie peuvent révéler des structures cachées et des relations entre les caractéristiques d'entrée. Parmi les exemples d'apprentissage non supervisé, citons le regroupement (Clustering), algorithmes de réduction de la dimensionnalité, et l'apprentissage règles associatives. [37]

4.5) Classification et régression :

La différence entre les méthodes de classification et de régression est un concept fondamental dans l'apprentissage automatique. Les méthodes de classification visent à placer un point de données dans un groupe parmi plusieurs, tandis que les méthodes de régression visent à calculer une valeur numérique. La figure 46 illustre cette différence. Dans le premier exemple, nous souhaitons uniquement tracer une ligne de séparation entre les cercles et les triangles, en différenciant une catégorie de l'autre. Dans le second exemple, nous voulons reproduire le plus fidèlement possible une valeur numérique continue.

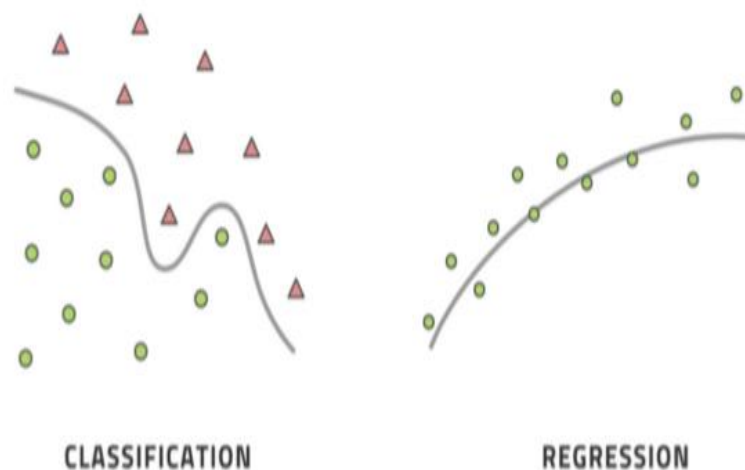


Figure 46: La différence entre les méthodes de classification et de régression [38]

Cette distinction est étroitement liée à la différence entre une variable catégorielle et une variable continue. Une variable catégorielle est égale à l'une de plusieurs valeurs, comme 1, 2 ou 3, où les valeurs indiquent l'appartenance à un certain groupe. La différence entre deux valeurs n'a donc aucune signification. Si le point de données A a la valeur 3 et le point de données B la valeur 2, tout ce que cela signifie, c'est que A appartient au groupe 3 et que B appartient au groupe 2.

Prendre la différence des valeurs, $3-2 = 1$ n'a aucune signification, c'est-à-dire que la différence entre les deux points de données n'appartient pas nécessairement au groupe 1 et que l'appartenance au groupe n'a même pas nécessairement de sens pour la différence. Une variable continue est différente en ce sens qu'elle mesure une quantité et que la prise en compte de la différence a une signification réelle.

Par exemple, si le point de données A a un débit de 1 tonne par heure et que le point de données B a un débit de 2 tonnes par heure, nous pouvons prendre la différence $2-1 = 1$ et la différence est de 1 tonne par heure de débit, ce qui est une quantité significative que nous pouvons comprendre. L'apprentissage automatique a de nombreuses méthodes et toutes les méthodes sont soit supervisées, soit non supervisées, et soit de classification, soit de régression. Toute tâche qui utilise l'apprentissage automatique peut également être divisée dans ces catégories. Ces deux regroupements constituent le premier point de départ du choix de la bonne méthode pour résoudre un problème donné. [38]

4.6) Concept de biais et de variance :

L'un des aspects les plus importants de l'apprentissage automatique concerne la qualité du modèle que l'on peut espérer, compte tenu de la qualité des données dont on dispose pour réaliser le modèle. Trois aspects du modèle sont importants ici. Le biais du modèle est une évaluation de l'écart entre le résultat moyen du modèle et le résultat moyen attendu. La variance du modèle est la mesure dans laquelle des entrées très proches les unes des autres donnent des sorties très éloignées les unes des autres. La complexité du modèle est généralement mesurée par le nombre de paramètres qui doivent être déterminés par un algorithme d'apprentissage automatique à partir des données empiriques fournies pour l'entraînement.

Les concepts de biais et de variance sont illustrés à la figure 47. Imaginez que vous lancez des fléchettes sur une cible. Si les fléchettes sont toutes proches les unes des autres, la variance est faible. Si les fléchettes sont, en moyenne, proches du centre de la cible, le biais est faible. Bien sûr, nous voulons un modèle avec un faible biais et une faible variance. Pour obtenir un tel modèle, nous choisissons d'abord un type de modèle, c'est-à-dire une formule avec des paramètres inconnus pour laquelle nous pensons qu'elle peut capturer la dynamique complète de notre ensemble de données si seulement les bonnes valeurs pour les paramètres sont trouvées. Un exemple de ce type de modèle est le modèle linéaire qui a naturellement deux paramètres pour une variable indépendante. Nous rencontrerons plus tard plusieurs modèles plus expressifs parce qu'ils ont plus de paramètres. Après avoir réfléchi à la ligne droite, nous pouvons facilement comprendre qu'un polynôme quadratique (trois paramètres pour une variable indépendante) peut modéliser un phénomène plus complexe qu'une ligne droite. [38]

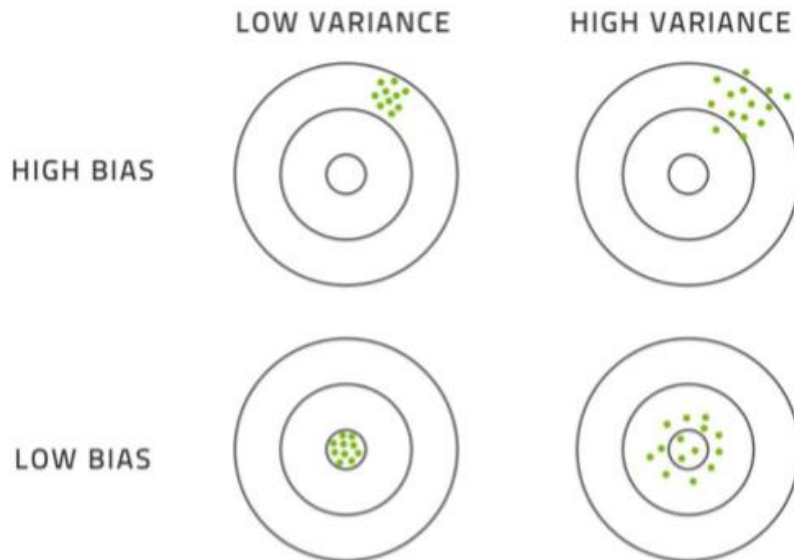


Figure 47: Les concepts de biais et de variance sont illustrés par le lancement de fléchettes sur une cible [38]

Lorsque nous formons les modèles d'apprentissage automatique, l'un des principaux objectifs est de les utiliser pour générer des prédictions pour les données non vues. Un modèle correctement formé devrait démontrer le même niveau de précision de prédiction pour les données non vues, que celui qu'il avait montré pour les données de formation.

Cependant, dans certains scénarios, le modèle peut ne pas être en mesure d'obtenir de bonnes performances sur les données non vues. Cela peut être dû à un sous-ajustement ou à un sur-ajustement du modèle.

- **Underfitting (Sous-ajustement)** : Lorsque le processus de formation est arrêté trop tôt, ou que le modèle n'est pas suffisamment exposé aux données de formation. Dans un tel scénario, le modèle n'est pas en mesure d'apprendre les schémas généraux présents dans les données d'apprentissage. Une mauvaise précision de prédiction à la fois sur les données d'apprentissage et sur les données non vues est un indicateur d'underfitting.
- **Overfitting (Sur-ajustement)** : On peut choisir d'entraîner un modèle en utilisant toutes les données disponibles dans un ensemble de données d'entraînement pour un plus grand nombre d'itérations afin qu'il apprenne tous les schémas présents dans les données d'entraînement. En conséquence, le modèle peut faire preuve d'une très grande précision sur les données d'apprentissage. Cependant, le modèle est tellement adapté aux données d'apprentissage qu'il perd sa capacité à généraliser au-delà des données d'apprentissage qu'il a vues. Dans de tels scénarios, le modèle présente une faible précision de prédiction pour les données non vues, même s'il était capable de générer des prédictions pour les données de formation avec une grande précision. Ce type de scénario indique un overfitting du modèle.

Il est important de s'assurer que le modèle ne souffre ni de sous-adaptation ni de sur-adaptation. Le modèle d'apprentissage automatique doit être entraîné de manière adéquate afin qu'il puisse générer des prédictions pour des données non vues avec une précision raisonnable. En d'autres termes, un modèle d'apprentissage automatique doit être capable de généraliser. [39]

4.7) Types de modèle Machine Learning :

Il existe de nombreux types de modèles d'apprentissage automatique qu'on peut les classer suivant deux classes :

- Apprentissage automatique supervisé
- Apprentissage automatique non supervisé

Dans cette section on va explorer les modèles utilisés pour notre étude plus spécifiquement quatre modèles supervisés.

4.7.1) K-nearest neighbors :

K-nearest neighbors également connu sous le nom de KNN est l'une des formes les plus simples d'algorithmes ML supervisés qui est utilisé pour les problèmes de classification et de régression. KNN est de type non-paramétrique ça veut dire il n'y a pas des hypothèses faites sur les données sous-jacentes. Cet algorithme fonctionne en se basant sur la proximité similaire en utilisant des calculs de distance par le biais d'une fonction de distance comme (L'euclydienne ou Manhattan) qui calcule la distance entre chaque instance et tous les observations (data points).

Ce modèle consiste à déterminer le nombre de voisins les plus proches également appelé K. Par exemple si $K=3$ trois des points les plus proches basés sur le calcul de la distance seront choisis pour déterminer où une instance sera affectée (dans un problème de classification). Donc, la sélection du K est l'une des étapes les plus importantes lors de développement de cet algorithme. Une approche très utilisée c'est la méthode du coude (Elbow method) qui consiste à entraîner un modèle avec différentes valeurs de K et choisir la valeur qui assure la plus grande précision de test à l'aide d'une technique de validation croisée (Cross validation).

Une note très importante sur K-nearest neighbors c'est qu'on doit normaliser les données avant de passer par ce modèle. Finalement les avantages de KNN sont :

- KNN est facile à mettre en œuvre.
- Contrairement à de nombreux autres algorithmes qui seront abordés plus loin dans ce chapitre, les deux principaux paramètres sont le nombre de K et la fonction de distance.

Et certains inconvénients sont comme suit :

- Il peut devenir intensive en termes de calcul, car la distance de chaque instance par rapport à toutes les observations d'entraînement doit être calculée, par conséquent L'utilisation de KNN peut être un défi dans les ensembles de données à haute dimension.
- L'algorithme ne fonctionne pas bien avec les caractéristiques catégorielles car le calcul de la distance dimensionnelle à l'aide d'une caractéristique catégorielle est difficile.

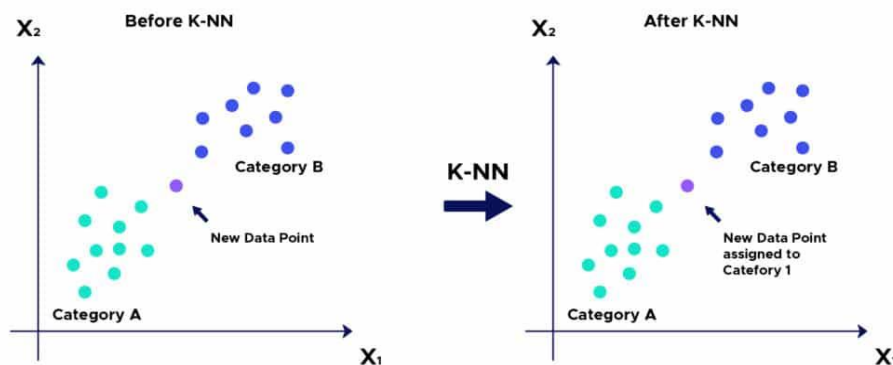


Figure 48: Illustration de l'algorithme KNN [41]

4.7.2) Machine à vecteurs de support (SVM) :

SVM est une autre forme d'algorithme ML supervisé qui peut être utilisé pour les problèmes de classification et de régression. Son objectif est d'essayer de maximiser la marge entre les vecteurs de support à travers un hyperplan de séparation (Cortes & Vapnik, 1995). La figure ci-dessous explique bien le principe de fonctionnement, un hyperplan de séparation (ligne noire pleine) est dessiné pour séparer les instances bleues des instances rouges et il sert à maximiser la marge des deux côtés, les vecteurs de support d'où vient le nom de cet algorithme sont les instances bleues et rouges les plus proches de l'hyperplan de séparation et la marge entre eux est représentée par la distance entre les lignes bleues et rouges en pointillés.

La figure 49 illustre la forme la plus simple d'une classification binaire pour les cas linéairement séparables et malheureusement ce n'est pas le cas pour la plupart des problèmes d'application du monde réel.

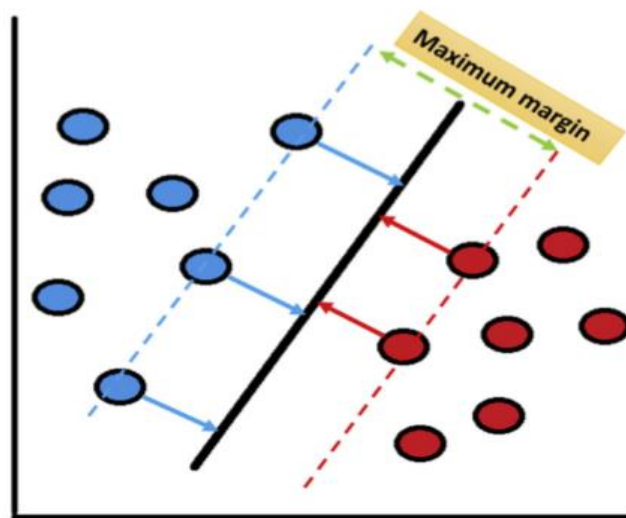


Figure 49: Machine à vecteurs de support pour la classification [33]

Afin de prendre en compte les cas lorsque les données ne sont pas linéairement séparables, on doit introduire deux concepts :

- Marge d'erreur : elle est utilisée pour trouver une ligne pour la séparation des classes, mais cette ligne tolère une ou quelques instances mal classées. Ce degré de tolérance est représenté par le paramètre "C" dans l'algorithme SVM. Comme le montre la figure 50, 'C' est essentiellement un paramètre de régularisation qui sert à assurer la grande marge possible pour une meilleure classification.
 - Plus C est grand = Plus petite largeur de la marge (marge plus petite) = Plus de pénalité encourue lors d'un mauvais classement = Plus grande précision d'entraînement et plus faible précision de test = Moins général
 - Plus C est petit = Plus grande largeur de la marge (marge plus grande) = Moins de pénalité encourue lors d'un mauvais classement = Plus faible précision d'entraînement et plus grande précision de test = Plus général.

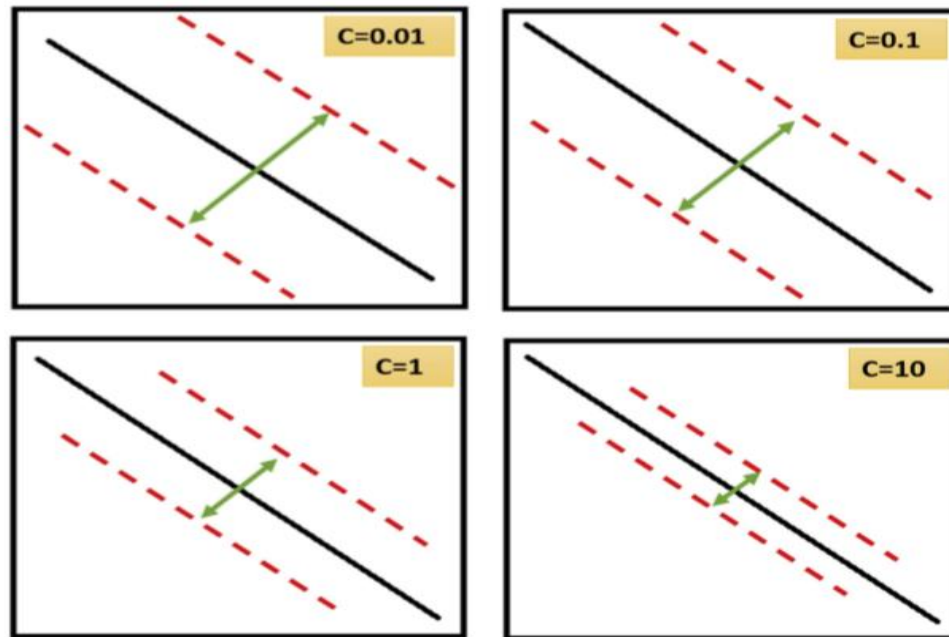


Figure 50:L'impact de l'augmentation de C sur la marge du SVM [33]

- L'astuce du kernel : lorsque les données ne sont pas linéairement séparables, les kernels utilisent les caractéristiques existantes et appliquent certaines fonctions de transformation pour créer de nouvelles caractéristiques. Par exemple, une base linéaire, polynomiale ou radiale. Dans la bibliothèque scikit-learn, de nombreux fonctions kernels peuvent être utilisées telles que "linéaire", "poly", "rbf", "sigmoïde", etc. Les plus populaires, fournissant généralement la plus grande précision, sont 'poly' et 'rbf'.
 Nous allons mentionner juste la fonction 'RBF' : Le kernel 'RBF' est un transformateur qui génère de nouvelles caractéristiques en utilisant la mesure de la distance entre tous les autres points et un ou plusieurs points ou centres spécifiques. Le type le plus populaire est la fonction de base radiale gaussienne. Un paramètre très important dans cette fonction est le gamma ' γ ' qui est utilisé pour contrôler l'influence des nouvelles caractéristiques sur la frontière de décision. Alors, il faut utiliser différentes valeurs de gamma et voir la différence lors du développement du modèle. La figure 51 illustre l'impact de l'augmentation du gamma sur les limites du SVM et c'est clair que le modèle a tendance à surajuster (overfit) par opposition à la construction d'un modèle généralisé.

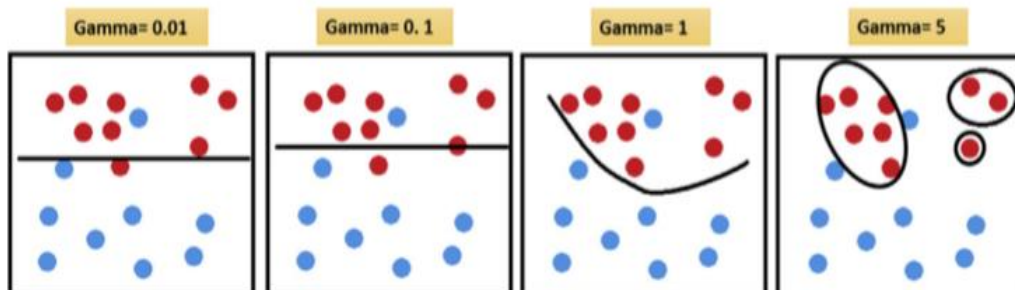


Figure 51:L'impact du gamma sur la frontière des SVM [33]

4.7.3) Arbre de décision :

L'arbre de décision est un autre algorithme ML supervisé qui peut être utilisé pour les problèmes de classification et de régression. Son principe de fonctionnement est simple, diviser les données en sous-arbres qui sont ensuite divisés en d'autres sous-arbres. Comme l'illustre la figure 52, il existe un nœud racine qui se trouve au niveau le plus élevé et représente essentiellement l'ensemble de la population, ensuite un nœud de décision également appelé nœud interne possède deux branches ou plus et finalement un nœud terminal également appelé nœud feuille qui est le nœud le plus bas et ne se divise plus. Notez que la division est aussi appelée 'splitting'.

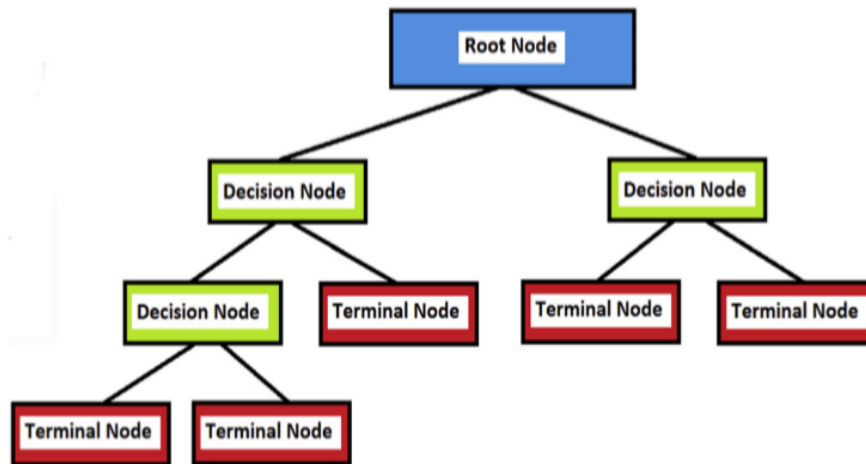


Figure 52: Illustration de l'arbre de décision [33]

4.7.3.1) Technique de sélection des attributs :

Dans un ensemble de données comportant N attributs (ou bien paramètres), le choix de placer ces dernières à la racine ou au nœud interne c'est difficile. Pour ça on devait suivre certains critères lors de la sélection des attributs :

- **Entropie :**

Représente une mesure de l'incertitude ou de la pureté en notant qu'une entropie élevée signifie une faible pureté. Elle se calcule par l'équation suivante :

$$E(S) = \sum_{i=1}^c P_i \log_2 P_i$$

Où c'est le nombre de classes, P_i est la probabilité d'une classe dans un ensemble de données.

- **Gain d'information :**

Une étape très importante lors de la construction d'un arbre de décision est de trouver un attribut qui assure le plus grand gain d'information et la plus petite entropie. Une autre façon pour expliquer, l'IG montre la façon dont un attribut sépare les exemples d'entraînement en fonction de leur classification cible.

Notez que l'IG favorise les partitions plus petites et qu'il peut être calculé comme suit :

$$IG(Y, X) = E(Y) - E(Y, X)$$

- **Indice de Gini :**

Contrairement au gain d'information, l'indice de Gini favorise les plus grandes partitions et il est calculé en soustrayant 1 de la somme des probabilités au carré de chaque classe, comme illustré ci-dessous. Veuillez noter que dans des exemples parfaitement classés, l'indice de Gini serait égal à 0.

$$Gini = 1 - \sum_{i=1}^c (P_i)^2 = 1 - (P(\text{Classe A})^2 + P(\text{Classe B})^2 + \dots + P(\text{Classe N})^2)$$

Où P_i est la probabilité qu'un élément soit classé sous une classe particulière. L'un des grands problèmes lors de l'utilisation de l'arbre de décision est le sur-ajustement (overfitting).

L'une des approches pour adresser ça est la technique de l'élagage (Pruning) qui consiste à éliminer les branches qui n'apportent pas beaucoup d'informations en termes de classification des instances ou de perturbation de la précision globale. Il est nécessaire d'utiliser la validation croisée dans l'arbre de décision pour optimiser le modèle et éviter le cas de sur-ajustement. Notez que la bibliothèque de scikit-learn utilise par défaut l'indice de Gini comme une technique de sélection des attributs.

4.7.4 Forêt aléatoire (Random Forest) :

La forêt aléatoire est un autre algorithme ML supervisé qui représente une extension de l'algorithme arbre de décision, plus spécifiquement c'est un ensemble d'arbres de décision et pour cette raison il est connu comme un algorithme d'ensemble. Par exemple, au lieu de construire un arbre de décision pour prédire un output donné, l'utilisation d'un seul arbre pourrait donner une valeur erronée en raison de la variance des prédictions. Une façon d'éviter cette variance est de prendre les prédictions de centaines ou de milliers d'arbres de décision et à utiliser la moyenne de ces arbres pour calculer la réponse finale. Par rapport aux problèmes de classification, la forêt aléatoire donne des prévisions en se basant sur le vote majoritaire de la classe prédite. Or, pour les problèmes de régression, la forêt aléatoire utilise la moyenne des arbres décisionnels pour la prédiction finale. La figure 53 illustre la différence entre un arbre de décision unique et une forêt aléatoire, qui consiste en un ensemble d'arbres de décision.

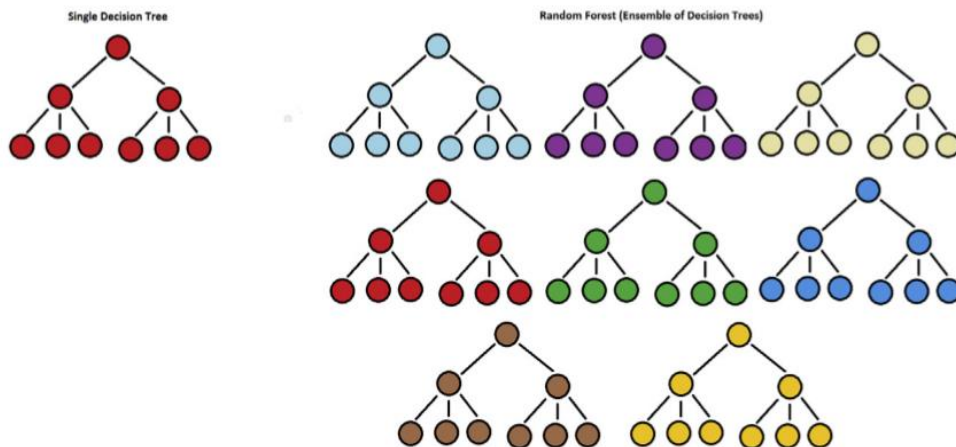


Figure 53: Arbre de décision versus forêt aléatoire [33]

On a deux types majeurs de combinaison d'arbres de décision multiples en un seul, à savoir :

- Bagging : aussi appelée " agrégation bootstrap " et utilisée dans l'algorithme de forêt aléatoire, elle utilise une moyenne de nombreux modèles indépendants construits de manière que les arbres de décision sont formés sur la base d'un échantillonnage aléatoire de sous-ensembles de données et l'échantillonnage est effectué avec remplacement. Parmi ces avantages est qu'elle sert à la réduction de la variance du modèle ce qu'on appelle généralement le sur-ajustement (overfitting).
- Boosting : aussi considéré comme une technique d'ensemble mais cette fois ci les modèles sont construits de manière séquentielle plutôt qu'indépendante afin d'assurer plus de poids aux prédictions incorrectes. Par opposition au Bagging qui utilise une moyenne pondérée égale, cette technique applique plus de poids sur les modèles avec une meilleure performance ça veut dire les échantillons qui ont été prédits de manière inexacte reçoivent un poids plus élevé, ce qui conduit à les échantillonner plus souvent. Une autre différence se manifeste est que le boosting réduit le biais non pas la variance d'un modèle.

4.8) Métriques d'évaluation des modèles :

L'objectif principal de l'apprentissage automatique est de construire des modèles capables d'être généralisés sur un ensemble de données non vues ce qui rend nécessaire d'avoir des mesures pour évaluer la précision de prédiction. Afin d'améliorer la performance de prédiction d'un modèle, il faut d'abord des métriques d'évaluation. Ensuite, ces métriques en conjonction avec une méthode d'évaluation peuvent être utilisées pour choisir les meilleures stratégies pour améliorer la qualité du modèle.

Certaines techniques d'amélioration des modèles comprennent, sans s'y limiter, (i) l'optimisation des hyperparamètres pour un modèle spécifique, (ii) l'entraînement de différents modèles, et (iii) finalement, la modification du pourcentage des portions d'entraînement et test pendant le partitionnement des données.

Dans cette section, on va aborder juste les métriques d'évaluation pour la classification ce qui est adaptable à notre approche de sélection de meilleure méthode de contrôle du sable :

- Exactitude (Accuracy) : La première métrique d'évaluation définie comme le rapport entre la prédiction correcte et la prédiction totale.

$$\text{Acc} = \frac{\text{Nombre de prdictions correctes}}{\text{Nombre total de prdictions}}$$

L'utilisation de cette mesure lors de l'évaluation d'un modèle fournisse une compréhension générale sur la prédiction mais malheureusement pose un problème majeur dans la classification à cause de la distribution déséquilibrée des classes, c'est pourquoi d'autres mesures sont utilisées simultanément pour supprimer l'incertitude liée au déséquilibre.

Afin de comprendre les autres métriques, il faut savoir précédemment certaines notions :

True positive (TP) : le modèle prédit correctement la valeur 1.

True négative (TN) : le modèle fait une prédiction correcte pour des valeurs 0.

False positive (FP) : le modèle fait une prédiction erronée égale à 1.

False négative (FN) : le modèle fait une prédiction erronée égale à 0.

Pour simplifier les choses, on considère la production du sable comme notre sujet et bien 1 représente l'existence de production du sable et 0 l'absence du sable.

- Précision : définie comme le rapport entre les positifs corrects ou TP et tous les positifs qui ont été prédits par le modèle.

$$\text{Précision} = \frac{TP}{TP+FP}$$

- Rappel : il est défini comme la proportion de toutes les valeurs positives qui sont des positifs correctement prédits. Le rappel identifie la capacité du modèle à prédire les valeurs positives.

$$\text{Rappel} = \frac{TP}{TP+FN}$$

- F1 score : Dans certains cas, l'objectif est d'avoir une précision élevée alors que dans d'autres cas, avoir un rappel élevé est une nécessité. Afin d'avoir un équilibre entre la précision et le rappel, une autre métrique, définie comme "score F1", est une moyenne harmonique de la précision et du rappel. [37]

$$\text{F1 score} = 2 * \frac{\text{Précision} * \text{Rappel}}{\text{Précision} + \text{Rappel}}$$

Partie pratique

Chapitre 5 : Prétraitement des données et développement des modèles

5.1) Description de l'ensemble des données :

L'étape fondamentale dans la résolution de tout problème d'apprentissage automatique est la collecte des données. Les données pour cette étude ont été recueillies à partir d'un article intitulé « Comparaison de la fiabilité et du taux de défaillance des complétions de contrôle du sable avec une base de données de plusieurs milliers de puits. ». Cet article comporte une base des données Excel qui contient des informations sur la fiabilité/défaillance de contrôle du sable pour plus de 2000 puits. Cela permet d'étudier les performances des complétions et de comparer leur fiabilité, tenant compte que cet ensemble des données peuvent triées par zone géographique. De plus, au cours de notre exploration de cette base de données, nous avons découvert que différents types de paramètres existent, tels que les propriétés du réservoir, les données de complétion, les propriétés de caractérisation du sable et les données de fiabilité.

Comme notre approche vise à sélectionner la meilleure méthode de prévention du sable pour un puits donné, Les données utilisées doivent être acquises avant que la conception de la complétion ne soit installée. A cet effet, nous avons effectué une première sélection sur le jeu des données suivant ce critère et les résultats sont présentés dans le tableau 7 :

Tableau 7: Les principaux paramètres dans l'ensemble des données

| Les données | Signification |
|---|--|
| Geographical Area | Zone géographique |
| Produced Fluid | Huile ou Gaz |
| Well Type | Puit à trou ouvert ou trou tubé |
| Screen Type | Type de crépine |
| Sorting coefficient D10/D90 | Coefficient du tri |
| Mass fraction of fines smaller than 44 micron | Teneur en fines (moins de 44 μ m) |
| Well deviation through pay zone_angle | L'angle de déviation à travers la section du réservoir |
| TVT_Pay zone | Épaisseur verticale réelle de la section du réservoir |
| Measured Pay Length | Longueur mesurée de la section du réservoir |
| Top Perf Depth_TVD | Profondeur de la perforation supérieure en termes de TVD |
| Perforation Method | Méthode de perforation |
| Perf differential pressure | Pression différentielle de la perforation |
| BHT | Température de fond du puits |
| BHP | Pression de fond du puits |
| Maximum Drawdown | Drawdown maximal |
| Max Oil Rate | Débit d'huile maximal |
| Max Gas Rate | Débit de Gaz maximal |
| Permeabiity | Perméabilité |
| kh | Perméabilité *Épaisseur de zone productrice |
| PI | Indice de productivité |
| Total Skin | Skin total |
| Fines Production_safe | Teneur en fines (qui n'endommagent pas la production) |
| Screen Size | Taille de crépine |

Selon notre ensemble des données, chaque technique de contrôle du sable est utilisée dans un certain nombre de puits comme indiqué dans le tableau 8. Chaque méthode représente une sortie pour notre modèle ce qui rend notre approche une classification multi-classes des sorties.

Il existe différents types de crépines selon l'industrie, et nous avons deux options dans notre base de données : les crépines extensibles (ESS) et les crépines seules (SAS). Pour éviter le problème des données déséquilibrées dans nos résultats, nous avons combiné les deux classes mentionnées ci-dessus en une seule classe appelée Crépines (Screens) :

Tableau 8: Nombre totale des puits pour chaque type de complétion

| Type de complétion | Nombre totale du puits |
|------------------------|------------------------|
| Screenless Fracs | 25 |
| Screens | 391 |
| Cased hole gravel pack | 379 |
| Open hole gravel pack | 187 |
| High rate water pack | 186 |
| Frac pack | 819 |
| Total | 1987 |

5.2) Prétraitement des données :

5.2.1) Analyse statistique :

La première étape avant de traiter l'ensemble des données est d'effectuer une analyse statistique qui sert à déterminer le nombre d'échantillons, la moyenne, l'écart type, les percentiles (25% ; 50% ; 75%) et les valeurs minimums et maximums de toutes les caractéristiques. Le tableau 9 indique les plages des différents paramètres, par exemple le coefficient de tri (Sorting coefficient D10/D95) l'un des paramètres clés dans la sélection des méthodes de contrôle du sable varie de 0.5 à 25. Pour l'angle de déviation à travers la section du réservoir, il est entre 0° et 108° ça veut dire que les puits disponibles sont de type verticaux et fortement déviés. Comme dernier point, le tableau ci-dessous mentionne le nombre des points de chaque variable numérique et peut être utilisé comme indice de la présence des valeurs manquantes.

Tableau 9: Description statistique de l'ensemble des données

| Paramètres | Nombre d'échantillons | Moyenne | Std | Min | 25% | 50% | 75% | Max |
|---------------------------------------|-----------------------|---------|---------|-----|-------|------|--------|-------|
| Sorting coefficient D10/D95 | 133 | 10.09 | 5.82 | 0.5 | 5 | 10 | 12 | 25 |
| Mass fraction of fines less than 44µm | 140 | 13.68 | 11.96 | 1 | 4 | 10 | 16 | 70 |
| Well deviation through pay zone_angle | 1056 | 45.22 | 32.11 | 0 | 19 | 43.5 | 76.12 | 108 |
| TVT_pay zone | 366 | 102.04 | 116.51 | 1 | 31 | 62 | 120.75 | 801 |
| Measured pay length | 798 | 620.75 | 1171.07 | 3 | 46.25 | 130 | 728.50 | 20395 |

| Paramètres | Nombre d'échantillons | Moyenne | Std | Min | 25% | 50% | 75% | Max |
|----------------------------|-----------------------|----------|----------|------|------|------|---------|--------|
| Perf differential pressure | 61 | 457.98 | 551.65 | 23 | 200 | 300 | 500 | 4000 |
| BHT | 829 | 256.28 | 487.44 | 75 | 150 | 174 | 195 | 5559 |
| BHP | 1015 | 3810.24 | 2571.30 | 100 | 2121 | 3240 | 4655 | 14400 |
| Maximum drawdown | 178 | 587.98 | 627.34 | 0 | 100 | 400 | 843.750 | 3700 |
| Max oil rate | 604 | 7228.96 | 9913.90 | 0 | 449 | 2840 | 10314 | 70704 |
| Max gas rate | 513 | 19808.26 | 33026.24 | 0 | 721 | 5433 | 20720 | 180000 |
| Permeability | 416 | 526.91 | 895.22 | 0 | 70 | 250 | 671.5 | 8200 |
| kh | 198 | 65062.73 | 115231 | 26 | 4644 | 2522 | 73200 | 861549 |
| PI | 126 | 63.71 | 98.66 | -2.5 | 9.25 | 30.5 | 79.5 | 771 |
| Total skin | 150 | 21.49 | 52.55 | -4 | 1 | 4.75 | 18 | 367 |
| Screen size | 163 | 3.76 | 0.77 | 2 | 3.5 | 3.5 | 4.25 | 5.5 |

5.2.2) Traitement des valeurs manquantes :

D'abord, nous avons vérifié le taux des valeurs manquantes dans notre ensemble de données et nous avons trouvé 30825 en total. Par la suite notre traitement était subdivisé en deux étapes :

- **Première étape :** Nous avons utilisé la méthode « dropna » qui consiste simplement à supprimer les valeurs manquantes sur la base d'un critère nommé « tresh », par exemple dans notre cas nous avons choisi la valeur de « tresh » égale à 12 ce qui signifie que chaque ligne qui contient douze valeurs manquantes ou plus sera définitivement supprimée. Après avoir exécuté cette méthode nous avons obtenu 422 puits.
- **Deuxième étape :** Le reste de l'ensemble de données a été subdivisé basant sur la zone géographique afin d'assurer un traitement efficace des valeurs manquantes en utilisant la méthode de la moyenne qui consiste à remplacer une valeur manquante par la moyenne du paramètre correspondant.

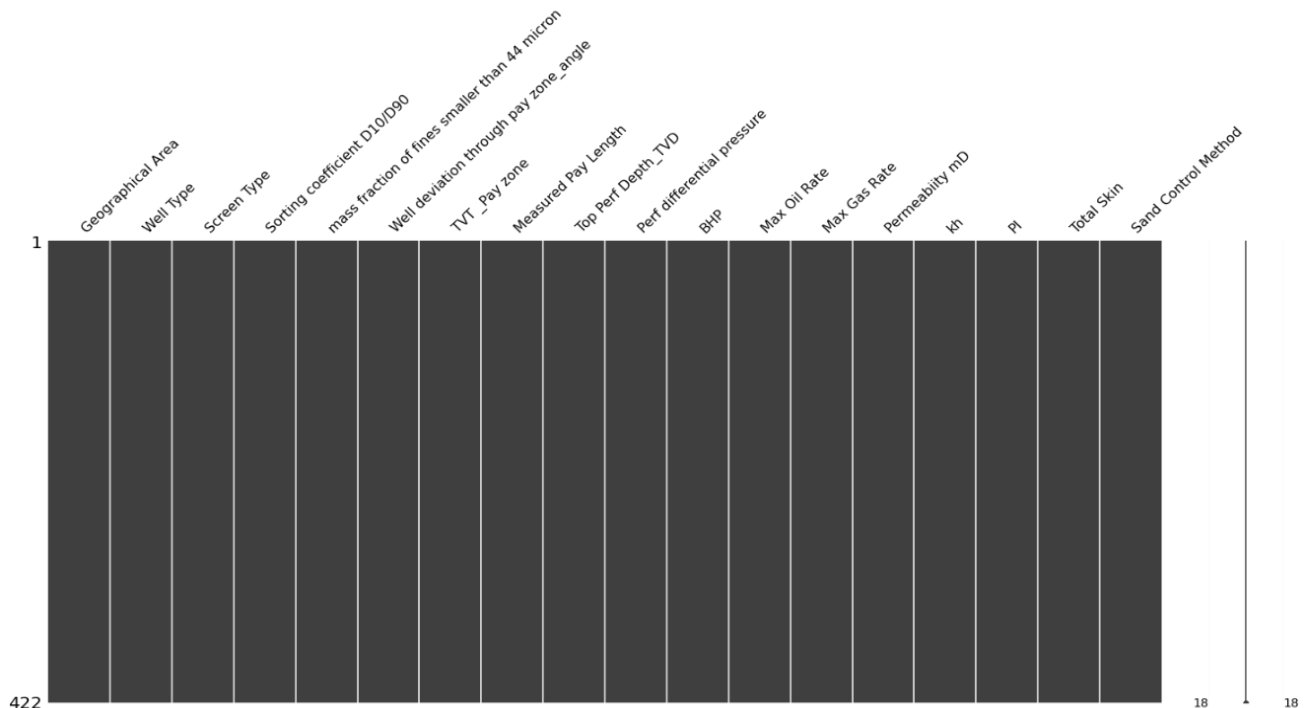


Figure 54: L'ensemble des données après traitement des valeurs manquantes

5.2.3) La sélection des principaux paramètres :

La sélection des caractéristiques est le processus qui consiste à réduire le nombre des variables d'entrée lors du développement d'un modèle prédictif. Il est souhaitable de réduire le nombre des paramètres d'entrée pour réduire le coût de calcul lors de la modélisation et dans certains cas, pour améliorer les performances du modèle.

Dans cette partie nous avons utilisé deux méthodes différentes à savoir la méthode « ANOVA » et la méthode « Chi Squared » l'une pour le traitement des variables numériques et l'autre pour les variables catégoriques :

- **ANOVA** : L'ANOVA (Analyse de la variance) nous aide à sélectionner les meilleurs paramètres numériques. Elle mesure la distance entre une valeur d'un variable donné et sa moyenne. La variance d'une caractéristique détermine son impact sur la variable de réponse. Si la variance est faible, cela signifie que cette caractéristique n'a pas d'impact sur la réponse et vice-versa. Le tableau 10 indique les valeurs de variance de chaque paramètre.

Tableau 10: Les variances des paramètres numériques

| Paramètre | Variance |
|---|----------|
| Measured Pay Length | 68.67 |
| Well deviation through pay zone_angle | 52.11 |
| TVT_Pay zone | 34.77 |
| Mass fraction of fines less than 44 μm | 23.72 |
| Total Skin | 23.58 |
| Sorting coefficient D10/D90 | 21.64 |
| Max Oil Rate | 18.85 |
| Permeability | 17.97 |
| kh | 17.03 |
| PI | 15.44 |
| Perf differential pressure | 13.87 |
| Max Gas Rate | 12.52 |
| Maximum Drawdown | 10.28 |
| Top perf Depth_TVD | 9.67 |
| BHP | 5.15 |
| Fines Production_safe | 1.71 |
| BHT | 0.72 |
| Screen Size | 0.05 |

D'après le tableau ci-dessus, nous avons choisi d'éliminer 4 variables sur la base des valeurs de variance faibles et n'ayant pas de relation avec les méthodes de contrôle du sable (Maximum Drawdown, Fines Production_safe, BHT, Screen Size).

- **Chi Squared** : Le test de chi squared est utilisé pour les caractéristiques catégorielles d'un ensemble de données. Nous calculons le score de chi square entre chaque entrée catégorielle et la sortie, par la suite nous sélectionnons le nombre souhaité des caractéristiques ayant les meilleurs scores de chi squared. Il détermine si l'association entre les variables catégorielles de l'échantillon reflète leur association réelle dans l'ensemble de données. Le tableau 11 dans la page suivante présente les scores de chi squared pour chaque paramètre catégorique.

Tableau 11: Les scores de chi squared pour les variables catégorielles

| Paramètre | Score de chi squared |
|--------------------|-----------------------------|
| Well Type | 205.58 |
| Screen Type | 79.46 |
| Perforation Method | 10.91 |
| Produced Fluid | 1.66 |

D'après le tableau ci-dessus nous avons choisi d'éliminer 2 variables catégorielles qui disposent les scores de chi squared les plus faibles (Produced Fluid et Perforation Method)

5.2.3.1) La carte thermique (Heatmap Correlation) :

Une carte thermique de corrélation est une carte qui montre une matrice de corrélation en 2D entre deux dimensions discrètes, en utilisant des cellules colorées pour représenter des données provenant généralement d'une échelle monochromatique. Les valeurs de la première dimension apparaissent comme les lignes du tableau tandis que celles de la seconde dimension comme une colonne. La couleur de la cellule est proportionnelle au nombre de mesures qui correspondent à la valeur de la dimension.

Les cartes thermiques de corrélation sont donc idéales pour l'analyse des données, car elles rendent les modèles facilement lisibles et compréhensibles et mettent en évidence les relations et les variations entre les données. Dans cette section nous avons utilisé cette carte pour identifier la colinéarité entre les variables d'entrées ce qui permet de sélectionner les paramètres les plus pertinents pour notre variable de sortie, par exemple si deux paramètres sont fortement corrélés (positivement ou négativement) ils rendront le modèle plus complexe cela nécessite d'éliminer l'un de ces paramètres.

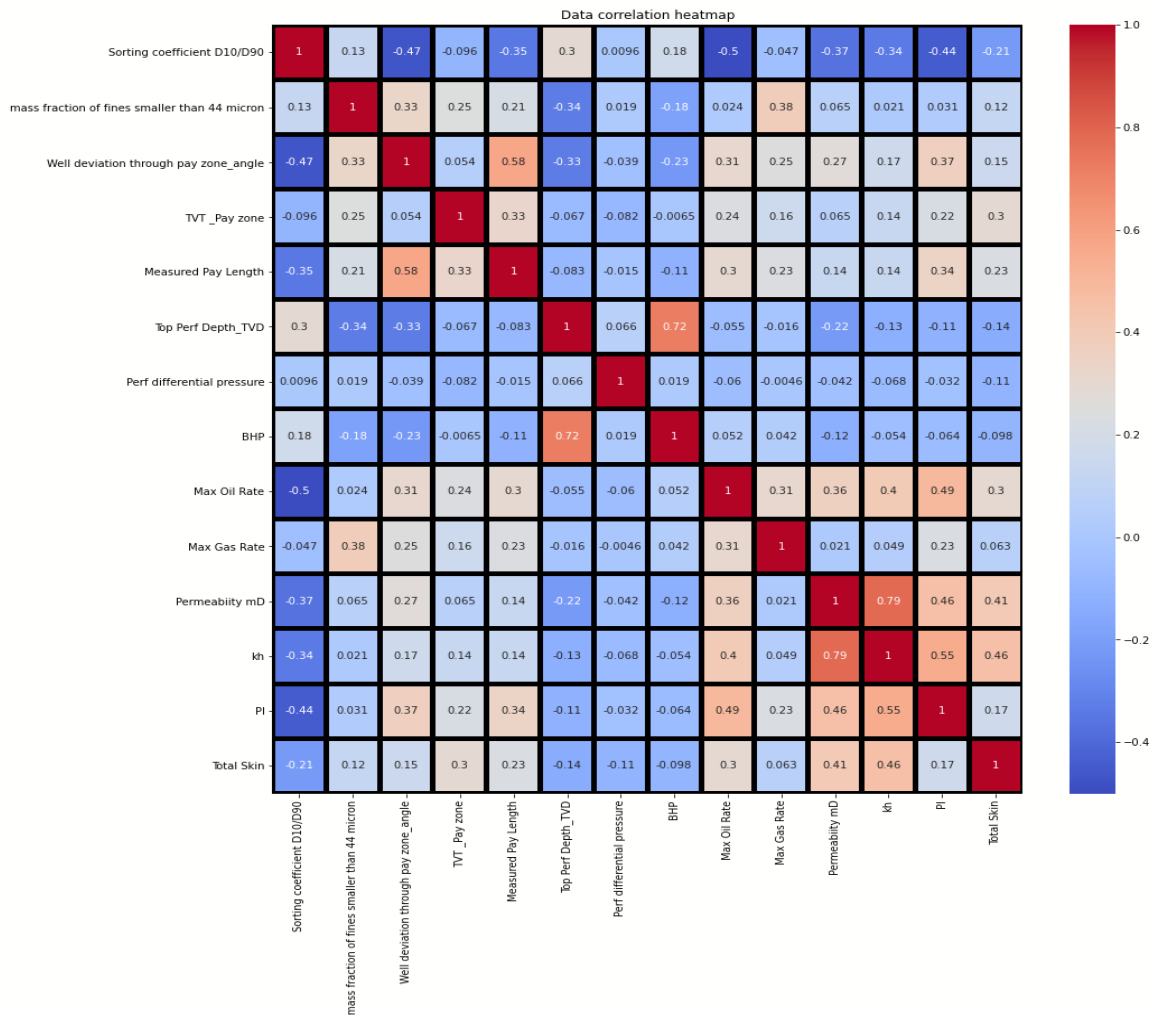


Figure 55: La carte thermique de corrélation pour notre ensemble de données

La figure 55 représente la carte thermique de notre ensemble de données, sur laquelle nous avons observé les éléments suivants :

- Le coefficient de corrélation entre « Measured Pay Length » et « Well deviation through pay zone_angle » est de 0.58 ce qui indique la forte relation entre eux.
- Pour le « Top Perf Depth_TVD » et « BHP » ils sont corrélés avec un coefficient élevé de 0.72 et d'un point de vue technique c'est évident.
- Pour la « Perméabilité » et « kh », déjà on sait qu'ils sont liés physiquement et le coefficient de 0.79 confirme cette réalité.

5.2.3.2) Tracé en paires de Seaborn :

Les graphes en paires de seaborn permettent d'analyser et de visualiser les relations entre les différents paramètres d'entrés et même les interactions de ces dernières avec la sortie (les méthodes de contrôle du sable).

À travers ces tracés mentionnés dans la page suivante (Figure 57), nous avons déduit les points suivants :

- « BHP » évolue linéairement avec « Top perf depth_TVD » comme indiqué dans la figure 56.
- « Perf differential pressure » n'a aucune relation avec les autres paramètres, et d'un point de vue physique elle n'est pas considérée comme un paramètre de sélection des méthodes de prévention du sable.

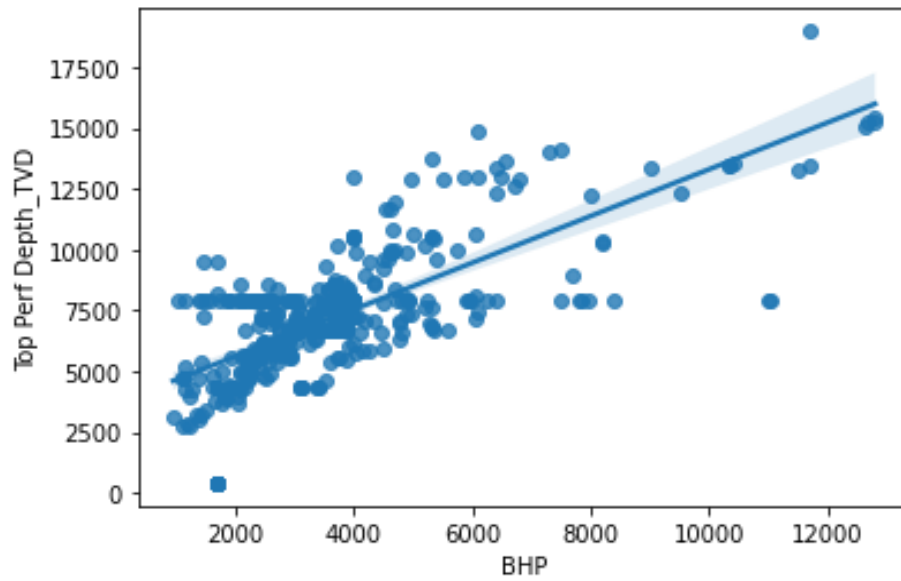


Figure 56:Tracé de courbe Top perf depth_TVD vs BHP

Après avoir analysé toutes les relations de colinéarité et explorer les tracés de seaborn, nous avons choisi d'éliminer les paramètres suivants puisqu'ils ont moins d'influence lors de la sélection des méthodes de contrôle du sable :

- Measured Pay Length
- Top Perf Depth_TVD
- kh
- Perf differential pressure



Figure 57: La représentation graphique de la relation par paire de l'ensemble des données

5.2.4) Traitement des valeurs aberrantes :

Tout modèle d'apprentissage automatique doit nécessairement subir un traitement des valeurs aberrantes afin d'assurer que chaque variable est comprise dans un intervalle qui correspond à son sens physique. Dans cette section, nous avons commencé par localiser les anomalies dans l'ensemble de données en utilisant le diagramme en boîte qui montre explicitement la plage de chaque paramètre (variable numérique), ce qui permet de voir clairement les valeurs erronées sous forme de petits hexagones. La figure 58 montre le taux d'anomalie pour chacune des variables suivantes (TVT Pay zone ; Permability ; PI ; Total skin) :

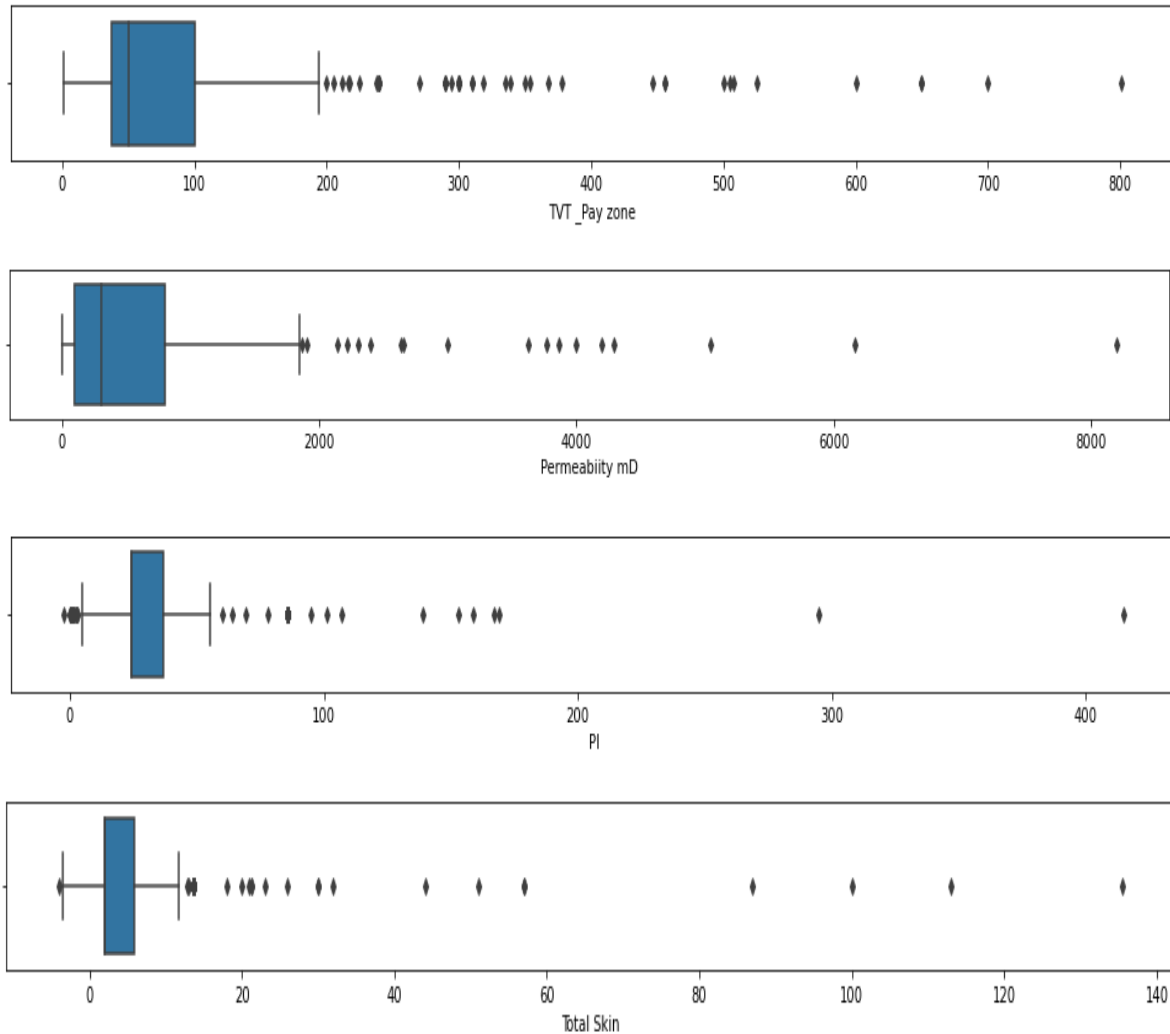


Figure 58: Diagramme en boîte de l'ensemble des données

À l'aide de l'analyse statistique des données et le diagramme en boîte ci-dessus, nous avons effectué un conditionnement de plage de manière d'assurer que l'intervalle de chaque paramètre mentionné ci-dessus est correct d'un point de vue technique. Les intervalles des différents paramètres sont comme suit :

- $5 \text{ ft} < \text{TVT_Pay zone} < 200 \text{ ft}$
- $100 \text{ md} < \text{Permeability} < 1500 \text{ md}$
- $5 \text{ bbl/ day/ psi} < \text{PI} < 100 \text{ bbl/ day/ psi}$
- $-4 < \text{Total Skin} < 30$

Tout simplement, nous n'avons pas utilisé les méthodes traditionnelles des valeurs aberrantes comme (Percentile, One class svm, Local outlier factor) puisque ces techniques ne tiennent pas compte de sens physique des données, en outre elles fournissent une faible performance du modèle. Il est donc préférable de s'en remettre au sens d'ingénierie.

5.2.5) Normalisation des données :

La normalisation de l'ensemble des données, comme mentionné dans le chapitre 4, consiste à s'assurer que la gamme des paramètres est similaire et comprise entre [0,1]. Dans ce but, la méthode "Min Max Scaler" fournie dans la bibliothèque "scikit -learn" a été utilisée pour effectuer cette transformation, comme le montre le tableau 12 :

Tableau 12: Description statistique des données après la normalisation

| Paramètres | Nombre d'échantillons | Moyenne | Std | Min | 25% | 50% | 75% | Max |
|---------------------------------------|-----------------------|---------|------|-----|-------|------|------|-----|
| Sorting coefficient D10/D90 | 267 | 0.44 | 0.13 | 0 | 0.47 | 0.51 | 0.51 | 1 |
| Mass fraction of fines less than 44µm | 267 | 0.08 | 0.11 | 0 | 0.046 | 0.05 | 0.07 | 1 |
| Well deviation through pay zone angle | 267 | 0.41 | 0.24 | 0 | 0.27 | 0.35 | 0.59 | 1 |
| TVT_ pay zone | 267 | 0.29 | 0.19 | 0 | 0.16 | 0.24 | 0.37 | 1 |
| BHP | 267 | 0.22 | 0.17 | 0 | 0.11 | 0.20 | 0.26 | 1 |
| Max oil rate | 267 | 0.10 | 0.14 | 0 | 0.04 | 0.06 | 0.12 | 1 |
| Max gas rate | 267 | 0.15 | 0.18 | 0 | 0.04 | 0.12 | 0.16 | 1 |
| Permeability | 267 | 0.34 | 0.31 | 0 | 0.17 | 0.18 | 0.67 | 1 |
| PI | 267 | 0.33 | 0.25 | 0 | 0.23 | 0.24 | 0.32 | 1 |
| Total skin | 267 | 0.22 | 0.12 | 0 | 0.17 | 0.18 | 0.18 | 1 |

5.3) Développement des modèles :

L'objectif de tous les modèles d'apprentissage automatique est de générer des prédictions pour des données non observables, afin de pouvoir les généraliser à des scénarios futurs. Le développement de tels modèles est divisé en deux phases : l'entraînement, au cours duquel le modèle va mémoriser et s'adapter à toutes les données, et l'évaluation, au cours de laquelle le modèle sera testé sur des nouvelles données afin de déterminer ses performances à long terme.

Pour notre cas nous avons classé l'ensemble des données de manière aléatoire comme suit :

- Le groupe d'entraînement (70% de l'ensemble des données, 187 puits) et le groupe du test (30% de l'ensemble des données, 80 puits)

Il faut noter que 4 paramètres ont été éliminés en plus inclus :

- « Geographical area » puisque notre modèle va être indépendant de n'importe quelle région.
- « Screen Type » parce que dans la réalité après avoir sélectionné les crépines seules (SAS) comme une technique de prévention du sable, ils utilisent des essais de rétention du sable pour choisir la meilleure crépine possible.
- « PI » et « Total skin », malgré nous savons qu'ils sont très importants dans le processus de sélection mais malheureusement ils réduisent la performance du modèle.

Finalement, les paramètres utilisés pour le développement de notre modèle sont 9 y compris (Sorting coefficient D10/D95; Mass fraction of fines less than 44 micron; Well deviation through pay zone angle; TVT_ pay zone; BHP; Max oil rate; Max gas rate; Permeability; Well Type.

Il est important de mentionner que tous les modèles ont été construits à l'aide du langage de programmation python 3.8 et de ses librairies.

5.3.1) La création du modèle de sélection des méthodes de contrôle du sable :

5.3.1.1) Sélection du modèle :

Afin de trouver le modèle le plus approprié et le plus robuste pour nos données de contrôle du sable, plusieurs modèles ont été construits et évalués. Tout d'abord, les algorithmes de classification utilisés sont : K-Nearest Neighbors (KNN), la forêt aléatoire (Random Forest), la machine à vecteurs de support (Support Vector Machine) et l'arbre de décision (Decision Tree).

En second, l'évaluation de la performance des modèles a été basé sur les métriques d'évaluation, à savoir le score f1 (f1_score), la précision (precision) et le rappel (recall).

Les résultats sont présentés dans le tableau 13 et les diagrammes à barres ci-dessous :

Tableau 13: Analyse des performances des modèles intelligents implémentées pour prédire la méthode appropriée de contrôle du sable

| Modèle | Données d'entraînement | | | Données de test | | |
|------------|------------------------|-----------|--------|-----------------|-----------|--------|
| | F1_score | Precision | recall | F1_score | Precision | recall |
| KNN | 0.8886 | 0.8905 | 0.8924 | 0.8665 | 0.8693 | 0.8641 |
| RF | 1 | 1 | 1 | 0.8920 | 0.9099 | 0.9012 |
| SVM | 0.7996 | 0.7553 | 0.8548 | 0.7783 | 0.7281 | 0.8395 |
| DT | 1 | 1 | 1 | 0.8578 | 0.8733 | 0.8641 |

A partir des résultats présentés, Random Forest était le plus performant à l'ensemble de données en assurant les métriques d'évaluation les plus élevés. Par conséquent, ce modèle est sélectionné pour la suite du développement (optimisation) et de la discussion.

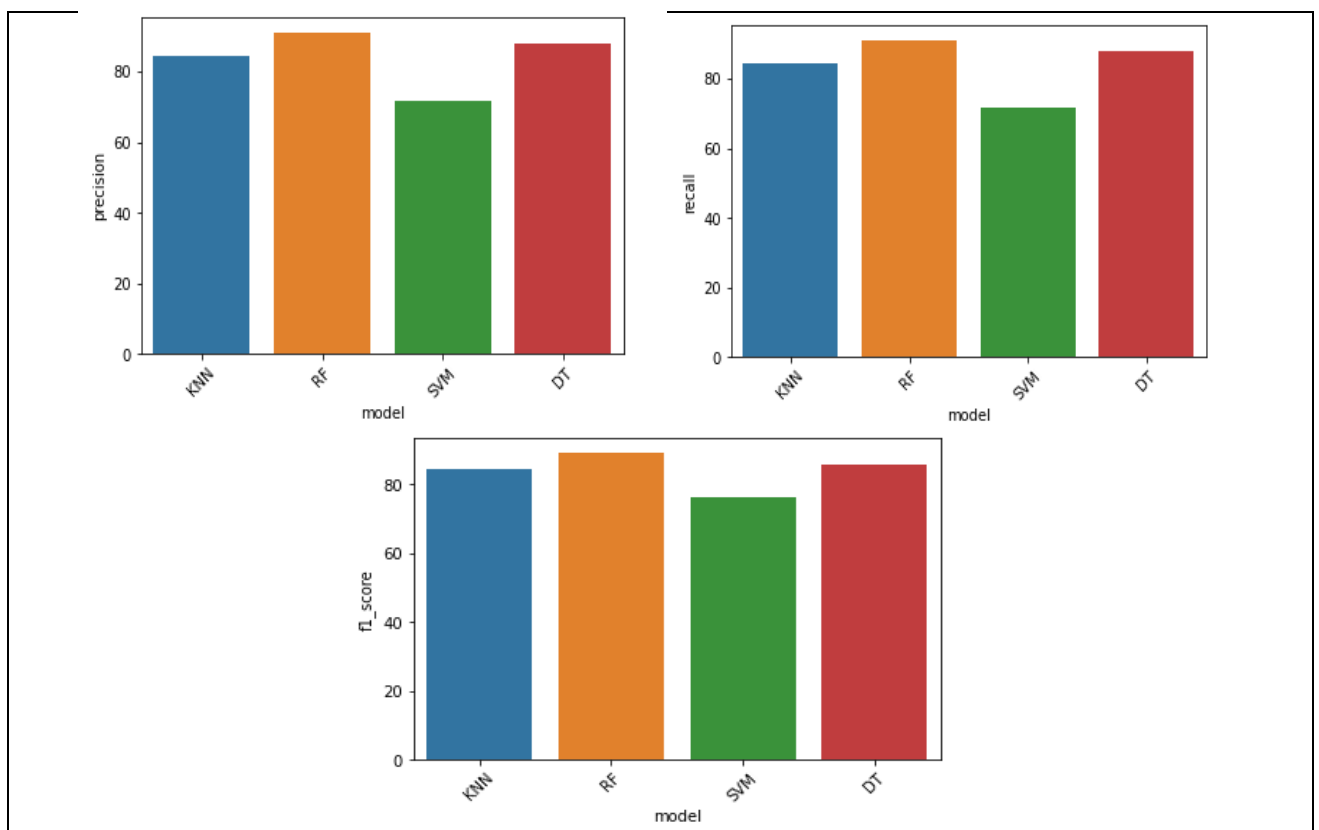


Figure 59: Diagrammes à barres de f1_score, precision, recall des quatres modèles

5.3.1.2) Optimisation du modèle :

Afin d'obtenir un modèle robuste qui est capable d'être généralisée pour des futures prédictions, il faut l'optimiser en utilisant le meilleur ensemble d'hyperparamètres.

Nous entamons toujours par générer une grande variété de combinaisons d'hyperparamètres, par la suite nous choisissons les meilleures parmi eux pour assurer les meilleures performances du modèle. Il existe deux techniques pour réaliser cette étape : Recherche par Grille (Grid Search), Recherche Aléatoire (Random Search). Pour notre cas nous avons utilisé la deuxième technique avec un schéma de validation croisée à trois reprises sur notre ensemble des données.

Les hyperparamètres de modèle Random Forest sont mentionnés ci-dessous, par la suite la meilleure combinaison des hyperparamètres pour notre modèle est indiquée dans le tableau 14 :

- **Nombre d'arbres** : représente le nombre d'arbres dans la forêt.
- **Minimum d'échantillons_feuille** : ce paramètre est défini comme "le nombre minimum d'échantillons requis pour être à un nœud de feuille", il est utilisé pour contrôler l'overffiting.
- **Minimum d'échantillons divisés** : il est défini comme "Le nombre minimum d'échantillons requis pour diviser un noeud interne".
- **Variables maximales** : défini comme "Le nombre de variables à prendre en compte lors de la recherche de la meilleure division.
- **Profondeur maximale** : elle est définie comme la "profondeur maximale de l'arbre" et la valeur par défaut est "None", ce qui signifie que les nœuds sont développés jusqu'à ce que toutes les feuilles soient pures.
- **Critère** : est utilisé pour mesurer la qualité d'un fractionnement.

Tableau 14: Sélection des hyperparamètres pour notre modèle Random Forest

| Modèle | Hyperparamètres | Intervalle | Meilleurs valeurs |
|----------------------|--------------------------------|--------------------------------------|-------------------|
| Random Forest | Nombre d'arbre | 10, 30, 50,70,90,110,130,150,170,190 | 110 |
| | Minimum d'échantillons_feuille | 0.005,0.015,0.025,0.035,0.045 | 0.025 |
| | Minimum d'échantillons divisés | 0.005,0.015,0.025,0.035,0.045 | 0.045 |
| | Variabes maximales | 'sqrt' ou 'log2' | log2 |
| | Profondeur maximale | 5,15,25,35,45, None | 45 |
| | Critère | 'gini' ou 'entropie' | entropie |

Il faut noter que nous avons appliqué la méthode SMOTE qui tient compte du déséquilibre de différentes méthodes de contrôle du sable dans notre ensemble des données avant d'exécuter le modèle.

D'après la comparaison entre les résultats d'entraînement et du test dans les tableaux 13 et 15 nous remarquons que l'utilisation des hyperparamètres optimaux a permet de réduire la variance du modèle (moins de overffiting).

Tableau 15: Analyse de la performance du Random Forest après l'optimisation

| Modèle | Données d'entrainement | | | Données de teste | | |
|----------------------|------------------------|-----------|--------|------------------|-----------|--------|
| | F1_score | Précision | Rappel | F1_score | Précision | Rappel |
| Random Forest | 0.9680 | 0.9680 | 0.9680 | 0.9568 | 0.9580 | 0.9568 |

Puisque notre étude vise à sélectionner la méthode de contrôle du sable approprié pour un puits donné, il est nécessaire d'assurer que notre modèle d'apprentissage automatique possède une bonne performance par rapport aux différentes techniques de contrôle disponibles. Les résultats présentés dans le rapport de classification ci-dessous indiquent que le modèle n'a pas tendance à prédire une méthode plutôt qu'une autre :

Tableau 16: Le rapport de classification du Random Forest

| Métriques d'évaluation Méthodes de contrôle | Précision | Rappel | F1_score | Nombre de puits dans la classe |
|--|-----------|--------|----------|--------------------------------|
| CHGP | 0.98 | 0.98 | 0.98 | 45 |
| FP | 0.96 | 0.94 | 0.95 | 52 |
| HRWP | 0.96 | 0.98 | 0.97 | 54 |
| OHGP | 0.98 | 0.91 | 0.94 | 55 |
| SCREENS | 0.91 | 0.98 | 0.94 | 49 |
| Exactitude | 0.96 | | | 255 |
| Moyenne macro | 0.96 | 0.96 | 0.96 | 255 |
| Moyenne pondérée | 0.96 | 0.96 | 0.96 | 255 |

Ce qui peut être analysé à partir du tableau 16, les métriques d'évaluation (précision, rappel, f1_score) ont des valeurs très proches ce qui montre la bonne performance du modèle en tant qu'outil de sélection très puissant.

La colonne 'Nombre de puits dans la classe' montre que le problème du déséquilibre de la classe a été résolu.

5.3.1.3) Validation du modèle :

La validation croisée est l'utilisation de diverses techniques pour évaluer la capacité d'un modèle d'apprentissage automatique à généraliser lors du traitement d'ensembles de données nouveaux et inconnus. Donc c'est une étape importante pour garantir la précision du modèle avant de le déployer dans un environnement réel.

Dans cette étape nous avons utilisé un schéma de validation croisée à 03 reprises plus précisément la technique « Straified K Fold ». La figure 60 illustre les résultats en termes de : f1 macro, précision macro, rappel macro tenant compte que ces métriques représentent la moyenne des trois reprises.

Tableau 17: Analyse de la performance de Random Forest lors de la validation

| Métriques d'évaluation | F1 macro % | Précision macro % | Rappel macro % |
|------------------------|------------|-------------------|----------------|
| Score | 88.85 | 89.14 | 89.38 |

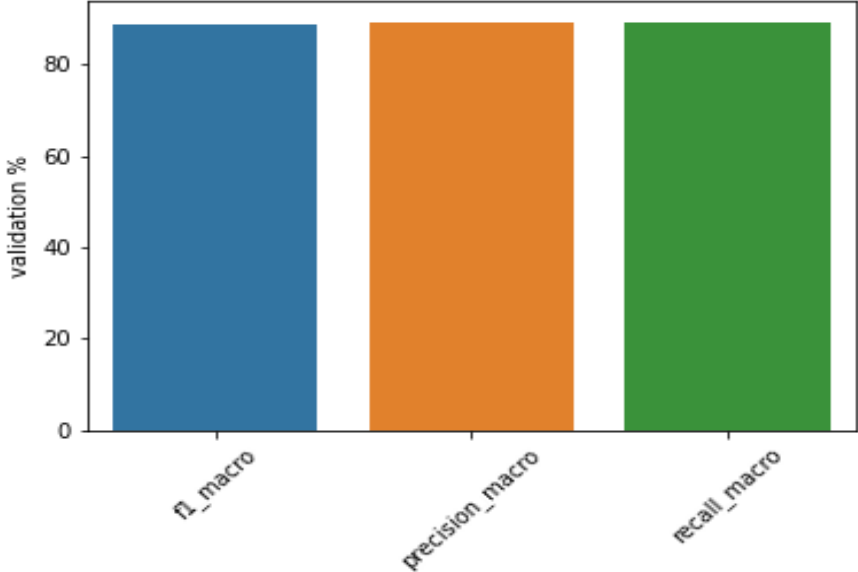


Figure 60:Diagramme à barres de f1_macro, précision_macro, rappel_macro du modèle sélectionné

Conclusion

Le contrôle du sable est une partie fondamentale de la conception de la complétion des puits ; la plupart des champs qui produisent du sable nécessitent une certaine forme de contrôle du sable. Heureusement, différentes techniques existent pour résoudre ce problème en fonction de la complexité de chaque région. Cependant, le fait que le processus de sélection de la meilleure technique possible reste complexe est ce qui nous a motivés à envisager le développement d'une nouvelle solution.

L'objectif de cette étude était de construire un modèle d'apprentissage automatique qui permet de sélectionner la meilleure méthode de prévention du sable. Une base de données adoptée à partir de la littérature publiée a été utilisée pour construire quatre modèles intelligents incluant K-nearest neighbors, Random Forest, SVM, Decision Tree en utilisant le langage de programmation Python. A partir des résultats obtenus, les conclusions suivantes peuvent être tirées :

1. Le modèle Random Forest après l'optimisation présente une valeur F1 score de 0.9568, une valeur Précision de 0.9580 et une valeur Rappel de 0.9568.
2. La matrice de confusion indique que le modèle assure une bonne performance par rapport à différentes techniques de contrôle disponibles (CHGP, FP, HRWP, OHGP, Screens)

La solution proposée dans cette thèse représente un outil de décision puissant. De plus, elle constitue un système de recommandation potentiel qui aidera les ingénieurs de complétion à faire le meilleur choix possible.

En raison d'un manque de données, certaines limites dans notre recherche sont apparues, que nous espérons aborder dans les occasions futures (thèse de doctorat et articles). Par conséquent, nos principales perspectives d'avenir sont les suivantes :

1. Examiner l'impact économique sur le choix de la meilleure méthode de contrôle du sable en incorporant de nouveaux paramètres tels que les coûts d'installation, etc.
2. Créer un modèle d'apprentissage automatique capable d'évaluer le succès de la méthode choisie et d'identifier les échecs sur la base de données de fiabilité.

Bibliographie

1. Calizaya Castillo. Luis Augusto. Completion Technology for Unconsolidated Formations. 1995.
2. K. William, P.E. Ott, and D.Woods. Joe. Word Oil Modern Sandface Completion Practices Handbook. Houston, 2003: World Oil magazine Gulf Publishing Company.
3. Han. Gang, et al. A Comprehensive Study of Sanding Rate From a Gas Field From Reservoir to Completion Production and Surface Facilities. 2011: Society of Petroleum Engineers.
4. Ni.Qinglai , Mohamad-Hussein.Assef. Numerical modeling of onset and rate of sand production in perforated wells. 2018 : Journal of Petroleum Exploration and Production Technology.
5. H.H. Abass, H.A. Nasr-El-Din, M.H. BaTaweel. Sand Control: Sand Characterization, Failure Mechanisms, and Completion Methods. Texas, 2002: Society of Petroleum Engineers.
6. Matanovic. Davorin, Cikes. Marin, Moslavac.Bojan. Sand Control in Well Construction and Operation.2012 : Springer Environmental Science and Engineering.
7. Baghdadi. Faical, et al. An Innovative Approach in Sand Onset Prediction at Different Water Cuts and Estimating Sand Production for Future Field Life. Beijing, 2019: International Petroleum Technology Conference.
8. Pasic. Borivoje, Gaurina-Medimurec. Nediljka, Matanovi. Davorin. Wellbore Instability: Causes and Consequences. Croatia, 2007 : University of Zagreb, Faculty of Mining Geology and Petroleum Engineering.
9. Rahman.Khalil, Khaksar. Abbas, Kayes. Toby. An Integrated Geomechanical and Passive Sand-Control Approach to Minimizing Sanding Risk From Openhole and Cased-and-Perforated Wells. 2010 : SPE Drilling &Completion.
10. X. Chen, C.P. Tan, C.M. Haberfeld. A comprehensive practical approach for wellbore instability management. 2002 : SPE Drilling & Completion.
11. Bowes. Colin, Procter. Ray. Drillers Stuck Pipe Handbook. Scotland, 1997 : Procter & Collins Ltd.
12. Subbiah. Surej Kumar, et al. Root cause of sand production and methodologies for prediction.2020 : Petroleum, <https://doi.org/10.1016/j.petlm.2020.09.007>.
13. Udebhulu Okhiria. Dickson . Mechanistic Models for Predicting Sand Production. Nigeria, 2014 : Department of Petroleum Engineering, The African University of Science and Technology.
14. Twynam. Allan, et al. Practical Approaches to Sand Management. 2004 : Oilfield Review.
15. H.H. Abass, A.H. Habbtar, A. Shebatalhamd. Sand Control during Drilling, Perforation, Completion and Production. Bahrain, 2003 : Society of Petroleum Engineers.
16. Sung Hyun. Kim. A Predictive Model for Sand Production in Poorly Consolidated Sands. 2010 : Faculty of the Graduate School of The University of Texas at Austin.
17. Olamuyiwa Oladoyin. Kolawole. Formation Susceptibility to Wellbore Instability and Sand Production in the Hajdúszoboszló field, Pannonian Basin: Hungary. Miskolc, Hungary, 2018 : University of Miskolc, Faculty of Earth Science and Engineering Petroleum and Natural Gas Institute.
18. Babs. Oyeneyin. Sand Control Completion Strategy. Integrated Sand Management For Effective Hydrocarbon Flow Assurance Volume 63. 2015 : Elsevier, Developments in Petroleum Science.
19. Bellarby. Jonathan. Sand Control/Well Completion Design. Aberdeen UK, 2009 : Elsevier,Developments in Petroleum Science.

Bibliographie

20. Syltøy. Christer. New Generation Expandable Sand Screens. Stavanger,2014 : Faculty of Science and Technology, University of Stavanger.
21. DuneFront Limited. [Online] London, 2012. <https://www.website.dunefront.com/applications>.
22. Reyes. Raymond, Sipi. Aluyah. Optimal Sand Control Design & Technique Selection: A Simplified Practical Guidance Tool. Malaysia,2018 : Offshore Technology Conference.
23. Waltman. B, et al. Technologies or Intangibles: What Drives Sand Control Selection. Texas, 2010 : Offshore Technology Conference.
24. Tiffin. D.L, et al. New Criteria for Gravel and Screen Selection for Sand Control. Los Angeles, 1998 : Society of Petroleum Engineers.
25. Price-Smith. C, et al. Design Methodology for Selection of Horizontal Openhole Sand Control Completions Supported by Field Case Histories. 2003 : SPE Drilling & Completion.
26. Latiff. Amir. Assessment and Evaluation of Sand Control Methods for a North Sea Field. 2011 : Imperial College London,Department of Earth Science and Engineering Center for Petroleum Studies.
27. Farrow. Chris, Munro. David, McCarthy. Thomas. Screening Methodology for Downhole Sand Control Selection. Australia,2004 : Society of Petroleum Engineers.
28. Mathisen. A.M, Aastveit. G.L, Alteras. E. Successful Installation of Stand Alone Screen in More than 200 Wells-The importance of Screen Selection Process and Fluid Qualification. Netherlands,2007 : Society of Petroleum Engineers.
29. Bryne. Michael, Slayter. Adrian, McCurdy. Philip. Improved Selection Criteria for Sand Control- When Are "Fines" Fines. USA, 2010 : Society of Petroleum Engineers.
30. Chan. Keng.Seng, et al. Production Integrated Sand Control Benchmark for Field Development. China, 2013 : International Petroleum Technology Conference.
31. Khamehchi. Ehsan, Ameri. Omid, Alizadeh. Ali. Choosing an Optimum Sand Control Method. 2015 : Egyptian Petroleum Research Institute.
32. Hu. Shaofeng, et al. Laboratory Investigation on Optimum Selection of a Sand Control Method for an Interbedded Sandstone and Mudstone Reservoir. 2020 : Hindawi Geofluids.
33. Belyadi. Hoss, Haghghat. Alireza. Machine Learning Guide for Oil and Gas Using Python A Step-by-Step Breakdown with Data, Algorithms, Codes, and Applications. 2021: Gulf Professional Publishing.
34. Berrehal. Badr Eddine. Conception et réalisation de la solution intelligente Deep Drill. Boumerdes, 2021 : Université de M'Hamed Bougara Boumerdes, Faculté des Hydrocarbures et de la Chimie (FHC).
35. Stewart. Matthew. The limitations of Machine Learning.[Online]. Disponible sur: <https://towardsdatascience.com/the-limitations-of-machine-learning>.
36. Wes. McKinney. Python for Data Analysis. United States of America, 2013 : O'Reilly Media, Inc.
37. Narayan Pandey. Yogendra, et al. Machine Learning in the Oil and Gas Industry Including Geosciences, Reservoir Engineering, and Production Engineering with Python. USA, 2020 : Apress Media LLC..
38. Bangert. Patrick. Machine Learning and Data Science in the Oil and Gas Industry Best Practices, Tools, and Case Studies. 2021 : Gulf Professional Publishing.
40. Géron. Aurélien. Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow Concepts, Tools, and Techniques to Build Intelligent Systems. United States of America, 2019 : O'Reilly Media Inc.
41. DataScientest. Qu'est-ce que l'algorithme KNN. 2020. Disponible sur: <https://datascientest.com/knn>.

Annexes

1) Prédiction de production du sable :

La production de sable est un phénomène très fréquent dans les puits de pétroliers et gaziers, qui nécessite une ou plusieurs solutions appropriées selon la région. Comme nous avons discuté des méthodes de contrôle du sable à l'aide d'algorithmes d'apprentissage automatique, nous avons envisagé d'utiliser cet outil pour prédire l'occurrence de la production de sable, en gardant à l'esprit que dans l'industrie, les ingénieurs se basent sur des études géomécaniques pour déterminer si un puits donné produit du sable ou non.

1.1) Description des données :

Les données utilisées dans cette partie représentent 31 puits de gaz localisés dans la région de bassin adriatique en Italie, c'est une zone géographique qui rencontre beaucoup de production du sable. Les paramètres disponibles sont les suivants :

Tableau An 1 : Les principaux paramètres dans le jeu des données

| Paramètres | Signification |
|-------------------------------------|---|
| TVD (m) | Profondeur verticale réelle |
| TT ($\mu\text{sec} / \text{ft}$) | Temps de déplacement sonique |
| COH (Kg/cm^2) | Force de cohésion |
| Q gas (Ksm^3/d) | Débit du gaz |
| Q water (l/d) | Débit d'eau |
| LGR ($1 / \text{Ksm}^3/\text{d}$) | Rapport gaz liquide |
| BHSP (Kg/cm^2) | Pression statique du fond |
| BHFP (Kg/cm^2) | Pression dynamique du fond |
| DD (Kg/cm^2) | Drawdown |
| CTD (Kg/cm^2) | Drawdown critique total au début de production du sable |
| EOVS (Kg/cm^2) | Contrainte effective verticale |
| SPF | Tirs par pied |
| Hperf (m) | Intervalle de perforation |

Ces données ont été obtenues à partir de la littérature publiée, plus précisément de l'article (SPE 28066) intitulé « Mécanique des roches à l'échelle du bassin : Observations sur le terrain de la production de sable ». Il convient de noter que 31 puits sont insuffisants pour développer un modèle d'apprentissage automatique car il en faut beaucoup plus pour entraîner ce modèle afin de le généraliser pour des prédictions futures ; ainsi, cette section ne fait qu'introduire le concept « d'apprentissage automatique et d'intelligence artificielle » dans le problème de la production de sable.

1.2) Sélection des paramètres principaux :

Tout d'abord, il faut savoir qu'on a un problème de classification binaire où nous avons essayé de prédire l'existence de production du sable ou non :

(1 : Production du sable observée ; 0 : Production du sable non observée).

Pour le but de sélectionner les variables les plus pertinents pour notre variable de sortie, nous avons utilisé la carte thermique de corrélation et les graphiques en paires de seaborn.

1.2.1) La carte thermique (Corrélations de Pearson) :

D'après la figure mentionnée ci-dessous, nous avons remarqué les points suivants :

- ‘TVD’ et ‘COH’ sont fortement corrélés avec un coefficient de 0.93.
- ‘TT’ et ‘COH’ sont négativement corrélés avec un coefficient assez élevé de (-0.78).
- ‘Qw’ et ‘LGR’ sont positivement corrélés avec un coefficient de 0.74 et c’est évident puisque le rapport gaz liquide est égal au débit du gaz sur le débit d’eau.
- ‘EOVS’ et ‘COH’ sont fortement corrélés avec un coefficient élevé de 0.83.

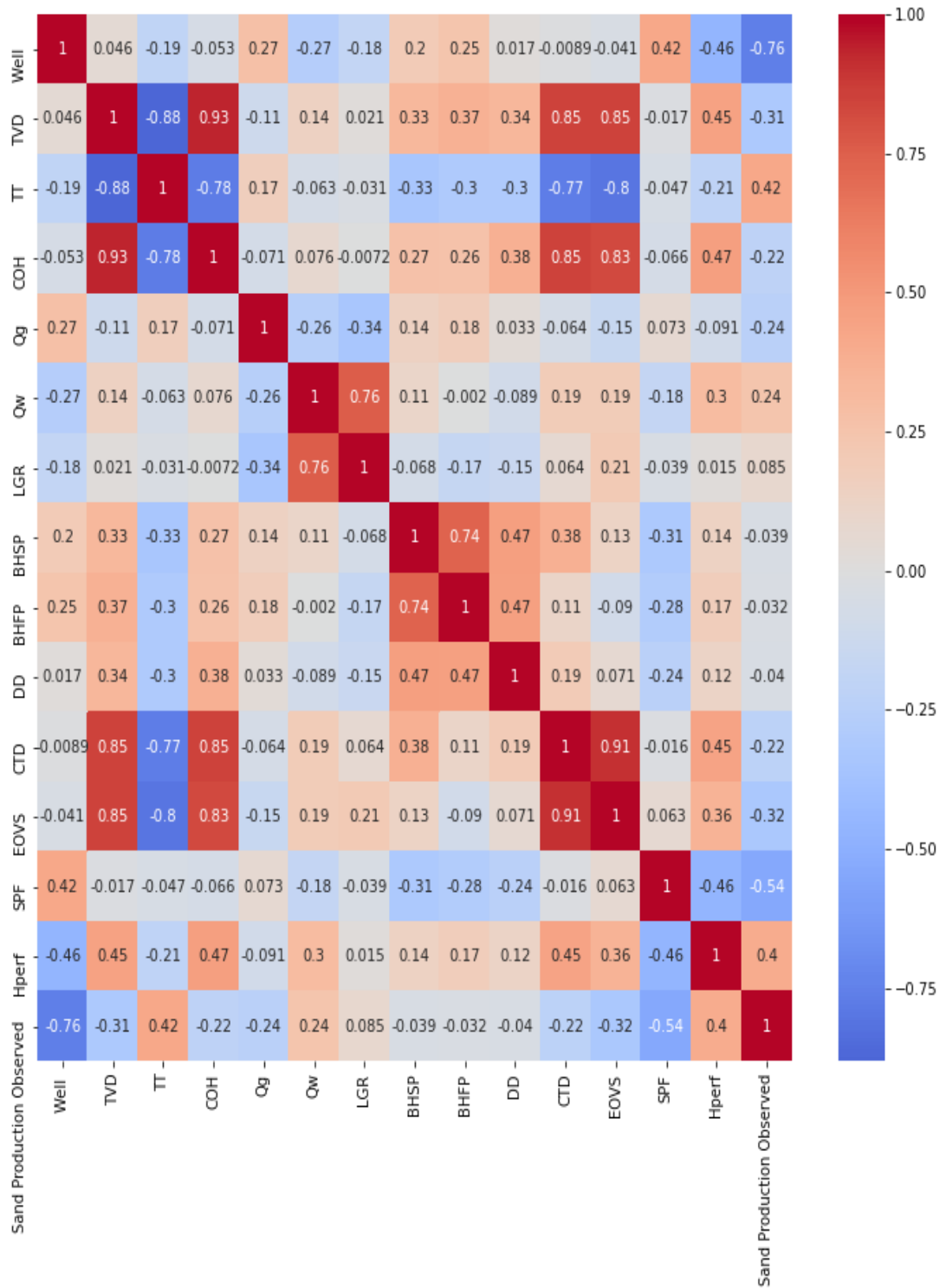


Figure An 1 : La carte thermique de corrélation pour l’ensemble des données de production du sable

1.2.2) Les graphiques en paires :

Les courbes en paires confirment les corrélations obtenues à partir de la carte thermique par exemple :

- ‘TVD’ et ‘COH’ : nous avons constaté que chaque fois la profondeur augmente la consolidation de la roche représentée par la force cohésive (entre les grains) augmente d’une façon logarithmique.

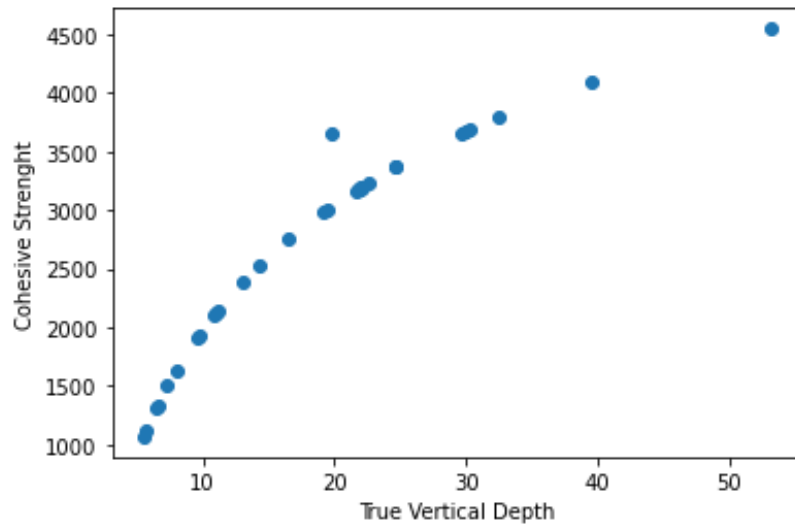


Figure An 2 : Tracé de courbe COH vs TVD

- ‘TT’ et ‘COH’ : nous avons remarqué que le temps de déplacement des ondes soniques diminue dans le cas des roches très consolidées (COH élevé), tout simplement parce que la forte cohésion entre les grains ne permet pas le passage des ondes soniques ce qui résulte en temps de déplacement faible.

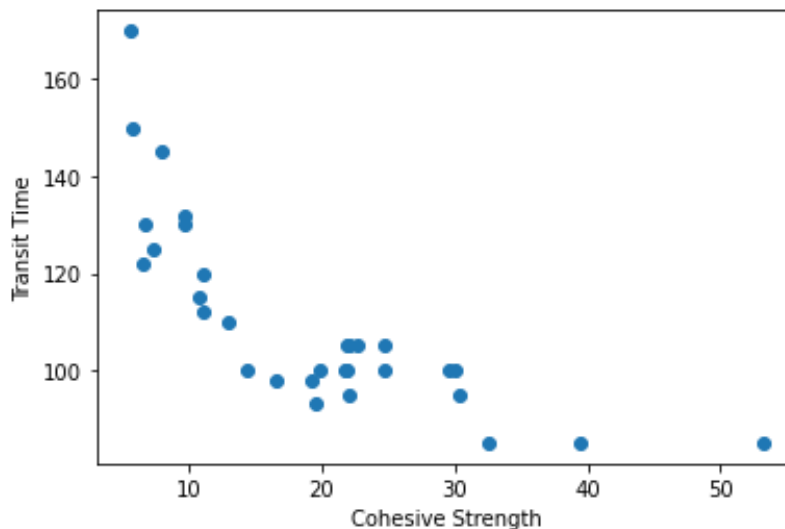


Figure An 3 : Tracé de courbe TT vs COH

- ‘EOVS’ et ‘COH’ : La roche devient plus consolidée lorsque la contrainte verticale effective est assez élevée, puisque cette dernière représente le poids des couches superposés. Donc dans les profondeurs élevées, la roche est généralement assez consolidée pour empêcher la production du sable.

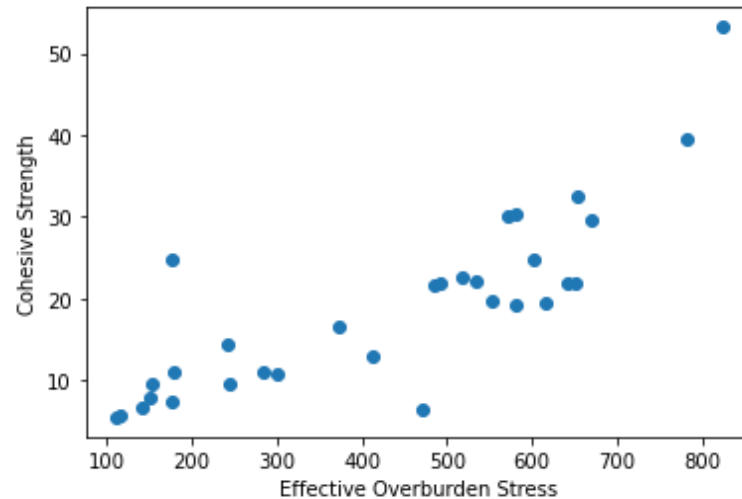


Figure An 4 : Tracé de courbe COH vs EOVS

Après avoir analysé tous les paramètres, nous avons choisi les suivants basant sur la colinéarité entre les variables d'entrée et la relation de ces dernières avec la production du sable :

- 'TT' ; 'COH' ; 'LGR' ; 'BHSP' ; 'CTD' ; 'SPF' ; 'Hperf'.

1.3) Développement du modèle :

Les variables mentionnées ci-dessus ont été utilisées pour le développement du modèle prédictif de production du sable. Quatre modèles de classification ont été utilisés à savoir : K-Nearest Neighbors ; Random Forest ; Support Vector Machine ; Decision Tree.

Pour éviter le problème de déséquilibre dans notre variable de sortie (Production du sable Observée), nous avons utilisé la méthode de SMOTE.

Les résultats obtenus ci-dessous sont ceux du modèle Random Forest, qui a obtenu les meilleures performances sur l'ensemble de données :

Tableau An 2 : La matrice de confusion du Random Forest

| Variable de sortie | Précision | Rappel | F1-score |
|----------------------------------|-----------|--------|----------|
| Production du sable non observée | 0.89 | 1 | 0.94 |
| Production du sable observée | 1 | 0.83 | 0.91 |

Ces métriques d'évaluation confirment la bonne performance du modèle ; cependant, comme indiqué précédemment, plus de données sont nécessaires pour développer un modèle généralisé pour les prédictions futures.