

N° Ordre...../FHC/UMBB/2023

République Algérienne Démocratique Et Populaire  
Ministère de l'enseignement Supérieur et de la Recherche Scientifique



**Université M'Hamed Bougara - Boumerdes**  
**Faculté des Hydrocarbures et de la Chimie**

**Département Gisements Miniers et Pétroliers**



**Mémoire de Fin d'Etude**  
**En vue de l'obtention du Diplôme**

**MASTER**

**Présenté par**

**YAHIAOUI SOUHILA**

**TAMAZIRT NESSERINE**

**Filière : Hydrocarbures**

**Spécialité : Génie pétrolier – Production des Hydrocarbures**

---

**APPORT DU BACK ALLOCATION POUR LE CONTROL DE LA  
PRODUCTION DES CHAMPS PÉTROLIERS**

---

Devant le jury composé :

MENAD Nait Amar	Docteur	SH DLAB-B	Président
ZIOUANI Sara	Docteur	SH IAP-B	Examineur
SAIFI Redha	MAA	UMB-B	Examineur
ZERAIBI Noureddine	Professeur	UMB-B	Encadreur
REDOUANE Kheireddine	Docteur	CRD-B	Co-encadreur

## Résumé

Cette étude se concentre sur l'utilisation de la back-allocation comme solution efficace pour corriger les débits calculés en l'absence de jaugeage individuel et détecter les anomalies dans la durée de vie des puits. Nous examinons les différentes corrélations théoriques existantes dans la littérature pour estimer les débits à travers les duses, tant pour le régime critique que non critique. Pour améliorer les performances des modèles théoriques nous utilisons les techniques de machine Learning telle que le random Forest et l'ANN et ce pour effectuer une analyse de sensibilité des paramètres à l'échelle du champ pétrolier. De même la meilleure équation de chaque régime a été réadaptée à des intervalles spécifiques de GOR, une autre alternative a été abordée à travers laquelle en évaluant la performance des puits individuellement par Le modèle LSTM.

Pour corriger les débits calculés, on détermine des coefficients  $k$  spécifiques à chaque puits, déterminés à partir de l'historique de production. Cette correction vise à améliorer la précision des estimations et à se rapprocher des débits réels. On cherche également à détecter les anomalies des débits à l'aide des deux méthodes différentes  $z$  score et LSTM

## الملخص

تركز هذه الدراسة على استخدام إعادة التخصيص العكسي كحل فعال لتصحيح التدفقات المحسوبة في حالة عدم توفر قياسات فردية، واكتشاف الشوائب في عمر الآبار. نحن ندرس النماذج النظرية المختلفة الموجودة لتقدير معدلات التدفق عبر البئر، سواء في النظام الحرج أو غير الحرج. لتحسين أداء النماذج النظرية، نستخدم تقنيات التعلم الآلي مثل الشبكات العصبية الاصطناعية لإجراء تحليل الحساسية على مستوى حقل النفط. بالإضافة إلى ذلك، تم تعديل أفضل معادلة لكل نظام في فترات زمنية محددة، تم التطرق إلى بديل آخر حيث يتم تقييم أداء الآبار بشكل فردي باستخدام نموذج ذاكرة طويلة وقصيرة المدى.

لتصحيح معدلات التدفق المحسوبة، يتم تحديد معاملات محددة لكل بئر، وذلك استنادًا إلى سجل الإنتاج التاريخي. يهدف هذا التصحيح إلى تحسين دقة التقديرات والاقتراب من معدلات التدفق الفعلية. كما نسعى أيضًا لاكتشاف الشوائب في معدلات التدفق باستخدام طريقتين مختلفتين: الترتيب القياسي ونموذج ذاكرة طويلة وقصيرة المدى.

## Abstract

This study focuses on the use of back-allocation as an effective solution to correct calculated flow rates in the absence of individual gauging and detect anomalies in well lifespan. We examine the different theoretical correlations existing in the literature to estimate flow rates through the wells, both for critical and non-critical regimes. To improve the performance of theoretical models, we employ machine learning techniques such as random forest and ANN to perform a sensitivity analysis of parameters at the field scale. Similarly, the best equation for each regime has been re-adapted to specific intervals of Gas-Oil Ratio (GOR), and an alternative approach has been addressed by evaluating the performance of individual wells using the LSTM model.

To correct the calculated flow rates, specific coefficients "k" are determined for each well, derived from production history. This correction aims to improve the accuracy of estimations and approach the actual flow rates. We also seek to detect flow rate anomalies using two different methods : z-score and LSTM.

## Remerciements

Nous sommes profondément reconnaissants envers Allah, le Clément et Miséricordieux, qui nous a accordé la force et la patience nécessaires pour mener à bien ce modeste travail.

Nous tenons à exprimer notre sincère gratitude envers nos parents et nos familles pour leur soutien constant et leurs encouragements tout au long de notre parcours. Leur présence bienveillante et leur encouragement ont été une source d'inspiration et de motivation inépuisable.

Nous souhaitons exprimer nos vifs remerciements à notre promoteur PR. ZERAIBI Noureddine, pour sa disponibilité, son sérieux et ses conseils judicieux. Sa présence et ses orientations éclairées ont été d'une valeur inestimable pour la réalisation de ce travail.

Nous adressons également nos sincères remerciements à notre Co-promoteur, le Dr. REDOUANE Kheireddine pour son investissement en temps, ses précieux conseils et ses orientations impeccables.

Nous tenons à exprimer notre profonde reconnaissance envers l'ensemble des enseignants de la FHC qui ont contribué à notre formation. Leur expertise, leur dévouement et leur passion pour l'enseignement ont largement contribué à notre développement académique.

Enfin, nous souhaitons exprimer notre gratitude à nos amis qui nous ont encouragés tout au long de notre parcours. Leur soutien indéfectible et leur amitié ont été d'une grande importance pour nous.

Nous tenons à remercier du fond du cœur toutes les personnes mentionnées ci-dessus, ainsi Ms AMROUN Salim, Ms BERKANE Salem qui ont contribué de près ou de loin à notre réussite. Leurs encouragements, conseils et soutien ont été d'une valeur inestimable, et nous ne saurions jamais assez les remercier.

# Table des matières

Résumé .....	i
Remerciements .....	iii
Liste des figures.....	vii
Liste des tableaux .....	ix
Nomenclature.....	x
Introduction générale.....	1
<b>CHAPITRE 1 : LES ECOULEMENTS MULTIPHASIQUES .....</b>	<b>3</b>
1.1 Introduction.....	4
1.2 Types d'écoulements multiphasiques.....	4
1.3 Écoulement de liquide monophasique .....	5
1.4 Écoulement diphasique .....	7
1.4.1 Les propriétés d'un écoulement diphasique .....	8
1.4.1.1 Holdup Liquide (HL) (fraction volumique).....	8
1.4.1.2 Vitesse superficielle .....	9
1.4.1.3 Masse volumique .....	9
1.4.1.4 Viscosité.....	9
1.4.2 Les Configurations d'écoulements diphasiques .....	10
1.4.2.1 Ecoulement vertical .....	10
1.4.2.2 Ecoulement horizontal .....	11
1.4.3 Modélisation des écoulements diphasiques.....	12
1.5 Conclusion .....	20
<b>CHAPITRE2 : MODELISATION DES ECOULEMENTS A TRAVERS LES</b>	
<b>RESTRICTION .....</b>	<b>21</b>
2.1 Introduction.....	22
2.2 Description de la Duse .....	22
2.3 Régime d'écoulements : critique et sous-critique.....	24
2.3.1 Écoulements critique .....	24
2.3.2 Écoulements sous critique .....	25
2.4 Modélisation des écoulements à travers les Duses .....	25
2.4.1 Modélisation des écoulements monophasiques.....	26
2.4.2 Modélisation des écoulements diphasiques .....	27
2.4.2.1 Les corrélations existantes dans le régime critique.....	28
2.4.2.2 Les corrélations pour Régime sous critique.....	30
2.5 Conclusion .....	32

CHAPITRE 3 : CALCUL ET OPTIMISATION DE LA PRODUCTION .....	33
3.1 L’algorithme de Régression non linéaire .....	34
3.2 Métriques d’évaluation d’un modèle .....	35
3.3 Résultats et discussions .....	36
3.3.1 Présentation des Données de production du champ de Hassi Messaoud : ....	36
3.3.2 Préparation des données .....	36
3.3.3 Méthodologie.....	38
3.3.4 Résultats de l'adaptation des modèles théoriques .....	38
3.3.4.1 Régime critique.....	38
3.3.4.2 Régime sous critique.....	41
3.3.5 Conclusion.....	44
3.4 Machine Learning (L’apprentissage automatique) .....	45
3.4.1 Les types de machine Learning .....	46
3.4.2 Réseau de neurones .....	47
3.4.2.1 Architecture du réseau de neurones .....	47
3.4.2.2 Random Forest .....	51
3.4.2.3 Réseau de neurones récurrent (RNN) .....	54
3.5 Résultat et discussion pour l’étude de sensibilités par les modèles intelligents ...	56
3.5.1 Méthodologie.....	56
3.5.2 Résultats.....	57
3.5.2.1 Analyse de sensibilité globale des inputs sur le champ complet .....	57
3.5.2.2 Etude de sensibilité suivant les intervalles de GOR .....	59
3.5.2.3 Etude de sensibilité de chaque puits .....	63
CHAPITRE 4 : APPLICATION DU BACK ALLOCATION.....	66
4.1 Introduction.....	67
4.2 Définition du Back allocation .....	67
4.3 Les structures back allocation.....	68
4.3.1 Structure à un niveau .....	68
4.3.2 Une structure d'allocation à plusieurs niveaux .....	68
4.4 Les types de back allocation : .....	69
4.5 Objectifs de back allocation.....	70
4.6 Mesure de Back allocation.....	71
4.6.1 La tête de puits de production.....	71
4.6.2 Manifolds et pipelines .....	74

4.6.3	Stockage.....	74
4.6.4	Télémetrie/SCADA .....	75
4.7	Une nouvelle approche de back allocation .....	76
4.8	Outil mathématique.....	77
4.9	Résultats et discussions.....	78
4.9.1	Présentation du dataset de production du champ de Hassi Messaoud.....	78
4.9.1.1	Description du champ .....	78
4.9.1.2	Description du data .....	79
4.9.1.3	Préparation des data :.....	79
4.9.2	Optimisation de débit des puits par la nouvelle méthode de back-allocation	80
4.9.2.1	Méthodologie .....	80
4.9.2.2	Résultats de back allocation.....	80
4.10	Conclusion .....	85
	Conclusion générale .....	86
	REFERENCE .....	xi
	Annexe.....	xiii

## Liste des figures

Figure 1- 1Écoulement dans un tuyau cylindrique incliné [5] .....	5
Figure 1- 2 Diagramme du Moody .....	7
Figure 1- 3 Ecoulement liquide-gaz co-courants [4] .....	10
Figure 1- 4 configurations d'écoulements diphasiques en conduites horizontales [4] .....	11
Figure 1- 5 coefficient de corrélation de la vitesse c [1] .....	14
Figure 1- 6 holdup-factor corrélation [1] .....	14
Figure 1- 7 corrélation pour facteur de correction secondaire [1] .....	15
Figure 1- 8 nombre de régime d'écoulement [1] .....	16
Figure 1- 9 nombre de vitesse de glissement de la bulle [1] .....	17
Figure 1- 10 nombre de vitesses de glissement des slug [1] .....	18
Figure 1- 11 correction du frottement de la bulle [1] .....	19
Figure 2- 1 schéma d'écoulement à travers une Duse .....	22
Figure 2- 2 Courbe de performance typique d'une Duse [4] .....	25
Figure 3- 5 cross plot de TRAINING du modèle de Towailib et Marhoum .....	41
Figure 3- 6 cross plot de TEST du modèle de Towailib et Marhoum .....	41
Figure 3- 9 cross plot de TRAINING du modèle de extension Gilbert 1 .....	43
Figure 3- 10 cross plot de TEST du modèle de extension Gilbert 1 .....	43
Figure 3- 12 cross plot de TEST du modèle de extension Gilbert 2 .....	44
Figure 3- 13 cross plot de TRAINING du modèle de extension Gilbert 3 .....	44
Figure 3- 14 cross plot de TEST du modèle de extension Gilbert 3 .....	44
Figure 3- 15 Les types de machine learning[46] .....	46
Figure 3- 16 The Neural Network Model .....	48
Figure 3- 17 Artificial Neuron [14] .....	48
Figure 3- 18 fonctions d'activations [17] .....	50
Figure 3- 19 fonctionnement des réseaux de neurones [17] .....	51
Figure 3- 20 les Le nœud de l'arbre de décision[27] .....	52
Figure 3- 21 le Fonctionnement de random forest [22] .....	54
Figure 3- 22 lstm modèle [19] .....	55
Figure 3- 23 R2 en fonction de variation des intervalle de GOR pour le régime critique ..	60
Figure 3- 24 R2 en fonction du changement des intervalle de GOR pour le régime sous critique .....	62
Figure 3- 25 cross plot de TEST du modèle LSTM .....	64
Figure 3- 26 cross plot de TRAINING du modèle LSTM .....	64
Figure 3- 27 prédiction du débit avec LSTM .....	65
Figure 4- 1 allocation à un niveau (back to Field A et B) [11] .....	68
Figure 4- 2 allocation à plusieurs niveaux (back to (field A +field B) et feild c) .....	69
Figure 4- 3 tête de puits [13] .....	73
Figure 4- 4 bac de stockage [25] .....	75
Figure 4- 5 schéma représentatif de la structure de back allocation pour un seul satellite .	76
Figure 4- 6 le champs de HMD .....	78
Figure 4- 7 intervalle du coefficient K pour le puits md409 .....	81



Figure 4- 8	intervalle du coefficient K pour le puits md409 du 01-10-2022 au31-10-2022	82
Figure 4- 9	débit corrigé par la back allocation	83
Figure 4- 10	détection des anomalies par z_score	84
Figure 4- 11	détection des anomalies par LSTM	85
Figure AN 1	cross plot training Gilbert( $GOR \leq 4000$ ), régime critique	xiii
Figure AN 2	cross plot test Gilbert( $GOR \leq 4000$ ), régime critique	xiii
Figure AN 3	cross plot training Gilbert( $4000 < GOR \leq 15000$ ), régime critique	xiii
Figure AN 4	cross plot test Gilbert( $4000 < GOR \leq 15000$ ), régime critique	xiii
Figure AN 5	cross plot training Gilbert( $15000 < GOR \leq 35000$ ), régime critique	xiv
Figure AN 6	cross plot test Gilbert( $15000 < GOR \leq 35000$ ), régime critique	xiv
Figure AN 7	cross plot training Gilbert( $35000 < GOR \leq 60000$ ), régime critique	xiv
Figure AN 8	cross plot test Gilbert( $35000 < GOR \leq 60000$ ), régime critique	xiv
Figure AN 9	cross plot training Gilbert( $60000 < GOR \leq 100000$ ), régime critique	xiv
Figure AN 10	cross plot test Gilbert( $60000 < GOR \leq 100000$ ), régime critique	xiv
Figure AN 11	cross plot training Gilbert( $GOR > 100000$ ), régime critique	xv
Figure AN 12	cross plot test Gilbert( $GOR > 100000$ ), régime critique	xv
Figure AN 13	cross plot training Gilbert( $GOR \leq 1000$ ), régime sous critique	xv
Figure AN 14	cross plot test Gilbert( $GOR \leq 1000$ ), régime sous critique	xv
Figure AN 15	cross plot training Gilbert( $1000 < GOR \leq 10000$ ), régime sous critique	xvi
Figure AN 16	cross plot test Gilbert( $1000 < GOR \leq 10000$ ), régime sous critique	xvi
Figure AN 17	cross plot training Gilbert( $10000 < GOR \leq 30000$ ), régime sous critique	xvi
Figure AN 18	cross plot test Gilbert( $10000 < GOR \leq 30000$ ), régime sous critique	xvi
Figure AN 19	cross plot training Gilbert( $GOR > 30000$ ), sous régime critique	xvi
Figure AN 20	cross plot test Gilbert( $GOR > 30000$ ), régime sous critique	xvi

## Liste des tableaux

Tableau 2- 1 les paramètres empiriques des équations.....	29
Tableau 2- 2 les coefficients spécifiques de la corrélation de Towailib et Marhoun .....	29
Tableau 2- 3 les coefficients spécifiques de la corrélation de Attar et Abdul Majeed .....	30
Tableau 2- 4 les coefficients spécifiques de la corrélation de Elgibaly et Nashawi.....	31
Tableau 2- 5 les coefficients spécifiques de la corrélation de extension de Gilberte 1 .....	31
Tableau 2- 6 :Différentes plages des paramètres de données du champ de HMD en régime critique.....	37
Tableau 2- 7 Différentes plages des paramètres de données du champ de HMD en régime sous-critique.....	37
Tableau 3- 1 les nouvelles constantes d'équation de Gilbert .....	39
Tableau 3- 2 les nouvelles constantes d'équation de Attar et Abdul Majeed.....	39
Tableau 3- 3 les nouvelles constantes d'équation de Towailib & Marhoun .....	39
Tableau 3- 4 les différentes statistiques liées à l'entraînement et aux tests des modèles théoriques.....	39
Tableau 3- 5 les nouvelles constantes d'équation de Elgibaly et Nashawi .....	41
Tableau 3- 6 les nouvelles constantes d'équation de extension de gilbert 1 .....	41
Tableau 3- 7 les nouvelles constantes d'équation de extension de gilbert 2 .....	41
Tableau 3- 8 les nouvelles constantes d'équation de extension de gilbert 3 .....	42
Tableau 3- 9 les différentes statistiques liées à l'entraînement et aux tests des modèles théoriques.....	42
Tableau 3- 10 Les composants principaux d'un LSTM[14] .....	56
Tableau 3- 11 les résultats de training et de test des modèles de random forest. ....	57
Tableau 3- 12 les résultats de training et de test des modèles de ANN.....	58
Tableau 3- 13 les résultats de training et de test des modèles de random forest. ....	58
Tableau 3- 14 les résultats de training et de test des modèles de ANN.....	59
Tableau 3- 15 les nouveaux coefficients obtenus.....	60
Tableau 3- 16 les résultats de training et de test des nouveaux modèles.....	60
Tableau 3- 17 les nouveaux coefficients obtenus.....	61
Tableau 3- 18 les résultats de training et de test des nouveaux modèles.....	61
Tableau 3- 19 les résultats de training et de test de modèle LSTM.....	63
Tableau 4- 1 représente les nouvelles valeurs (minimale et maximale) de coefficient k pour le puit MD409.....	82

## Nomenclature

$g_c$	Facteur de conversion
$\psi$	Facteur de correction secondaire
$v_m$	La vitesse du mélange
$v_s$	Vitesse de glissement
$v_i$	Vitesse moyenne
$v_{is}$	Vitesse qu'aurait la phase i
$\rho_f$	La densité du mélange
$v_m$	Vitesse du mélange
$Q_M$	Débit mesuré
$Q_C$	Débit calculé
$\gamma_g$	Densité de gaz
$\gamma_m$	Densité du mélange
$\gamma_o$	Densité d'huile
$C_d$	Coefficient de décharge
$P_1, P_u, P_t$	Pression avant la duse(amont)
$P_2, P_p$	Pression après la duse
$Q_{corr}$	Somme des débits des puits corrigés par back allocation
$Q_j$	Somme des débits des puits mesurés
$d_1$	Diamètre intérieur du tube, ft
$d_2$	Diamètre de l'orifice, ft
$k_i$	Coefficient pour chaque puits
$q_g$	Débit de gaz
$\gamma_c$	Rapport de pression critique
$\emptyset$	Diamètre de la duse en 1/64 in.
$\Delta P$	Différence de pression entre l'amont et l'aval de la duse
ANN	Artificiel neural network

C	Coefficient d'écoulement
f	Facteur de frottement du mélange
g	Accélération de la gravité
HMD	Champ Hassi Massoud
IA	Intelligence artificielle
k	Rapport de chaleur spécifique
L	Longueur d'un tuyau cylindrique incliné
LSTM	Long Short Term Memory
MAE	Erreur absolue moyenne
ML	Machine Learning
MSE	Erreur quadratique moyenne
Nd	Nombre de diamètre de pipe
Ngv	Nombre de vitesse du gaz
NL	Nombre de viscosité du liquide
r , y	Rapport entre la pression amont et aval
R <sup>2</sup>	Coefficient de détermination
RAE	Erreur absolue relative
Re	Nombre de Reynolds
RMSE	Racine carrée de l'erreur quadratique moyenne
SAT	Satellite
SCADA	Supervisory Control and Data Acquisition
$\gamma$	Densité API
$\varepsilon$	Rugosité
$\theta$	Angle d'inclinaison
$\mu$	Viscosité dynamique
$\rho$	Densité de fluide monophasique
<i>GLR</i>	Rapport gaz-liquide de production Scf/bbl,
<i>GOR</i>	Rapport gaz-huile de production Scf/bbl
<i>NLv</i>	Nombre de vitesse du liquide

$Q$

Débit d'huile, bbl/day

## Introduction générale

Dans l'industrie pétrolière, la surveillance continue des puits est essentielle pour garantir une production optimale et détecter les éventuelles anomalies ou les problèmes potentiels dans leur durée de vie. Cependant, la réalisation d'un jaugeage individuel de chaque puits nécessite l'installation de dispositifs de mesure spécifiques à chaque puits, ce qui peut être coûteux en termes de matériel, d'installation et de maintenance. Dans les champs pétroliers qui comptent un grand nombre de puits, cette tâche devient encore plus complexe. De plus, le processus de jaugeage peut perturber la production normale du puits, entraînant des arrêts temporaires qui se traduisent par une diminution des revenus

Face à ces défis, les professionnels de l'industrie ont développé des techniques alternatives pour estimer les débits individuels des puits en l'absence de jaugeage précis. C'est là qu'intervient la back-allocation

L'objectif essentiel de la back-allocation est de contrôler les puits, de répartir les volumes de production de manière précise et de rapprocher les débits calculés de la réalité de la production. Elle permet de compenser l'absence de jaugeage individuel en utilisant des débits mesurés pour certains puits. Pour les puits non jaugés, les débits sont généralement estimés à partir de la modélisation physique ou l'approche basée sur les données

La modélisation physique est une approche utilisée pour estimer les débits de production en se basant sur les principes fondamentaux de la physique des procédés. Cependant, la complexité de cette méthode réside dans la nature multiphasique des écoulements, tels que l'huile, le gaz, l'eau et le sable. Par conséquent, aucune corrélation développée jusqu'à présent ne peut être considérée comme ayant une application universelle. Les corrélations théoriques sont établies en se basant sur des relations empiriques entre certaines variables mesurables, telles que la pression, le diamètre de la Duse, etc. Cependant, en raison de la complexité des écoulements multiphasiques et des variations entre les réservoirs, des adaptations spécifiques sont nécessaires pour obtenir des résultats précise à chaque champ pétrolier étudié.

L'approche basée sur les données est une méthode prometteuse pour estimer les débits de production en utilisant des algorithmes de machine Learning. Elle offre des résultats intéressants lorsqu'elle est appliquée à des ensembles de données suffisamment représentatifs. Les modèles intelligents, tels que les techniques d'apprentissage automatique et d'intelligence artificielle, peuvent être utilisés pour développer des modèles prédictifs plus précis. Ces modèles exploitent des données historiques des puits pour estimer les débits. L'approche basée sur les données permet également d'effectuer des études de sensibilité des données ce qui signifie qu'elle peut évaluer l'impact des différentes variables sur les débits de production, ce qui permet de mieux comprendre les facteurs qui influencent la production.

Lorsque les débits calculés seront disponibles, on procédera à leur correction en utilisant des coefficients  $k$  spécifiques à chaque puits., ce qui nous permettra d'ajuster les débits calculés et se rapprocher au maximum des débits réels de production.

Enfin, on cherchera à détecter les éventuelles anomalies avec l'utilisation des techniques statistiques telles que l'écart-type ou des méthodes de modélisation basées sur les réseaux de neurones, comme le modèle LSTM

Notre travail sera structuré de la manière suivante :

Il comporte quatre chapitres, les deux premiers chapitres sont théoriques où Le premier chapitre de notre étude sera consacré à la présentation des concepts fondamentaux liés à la description des écoulements multiphasiques On discutera également des différents types d'écoulements multiphasiques. Dans le deuxième chapitre, on se concentrera sur la modélisation des écoulements à travers les conduites. On expliquera les différentes approches utilisées pour représenter ces écoulements

Dans le troisième chapitre de notre étude, nous aborderons à la fois la partie théorique et la partie pratique. Dans la partie théorique, nous présenterons les concepts clés liés à l'utilisation de techniques de machine learning dans le domaine pétrolier. Nous expliquerons les principes fondamentaux des modèles mathématiques et des algorithmes de machine learning utilisés pour estimer les débits de production. Nous discuterons également des avantages et des limitations de chaque approche.

Dans la partie pratique de ce chapitre, nous appliquerons les connaissances théoriques acquises pour calculer les débits à travers la duse. Nous utiliserons à la fois les corrélations traditionnelles et les techniques de machine learning pour estimer les débits. Nous effectuerons également une étude de sensibilité des paramètres afin de comprendre leur impact sur les résultats de débit. Cette partie pratique nous permettra d'évaluer les performances et la précision des différentes approches utilisées.

Dans le quatrième chapitre, nous présenterons le concept de back-allocation de manière générale. Nous expliquerons en détail comment cette méthode est utilisée pour corriger les débits calculés en l'absence de jaugeage individuel. Nous discuterons des principes sous-jacents de la back-allocation et des différentes stratégies employées pour obtenir des estimations plus précises des débits réels.

Ensuite, nous appliquerons spécifiquement la méthode de back-allocation à notre champ pétrolier d'étude. Nous mettrons en œuvre les coefficients  $k$  spécifiques à chaque puits, pour corriger les débits calculés. Nous évaluerons la performance de cette méthode de correction et analyserons les résultats obtenus.

**CHAPITRE 1**

**LES ECOULEMENTS MULTIPHASIQUES**



## 1.1 Introduction

Les écoulements multiphasiques font référence aux écoulements dans lesquels plusieurs phases coexistent simultanément, tels que les gaz, les liquides ou les solides dispersés dans un fluide. Ces écoulements sont souvent rencontrés dans de nombreux domaines tels que la production pétrolière.

Les écoulements multiphasiques se produisent lorsqu'un système est traversé simultanément par un flux composé de deux phases ou plus. Le cas le plus simple d'un écoulement multiphasique est celui d'un écoulement à deux phases, qui implique généralement des solides et des liquides ou des gaz.

## 1.2 Types d'écoulements multiphasiques

Les écoulements multiphasiques peuvent prendre différentes formes en fonction des phases impliquées. La catégorie la plus courante d'écoulements multiphasiques est celle des écoulements diphasiques, qui englobe plusieurs types d'écoulements :

- Les écoulements gaz-solide, où des particules solides sont suspendues dans des gaz.
- Les écoulements liquides-liquides sont caractérisés par la présence de deux liquides immiscibles qui s'écoulent ensemble tel que les écoulements d'émulsions d'huile et d'eau.
- Les écoulements liquide-solide se produisent lorsque des particules solides sont dispersées ou suspendues dans un liquide. Ils sont fréquemment rencontrés dans le transport hydraulique de matières solides, comme les suspensions de sable dans l'eau.
- Les écoulements gaz-liquide, qui constituent probablement la forme la plus importante d'écoulement multiphasique et sont largement utilisés dans les applications industrielles. Ils impliquent l'interaction entre un gaz et un liquide, comme l'écoulement de gaz dans un liquide ou de liquide dans un gaz.

Les écoulements triphasiques ont également une importance pratique. Ils impliquent trois phases simultanément, tels que :

- Les écoulements gaz-liquide-solide où les trois phases : le gaz, le liquide et le solide, coexistent simultanément. Ce type d'écoulement peut se produire dans divers contextes industriels et naturels.
- Les écoulements solide-liquide-liquide se produisent lorsque trois phases sont présentes simultanément, à savoir un solide, un liquide et un autre liquide. Un exemple courant de ce type d'écoulement est lorsque du sable est mélangé avec du pétrole et de l'eau dans un oléoduc.
- Les écoulements gaz-liquide-liquide qui constituent le cas central de la présente étude où les écoulements sont respectivement du gaz naturel, du pétrole, et de l'eau.

Effectivement, le cas le plus complexe en termes d'écoulement multiphasique est celui qui implique quatre phases, à savoir le pétrole, l'eau, le gaz et le sable. Ce type d'écoulement est

souvent rencontré dans l'industrie pétrolière et gazière, en particulier dans les pipelines et les installations de production d'hydrocarbures.

Dans le contexte actuel, on se concentre spécifiquement sur les écoulements multiphasiques diphasiques impliquant les phases (pétrole-gaz naturel)

### 1.3 Écoulement de liquide monophasique

Pour évaluer l'efficacité et les performances d'un écoulements monophasiques liquides ou gazeux, les pertes de charge sont généralement déterminées en utilisant des équations empiriques ou des modèles théoriques basés sur la taille et la rugosité des conduites.

La relation entre la vitesse d'écoulement du fluide et la variation de pression le long de l'axe longitudinal du tuyau cylindrique est obtenue en effectuant un calcul de bilan d'énergie. Cette relation est basée sur la conservation de l'énergie entre deux points du système la Figure (1-1) montre la géométrie d'un tuyau cylindrique incliné d'une longueur  $L$  le long de l'axe longitudinal et d'un angle d'inclinaison  $\theta$ . Le fluide monophasique a une densité  $\rho$  et une viscosité dynamique  $\mu$ . Il s'écoule dans un champ gravitationnel avec une accélération  $g$ . où  $p$  est la pression. Nous avons exprimé le gradient de pression le long de l'axe longitudinal du tuyau comme la somme d'un terme d'énergie potentielle. [5]

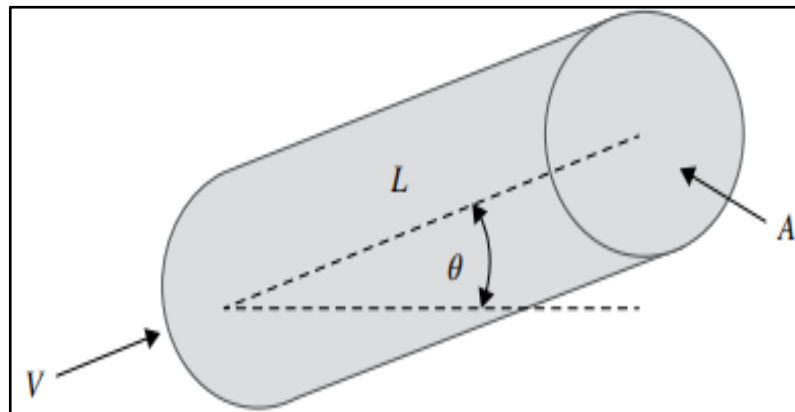


Figure 1- 1Écoulement dans un tuyau cylindrique incliné [5]

L'équation générale de gradient de pression en écoulement poly-phasique est

$$\left(\frac{dP}{dL}\right) = \left(\frac{dP}{dL}\right)_{ele} + \left(\frac{dP}{dL}\right)_f + \left(\frac{dP}{dL}\right)_{acc} \quad (1.1)$$

Avec :

- $\left(\frac{dP}{dL}\right)_{ele} = \frac{g}{g_c} \rho_f \sin \theta$  (Terme de perte par élévation)
- $\left(\frac{dP}{dL}\right)_f = \frac{f \cdot \rho_f \cdot v_m^2}{2 \cdot g_c \cdot d}$  (Terme de perte avec frottement)

- $\left(\frac{dP}{dL}\right)_{acc} = \frac{\rho_f v_m dv_m}{g_c dL}$  (Terme de perte par accélération)

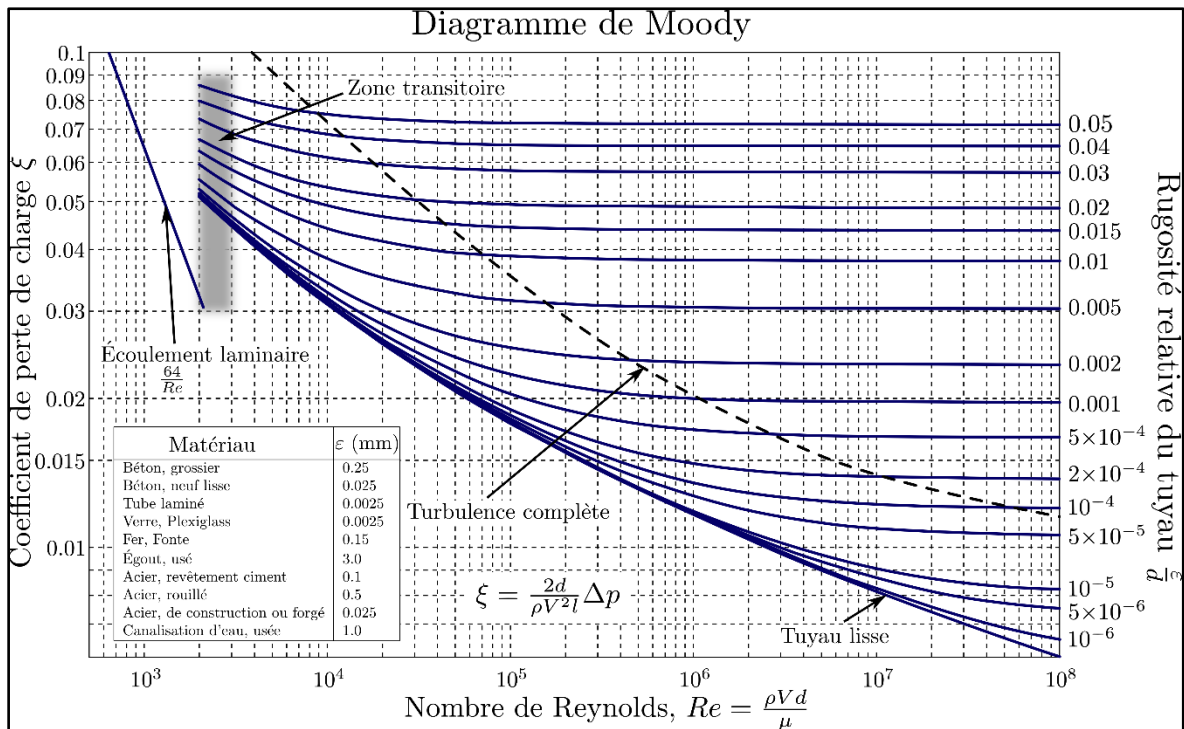
Et  $\rho_f$  : la densité moyenne,  $v_m$  : la vitesse moyenne,  $d$  : diamètre du pipe,  $g_c$  : facteur de conversion,  $g$  : accélération de la gravité (9,81 m/s<sup>2</sup>),  $\theta$  : Pente de la conduite.,  $f$  : facteur de frottement du mélange. Il est au fonction de  $Re$ ,  $\varepsilon$  et  $d$

Le coefficient de frottement  $f$  peut être évalué en fonction du nombre de Reynolds  $Re$  (défini comme le rapport entre la force d'inertie et la force visqueuse) et de la rugosité relative  $\varepsilon/D$

L'intervalle de nombre de Reynolds  $Re$  dépend du type d'écoulement du fluide et des caractéristiques géométriques du conduit. Cependant, généralement, on considère les plages de Reynolds suivantes pour déterminer le type d'écoulement :

- Écoulement laminaire :  $Re < 2\,000$  Dans cette plage, l'écoulement est caractérisé par des lignes de courant régulières et parallèles. Le coefficient de frottement peut être estimé de manière simplifiée à l'aide de la relation  $64/Re$ . Dans ce cas, il est indépendant du rapport de rugosité relative  $\varepsilon/D$ .
- Écoulement Transitoire :  $2\,000 < Re < 4\,000$  Dans cette plage, l'écoulement peut commencer à montrer des caractéristiques turbulentes, mais peut également rester laminaire selon les conditions spécifiques.
- Écoulement turbulent :  $Re > 4\,000$  À partir de cette plage, l'écoulement est généralement considéré comme turbulent, avec des tourbillons et des mouvements chaotiques du fluide. Pour ces écoulements, l'équation de Colebrook-White ou le diagramme de Moody sont utilisés pour estimer le coefficient de frottement.

La figure (1-2) illustre le diagramme de Moody qui représentant le facteur de frottement en fonction du nombre de Reynolds et de la rugosité relative



**Figure 1- 2 Diagramme du Moody**

L'équation de Colebrook-White, également connue sous le nom d'équation de Colebrook, est une équation couramment utilisée pour calculer le facteur de frottement dans les écoulements turbulents. Elle est applicable à l'ensemble de l'intervalle des écoulements non laminaire du diagramme de Moody.

L'équation de Colebrook est donnée par :

$$\frac{1}{\sqrt{f}} = -2 \log \left( \frac{\epsilon/D}{3.7} + \frac{2.51}{Re \sqrt{f}} \right) \quad (1.2)$$

## 1.4 Écoulement diphasique

« Un écoulement monophasique de liquide existe dans un puits de pétrole uniquement lorsque la pression à la tête de puits est supérieure à la pression de bulle de l'huile, ce qui est généralement rarement le cas. Cependant, il est pratique de partir de l'hypothèse d'un écoulement monophasique de liquide pour établir le concept d'écoulement des fluides dans les puits de pétrole, où l'écoulement multiphasique prédomine généralement. » [3]

Cependant, il convient de noter que dans un puits de pétrole un écoulement monophasique liquide existe dans un puits de pétrole uniquement lorsque la pression à la tête du puits est supérieure à la pression de pointe de bulle de l'huile, ce qui est généralement rarement le cas. En réalité, la pression à la tête du puits est généralement inférieure à la pression de bulle du pétrole, ce qui signifie que l'écoulement est multiphasique.

Lorsque la seconde phase est introduite dans un flux d'écoulement, cela peut entraîner des changements importants dans les propriétés du fluide global et compliquer l'analyse des équations du gradient de pression. Les différences de densité entre les phases peuvent provoquer une séparation, où les phases se séparent et s'écoulent à des vitesses différentes dans le tuyau. Cela nécessite l'utilisation de modèles et de techniques spécifiques pour caractériser et prédire le comportement de l'écoulement diphasique.

### 1.4.1 Les propriétés d'un écoulement diphasique

Les propriétés d'un écoulement diphasique dépendent des caractéristiques des phases individuelles ainsi que de leur interaction. Voici quelques-unes des propriétés couramment considérées dans l'analyse des écoulements diphasiques :

#### 1.4.1.1 Holdup Liquide (HL) (fraction volumique)

Appelée également taux de vide dans le cas particulier d'un écoulement liquide / vapeur, est le volume du liquide  $L$  dans un segment de la conduite divisé par le volume de ce segment à un instant (t)

$$H_L = \frac{\text{volume de liquide dans le segment}}{\text{volume du segment de la conduite}} \quad (1.3)$$

Le volume du segment est choisi grand devant les fluctuations que l'on veut lisser, mais petit devant les variations globales que souhaite étudier.

La valeur de  $HL$  ne peut pas être calculée analytiquement. Elle doit être déterminée à partir des corrélations empiriques, elle est en fonction des (propriétés de gaz et de liquide, configuration de l'écoulement, diamètre de conduite et l'inclinaison de la conduite).

Le gaz Holdup est :

$$H_g = 1 - HL \quad (1.4)$$

Holdup de non glissement  $\lambda L$  et  $\lambda g$ : Le Holdup liquide de non glissement  $\lambda L$  est défini comme le rapport de volume de liquide dans un élément de conduite qui existerait si le gaz et le liquide se déplaçaient à la même vitesse (pas de glissement) divisé par le volume de l'élément de conduite. Il peut être calculé directement à partir des débits connus de gaz et de liquide in situ.

$$\lambda_L = \frac{q_L}{q_L + q_g} \quad (1.5)$$

Où  $q_L$  est la somme des débits d'huile et d'eau in-situ et  $q_g$  est le débit de gaz in-situ.

Le Holdup gaz de non glissement  $\lambda g$  est défini comme suit :

$$\lambda_g = 1 - \lambda L = \frac{q_g}{q_g + q_L} \quad (1.6)$$

### 1.4.1.2 Vitesse superficielle

La vitesse superficielle d'une phase fluide est définie comme la vitesse que présenterait cette phase si elle s'écoulait seule à travers la section transversale totale de la conduite. Pour la phase i, on définit :

$$v_{is} = \frac{q_i}{A} \quad (1.7)$$

Son interprétation est la suivante : il s'agit de la vitesse qu'aurait la phase i si elle s'écoulait seule dans la conduite, on définit également la vitesse du mélange :

$$v_m = \frac{q}{A} = \sum v_{is} \quad (1.8)$$

Vitesse moyenne Il s'agit de la vitesse moyenne débitante de la phase i à travers la portion de conduite occupée en moyenne par la phase i.

$$v_i = \frac{q_i}{AH_i} \quad (1.9)$$

Vitesse de glissement

$$v_s = v_g - v_L = \frac{v_{sg}}{H_g} - \frac{v_{sL}}{H_L} \quad (1.10)$$

### 1.4.1.3 Masse volumique

Toutes les équations d'écoulement de fluide nécessitent qu'une valeur de la densité du fluide soit disponible. Or, La densité d'un mélange gaz/liquide en écoulement est difficile à évaluer en raison de la séparation gravitationnelle des phases et du glissement entre les phases.

$$\rho_s = \rho_L \cdot H_L + \rho_g \cdot H_g \quad (1.11)$$

$$\rho_k = \frac{\rho_L \cdot \lambda_L^2}{H_L} + \frac{\rho_g \cdot \lambda_g^2}{H_g} \quad (1.12)$$

### 1.4.1.4 Viscosité

La viscosité du fluide en écoulement est utilisée pour déterminer le nombre de Reynolds et d'autres nombres sans dimension qui sont utilisés dans certaines corrélations. Le concept de viscosité à deux phases est défini différemment par divers chercheurs.

$$\mu_s = \mu_L \cdot H_L + \mu_g \cdot H_g \quad (1.13)$$

$$\mu_k = \mu_L^{H_L} \cdot \mu_g^{H_g} \quad (1.14)$$

## 1.4.2 Les Configurations d'écoulements diphasiques

Lors de l'écoulement diphasique d'un mélange gaz-liquide, plusieurs facteurs influencent l'apparition et la nature du régime d'écoulement observé. Parmi ces facteurs, on peut citer les propriétés spécifiques du fluide, les conditions d'opération du système, les débits des phases, la géométrie de la conduite, ainsi que les effets thermiques induits par le flux de chaleur. La combinaison de ces différents paramètres peut donner lieu à une variété de régimes d'écoulement diphasique.

Un écoulement multiphasique en conduite peut présenter diverses organisations spatiales pour les phases. On distingue ainsi :

- Régimes dispersés
- Régimes séparés
- Régimes intermittents

Il existe différentes configurations d'écoulements diphasiques qui décrivent la répartition spatiale des phases dans un système donné :

### 1.4.2.1 Ecoulement vertical

Pour un écoulement vertical, les principales configurations suivantes sont observées, comme illustré dans la figure (1-3) où les deux fluides s'écoulent dans le même sens descendant en conduite verticale ((1) : Bulles ; (2) : Poches ; (3) : Film tombant ; (4) : Film tombant à bulles ; (5) : Churn ; (6) : Annulaire-dispersé).

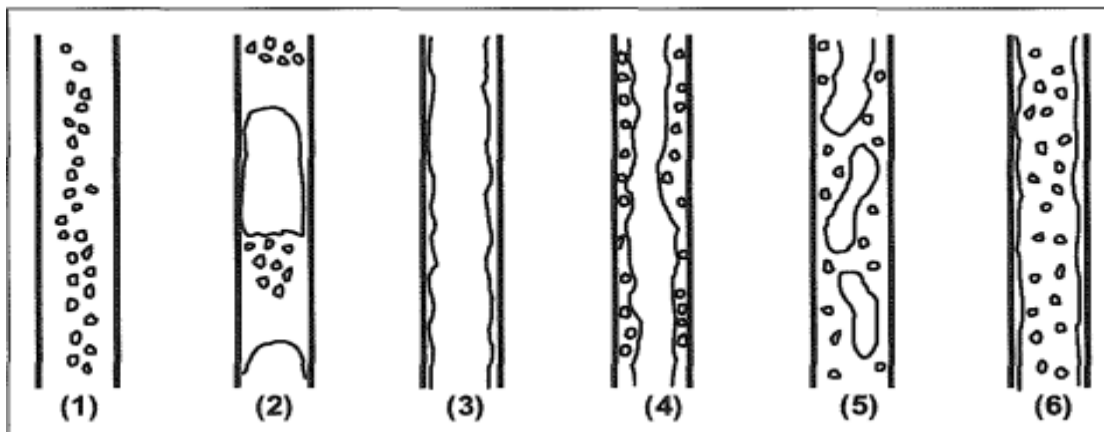


Figure 1- 3 Ecoulement liquide-gaz co-courants [4]

- Ecoulement monophasique : C'est un cas particulier qui se produit lorsque la pression est au-dessus de la pression de bulle où on a seulement l'écoulement d'une phase liquide.
- Ecoulement à bulles (Bubble flow) : Ici, la phase liquide est continue et une dispersion de bulles s'écoule dans le liquide, cela est dû à la chute de pression au-dessous de point de bulle. Les bulles ont des mouvements complexes et sont

généralement de petite taille. Plus l'écoulement est turbulent plus ces bulles vont fragmenter.

- Ecoulement par bouchon (slug flow) : Plus le fluide se déplace vers le haut plus la pression diminue, et dans ces conditions une grande quantité de gaz se libère de la solution et forme de grosses bulles appelées bulles de Taylor. Ces bulles se déplacent uniformément vers le haut et portent des gouttelettes d'huiles avec eux. Ce type d'écoulement est observé généralement pour des faibles vitesses de liquide.
- Ecoulement transitoire (Churn flow) : À des vitesses d'écoulement plus élevées, les bulles de l'écoulement de bouchon se brisent, ce qui conduit à un régime d'écoulement instable dans lequel il y a, dans les tubes à large alésage, un mouvement oscillatoire du liquide.
- Ecoulement annulaire (Annular flow) : Pour ce type d'écoulement, le liquide forme un film et occupe les parois du tubing, le liquide se déplace lentement vers le haut, par contre, la phase gazeuse occupe la partie centrale du tubing et se déplace à de grandes vitesses. Habituellement, une partie de la phase liquide est entraînée sous forme de petites gouttelettes dans le noyau ; à des débits élevés, il est également courant que des bulles de gaz soient entraînées dans le film liquide.
- Ecoulement à brouillard (Mist flow) : A des pressions très basses, le taux de gaz est très élevé, il surmonte la force de tension liquidetuyau et enlève le film liquide du tuyau. Dans ce cas, la phase gazeuse est la phase continue et la phase liquide est présente sous forme de gouttelettes dispersées.

#### 1.4.2.2 Ecoulement horizontal

L'effet de la pesanteur sera plus marqué lors d'un écoulement horizontal, créant des configurations qui diffèrent un peu de l'écoulement vertical. Les régimes sont représentés dans la figure(1-4) où elle illustre les configurations d'écoulements diphasiques en conduites horizontales ((1): Bulles; (2): Bulles allongées; (3): Stratifié lisse; (5): Stratifié ondulé; (5): Poches; (6): Annulaire).

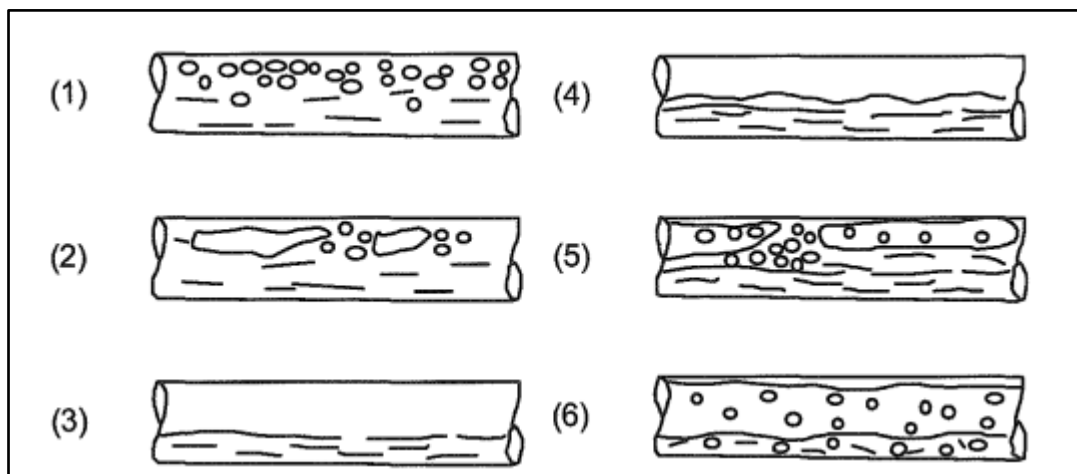


Figure 1- 4 configurations d'écoulements diphasiques en conduites horizontales [4]



- Écoulement à bulles : Dans cette configuration, des bulles de gaz (air) sont dispersées dans le liquide (eau) de manière relativement uniforme. Les bulles ont une forme sphérique et leur taille peut varier en fonction des conditions d'opération.
- Écoulement à bulles allongées : Ce régime d'écoulement est similaire à l'écoulement à bulles, mais les bulles de gaz ont une forme allongée. Cela peut être dû à l'effet de cisaillement exercé par le liquide sur les bulles lors de leur mouvement dans la conduite horizontale.
- Écoulement stratifié lisse : Dans cette configuration, le liquide et le gaz s'écoulent sous forme de couches distinctes et bien définies, sans présence de turbulence significative à l'interface entre les deux phases. Le liquide occupe généralement la partie inférieure de la conduite, tandis que le gaz est situé au-dessus.
- Écoulement stratifié ondulé : Cet écoulement stratifié présente des ondulations à l'interface entre le liquide et le gaz. Ces ondulations peuvent être causées par des instabilités de surface ou des variations de débit dans la conduite.
- Écoulement en poches : Dans cette configuration, de grandes poches de gaz sont emprisonnées dans le liquide. Les poches peuvent se former en raison de l'accumulation de gaz ou de la présence d'obstacles dans la conduite qui favorisent la rétention du gaz.
- Écoulement annulaire : Dans cette configuration, un film de liquide adhère à la paroi interne de la conduite, entourant un noyau central de gaz. L'interface entre le film liquide et le gaz est généralement bien définie.

### **1.4.3 Modélisation des écoulements diphasiques**

La forme générale de l'équation de gradient de pression pour un mélange diphasique peut être dérivée à partir de l'équation générale de gradient de pression en effectuant quelques adaptations spécifiques pour tenir compte de l'écoulement diphasique. En remplaçant chaque terme de l'équation générale par ceux appropriés à l'écoulement diphasique, on obtient une formulation spécifique pour le mélange diphasique.

Lors de l'étude des écoulements diphasiques, deux composantes principales contribuent au gradient de pression total : le holdup (ou taux de rétention) et le facteur de friction diphasique. La prédiction précise de ces deux paramètres est essentielle pour développer un modèle fiable du gradient de pression dans les écoulements diphasiques.

#### **Modèles d'écoulements**

On distingue trois catégories de corrélations pour estimer les pertes de dans les écoulements diphasiques :

##### **Catégorie a**

Dans cette catégorie, la configuration de l'écoulement n'est pas prise en compte et la vitesse de glissement entre les phases est supposée nulle. Cela signifie que les phases sont

considérées comme se déplaçant à la même vitesse, sans interaction significative entre elles. Les corrélations utilisées dans cette catégorie sont généralement basées sur des équations empiriques simplifiées, telles que l'équation de Darcy-Weisbach ou l'équation de Hazen-Williams. Où On remplace l'écoulement diphasique par un pseudo-fluide monophasique dont les propriétés (vitesse, masse volumique, viscosité) sont les moyennes du mélange obéissant aux lois d'un écoulement monophasique.

### Catégorie b

Dans cette catégorie, la configuration de l'écoulement n'est toujours pas prise en compte, mais la vitesse de glissement entre les phases est non nulle. Cela signifie qu'il y a une différence de vitesse entre les phases, ce qui peut entraîner des interactions et des pertes de charge supplémentaires. Les corrélations utilisées dans cette catégorie sont plus avancées que celles de la catégorie a et prennent en compte la vitesse de glissement entre les phases.

Hagedorn et Brown ont développé une méthode en obtenant des données expérimentales de chute de pression et de débit d'un puits instrumenté de 1500 pieds de profondeur. La rétention de liquide (liquid holdup) a été calculée sur la base de corrélations. Le liquide holdup est calculé comme suit :

1-calculer les nombres adimensionnels

$$N_{LV} = 1.938V_{sL} \sqrt[4]{\frac{\rho_L}{\sigma_L}} \quad (1.15)$$

$$N_{gV} = 1.938V_{sg} \sqrt[4]{\frac{\rho_L}{\sigma_L}} \quad (1.16)$$

$$N_d = 120.872d \sqrt{\frac{\rho_L}{\sigma_L}} \quad (1.17)$$

$$N_L = 0.15726\mu_L \sqrt[4]{\frac{1}{\rho_L \sigma_L^3}} \quad (1.18)$$

2-déterminer la valeur de  $CN_L$  à partir du graphe suivant ( $N_L$  vs  $CN_L$ )

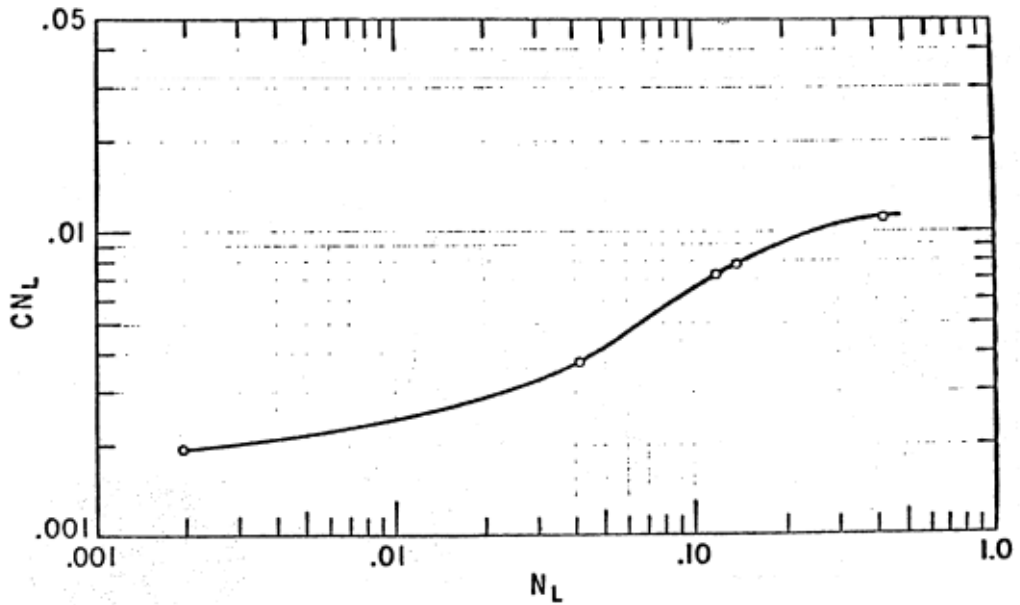


Figure 1- 5 coefficient de corrélation de la vitesse  $c$  [1]

3-déterminer la valeur de  $\frac{H_L}{\psi}$  à partir du graphe suivant  $\left( \frac{N_{LV}}{N_{gV}^{0.575}} \right) \left( \frac{P}{P_a} \right)^{0.1} \left( \frac{CN_L}{N_d} \right)$  vs  $\frac{H_L}{\psi}$

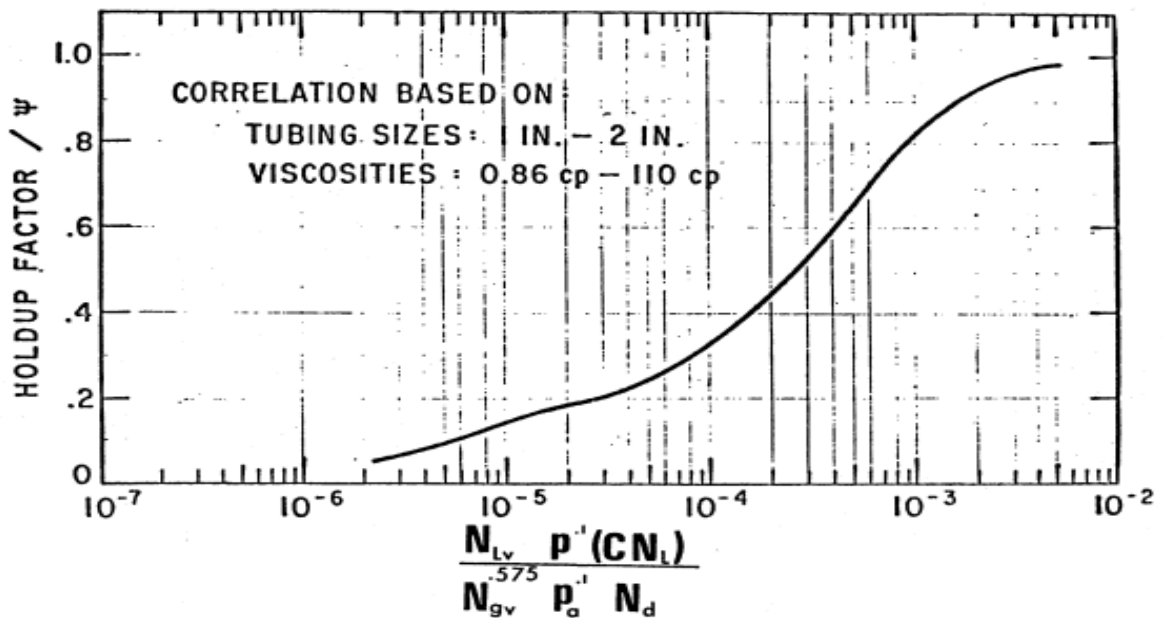


Figure 1- 6 holdup-factor corrélation [1]

4-déterminer la valeur de  $\psi$  à partir du graphe suivant  $\left( \frac{N_{gv} N_L^{0.380}}{N_d^{2.14}} \right)$  vs  $\psi$

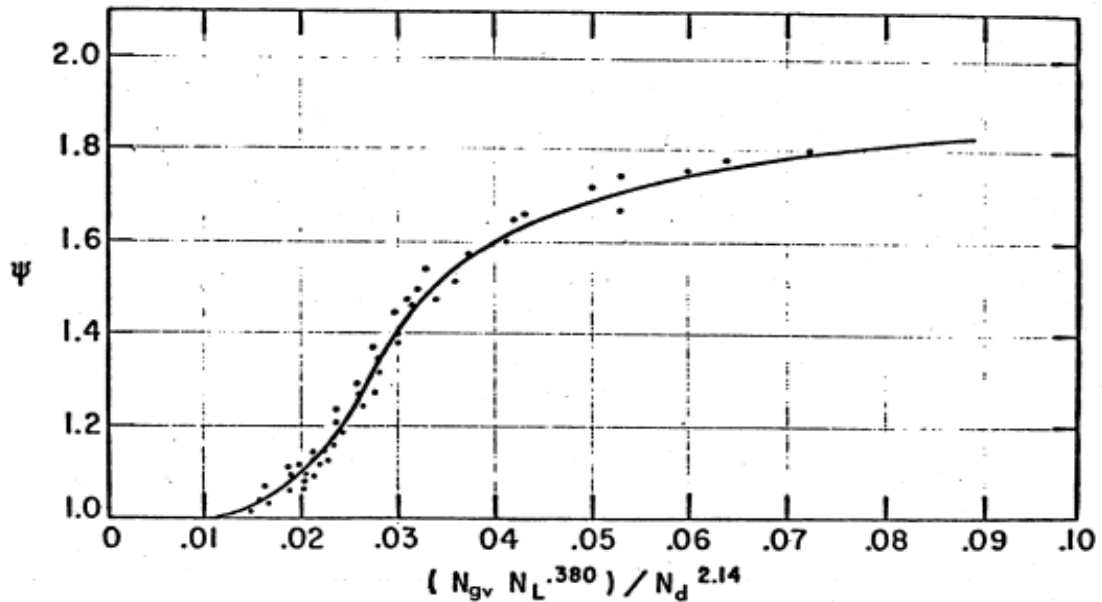


Figure 1- 7 corrélation pour facteur de correction secondaire [1]

5-calculer la valeur de la fraction volumique  $H_L = \left( \frac{H_L}{\psi} \right) \psi$  (1.19)

Le facteur de friction est corrélé avec le nombre de Reynolds diphasique en utilisant le diagramme standard de Moody

$$Re = \frac{\rho_s v_m d}{\mu_s} \quad (1.20)$$

6-Une fois les valeurs d'holdup, de densité du mélange et du facteur de friction sont déterminées, il est possible de calculer le gradient total de perte de pression.

### Catégorie c

Dans cette catégorie, la configuration de l'écoulement est prise en compte, ce qui signifie que la géométrie spécifique de la conduite est considérée. En plus de cela, la vitesse de glissement entre les phases est non nulle. Cela nécessite une modélisation plus détaillée des interactions entre les phases et des pertes de charge associées. Des approches plus complexes

Plusieurs auteurs ont développé des corrélations pour quelques paramètres utilisés dans le calcul de perte de pression, parmi eux Les corrélations de Duns et Ros sont le résultat d'une étude approfondie en laboratoire dans laquelle le Holdup de liquide et les gradients de pression ont été mesurés. Des corrélations ont été développées pour la vitesse de glissement (à partir de laquelle le Holdup peut être calculée) et le facteur de friction pour trois régimes d'écoulement

Les étapes qui ont été suivis par ces deux auteurs pour calculer la perte de pression sont :

1-calculer les nombres adimensionnels

$$N_{LV} = 1.938V_{sL}^4 \sqrt{\frac{\rho_L}{\sigma_L}} \quad (1.21)$$

$$N_{gV} = 1.938V_{sg}^4 \sqrt{\frac{\rho_L}{\sigma_L}} \quad (1.22)$$

$$N_d = 120.872d \sqrt{\frac{\rho_L}{\sigma_L}} \quad (1.23)$$

$$N_L = 0.15726\mu_L^4 \sqrt{\frac{1}{\rho_L\sigma_L^3}} \quad (1.24)$$

$$L_s = 50 + 36N_{Lv} \quad (1.25)$$

$$L_m = 75 + 84N_{Lv}^{0.75} \quad (1.26)$$

$L_1$  et  $L_2$  à partir du graphe suivant ( $N_d$  vs  $L_1, L_2$ )

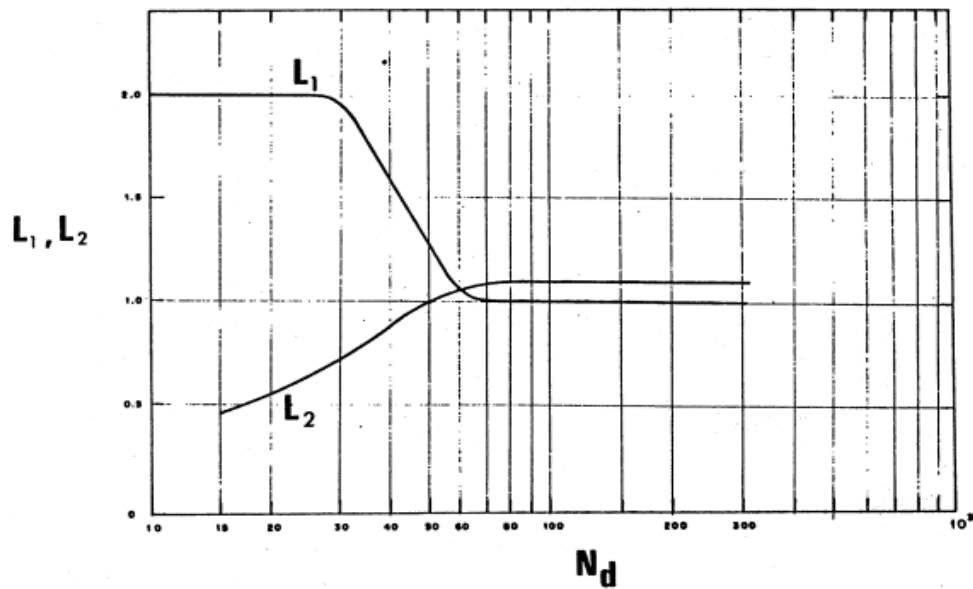


Figure 1- 8 nombre de régime d'écoulement [1]

2-déterminer le régime d'écoulement

- Bubble flow  $0 \leq N_{gv} \leq L_1 + L_2 N_{Lv}$
- Slug flow  $L_1 + L_2 N_{Lv} \leq N_{gv} \leq L_s$
- Transition flow  $L_s \leq N_{gv} \leq L_m$
- Mist flow  $N_{gv} \geq L_m$

3-Calculer la vitesse de glissement adimensionnelle  $S$  pour le bubble flow et le slug flow

- Bubble flow

$$S = F_1 + F_2 N_{Lv} + F_3' \left( \frac{N_{gv}}{1 + N_{Lv}} \right)^2 \quad (1.27)$$

$$F_3' = F_3 - \frac{F_4}{N_d} \quad (1.28)$$

$F_1, F_2$  et  $F_3$  sont déterminées à partir des graphes suivants

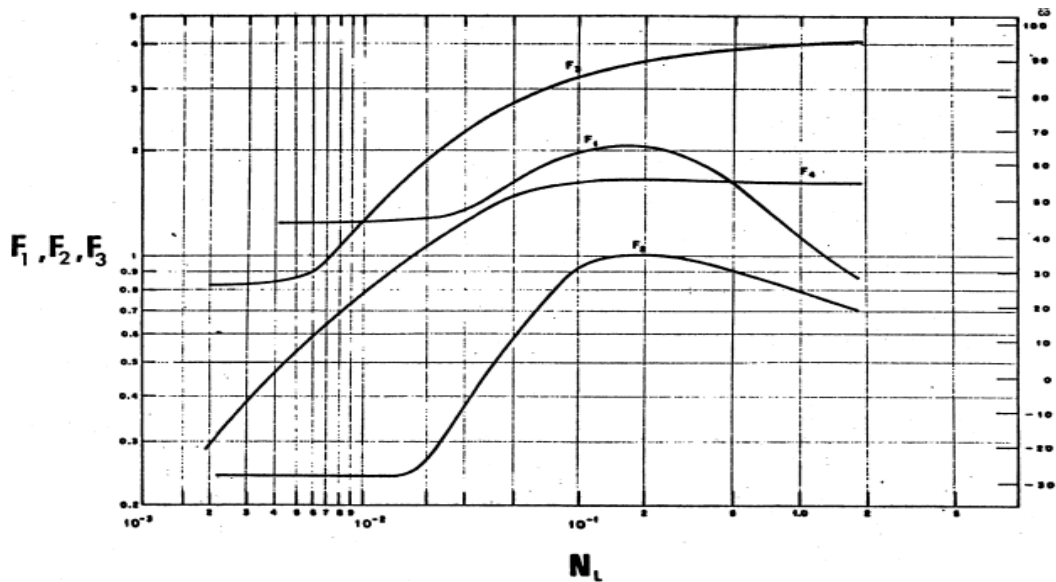


Figure 1- 9 nombre de vitesse de glissement de la bulle [1]

- Slug flow

$$S = (1 + F_5) \frac{N_{gv}^{0.982} + F_6'}{(1 + F_7 N_{Lv})^2} \quad (1.29)$$

$$F_6' = 0.029 N_d + F_6 \quad (1.30)$$

$F_5, F_6$  et  $F_7$  sont déterminées à partir des graphes suivants

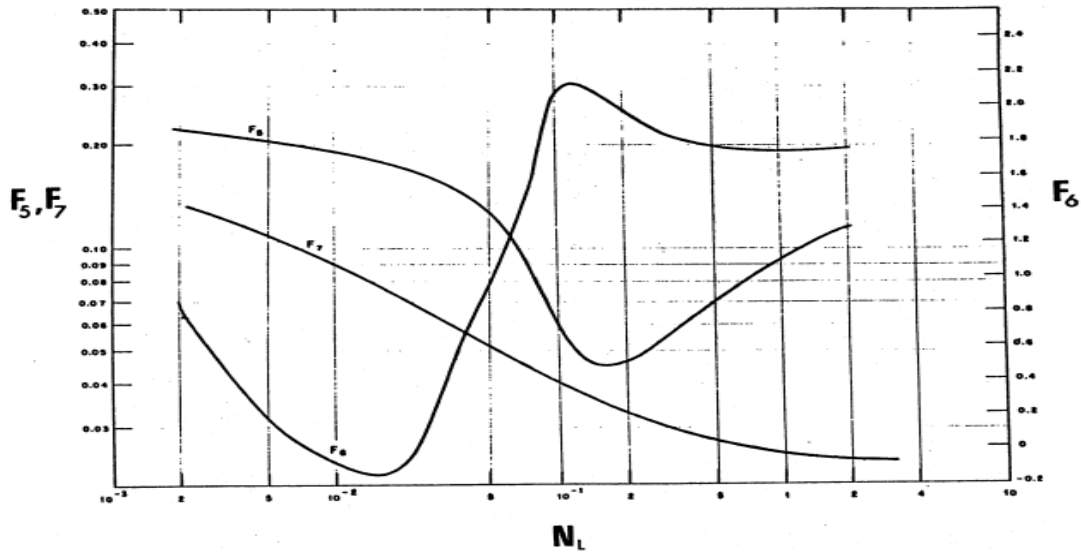


Figure 1- 10 nombre de vitesses de glissement des slug [1]

4-Calculer la fraction volumique pour le bubble flow, slug flow et mist flow

- Bubble flow et slug flow

$$v_s = \frac{S}{\left(\frac{\rho_L}{\sigma_L g}\right)^{1/4}} \quad (1.31)$$

$$H_L = \frac{v_s - v_m + \left[(v_s - v_m)^2 + 4v_m v_{sL}\right]^{1/2}}{2v_s} \quad (1.32)$$

- Mist flow

$$H_L = \lambda_L \quad (1.33)$$

5-Calculer le gradient de pression hydrostatique pour le bubble flow, slug flow et mist flow

$$\frac{dp}{dL} = \frac{g}{g_c} \left[ \rho_L H_L + \rho_g (1 - H_L) \right] \quad (1.34)$$

6-Calculer la perte de charge due au frottement

- Bubble flow et slug flow

$$\left(\frac{dp}{dL}\right)_f = \frac{f_m \rho_L v_{sL} v_m}{2g_c d} \quad (1.35)$$

$$f_m = f_1 f_2 / f_3 \quad (1.36)$$

$$f_3 = 1 + f_1 \sqrt{\frac{v_{sg}}{50v_{sL}}} \quad (1.37)$$

$f_1$  : Facteur de frottement à partir du diagramme de Moody

$f_2$  : à partir du graphe suivant

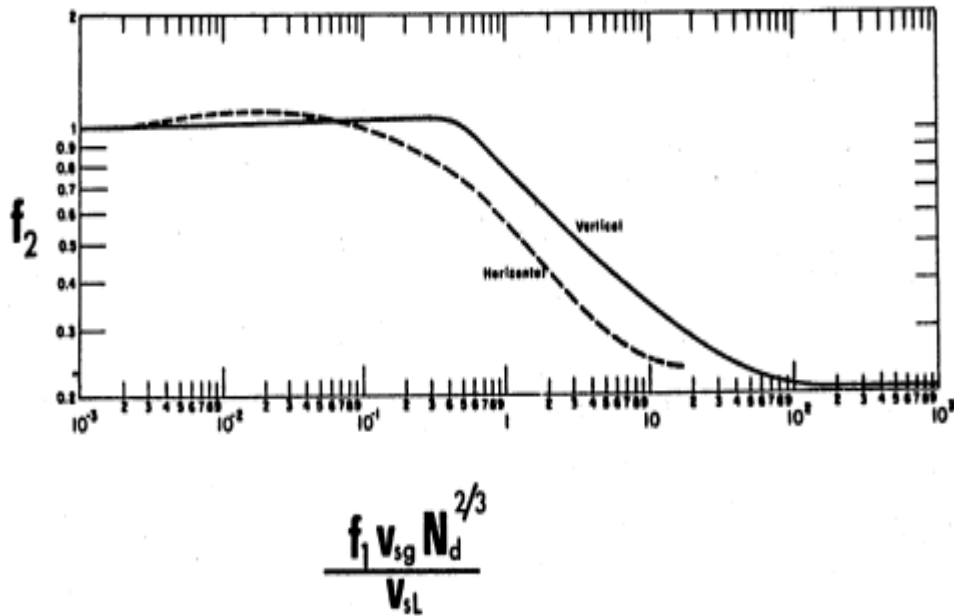


Figure 1- 11 correction du frottement de la bulle [1]

- Mist flow

$$\frac{dp}{dL} = \frac{f \rho_g v_{sg}^2}{2g_c d} \quad (1.38)$$

$f$  : facteur de frottement à partir du diagramme de Moody

Avec :

$$R_e = \frac{\rho_g v_{sg} d}{\mu_g} \quad (1.39)$$

#### 7-Calcul de la perte de charge total

- Bubble flow, slug flow, mist flow

$$\left(\frac{dp}{dL}\right)_t = \left(\frac{dp}{dL}\right)_{elev} + \left(\frac{dp}{dL}\right)_f \quad (1.40)$$

- Transition flow

$$\left(\frac{dp}{dL}\right)_t = A \left(\frac{dp}{dL}\right)_{t,slug} + B \left(\frac{dp}{dL}\right)_{t,mist} \quad (1.41)$$

$$A = \frac{L_m - N_{gv}}{L_m - L_s} \quad (1.42)$$



$$B = \frac{N_{gv} - L_s}{L_m - L_s} = 1 - A \quad (1.43)$$

La densité du gaz utilisé pour calculer le gradient hydrostatique en écoulement à brouillard (mist flow) est donnée comme suit :

$$\rho_g' = \frac{\rho_g N_{gv}}{L_m} \quad (1.44)$$

## 1.5 Conclusion

La modélisation des écoulements multiphasiques, sont des phénomènes complexes. Et représente une étape cruciale pour la conception, la compréhension et optimisation des systèmes et les processus associés.

Il est important de choisir la catégorie de corrélation appropriée en fonction de la configuration de l'écoulement et de la vitesse de glissement entre les phases pour obtenir des estimations précises des pertes de charge.

**CHAPITRE2**

**MODELISATION DES ECOULEMENTS A**

**TRAVERS LES RESTRICTION**

## 2.1 Introduction

L'écoulement des fluides à travers des restrictions est un phénomène fréquent et bien connu dans l'industrie pétrolière et gazière. Il se produit dans plusieurs situations, notamment lors de la mesure des liquides et/ou des gaz à l'aide de compteurs à orifice, du passage des fluides à travers des restrictions de surface ou de fond, des soupapes de sécurité souterraines, ainsi que de tout type de soupape de surface.

Dans le cadre de ce mémoire, on se concentre sur l'écoulement à travers un type spécifique de restriction connue sous le nom de "duse"

## 2.2 Description de la Duse

« Une duse (choke) est une restriction à l'écoulement des fluides, ce qui entraîne une chute de pression. Il est largement utilisé dans l'industrie pétrolière pour contrôler et optimiser les débits de production. » [6]

Une Duse est une restriction de circulation des fluides placé à la sortie de la tête de puits qui entraîne une diminution de la pression. Elle est largement utilisée dans l'industrie pétrolière pour contrôler et optimiser les taux de production. Elle remplit également d'autres fonctions telles que la fourniture d'une contre-pression au réservoir pour éviter les dommages à la formation, les fluctuations de pression excessives et le reflux.

Il existe deux types d'étrangleurs de tête de puits sont couramment utilisés : les Duses positives (fixes) et les Duses réglables.

Les duses positives offrent une restriction fixe et constante, tandis que les duses réglables permettent d'ajuster la restriction et donc le débit de production. Le choix entre les deux dépend des besoins spécifiques du puits, des contraintes de production et des objectifs opérationnels.

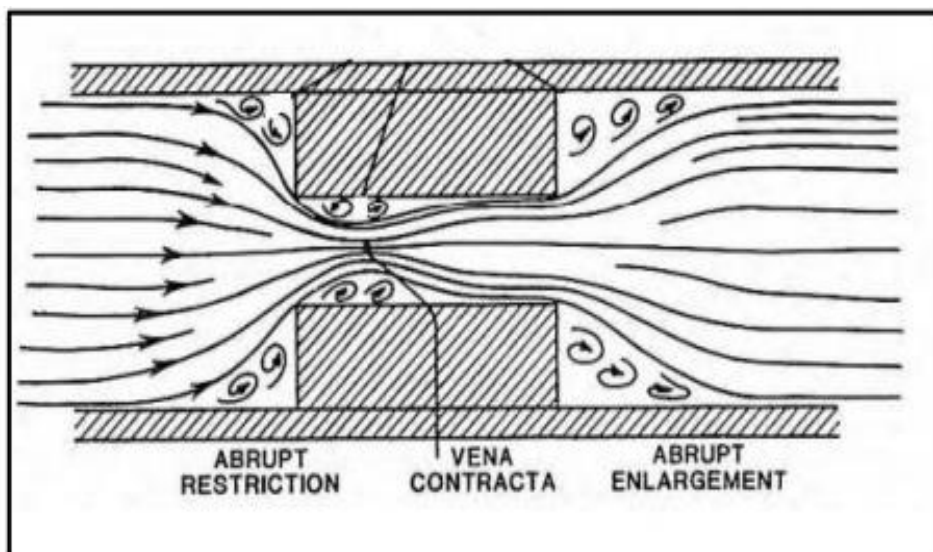


Figure 2- 1 schéma d'écoulement à travers une Duse

Cette représentation schématique Dans la figure (2-1), permet de visualiser les principales étapes de l'écoulement à travers une Duse :

- Approche de l'orifice : Le fluide s'approche de l'orifice de la duse il quitte la paroi du tube et commence à former un jet. Cette région où le fluide se sépare de la paroi est appelée la zone d'approche.
- Contraction du jet : Une fois que le fluide quitte la paroi du tube, il se contracte, c'est-à-dire qu'il se rétrécit pour former un jet plus étroit. Cette contraction se produit en raison de la diminution de la section transversale à mesure que le fluide passe à travers l'orifice de la duse. Le jet converge vers un point minimum, appelé la gorge ou "vena contracta". À ce stade, le jet atteint sa section la plus étroite.
- Élargissement du jet : Après avoir atteint la gorge, le jet s'élargit à nouveau en se rapprochant de la paroi de l'alésage de l'étranglement. Cela se produit car la pression du fluide diminue à mesure qu'il se déplace à travers la gorge, ce qui entraîne une expansion du jet.
- Une zone de turbulence se forme à l'entrée et à la sortie de l'étrangleur, ce qui contribue aux pertes de pression. Lorsque le fluide traverse ces zones, il rencontre des variations rapides de vitesse, de direction, de géométrie et aux interactions complexes entre le fluide et la paroi de l'étrangleur, créant ainsi des tourbillons

Les duses jouent plusieurs rôles essentiels dans l'industrie pétrolière et gazière. Leurs principales fonctions comprennent :

- Les duses sont utilisées pour réguler et contrôler le débit de fluide provenant du puits. Elles permettent de limiter le taux de production et de maintenir un débit constant et contrôlé. Cela permet d'optimiser l'exploitation du puits en évitant un écoulement excessif ou incontrôlé.
- Produire le réservoir au taux le plus optimal : Les duses sont calibrées pour atteindre un taux de production optimal du réservoir. En ajustant la taille et les caractéristiques de la duse, les ingénieurs peuvent réguler le débit de fluide et maintenir une production cohérente. Cela permet de maximiser la récupération du pétrole ou du gaz tout en préservant l'intégrité du réservoir.
- Empêcher la formation de cônes d'eau ou de gaz : Dans certains cas, les puits peuvent produire simultanément du pétrole, du gaz et de l'eau. Les duses sont utilisées pour contrôler la production de ces différents fluides afin d'éviter la formation de cônes d'eau ou de gaz indésirables. Cela permet de maximiser l'extraction du pétrole tout en minimisant la production d'eau ou de gaz non rentables.
- Maintenir la contre-pression : Les duses contribuent à maintenir une contre-pression dans le puits. Cela permet de stabiliser le flux de fluide et de prévenir l'affaissement de la pression du réservoir, ce qui pourrait entraîner une diminution de la production. En maintenant une contre-pression adéquate, les duses favorisent une production plus stable et efficace.
- Protéger la formation et les équipements de surface contre les fluctuations anormales de pression : Les duses jouent un rôle essentiel dans la protection de la formation géologique ainsi que des équipements de surface. En contrôlant le débit et la pression

du fluide, elles contribuent à prévenir les fluctuations anormales de pression qui pourraient endommager la formation souterraine ou les équipements de surface.

### 2.3 Régime d'écoulements : critique et sous-critique

Le régime d'écoulement dans les duses peut être classé en deux catégories principales : le régime critique et le régime sous-critique. Ces régimes décrivent le comportement de l'écoulement en fonction des conditions de pression, de débit et des propriétés du fluide

#### 2.3.1 Écoulements critique

« L'onde sonore et l'onde de pression sont toutes deux des ondes mécaniques. Lorsque la vitesse d'écoulement du fluide dans un étranglement atteint la vitesse de déplacement du son dans le fluide dans les conditions in situ, l'écoulement est appelé "écoulement sonique(critique)". Dans des conditions d'écoulement sonique, l'onde de pression en aval de l'étranglement ne peut pas remonter à travers l'étranglement parce que le fluide se déplace dans la direction opposée à la même vitesse. Par conséquent, il existe une discontinuité de pression au niveau de l'étranglement, c'est-à-dire que la pression en aval n'affecte pas la pression en amont. En raison de la discontinuité de pression au niveau de la duse, toute modification de la pression en aval ne peut être détectée par le manomètre en amont. Bien entendu, tout changement de la pression en amont ne peut pas non plus être détecté par le manomètre en aval. Ce flux sonique fournit une caractéristique unique d'étranglement qui stabilise le taux de production du puits ». [3]

Pendant l'écoulement critique, le débit à travers la Duse atteint une valeur maximale par rapport aux conditions dominantes en amont, et puisque aucune perturbation en aval ne peut être transmise en amont, cela implique que la diminution de la pression aval ne va pas augmenter le débit.

L'existence ou non d'un écoulement critique ou sous critique au niveau d'une Duse dépend du rapport entre la pression aval et la pression amont. Si ce rapport de pression  $y$  est inférieur à un rapport de pression critique  $y_c$ , il existe un écoulement sonique (critique). Si ce rapport de pression est supérieur ou égal ce rapport, il y a un écoulement subsonique (sous-critique).

Le rapport de pression critique  $y_c$  à travers les duses est exprimé comme suit:

$$y_c = \left(\frac{2}{k+1}\right)^{\frac{k}{k-1}} = 0.55 \quad (2.1)$$

Où :

$k = C_p/C_v$  est le rapport de chaleur spécifique. La valeur de  $k$  est de 1,4 pour l'air et de 1,28 pour le gaz naturel. Le rapport de pression critique est donc de 0,528 pour l'air et de 0,549 pour le gaz naturel.

### 2.3.2 Écoulements sous critique

Dans le régime sous-critique, l'écoulement du fluide se fait à des vitesses inférieures à la vitesse du son dans le fluide. Cela signifie que l'écoulement est subsonique, avec des variations de pression et de vitesse relativement lentes et graduelles le long de l'écoulement.

Les variations de pression ou bien les pertes de pression peuvent entraîner une diminution significative du débit.

La figure ci-dessous représente une courbe typique montrant la variation du débit en fonction du rapport de pression Poutlet/Pupstream.

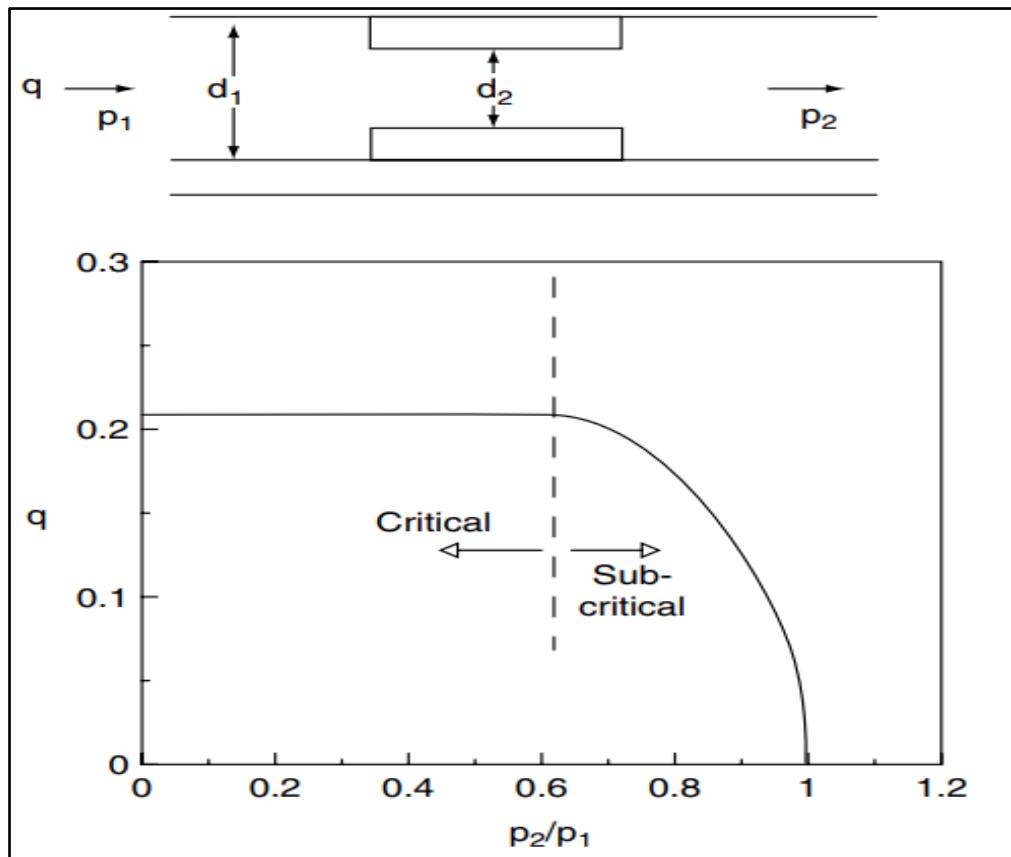


Figure 2- 2 Courbe de performance typique d'une Duse [4]

### 2.4 Modélisation des écoulements à travers les Duses

La chute de pression à travers une Duse dépend du débit du fluide. Par conséquent, la mesure indirecte du débit peut être réalisée en mesurant la chute de pression à travers la Duse. Cette méthode est couramment utilisée pour les écoulements monophasiques et peut être réalisée avec précision. Cependant, pour les écoulements multiphasiques, où plusieurs phases de fluide (liquide, gaz) coexistent, la mesure du débit devient plus difficile.

Dans le cas des écoulements multiphasiques, où plusieurs phases de fluide coexistent, il est plus complexe de mesurer le débit en utilisant uniquement la chute de pression à travers une

duse. Les interactions entre les différentes phases de fluide rendent les calculs plus difficiles et nécessitent souvent l'utilisation d'équations et de modèles spécifiques pour estimer le débit.

### 2.4.1 Modélisation des écoulements monophasiques

Les équations de pression pour l'écoulement de gaz à travers un étranglement sont dérivées d'un processus isentropique. Cela s'explique par le fait qu'il n'y a pas de temps pour le transfert de chaleur (adiabatique) et la perte par frottement est négligeable (en supposant qu'elle soit réversible) au niveau des duses

#### Dérivation de l'équation

En utilisant l'équation de Bernoulli et en assumant que les frottements sont négligeables pour un écoulement à travers les duses d'un point amont 1 à un point en aval 2

$$\int_{P_1}^{P_2} \frac{dp}{\rho} + \int_{V_1}^{V_2} V * dV = 0 \quad (2.2)$$

Dans un écoulement de gaz, la vitesse  $V_2$  est généralement plus élevée que la vitesse  $V_1$ , donc  $V_2^2 \gg V_1^2$

$$\int_{P_1}^{P_2} \frac{dp}{\rho} + \frac{V_2^2}{2} = 0 \quad (2.3)$$

La densité varie à travers la duse. Une hypothèse couramment utilisée est que ce processus est adiabatique (avec un exposant  $k$ , le rapport entre les chaleurs spécifiques du gaz) :

$$p * \rho^k = C \quad (2.4)$$

Ou  $C$  est une constante.

$$C^{\frac{1}{k}} * \int_{P_1}^{P_2} \frac{dp}{p^{\frac{k}{k-1}}} + \frac{V_2^2}{2} = 0 \quad (2.5)$$

$$C^{\frac{1}{k}} * \frac{k}{k-1} * \left( P_2^{\frac{k-1}{k}} - P_1^{\frac{k-1}{k}} \right) + \frac{V_2^2}{2} = 0 \quad (2.6)$$

La constante  $C$  est exprimée en fonction des conditions d'entrée :

$$C^{\frac{1}{k}} = \frac{p_1^{1/k}}{\rho_1} \quad (2.7)$$

Avec le rapport de pression  $y$  est donné par  $p_2/p_1$  :

$$V_2 = \sqrt{2 * \frac{Z_1 * R * T_1}{M_W} * \frac{k}{k-1} * \left( 1 - y^{\frac{k-1}{k}} \right)} \quad (2.8)$$

En terme de débit

$$q_g^2 = A_2 * C_D * \sqrt{2 * \frac{Z_1 * R * T_1}{M_W} * \frac{k}{k-1} * \left( 1 - y^{\frac{k-1}{k}} \right)} \quad (2.9)$$

$C_d$  est un facteur qui compte tenu de la section effective d'écoulement dans la duse.

Le débit volumique local au point 2 est lié au taux aux conditions standards par

$$q_g^2 * \rho_2 = \bar{q}_g * \rho_{sc} \quad (2.20)$$

Donc, on obtient :

$$\bar{q}_g = \frac{\rho_2 * A_2 * C_D}{\rho_{sc}} * \sqrt{2 * \frac{Z_1 * R * T_1}{M_W} * \frac{k}{k-1} * \left(1 - y^{\frac{k-1}{k}}\right)} \quad (2.21)$$

Et

$$\frac{p_2^{\frac{1}{k}}}{\rho_2} = \frac{p_1^{\frac{1}{k}}}{\rho_1} \quad (2.22)$$

Enfin

$$\bar{q}_g = \left(\frac{P_1 * T_{sc} * A_2 * C_d}{P_{sc}}\right) * \sqrt{2 * \frac{R}{Z_1 * T_1 * M_W} * \frac{k}{k-1} * \left(y^{\frac{2}{k}} - y^{\frac{k+1}{k}}\right)} \quad (1.23)$$

#### **Pour le régime sous critique :**

Dans le régime subsonique, le débit de gaz est donné par l'expression suivante

$$\bar{q}_g = \left(\frac{P_1 * T_{sc} * A_2 * C_d}{P_{sc}}\right) * \sqrt{2 * \frac{R}{Z_1 * T_1 * M_W} * \frac{k}{k-1} * \left(y^{\frac{2}{k}} - y^{\frac{k+1}{k}}\right)} \quad (2.24)$$

#### **Pour le régime critique :**

Dans le régime sonique, le débit du gaz atteint sa valeur maximale. Dans ce cas, il faut utiliser le rapport de pression critique  $y_c$  au lieu du rapport de pression réel  $y$  dans l'équation précédente

$$\bar{q}_g = \left(\frac{P_1 * T_{sc} * A_2 * C_d}{P_{sc}}\right) * \sqrt{2 * \frac{R}{Z_1 * T_1 * M_W} * \frac{k}{k-1} * \left(y_c^{\frac{2}{k}} - y_c^{\frac{k+1}{k}}\right)} \quad (2.25)$$

### **2.4.2 Modélisation des écoulements diphasiques**

L'écoulement multiphasique à travers les duses est un processus complexe qui nécessite l'utilisation de modèles et de corrélations pour prédire le débit d'écoulement. L'incorporation de données de terrain réelles permet d'améliorer l'exactitude de ces modèles en les calibrant et en les validant pour des conditions spécifiques. Cela permet d'obtenir des prédictions plus



fiables et précises pour l'écoulement à travers les étranglements dans l'industrie du pétrole et du gaz.

Les études sur les duses de tête de puits et leurs impacts sur les débits de production ont commencé avec Tangren et al. (1949). Leur analyse s'est concentrée sur le débit à travers les restrictions et a spécifiquement examiné les conditions de débit critique. Par la suite, plusieurs auteurs ont développé des corrélations pour décrire les régimes critique et sous-critique des chokes de tête de puits.

Ces corrélations peuvent être utilisées pour prédire les débits de production en fonction des caractéristiques du puits et des conditions de production. Elles sont souvent basées sur des données empiriques et des modèles mathématiques.

Dans cette section, on présentera quelques corrélations améliorées pour le calcul du débit d'écoulement à travers les duses, tant pour le régime d'écoulement critique que sous-critique.

#### **2.4.2.1 Les corrélations existantes dans le régime critique**

##### **La corrélation de Gilbert**

La corrélation développée par Gilbert est une équation qui permet de prédire le débit massique à travers une conduite pour un régime d'écoulement critique. Gilbert a utilisé 268 points de données provenant du champ Ten Section en Californie pour développer cette corrélation. Il a établi une relation entre la pression en amont de la duse, le diamètre de la duse et le débit de liquide produit afin de prédire le débit.

La corrélation de Gilbert est décrite sous la forme :

$$Q = a \frac{P_U \cdot \phi^b}{GLR^c} \quad (2.26)$$

Dans cette équation,  $P_U$  représente la pression en amont (tête de puits) exprimée en psia,  $Q$  est le taux de liquide brut produit en bbl/jour,  $GLR$  est le rapport gaz-liquide de production exprimé en  $Scf/bbl$ , et  $\phi$  est le diamètre de l'étranglement en  $1/64$  in.

Les constantes  $a$ ,  $b$  et  $c$  sont des paramètres empiriques liés aux propriétés des fluides. Gilbert, sur la base des données de production du champ Ten Section en Californie, a déterminé les valeurs de ces constantes comme suit :  $a = 10$ ,  $b = 0,546$  et  $c = 1,89$ . Cependant, d'autres chercheurs ont proposé différentes valeurs pour ces constantes. Par exemple, Baxendell (1957), Ros (1960), Achong (1961), Owolabi et al (1991), Okon et al (2015), et Pilehvari (1980) ont suggéré d'autres valeurs. Un résumé de ces différentes valeurs est présenté dans le tableau (2-1)

**Tableau 2- 1 les paramètres empiriques des équations**

Corrélation	a	b	c
Gilbert	10	0.546	1.89
OWOLABI et AL	35.72	0.289	1.83
OKON et AL	5.1474	0.5048	1.7093
Ros	17.4	0.5	2
Baxendell	9.56	0.546	1.93
Achong	3.82	0.65	1.88
Pilehvari	46.67	0.313	2.11

Il existe une autre forme de l'équation de Gilbert, Cette forme introduit une puissance exponentielle à l'équation originale, La forme exponentielle de l'équation permet de prendre en compte les interactions complexes entre les différents paramètres et d'améliorer la précision des prédictions.

$$Q = \left( a \frac{P_U \cdot \phi^b}{GLR^c} \right)^e \quad (2.27)$$

#### La corrélation de Towailib et Marhoun

La corrélation développée par Towailib et Marhoun est une autre corrélation utilisée pour prédire le débit d'écoulement critique à travers une duse. Cette corrélation est similaire à celle de Gilbert, mais elle remplace le rapport gaz-liquide de production (GLR) par la densité du mélange. Elle a été basée sur l'analyse de 3554 données de tests de production provenant de dix champs du Moyen-Orient. La corrélation de Towailib et Marhoun s'applique spécifiquement aux conditions d'écoulement critique.

La corrélation de Towailib et Marhoun est exprimée par l'équation suivante :

$$Q = \frac{P_u^b \cdot \phi^c}{a \cdot \gamma_m^d} \quad (2.28)$$

Où  $\gamma_m$  est la densité du mélange est calculée à l'aide de l'équation

$$\gamma_m = \gamma_o + 2.18 * 10^{-4} R_p \gamma_G \quad (2.29)$$

Le tableau (2-2) présente les coefficients spécifiques de la corrélation de Towailib et Marhoun

**Tableau 2- 2 les coefficients spécifiques de la corrélation de Towailib et Marhoun**

Corrélation	a	b	c	e
Towailib et Marhoun	545.5	0.9813	2.0704	1.4637

### La corrélation de Attar et Abdul Majeed

La corrélation développée par Attar et Abdul Majeed est une autre corrélation utilisée pour prédire le débit d'écoulement critique à travers une duse. Cette corrélation a été basée sur des données provenant des champs de l'est de Bagdad. Elle présente des similitudes avec la corrélation de Gilbert, mais avec l'introduction de la densité de l'huile dans le dénominateur de la formule.

La corrélation de Attar et Abdul Majeed s'exprime comme suit :

$$Q = a \frac{P_u^b \cdot \phi^c}{\gamma^d \cdot GLR^e} \quad (2.30)$$

$\gamma$  Représente La densité API est calculée à l'aide de l'équation :

$$API = \frac{141}{\gamma_o} - 131.5 \quad (2.31)$$

Les constantes spécifiques de la corrélation de Attar et Abdul Majeed ont été déterminées en se basant sur les données des champs de l'est de Bagdad. Le tableau (2-3) présente les valeurs des constantes  $a$ ,  $b$ ,  $c$  et  $d$  pour cette corrélation.

**Tableau 2- 3 les coefficients spécifiques de la corrélation de Attar et Abdul Majeed**

Corrélation	a	b	c	d	e
Attar et Abdul Majeed	0.033567	0.8756	1.796	0.43957	0.2693

#### 2.4.2.2 Les corrélations pour Régime sous critique

##### La corrélation de Elgibaly et Nashawi

La corrélation développée par Elgibaly et Nashawi est une formule utilisée pour estimer le débit d'écoulement non critique à travers une duse. Elle diffère de la corrélation de Gilbert en ce qu'elle remplace la pression en amont par la différence de pression entre l'amont et l'aval de la duse, ce qui prend en compte les effets de la pression différentielle sur le débit.

L'équation de la corrélation de Elgibaly et Nashawi est donnée par :

$$Q = \left( \frac{a \cdot \Delta P \cdot \phi^b}{GOR^c} \right)^{\frac{1}{2}} \quad (2.32)$$

$\Delta P$  représente la différence de pression entre l'amont et l'aval de la duse

Les constantes  $a$  et  $b$  sont spécifiques à cette corrélation et ont été déterminées à partir de données expérimentales ou d'études spécifiques. Le tableau (2-4) présente les valeurs des constantes  $a$  et  $b$  pour la corrélation de Elgibaly et Nashawi.

**Tableau 2- 4 les coefficients spécifiques de la corrélation de Elgibaly et Nashawi**

Corrélation	a	b	c
Elgibaly et Nashawi	1204.8193	2.747	1.13501

## Extensions de l'équation de Gilbert

### 1<sup>ère</sup> forme

L'extension de l'équation de Gilbert proposée par Mohammad Ali Kargarpour dans son article "Oil and gas well rate estimation by choke formula : semi-analytical approach" introduit des considérations supplémentaires pour améliorer la précision du calcul du débit à travers une duse dans le régime non critique.

L'objectif de cette extension est d'améliorer la précision du calcul en prenant en compte la chute de pression à travers l'étrangleur, en plus des autres facteurs déjà considérés dans l'équation de Gilbert d'origine.

L'équation proposée par Mohammad Ali Kargarpour s'exprime comme suit :

$$Q = a \frac{P_u^b \cdot \Delta P^c \cdot \phi^e}{GOR^f} \quad (2.33)$$

Les constantes  $a$ ,  $b$ ,  $c$ ,  $e$  et  $f$  sont des coefficients empiriques présentés dans le tableau(2-5) suivant :

**Tableau 2- 5les coefficients spécifiques de la corrélation de extension de Gilberte 1**

Corrélation	a	b	c	e	f
extension de Gilberte 1	0.098	0.41	0.44	2	0.42

### 2<sup>ème</sup> forme

Deux fonctions ont été proposées pour améliorer les performances sur le champ de Hassi Messaoud *HMD*. La première fonction dépend de la pression amont  $P_u$ , de la pression aval  $P_d$ , du diamètre de la duse et du *GOR* (Gas-Oil Ratio). La deuxième fonction, quant à elle, dépend de la chute de pression, du diamètre de la duse et du *GOR*.

Extension de Gilbert 2 proposée s'exprime comme suit :

$$Q = a \cdot \Delta P^b \cdot \phi^c \cdot GOR^e \quad (2.34)$$

Extension de Gilbert 3 proposée s'exprime comme suit :

$$Q = a \cdot (p_t^2 - p_p^2) \cdot \phi^c \cdot GOR^e \quad (2.35)$$

Où Le débit  $Q$  est exprimé en *STB/day*, le *GOR* en *scf/STB*, la pression amont  $P_t$ , de la pression aval  $P_p$  en psi et le diamètre en *1/64 inch*.

## **2.5 Conclusion**

Les duses, jouent un rôle fondamental dans la gestion efficace du débit des puits dans l'industrie pétrolière. Pour estimer avec précision les débits à travers ces conduites, des corrélations spécifiques ont été établies dans la littérature. Chaque corrélation est élaborée en se basant sur des caractéristiques particulières de champs pétroliers spécifiques, ce qui rend essentielle leur adaptation aux conditions propres à chaque champ.

## **CHAPITRE 3**

# **CALCUL ET OPTIMISATION DE LA PRODUCTION**

### 3.1 L'algorithme de Régression non linéaire

« La régression non linéaire est une procédure qui détermine les valeurs des paramètres minimisant la somme des carrés des distances entre les points de données et la courbe pour l'équation d'intérêt. En d'autres termes, étant donné le point de données  $y$  et le point correspondant sur la courbe  $\hat{y}$ , la régression non linéaire minimise la somme des carrés des différences entre eux :  $SS = \text{somme} [(y - \hat{y})^2]$ .»[2]

L'algorithme `curve_fit` de la bibliothèque SciPy de Python est utilisé pour effectuer des ajustements de courbes non linéaires. Il est basé sur la méthode des moindres carrés et cherche à trouver les paramètres optimaux d'une fonction mathématique qui représente au mieux les données fournies.

L'algorithme `curve_fit` utilise des méthodes d'optimisation numérique pour ajuster les paramètres de la fonction mathématique afin de minimiser l'erreur entre les valeurs observées et prédites. Il cherche à trouver les paramètres qui donnent la meilleure correspondance possible entre les données et la courbe ajustée.

Le fonctionnement général de l'algorithme `curve_fit`

- Équation de la fonction mathématique : La fonction mathématique à ajuster est généralement définie par une équation avec des paramètres inconnus.
- Estimation initiale des paramètres : Avant de commencer l'ajustement, il est nécessaire de fournir une estimation initiale des paramètres. Ces estimations peuvent être basées sur des connaissances préalables ou des valeurs approximatives obtenues à partir des données.
- Calcul de l'erreur : À chaque itération de l'algorithme, les valeurs prédites par la fonction avec les paramètres actuels sont comparées aux valeurs observées. L'erreur est calculée comme la différence entre les valeurs observées et prédites.
- Méthodes d'optimisation : Différentes méthodes d'optimisation peuvent être utilisées pour ajuster les paramètres telles que la descente de gradient ou la méthode de Gauss-Newton. Le critère utilisé pour toutes les méthodes est le moindre carré, qui vise à minimiser la somme des carrés des erreurs.
- Mise à jour des paramètres : Une fois que l'erreur est calculée, les paramètres sont ajustés dans le sens qui réduit cette erreur. Les techniques d'optimisation utilisent des algorithmes spécifiques pour déterminer la direction et l'amplitude des ajustements à effectuer sur les paramètres.
- Convergence : L'algorithme itère jusqu'à ce qu'il atteigne une convergence, c'est-à-dire lorsque les paramètres ajustés donnent une bonne correspondance entre les

valeurs observées et prédites. La convergence est généralement déterminée par des critères prédéfinis, tels que la tolérance de l'erreur ou le nombre maximum d'itérations.

### 3.2 Métriques d'évaluation d'un modèle

Les métriques d'évaluation d'un modèle sont des indicateurs statistiques qui permettent de mesurer la performance d'un modèle d'apprentissage automatique. Il existe de plusieurs métriques d'évaluation en fonction du type de modèle et du type de problème.

Pour Les tâches de régression ou de prédiction, Il existe quatre métriques couramment utilisées pour évaluer la performance :

- **Le coefficient de détermination ( $R^2$ )** : est la métrique la plus couramment utilisée pour évaluer la qualité de l'ajustement d'un modèle. C'est une mesure de la proportion de la variance totale de la variable dépendante qui est expliquée par le modèle.  $R^2$  prend des valeurs entre 0 et 1 où une valeur proche de 1 indique que le modèle ajuste très bien les données.

$$R^2 = 1 - \frac{\sum(y_i - \hat{y})^2}{\sum(y_i - \bar{y})^2} \quad (3.1)$$

Avec :

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad (3.2)$$

Où :

$y_i$  : chaque point se données

$\hat{y}$  : la valeur prédite par le modèle

$\bar{y}$  : la valeur moyenne des données

$n$  : le nombre de points

- **Root Mean Squared Error (RMSE)** La racine carrée de l'erreur quadratique moyenne : Elle mesure l'écart quadratique moyen entre les prévisions du modèle et les valeurs réelles. Elle indique à quel point les valeurs prédites sont proches des valeurs réelles, donc une valeur de RMSE plus faible signifie que le modèle a une meilleure performance.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y} - y_i)^2} \quad (3.3)$$

- **L'erreur absolue moyenne (MAE)** : mesure la moyenne des écarts absolus entre les valeurs prédites et les valeurs réelles de la variable dépendante. Elle donne une



idée de la distance moyenne entre les prévisions et les observations réelles. MAE Plus petite indique une meilleure performance du modèle

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y} - y_i| \quad (3.4)$$

- **L'erreur relative absolue (RAE)** : est une mesure qui évalue l'erreur absolue d'une prédiction ou d'une estimation en fonction de la valeur réelle. Elle est calculée en prenant la différence absolue entre la prédiction et la valeur réelle, puis en la divisant par la valeur réelle. Cela donne une mesure de l'erreur en termes relatifs par rapport à la valeur réelle.

$$RAE = \frac{1}{n} \sum_{i=1}^n \frac{|\hat{y} - y_i|}{\hat{y}} \quad (3.5)$$

### 3.3 Résultats et discussions

#### 3.3.1 Présentation des Données de production du champ de Hassi

##### Messaoud :

Le dataset se présente sous forme d'un tableau Excel (08 colonnes et 24455 lignes). Il comprend les données brutes de production de 609 puits du champ de Hassi Messaoud, couvrant la période de 2004 à 2023. Ces données sont obtenues à partir d'opérations de jaugeage soit par Le jaugeage conventionnel ou par les nouvelles méthodes de jaugeage telles que le *VX* (Virtual Flowmetering). Cependant, le nombre de jaugeage qui s'effectue sur les puits peut varier d'un puits à un autre en fonction de divers facteurs, tels que les problèmes rencontrés lors de la production de certains puits, la disponibilité des installations et du personnel, ainsi que les conditions économiques du moment

Le dataset brut contient initialement 24455 lignes de données et comprend les variables suivantes : nom du puits, la date du jaugeage, le diamètre de la duse, la pression en amont de puits, la pression en aval, *GOR* (Gas-Oil Ratio), le débit de production d'huile, le débit de production d'eau et la densité du mélange produit.

#### 3.3.2 Préparation des données

La préparation des données est une étape essentielle pour tirer le maximum de valeur des données brutes et concevoir des modèles performants et pour éviter les résultats biaisés. Elle comprend plusieurs étapes pour traiter les valeurs manquantes, les erreurs de frappe, les outliers et les valeurs erronées.

Pour préparer les données à des fins d'analyse, On a suivi les étapes suivantes :

- Identification des valeurs manquantes dans les différentes variables et Elimination des lignes contenant ces valeurs.
- Eliminations les lignes où certaines variables contiennent des valeurs nulles telle qu'une pression flow ligne nulle, un diamètre de la duse nul
- Identification des outliers, qui sont des valeurs aberrantes ou extrêmes comme des valeurs de diamètres de la duse ou de pressions très grandes qui peuvent indiquer soit

une erreur de mesure ou un résultat de variabilité présente dans les données et suppression de ces outliers,

- Eliminations des lignes où Une pression de flow line supérieure à la pression de tête.
- Création de nouvelle variable en dérivant le rapport entre la pression du flow line(aval) et la pression de tête(amont)  $P_d/P_u$ . Cette nouvelle variable peut fournir des informations supplémentaires pertinentes pour diviser l'ensemble de données en deux catégories : le régime critique et le régime sous-critique pour le champ de HMD. Le régime critique est défini lorsque ce rapport est inférieur ou égale à 0.75. Pour le régime sous critique ce rapport est strictement supérieur à 0.75.
- Une autre variable pertinente à créer est la chute de pression à travers la duse  $\Delta P$ , qui est calculée en soustrayant la pression du flow line de la pression de tête  $P_d - P_u$ . Cette nouvelle variable permet de quantifier la différence de pression entre l'entrée et la sortie de la duse
- Partitionnement des données : Division des données en ensembles d'apprentissage (données d'entraînement représente 80% des données initiales), et de test (données de test) représente 20% des données initiales pour l'évaluation des modèles.
- L'ensemble d'apprentissage est utilisé pour créer et ajuster le modèle, tandis que l'ensemble de test est utilisé pour évaluer la performance et la capacité de prédiction du modèle sur de nouvelles données.

Les deux tableaux ci-dessous présentent les plages de valeurs des paramètres utilisées dans les régimes critique et sous-critique, ainsi que les valeurs maximales, minimales et moyennes de chaque paramètre après le traitement des données.

**Tableau 2- 6 :Différentes plages des paramètres de données du champ de HMD en régime critique**

	DUSE(1/64in)	QO(bbl/day)	PP(psi)	PT(psi)	GOR(scF/STB)
mean	15.374759	603.931412	277.013479	666.303073	8904.78
min	6	1.5096	14.5038	66.71748	111.56
25%	12.7	297.3912	203.0532	384.3507	2110.88
50%	15	499.6776	234.96156	525.03756	4834.81
75%	17	797.0688	295.87752	807.86166	10506.47
max	40	6461.088	784.075428	6294.6492	1 259 938.00

**Tableau 2- 7 Différentes plages des paramètres de données du champ de HMD en régime sous-critique.**

	DUSE(1/64 in)	QO((bbl/day)	PP(psi)	PT(psi)	GOR(scF/STB)
mean	16.931747	389.723768	328.113453	390.461754	9 756.00
min	6	1.5096	101.5266	123.427338	5.61
25%	13	170.5848	224.8089	268.3203	2 401.39
50%	16	298.9008	268.3203	320.53398	5 730.38
75%	20	499.3002	481.52616	551.1444	11 168.43
max	50.8	3441.888	826.7166	1068.93006	1 622 015

### 3.3.3 Méthodologie

Dans le cadre de notre étude visant à améliorer la capacité de calcul du débit sur le champ Hassi Massoud, nous utilisons les corrélations sélectionnées. Ces équations sont des modèles mathématiques permettant d'estimer le débit en fonction de différents paramètres. Les coefficients de ces équations ont été déterminés à partir de différents champs.

Les corrélations sélectionnées sont adaptées sur le champs de *HMD* pour déterminer la corrélation qui offre la meilleure prédiction en termes de précision.

Afin d'adapter ces corrélations aux données spécifiques du champ *HMD*, nous utilisons l'algorithme de Régression non linéaire curve fit en Python pour ajuster les courbes et estimer les paramètres optimaux des équations choisies en fonction des données de champ *HMD*, et améliorant ainsi la précision de la prédiction du débit.

Pour le régime critique on choisit :

- La corrélation de Gilbert (2.27)
- La corrélation de Towailib et Marhoun (2.28)
- La corrélation de Attar et Abdul Majeed (2.30)

Pour le régime sous critique on choisit :

- La corrélation de Elgibaly et Nashawi (2.32)
- Extensions de l'équation de Gilbert 1ere forme (2.33) et 2eme forme (2.34) et (2.35)

Pour évaluer les performances de chaque équation, nous utiliserons plusieurs métriques d'évaluation, notamment : RMSE (Root Mean Square Error) ,  $R^2$  (Coefficient de détermination), RAE (Relative Absolute Error), MAE (Mean Absolute Error)

En évaluant les performances des équations à l'aide de ces métriques, Cela nous permettra de choisir l'équation qui présente la meilleure adéquation aux données et qui est la plus adaptée pour prédire le débit en fonction des paramètres spécifiés.

### 3.3.4 Résultats de l'adaptation des modèles théoriques

Cette section mettra en évidence les résultats obtenus à la suite de l'adaptation des modèles théoriques existants pour le régime critique et le régime sous critique

#### 3.3.4.1 Régime critique

Pour cette étude, un total de 19239 lignes de données a été utilisés. Parmi ces données, 80% ont été réservées pour le processus d'entraînement des modèles, tandis que les 20% restants ont été utilisés pour les tests de validation.

Les tableaux (3-1), (3-2) et (3-3) affichent les nouvelles constantes des corrélations adaptées, tandis que le tableau (3-4) présente les différentes statistiques liées à l'entraînement et aux tests des modèles théoriques

**Tableau 3- 1les nouvelles constantes d'équation de Gilbert**

modèle	a	b	c	e
<b>GILBERT</b>	12.24	1.51	0.59	0.79

**Tableau 3- 2 les nouvelles constantes d'équation de Attar et Abdul Majeed**

modèle	a	b	c	e	f
<b>Attar et Abdul Majeed</b>	1.099	0.798	1.206	0.478	-0.501

**Tableau 3- 3 les nouvelles constantes d'équation de Towailib & Marhoun**

modèle	a	b	c	e
<b>Towailib &amp; Marhoun</b>	0.093	0.422	0.337	1.839

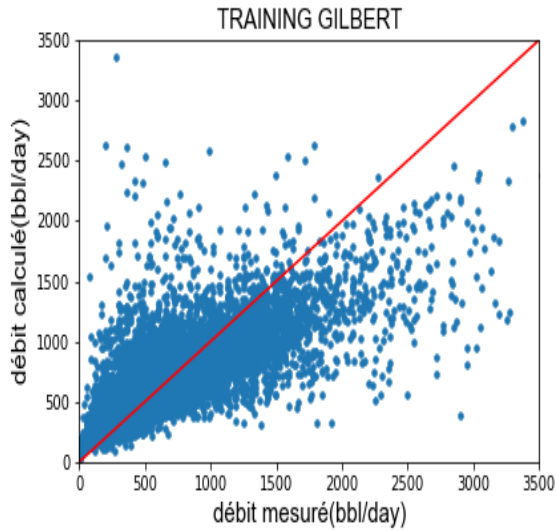
**Tableau 3- 4 les différentes statistiques liées à l'entraînement et aux tests des modèles théoriques**

modèle	train data				test data			
	R <sup>2</sup>	MAE	RMSE	RAE	R <sup>2</sup>	MAE	RMSE	RAE
<b>GILBERT</b>	0.62	169.05	273.31	0.43	0.58	174.01	280.13	0.4
<b>Attar et Abdul Majeed</b>	0.61	168.57	274.69	0.44	0.56	166.72	288.63	0.44
<b>Towailib &amp; Marhoun</b>	0.12	306.90	417.27	1.11	0.1	308.03	429.93	1.02

### Interprétation des résultats

- L'équation de *Gilbert* présente les meilleures performances avec un R<sup>2</sup> de 0.62 pour le train et 0.58 pour le test , une MAE de 174.01 et une RAE de 0.4 .
- L'équation de *Attar et Abdul Majeed* présente performance légèrement moins efficace par rapport L'équation de Gilbert avec un R2 de 0.56 , une MAE de 166.72et une RAE de 0.44
- Cela apparaît clairement dans les cross-plots, où les points de données se trouvent relativement proches de la ligne parfaite. Cette proximité indique une bonne correspondance entre les prédictions du modèle et les valeurs réelles
- Par contre, l'équation *Towailib & Marhoun* présente performance moins efficace. Elle a un faible R<sup>2</sup> (0.1) sur les ensembles du test, De plus, les valeurs de MAE, RMSE et RAE sont élevées, indiquant des erreurs de prédiction importantes et une

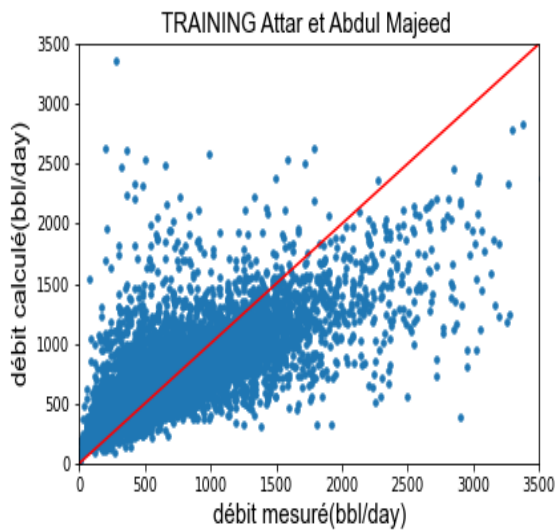
faible précision. Les graphiques (3-5) et (3-6) correspondant pour cette équation montre une distribution éloignée de la ligne parfaite



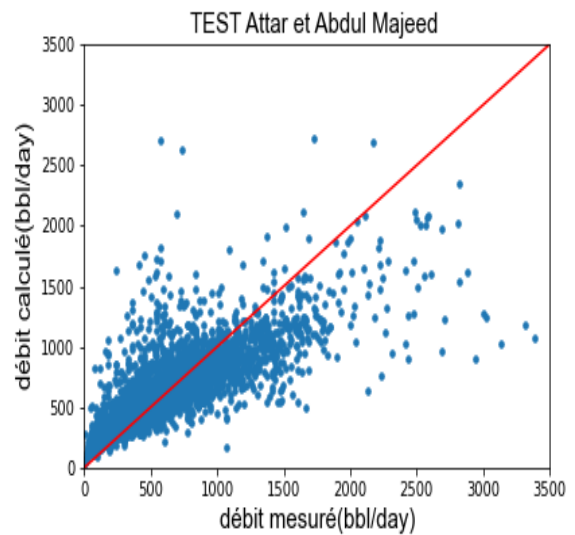
**Figure 3- 1 cross plot de training du modèle de GILBERT**



**Figure 3- 2 cross plot de TEST du modèle de GILBERT**



**Figure 3- 3 cross plot de TRAINING du modèle de Attar et Majeed**



**Figure 3- 4 cross plot de TEST du modèle de Attar et Majeed**

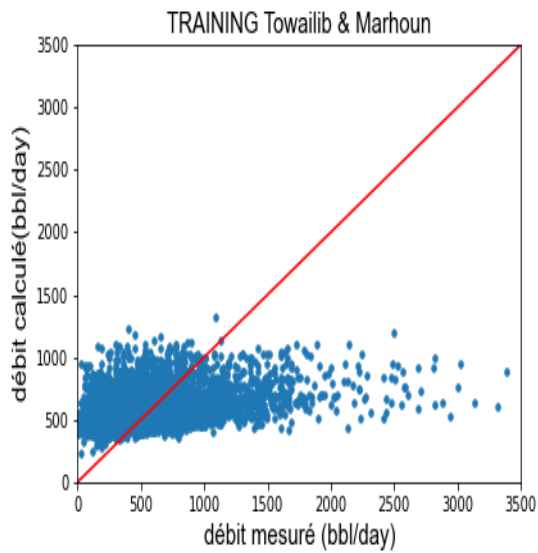


Figure 3- 5 cross plot de TRAINING du modèle de Towailib et Marhoun

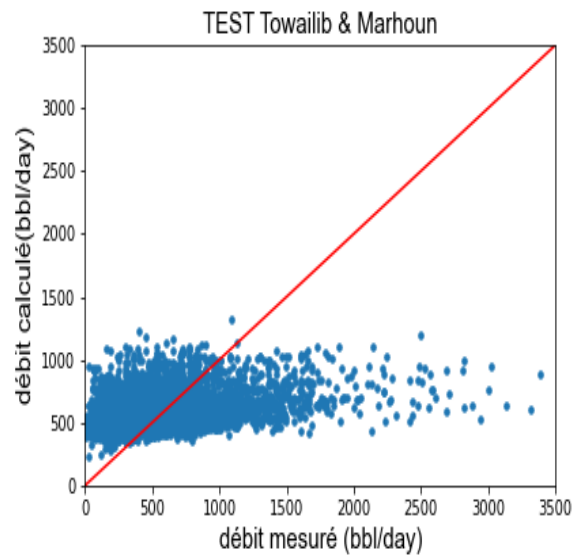


Figure 3- 6 cross plot de TEST du modèle de Towailib et Marhoun

### 3.3.4.2 Régime sous critique

Un total de 5214 points de données a été utilisés pour créer ce modèle, avec 80 % des données utilisées pour l'entraînement et 20% pour le test.

Les tableaux suivants fournissent les nouveaux coefficients du modèle ainsi que les différentes mesures d'évaluation utilisées pour évaluer sa performance.

Tableau 3- 5 les nouvelles constantes d'équation de Elgibaly et Nashawi

modèle	a	b	c
Elgibaly et Nashawi	159.05	2.77	0.59

Tableau 3- 6 les nouvelles constantes d'équation de extension de gilbert 1

modèle	a	b	c	e	f
extension de gilbert 1	3.85	0.41	0.29	1.28	0.31

Tableau 3- 7 les nouvelles constantes d'équation de extension de gilbert 2

modèle	a	b	c	e
extension de gilbert 2	12.93	0.48	1.39	-0.29

**Tableau 3- 8 les nouvelles constantes d'équation de extension de gilbert 3**

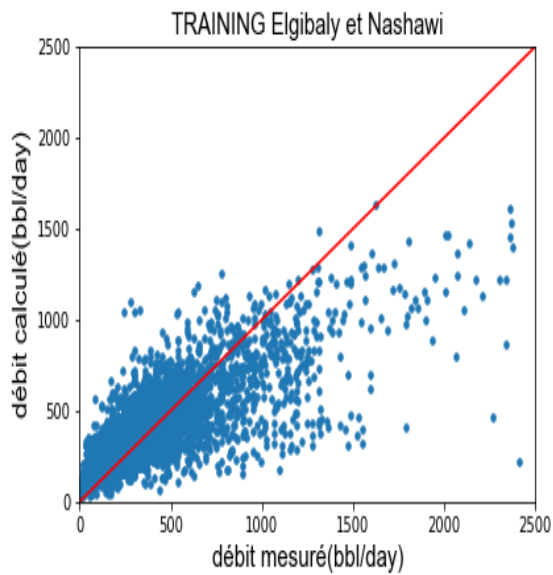
modèle	a	b	c	e
extension de gilbert 3	0.65	-5.84	1.54	-0.24

**Tableau 3- 9 les différentes statistiques liées à l'entraînement et aux tests des modèles théoriques**

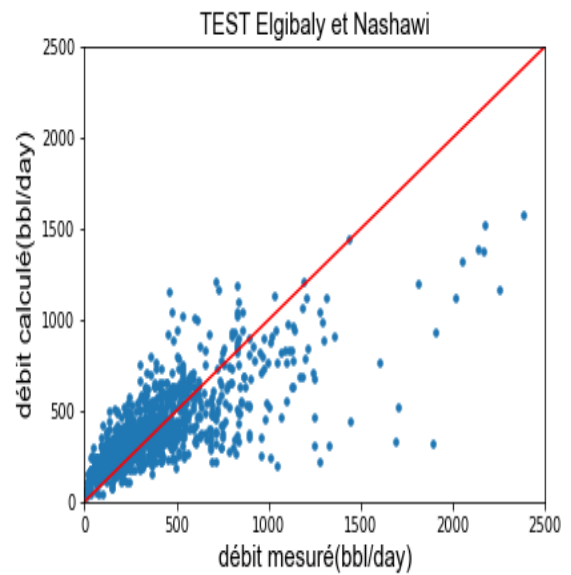
modèle	train data				test data			
	R <sup>2</sup>	MAE	RMSE	RAE	R <sup>2</sup>	MAE	RMSE	RAE
Elgibaly et Nashawi	0.54	146.77	229.19	0.65	0.55	146.77	215.87	0.6
extension de gilbert 1	0.56	144.4	225	0.64	0.57	143.1	210.79	0.58
extension de gilbert 2	0.54	146.9	229.17	0.65	0.54	146.82	215.88	0.61
extension de gilbert 3	0.42	170.59	257.05	0.84	0.42	167.26	244.42	0.78

### **Interprétation des résultats**

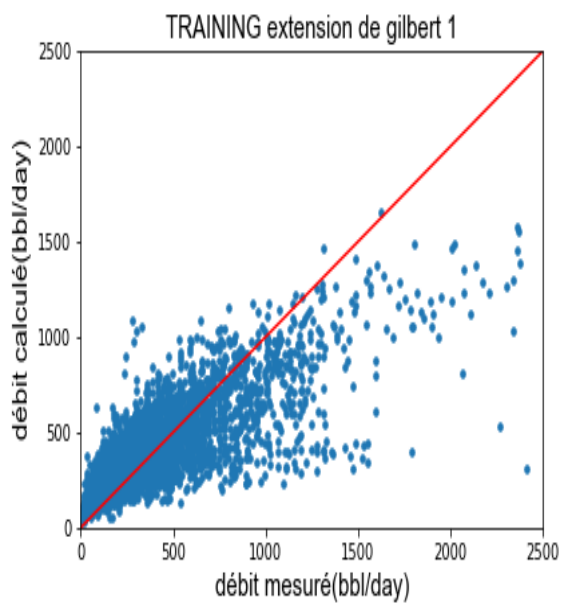
- Les corrélations *Elgibaly & Nashawi*, *extension de gilbert\_1* et *extension de gilbert 2* semblent, avoir des performances acceptables, avec des valeurs de R<sup>2</sup> (coefficient de détermination) autour de 0.54 à 0.57. Les erreurs absolues moyennes (MAE) se situent autour de 143 à 147, Les erreurs absolues relatives (RAE) sont d'environ 0.58 à 0.61. *L'Extension de Gilbert\_1* Semble présenter les performances les plus élevées parmi les quatre équations
- L'équation *extension de Gilbert\_3* présente des performances légèrement inférieures, avec un R<sup>2</sup> de 0.42, une MAE de 167.26 et une RAE de 0.78 . Et des erreurs absolues plus élevées
- En générale les équations *Elgibaly & Nashawi*, *extension de gilbert\_1* et *extension de gilbert\_2* ont des performances similaires, avec une précision modérée. *Extension de gilbert\_1* semble légèrement meilleure, tandis que l'équation *extension de Gilbert\_3* présente des performances inférieures avec des erreurs importantes dans les prédictions. Les graphiques permettent de visualiser la dispersion des points par rapport à la ligne de régression et d'évaluer la précision des modèles.



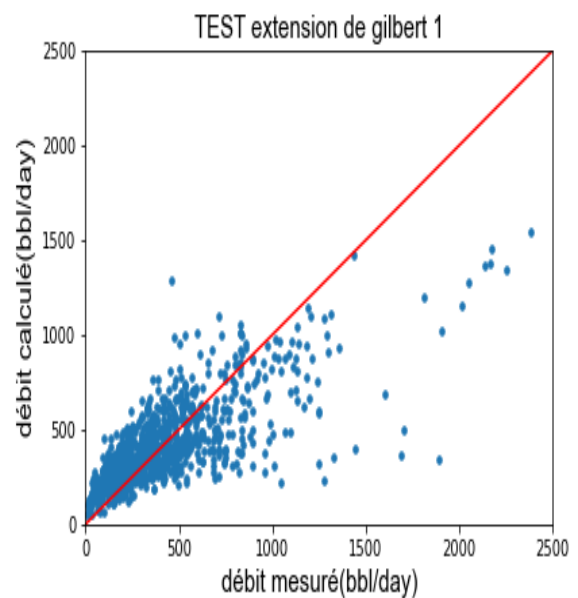
**Figure 3- 7 cross plot de TRAINING du modèle de Elgibaly et Nashawi**



**Figure 3- 8 cross plot de TEST du modèle de Elgibaly et Nashawi**

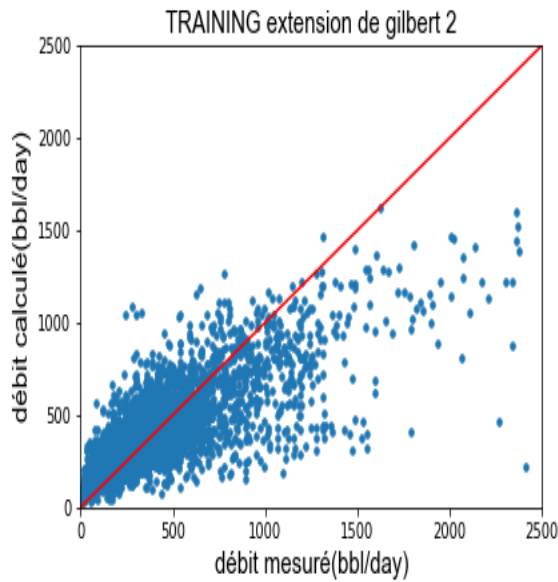


**Figure 3- 9 cross plot de TRAINING du modèle de extension Gilbert 1**

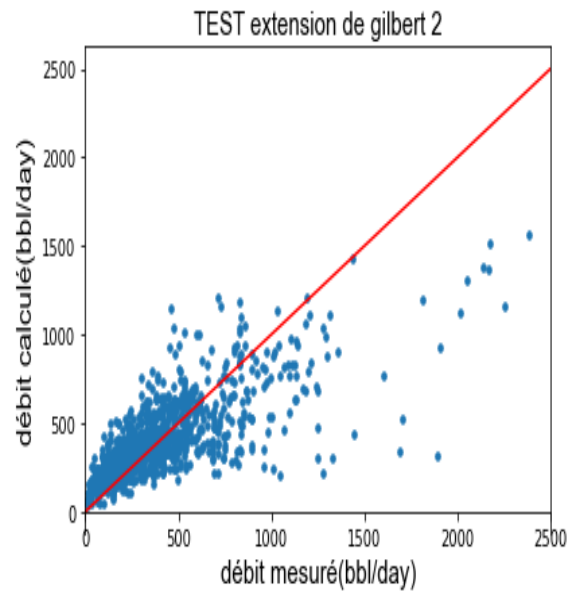


**Figure 3- 10 cross plot de TEST du modèle de extension Gilbert 1**

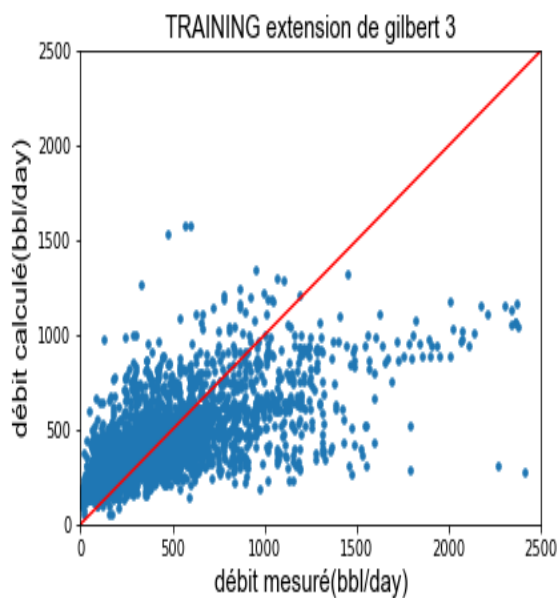




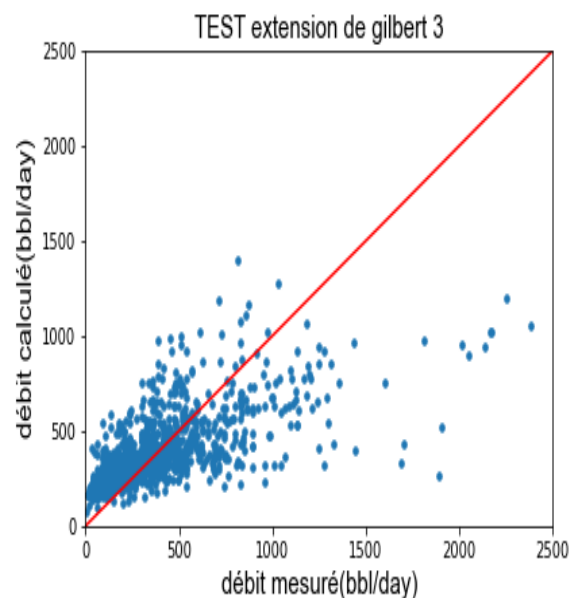
**Figure 3- 11 cross plot de TRAINING du modèle de extension Gilbert 2**



**Figure 3- 12 cross plot de TEST du modèle de extension Gilbert 2**



**Figure 3- 13 cross plot de TRAINING du modèle de extension Gilbert 3**



**Figure 3- 14 cross plot de TEST du modèle de extension Gilbert 3**

### 3.3.5 Conclusion

D'après les résultats obtenus dans cette partie, les conclusions suivantes peuvent être tirées :

- Pour le régime critique ; et en comparant les corrélations théoriques adaptés, il paraît que la corrélation de *Gilbert* arrive à donner une estimation acceptable du débit de production d'huile au champ de *HMD* avec des valeurs de  $R^2$  relativement élevées et des erreurs de prédiction relativement faibles. En revanche, l'équation de

*Towailib & Marhoun* ne semble pas bien fonctionner pour expliquer les données et produire des prédictions précises.

- Pour le régime non critique ; De façon générale, les modèles théoriques adaptés pour le régime non critique ont montré de bonnes performances, plus particulièrement le modèle *Extension Gilbert\_1* a enregistré relativement le minimum d'erreurs. L'équation *Extension Gilbert\_3* présente des performances légèrement inférieures.
- En général, les équations de Gilbert pour le régime critique, et les équations *extension de gilbert \_1* pour le régime non critique semblent fournir de meilleures estimations du débit d'huile

### **3.4 Machine Learning (L'apprentissage automatique)**

« Les méthodes d'apprentissage automatique fournissent une relation entre les variables d'entrée et prédisent la sortie. En apprentissage automatique, le comportement physique du système n'est pas perturbé. Les données associées aux industries du pétrole et du gaz sont énormes et le processus de corrélation des données est très complexe. » [9]

Dans les industries pétrolière et gazière, l'utilisation de l'apprentissage automatique est devenue important pour analyser et exploiter efficacement les données collectées à la fois en surface et en subsurface. Cette approche permet d'extraire des informations précieuses à partir de ces données sans perturber le comportement physique du système. Elle offre ainsi la possibilité d'explorer en profondeur les données, de détecter des modèles et des relations cachées, et de prédire des résultats en fonction des variables d'entrée.

Le machine Learning (ML) est une forme d'intelligence artificielle (IA) qui est axée sur la création de systèmes qui apprennent, ou améliorent leurs performances, en fonction des données qu'ils traitent. L'intelligence artificielle est un terme large qui désigne des systèmes ou des machines simulant une forme d'intelligence humaine[20]. ML est un ensemble d'algorithmes et de techniques utilisés pour créer des systèmes informatiques qui apprennent des données afin de faire des prédictions et des inférences [14]. L'apprentissage automatique se divise généralement en trois grandes catégories L'apprentissage supervisé, L'apprentissage non supervisé et L'apprentissage par renforcement

L'apprentissage automatique présente de nombreux avantages dans le domaine pétrolier, il permet du prédire la production de pétrole et de gaz, ce qui aide les entreprises à planifier et à optimiser leurs opérations. Egalement elle est utilisé pour la détection des anomalies dans les équipements, ce qui permet de prévenir les défaillances et d'améliorer la maintenance préventive.

Dans notre chapitre, nous utilisons le Machine Learning pour étudier la sensibilité de chaque paramètre agissant et contrôlant l'évolution du débit des puits pétroliers. L'objectif est d'améliorer la précision des prédictions en comprenant comment chaque paramètre contribue à la variation du débit.

### 3.4.1 Les types de machine Learning :

Machine Learning peuvent être classées en trois groupes qui sont utilisés en fonction des caractéristiques des données et des objectifs de la tâche à accomplir.

La figure (3-15) représentant les différents types d'apprentissage automatique :

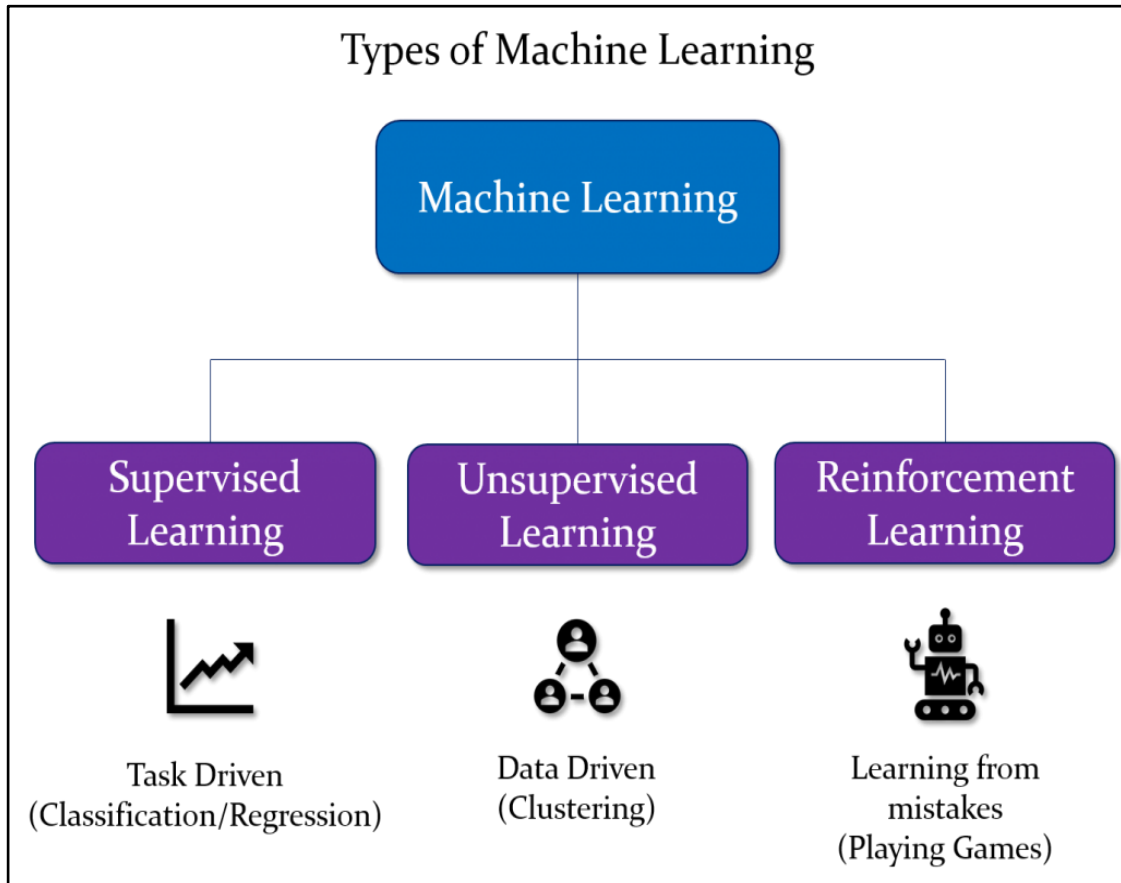


Figure 3- 15 Les types de machine Learning[26]

**L'apprentissage supervisé** : est une méthode d'apprentissage automatique qui implique la création d'un modèle prédictif en reliant des variables d'entrée ( $x$ ) à des variables de sortie ( $y$ ) à l'aide d'un algorithme. Cette technique est particulièrement utile pour résoudre des problèmes de classification et de régression en produisant des modèles prédictifs pertinents.

Il existe deux types d'algorithmes d'apprentissage supervisé.

- La classification : on cherche à classer un objet dans différentes classes, c'est-à-dire que l'on cherche à prédire la valeur d'une variable discrète (qui ne prend qu'un nombre fini de valeurs). Par exemple : prédire si une tumeur est maligne ( $y = 1$ ) ou bénigne ( $y = 0$ ) selon la taille de la tumeur ( $x_1$ ) et l'âge du patient ( $x_2$ ) ...

- La régression : Dans le cadre d'une analyse de régression, l'objectif est de prédire la valeur d'une variable continue qui peut prendre une gamme infinie de valeurs. Donc il est possible de créer un modèle qui permet de prédire la valeur de la variable dépendante (output) en fonction des variables indépendantes (input) Le modèle résultant peut-être utilisé pour effectuer des prévisions et des analyses prédictives dans divers contextes.

**L'apprentissage non supervisé** : est une technique d'apprentissage automatique utilisée lorsque seules les données d'entrée (x) sont disponibles, sans données de sortie correspondantes (y). L'objectif de l'algorithme dans ce cas est de découvrir de manière autonome des structures ou des associations pertinentes au sein des données

**L'apprentissage par renforcement** : est une méthode d'apprentissage automatique qui implique un programme informatique interagissant avec un environnement dynamique pour atteindre un objectif spécifique. Elle utilise des algorithmes qui explorent les différentes actions possibles et apprennent à prendre des décisions en fonction de leur efficacité pour atteindre l'objectif fixé. Cette méthode est très utile pour la résolution de problèmes complexes dans des environnements dynamiques, tels que la robotique, l'automatisation industrielle, et les jeux vidéo.

### 3.4.2 Réseau de neurones

Les réseaux de neurones sont des modèles de Machine Learning qui ont connu une forte croissance ces dernières années en raison de leur capacité à apprendre des représentations de données complexes et des fonctions mathématiques. Ils traitent les données d'une manière inspirée par le cerveau humain. Ils sont composés d'un grand nombre de neurones artificiels organisés en couches, chacune ayant des connexions avec les couches voisines.

#### 3.4.2.1 Architecture du réseau de neurones

Dans un réseau de neurones, chaque neurone prend une entrée, multiplie cette entrée par un poids spécifique, puis somme les produits pondérés pour obtenir une valeur. Cette valeur est ensuite transformée à l'aide d'une fonction d'activation et transmise aux neurones de la couche suivante. Ce processus se répète pour chaque couche du réseau

Dans un réseau de neurones, plusieurs couches de neurones peuvent être connectées les unes aux autres pour former une architecture complexe. Comme c'est schématisé dans la figure (3-16) Les données d'entrée (input layer) sont transmises à une couche cachée (hidden layer 1) où les calculs et les transformations ont lieu. Ensuite, les sorties de cette couche cachée sont modifiées par une autre couche cachée (hidden layer 2) avant de produire la sortie finale (output layer).

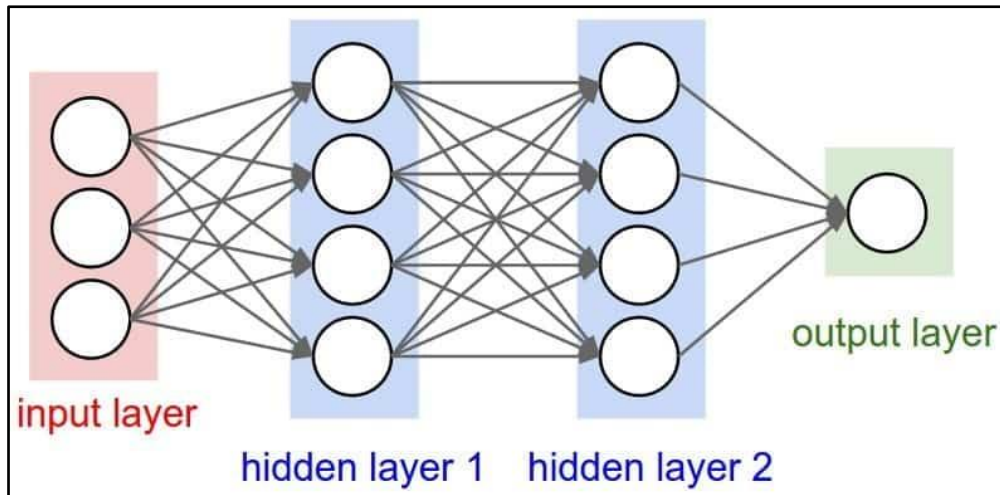


Figure 3- 16 The Neural Network Model

**Les couches de neurones :** font référence à la structure organisée des neurones dans le réseau. Chaque couche est composée de plusieurs neurones :

- Couche d'entrée (input layer) : cette couche contient des nœuds, elle sert d'intermédiaire pour transmettre les signaux d'entrée vers les nœuds qui suivent
- Les couches cachées : (hidden layer) Ces couches cachées reçoivent des entrées de la couche d'entrée ou d'autres couches cachées (En fonction de la complexité du problème à résoudre), Chaque couche cachée traite la sortie de la couche précédente, en la modifiant et en la transmettant à la couche suivante pour une analyse plus approfondie.
- La couche de sortie (output layer) est responsable de la production du résultat final du réseau neuronal artificiel. Elle peut être constituée d'un ou plusieurs nœuds de sortie.

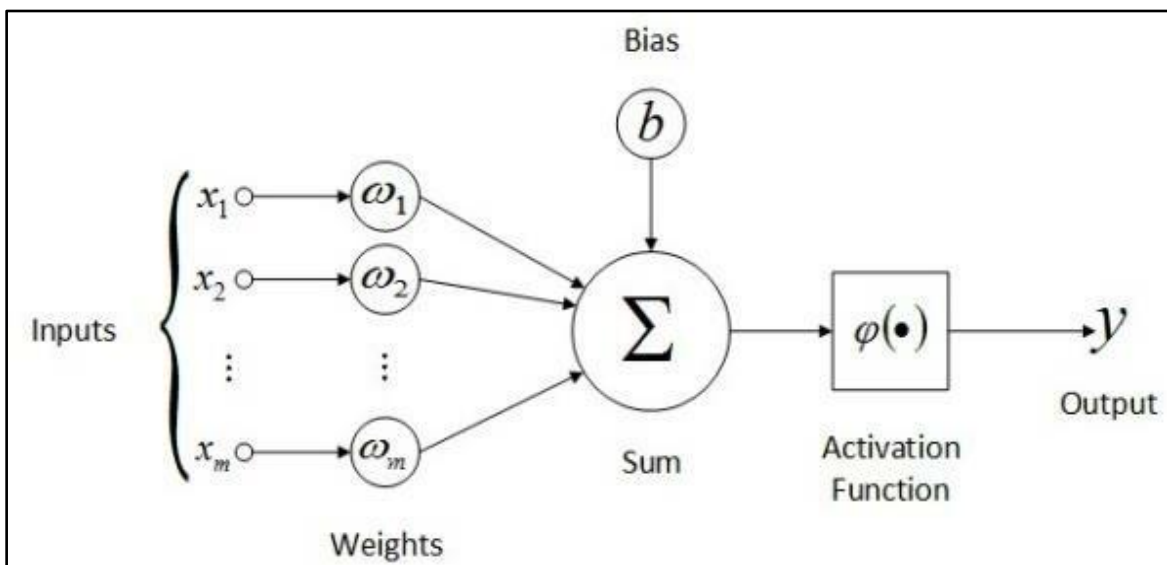


Figure 3- 17 Artificiel neurones [14]

**features (inputs)** : sont les caractéristiques d'entrée qui désignent les variables ou les informations qui sont fournies en entrée de modèle. Ces caractéristiques sont utilisées pour entraîner le modèle.

**Poids (weights)** représente les connexions entre un nœud et un autre. Ils sont modifiés itérativement à chaque passage de données d'entraînement dans le réseau, jusqu'à ce que l'erreur soit suffisamment faible ou que le nombre d'itérations atteigne un maximum prédéfini.

**La fonction d'activation** sont des fonctions mathématiques. Elles sont des éléments clés pour les réseaux de neurones. Elles déterminent la sortie d'un neurone en fonction de son entrée. Les fonctions d'activation peuvent être linéaires ou non linéaires, et elles sont souvent utilisées pour introduire de la non-linéarité dans le modèle de réseau neuronal. Elles ont un impact important sur la précision du modèle. Il existe plusieurs fonctions d'activation couramment utilisées dans les réseaux de neurones, telles que :

- La fonction d'activation linéaire est une fonction mathématique simple qui calcule une somme pondérée des entrées et ajoute un biais pour produire une sortie linéaire. Elle produit une valeur réelle entre les plages  $(-\infty, +\infty)$ . Elle est utile pour la régression. Généralement elle est utilisée comme nœud de sortie

$$z = x * w + b \quad (3.6)$$

$z, x, w$  et  $b$  Représentent respectivement : la sortie, l'entrée, le poids et le biais

- La fonction d'activation Sigmoidale, également connue sous le nom de fonction logistique, est utilisée pour produire une sortie dans la plage de 0 à 1 en fonction de l'entrée. Elle est donc très utilisée pour les classifications binaires

$$\sigma(z) = \frac{1}{1+e^{-z}} \quad (3.7)$$

- La fonction Tanh est une fonction non linéaire qui produit des valeurs dans la plage  $[-1, 1]$ . Elle est similaire à la fonction Sigmoidale mais elle est centrée sur zéro,

$$G(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \quad (3.8)$$

- La fonction d'activation ReLu (Rectified Linear Unit) est une fonction linéaire qui renvoie 0 pour toutes les valeurs d'entrée négatives et renvoie directement l'entrée pour toutes les valeurs d'entrée positives. Elle est principalement utilisée dans les couches cachées des réseaux de neurones. Cette fonction converge plus rapidement, optimise et produit la valeur souhaitée plus rapidement.

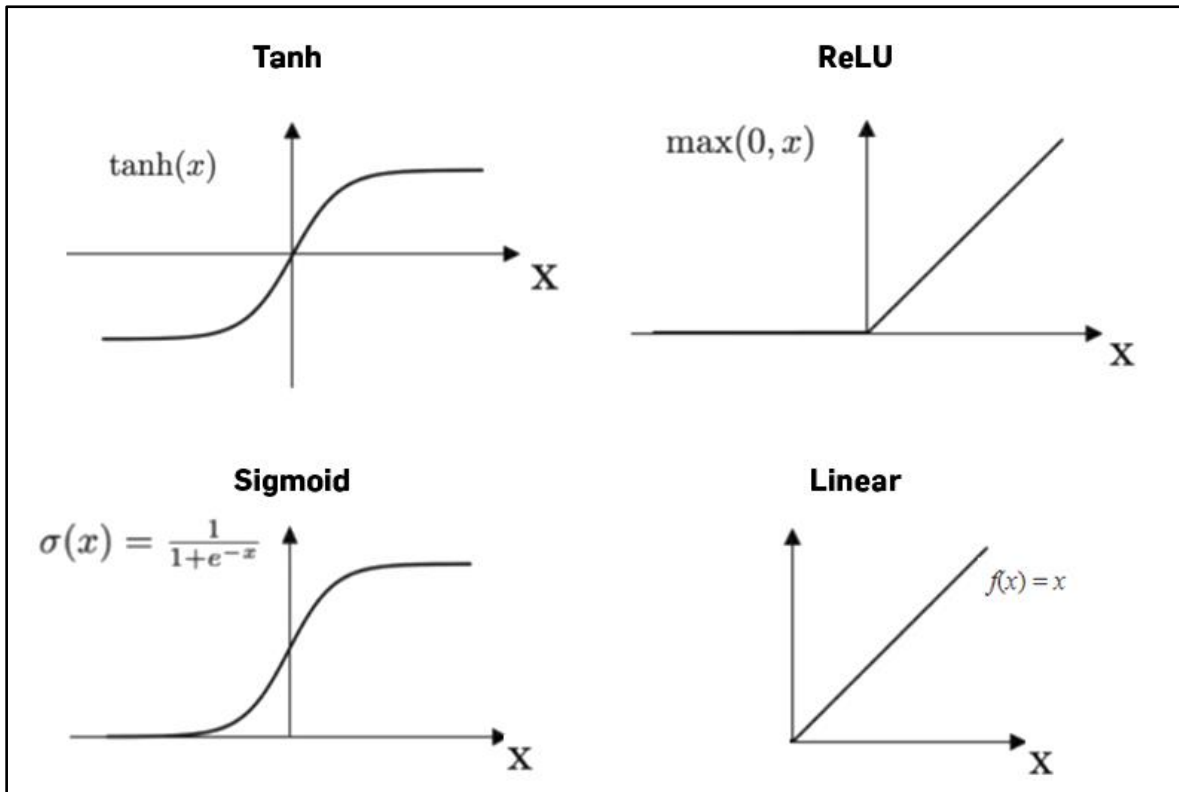


Figure 3- 18 fonctions d'activations [17]

### Fonctionnement des réseaux de neurones

Le concept de réseau de neurones repose sur plusieurs étapes qui sont représentées dans la figure :

- Initialiser le network : les valeurs des poids et des biais sont initialisées Pour chaque neurone dans une couche, multiplier la valeur d'entrée par le poids.
- Ensuite, pour chaque couche, additionner tous les poids des neurones et ajouter un biais.
- Appliquer la fonction d'activation sur cette valeur pour calculer une nouvelle sortie.
- Une fois la valeur de sortie est obtenue, la fonction de perte est calculée. (Si l'erreur est inférieure à un seuil prédéfini, le réseau est formé et l'algorithme se termine)
- Back propagation pour mis à jours les poids grâce aux algorithmes d'optimisation. Dans le but de réduire l'erreur

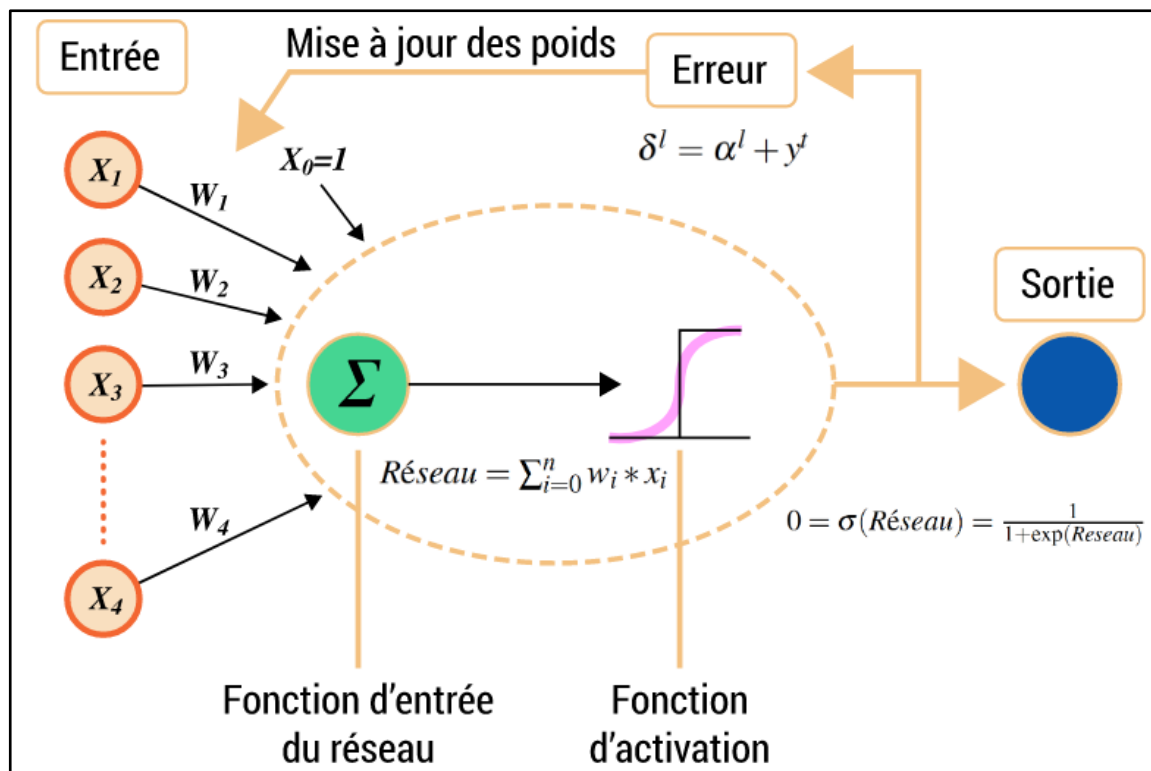


Figure 3- 19 fonctionnement des réseaux de neurones [17]

### 3.4.2.2 Random Forest

« La forêt aléatoire est un algorithme d'apprentissage supervisé. La "forêt" qu'elle construit est un ensemble d'arbres de décision, généralement entraînés avec la méthode du bagging. L'idée générale de la méthode du bagging est qu'une combinaison de modèles d'apprentissage augmente le résultat global. » [21]

Random Forest est un algorithme d'apprentissage automatique flexible et facile à utiliser qui produit un excellent résultat la plupart du temps. C'est également l'un des algorithmes les plus utilisés, en raison de sa simplicité et de sa polyvalence (il peut être utilisé à la fois pour des tâches de classification et de régression).

#### Architecture de random forest

L'architecture de la Random Forest est basée sur l'ensemble de plusieurs arbres de décision, qui sont combinés pour former un modèle prédictif plus puissant. Chaque arbre de décision dans la forêt est construit de manière indépendante, en utilisant un échantillon aléatoire des données d'entraînement

Les arbres de décision sont les composants de base d'un algorithme de random forest. Un arbre de décision est une technique d'aide à la décision qui forme une structure en forme d'arbre. Une compréhension des arbres de décision nous aidera à comprendre le fonctionnement des algorithmes de random forest.



Un arbre de décision se compose de trois composants : les nœuds de décision, les nœuds feuilles et un nœud racine. Un algorithme d'arbre de décision divise un ensemble de données d'entraînement en branches, qui se subdivisent ensuite en d'autres branches. Cette séquence se poursuit jusqu'à ce qu'un nœud feuille soit atteint. Le nœud feuille ne peut pas être subdivisé davantage.

- Nœuds de décision : Les nœuds de décision sont des points de bifurcation dans l'arbre où une condition ou une règle est évaluée pour déterminer quelle branche suivre. Chaque nœud de décision est associé à un attribut et à une valeur seuil correspondante. Lorsqu'un exemple atteint un nœud de décision, la condition est vérifiée et l'exemple est dirigé vers la branche appropriée en fonction du résultat de la condition.
- Nœuds de feuille : Les nœuds de feuille sont les extrémités de l'arbre où aucune autre bifurcation n'est possible. Ils représentent les résultats finaux ou les prédictions du modèle. Chaque nœud de feuille est associé à une classe (dans le cas de la classification) ou à une valeur (dans le cas de la régression). Lorsqu'un exemple atteint un nœud de feuille, la classe ou la valeur correspondante est attribuée à cet exemple.
- Nœud racine : Le nœud racine est le point de départ de l'arbre et se situe au sommet de la hiérarchie. Il représente la première condition ou règle qui est évaluée pour diviser les exemples en sous-groupes. À partir du nœud racine, les exemples sont dirigés vers les nœuds de décision appropriés en fonction des résultats des tests de condition

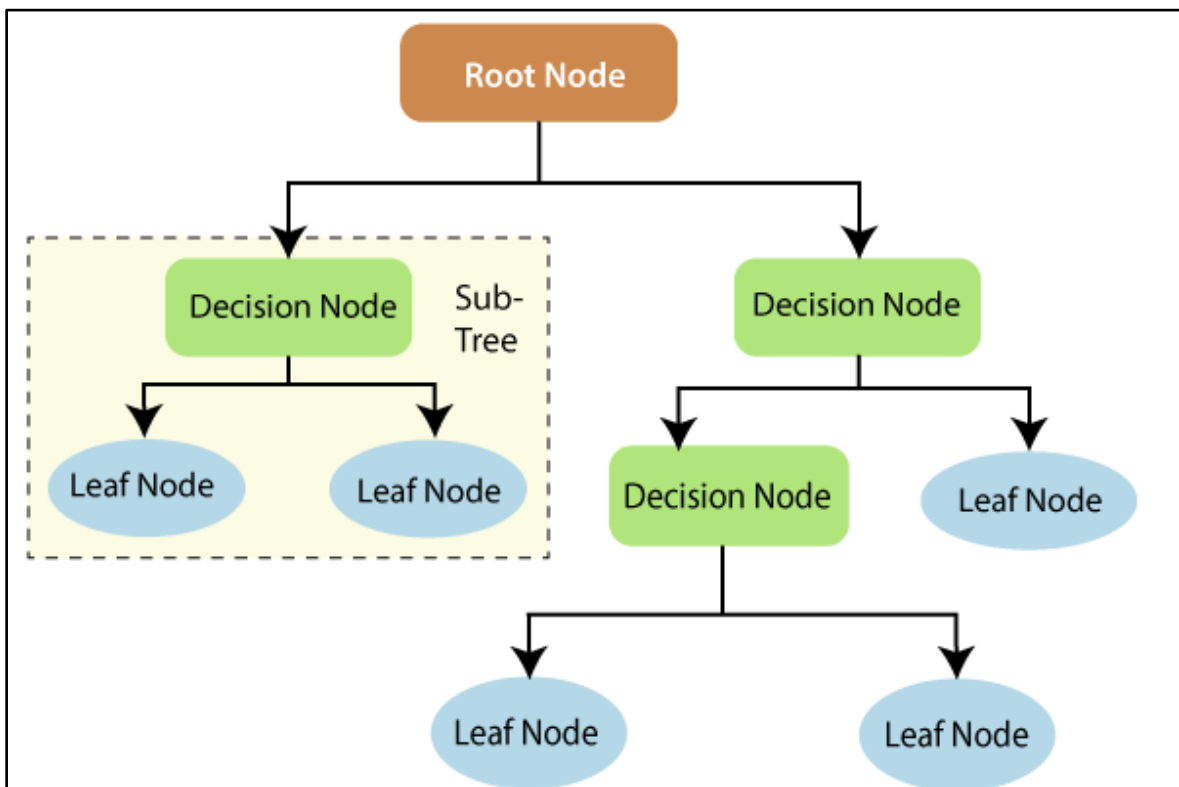


Figure 3- 20 les Le nœud de l'arbre de décision[27]

## **Hyperparamètres de la Random Forest**

Les Hyperparamètres dans la random forest sont utilisés pour augmenter la puissance prédictive du modèle ou pour le rendre plus rapide.

### **Pour Augmenter la puissance prédictive :**

- `n_estimators`, qui correspond simplement au nombre d'arbres que l'algorithme construit avant de prendre la décision finale par vote majoritaire ou en faisant la moyenne des prédictions. En général, un nombre plus élevé d'arbres améliore les performances et rend les prédictions plus stables, mais cela ralentit également le calcul.
- `max_features`, qui correspond au nombre maximum de caractéristiques prises en compte par la random forest pour diviser un nœud. Scikit-learn propose plusieurs options, toutes décrites dans la documentation.
- `min_samples_leaf`, qui détermine le nombre minimum de feuilles requis pour diviser un nœud interne.

En utilisant ces différents Hyperparamètres, il est possible de régler et d'optimiser la random forest en fonction des besoins spécifiques du problème à résoudre.

## **Fonctionnement de random forest**

Le concept de réseau de neurones repose sur plusieurs étapes qui sont représentées dans la figure (3-21) :

- Sélection aléatoire d'échantillons : À partir de l'ensemble de données d'entraînement, des échantillons sont sélectionnés aléatoirement avec remplacement pour former un échantillon bootstrap. Cela crée plusieurs ensembles de données d'entraînement différents pour chaque arbre de décision.
- Construction des arbres de décision : Pour chaque ensemble de données d'entraînement, un arbre de décision est construit en utilisant des règles de partitionnement basées sur les caractéristiques et les valeurs cibles. Les arbres de décision sont construits de manière récursive en divisant l'ensemble de données en sous-ensembles jusqu'à ce que les feuilles de l'arbre représentent les classes finales ou les valeurs de régression.
- Prédiction par vote majoritaire : Une fois que tous les arbres de décision sont construits, la prédiction est effectuée en utilisant un vote majoritaire pour les problèmes de classification. Chaque arbre de décision vote pour une classe et la classe ayant le plus grand nombre de votes est choisie comme prédiction finale. Pour les problèmes de régression, la prédiction est calculée en prenant la moyenne des valeurs prédites de tous les arbres de décision.
- Gestion de l'overfitting : La random forest gère efficacement le problème de surajustement (overfitting) en utilisant la technique du bagging. En construisant

plusieurs arbres de décision indépendants sur des échantillons aléatoires, elle réduit la corrélation entre les arbres et améliore la généralisation des prédictions.

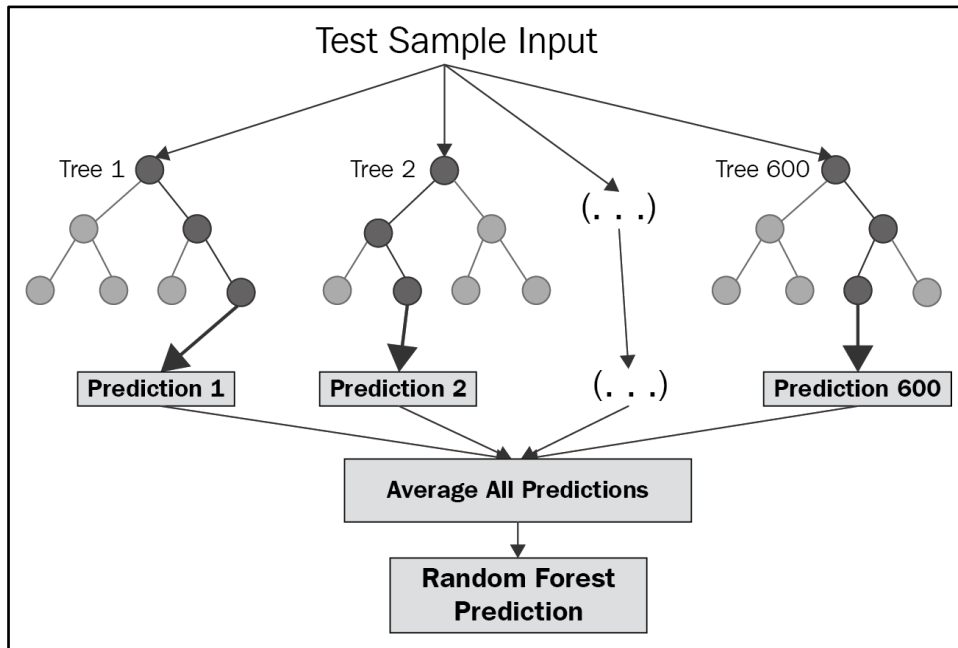


Figure 3- 21 le Fonctionnement de random forest [22]

### 3.4.2.3 Réseau de neurones récurrent (RNN)

Un réseau de neurones récurrents (RNN) est un type de réseau de neurones artificiel utilisé pour traiter des données séquentielles ou des données de séries temporelles. Contrairement aux réseaux de neurones classiques qui traitent chaque entrée indépendamment les uns des autres

Les RNN possèdent une mémoire interne qui leur permet de traiter les entrées successives en utilisant l'information de la ou des entrées précédentes pour influencer l'entrée et la sortie actuelles.

#### Long Short Term Memory (LSTM)

« La cellule de mémoire à long terme (LSTM), a été proposée en 1997 par Sepp Hochreiter et Jürgen Schmidhuber, et elle a été progressivement améliorée au fil des années par plusieurs chercheurs, tels qu'Alex Graves, Haşim Sak, Wojciech Zaremba, et bien d'autres encore ». [18]

LSTM est une architecture de réseau neuronal récurrent (RNN). Il a une "mémoire à long terme" qui leur permet de retenir des informations importantes sur les entrées précédentes pendant une période. Les LSTM ont des cellules de mémoire qui leur permettent de stocker des informations pendant une longue période de temps Ces cellules sont contrôlées par des portes (gates) qui régulent la quantité d'information qui est stockée et oubliée

Comme représente la figure (3-22) le LSTM contiennent plusieurs portes pour contrôler le flux d'informations dans la cellule, une porte d'oubli, une porte d'entrée et une porte de sortie.

Avec  $x, c, h$  représentent respectivement: input(entrée) , mémoire à long terme (memory cell) ,mémoire à court terme (hidden state)

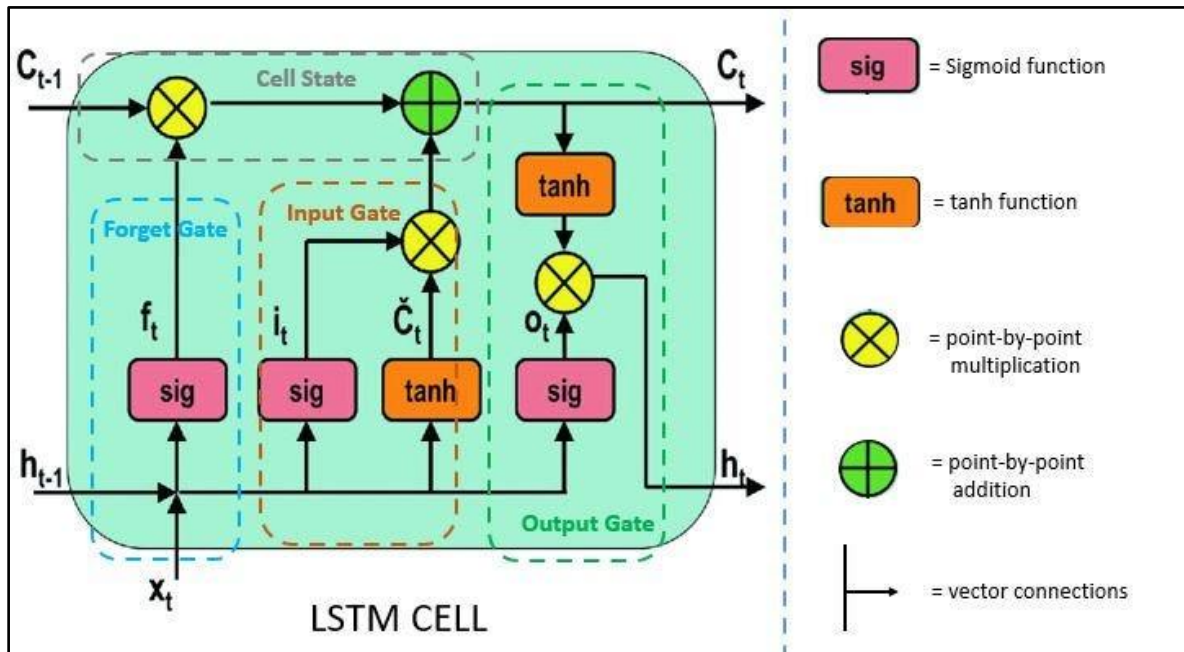


Figure 3- 22 LSTM modèle [19]

- La porte d'oubli (forget gate) : elle permet de contrôler l'oubli de l'information stockée dans la cellule de mémoire et permet de stocker des informations à long terme, à l'aide de la fonction d'activation sigmoïd. La porte d'oubli calcule une valeur entre 0 et 1 qui est utilisée pour décider quelles informations doivent être conservées ou oubliées. Si la valeur est proche de 0, cela signifie que l'information doit être oubliée, tandis que si elle est proche de 1, l'information est conservée.
- La porte d'entrée (Input gate) : elle décide quelles nouvelles informations seront stockées dans la mémoire à long terme. L'entrée actuelle et les informations sur l'état précédent passent par la fonction sigmoïde et par la modification et tanh pour la régulation du réseau, la sortie de sigmoïde se multiplie par la sortie de tanh. Dans le but de la sortie de sigmoïde identifiera les informations précieuses à retenir de la sortie de tanh.
- La porte de sortie (output gate) : c'est la troisième porte. Elle prend en compte l'entrée actuelle, la mémoire à court terme précédente et la mémoire à long terme nouvellement calculée. On multiplie la sortie sigmoïde par la sortie tanh. Pour décider quelle information doit être sortie de la cellule et pour produire la nouvelle mémoire à court terme qui sera transmise à la cellule au prochain pas de temps.

Le tableau (3-10) ci-dessous représente les formules clés des composants LSTM.

**Tableau 3- 10 Les composants principaux d'un LSTM[14]**

LSTM Component	Formula
Input gate layer: This decides which values to store in the cell state.	$i_t = \text{sigmoid}(w_i x_t + u_i h_{t-1} + b_i)$
Forget gate layer: As the name suggested this decides what information to throw away from the cell state.	$f_t = \text{sigmoid}(W_f x_t + U_f h_{t-1} + b_f)$
Output gate layer: Create a vector of values that can be added to the cell state.	$O_t = \text{sigmoid}(W_o x_t + u_o h_{t-1} + b_o)$
Memory cell state vector.	$c_t = f_t \circ c_{t-1} + i_t \circ * \text{hyperbolic tangent}(W_c x_t + u_c h_{t-1} + b_c)$

### 3.5 Résultat et discussion pour l'étude de sensibilités par les modèles intelligents

#### 3.5.1 Méthodologie

Les corrélations de littérature utilisées dans le chapitre précédent pour estimer le débit dans le régime critique et sous-critique sur le champ HMD ont donné des résultats généralement acceptables, mais parfois moyens. Pour améliorer la précision de ces estimations, en explorant différentes approches :

Dans la première partie on a entrepris une étude de sensibilité afin d'évaluer l'influence des paramètres tels que : la pression de tête  $P_t$ , le diamètre de la Duse  $d$ , le rapport gaz/huile produit  $GOR$ , la pression en aval de la Duse  $P_p$ , la chute de pression à travers la Duse  $\Delta p$  et le rapport de pression critique  $r = (P_p/P_t)$  pour le régime critique et le régime sous critique sur l'ensemble du champ pétrolier. On a examiné ces paramètres dans leur globalité pour comprendre leur impact global sur le débit.

Dans la deuxième partie, on a divisé le champ en différents intervalles en fonction du  $GOR$ , puis réadapter le meilleur modèle théorique retrouvé précédemment selon chaque intervalle. Ceci sera réalisé pour les deux régimes d'écoulement critique et sous critique Cette approche on a permis d'étudier séparément chaque intervalle de  $GOR$  afin d'améliorer la capacité prédictive du ces modèle en tenant compte des variations du  $GOR$  et d'obtenir des prédictions plus précises du débit de production dans différents scénarios de régime d'écoulement.

Dans la troisième partie, on a analysé chaque puits individuellement, en tenant compte de leur indépendance par rapport aux autres. Cette approche permettra de prendre en compte les caractéristiques spécifiques de chaque puits et d'optimiser les modèles en fonction de leurs comportements uniques.

Pour mener cette étude, on a utilisé différentes techniques d'apprentissage automatique telles que les réseaux de neurones artificiels  $ANN$ , la forêt aléatoire (random forest) et le modèle de mémoire à long terme  $LSTM$ . Ces méthodes d'apprentissage automatique on ont permis

de prendre en compte les relations complexes et les interactions non linéaires entre les paramètres et le débit.

On a évalué les performances des modèles à l'aide de mesures telles que : *RMSE* (Root Mean Square Error) ,  $R^2$  (Coefficient de détermination), *RAE* (Relative Absolute Error), *MAE* (Mean Absolute Error). Ces mesures nous ont permis de quantifier l'efficacité des modèles et de déterminer l'importance de chaque paramètre dans la prédiction du débit.

L'objectif ultime de cette étude est d'améliorer notre compréhension des relations entre les paramètres  $P_t, d, GOR, P_p, \Delta P$  et  $r$  et le débit dans le champ *HMD*. En utilisant des méthodes avancées d'apprentissage automatique, notre objectif est de fournir des estimations plus précises du débit et d'identifier les paramètres les plus influents dans ce processus.

### 3.5.2 Résultats

Dans cette section, on présentera les résultats obtenus à la suite de l'étude de sensibilité menée pour le régime critique et le régime sous-critique, en utilisant les mêmes données que celles présentées dans le chapitre précédent.

#### 3.5.2.1 Analyse de sensibilité globale des inputs sur le champ complet

Pour réaliser cette étude de sensibilité, on utilise les techniques d'apprentissage automatique suivantes les réseaux de neurones artificiels *ANN*, la forêt aléatoire (random forest)

##### ❖ Régime critique

Les tableaux(3-11) et(3-12) représentent respectivement les résultats de training et de test des modèles de random forest et de *ANN*.

**Tableau 3- 11 les résultats de training et de test des modèles de random forest.**

INPUTS	train data				test data			
	$R^2$	MAE	RMSE	RAE	$R^2$	MAE	RMSE	RAE
PP.PT. $\emptyset$	0.88	106.5	152.06	0.37	0.25	273.84	387.83	0.91
PP.PT. $\emptyset$ .GOR	0.96	50.87	87.29	0.1	0.71	134.86	233.74	0.28
PT. $\emptyset$ .GOR	0.95	53.46	91.44	0.11	0.7	139.92	242.31	0.28
PP.PT. $\emptyset$ .GOR. $\Delta P$	0.96	50.7	87.1	0.1	0.73	133.63	232.3	0.8
PP.PT $\emptyset$ .GOR. $\Delta P$ . $r$	0.96	50.53	86.77	0.1	0.73	133.56	231.1	0.28
$r$ . $\emptyset$ .GOR	0.93	72.46	112.14	0.16	0.55	188.32	298.34	0.44
$(PT^2-PP^2)^{0.5}$ . $\emptyset$ .GOR	0.96	54.54	92.86	0.11	0.69	142.78	245.72	0.29
PT*PP. $\emptyset$ .GOR	0.95	59.18	97.88	0.12	0.64	159.7	267.17	0.32
PP.PT. $\emptyset$ . $\Delta P$ . $r$	0.88	106.44	151.42	0.38	0.25	274.55	386.51	0.25

**Tableau 3- 12 les résultats de training et de test des modèles de ANN.**

INPUTS	train data				test data			
	R <sup>2</sup>	MAE	RMSE	RAE	R <sup>2</sup>	MAE	RMSE	RAE
PP.PT. $\emptyset$	0.15	304.39	412.72	1.11	0.16	298.19	410.74	1.06
PP.PT. $\emptyset$ .GOR	0.67	148.54	255.42	0.31	0.65	146.87	262.65	0.3
PT. $\emptyset$ .GOR	0.67	147.39	255.07	0.31	0.66	145.48	260.39	0.3
PP.PT. $\emptyset$ .GOR. $\Delta P$	0.67	150.07	255.96	0.3	0.66	149.02	261.85	0.3
PP.PT. $\emptyset$ .GOR. $\Delta P$ .r	0.67	150.67	254.85	0.34	0.66	149.69	262.19	0.32
r. $\emptyset$ .GOR	0.54	139.64	221.86	0.54	0.53	153.15	245.14	0.53
(PT <sup>2</sup> -PP <sup>2</sup> ) <sup>0.5</sup> . $\emptyset$ .GOR	0.63	159.43	269.87	0.34	0.62	0.62	273.47	0.32
PT*PP. $\emptyset$ .GOR	0.62	168.28	272.15	0.37	0.6	169.45	280.3	0.35
PP.PT. $\emptyset$ . $\Delta P$ .r	0.19	291.52	400.08	1.03	0.2	286.72	398.55	0.97

### Interprétation des résultats

- D'après les résultats du test les deux modèles Random Forest et ANN présentent la même sensibilité aux mêmes entrées cela suggère que les variables d'entrées ont un impact similaire sur les prédictions des deux modèles.
- les paramètres  $PT$ ,  $\emptyset$ ,  $GOR$  semblent avoir une plus grande influence sur les performances des modèles, tandis que l'ajout des paramètres  $\Delta P$  et  $r$  n'améliore pas significativement les résultats.

### ❖ Régime sous critique

Les tableaux (3-13) et (3-14) représentent respectivement les résultats de training et de test des modèles de random forest et de ANN.

**Tableau 3- 13 les résultats de training et de test des modèles de random forest.**

input	train data				test data			
	R <sup>2</sup>	MAE	RMSE	RAE	R <sup>2</sup>	MAE	RMSE	RAE
PP.PT. $\emptyset$	0.89	71.57	104.82	0.41	0.38	191.8	282.75	0.96
PP.PT. $\emptyset$ .GOR	0.95	40.01	67.33	0.14	0.69	114.28	198.6	0.36
PT. $\emptyset$ .GOR. $\Delta P$	0.95	39.71	67.44	0.14	0.7	113.37	195.91	0.32
PT. $\emptyset$ .GOR	0.95	41.23	68.32	0.14	0.68	119.03	202.19	0.38
PP.PT. $\emptyset$ .GOR. $\Delta P$	0.95	39.54	67.19	0.14	0.7	113.15	196.39	0.36
PP.PT. $\emptyset$ .GOR. $\Delta P$ .r	0.95	39.63	67.28	0.14	0.69	114.48	198.32	0.37
r. $\emptyset$ .GOR	0.93	49.74	80.46	0.18	0.58	144.53	232.65	0.48
(PT <sup>2</sup> -PP <sup>2</sup> ) <sup>0.5</sup> . $\emptyset$ .GOR	0.95	40.5	68.78	0.14	0.7	115.04	196.24	0.37
PT*PP. $\emptyset$ .GOR	0.95	41.7	68.58	0.15	0.68	120.62	203.38	0.39
PP.PT. $\emptyset$ . $\Delta P$ .r	0.89	71.4	104.9	0.41	0.4	189.98	277.46	0.95

Tableau 3- 14 les résultats de training et de test des modèles de ANN.

input	train data				test data			
	R <sup>2</sup>	MAE	RMSE	RAE	R <sup>2</sup>	MAE	RMSE	RAE
PP.PT. $\emptyset$	0.33	189.68	267.05	1.13	0.36	198.26	288.22	1.01
PP.PT. $\emptyset$ .GOR	0.67	112.32	185.82	0.43	0.68	117.59	202.58	0.37
PT. $\emptyset$ .GOR. $\Delta P$	0.67	112.33	185.66	0.39	0.68	117.60	202.59	0.33
PT. $\emptyset$ .GOR	0.65	118.66	191.56	0.48	0.66	123.6	208.81	0.4
PP.PT. $\emptyset$ .GOR. $\Delta P$	0.67	117.09	188.24	0.42	0.67	123.11	206.24	0.39
PP.PT. $\emptyset$ .GOR. $\Delta P$ .r	0.68	113.15	184.69	0.48	0.68	119.15	202.85	0.4
r. $\emptyset$ .GOR	0.54	139.64	221.86	0.51	0.54	153.15	245.14	0.49
(PT <sup>2</sup> -PP <sup>2</sup> ) <sup>0.5</sup> . $\emptyset$ .GOR	0.67	117.52	189.4	0.46	0.67	124.3	207.03	0.41
PT*PP. $\emptyset$ .GOR	0.64	124.1	195.27	0.48	0.64	132.07	214.98	0.65
PP.PT. $\emptyset$ . $\Delta P$ .r	0.33	188.99	266.96	1.11	0.36	197.55	288.12	1

### Interprétation des résultats

- ❖ D'après les résultats du test les deux modèles Random Forest et ANN présentent la même sensibilité aux mêmes entrées cela suggère que les variables d'entrées ont un impact similaire sur les prédictions des deux modèles.
- ❖ les paramètres  $PT$ ,  $\emptyset$ ,  $GOR$ ,  $\Delta P$  semblent avoir une plus grande influence sur les performances du modèle ANN.

#### 3.5.2.2 Etude de sensibilité suivant les intervalles de GOR

Dans le but d'améliorer les performances des modèles théoriques adaptés, une approche de subdivision en intervalles de  $GOR$  a été utilisée. Pour chaque régime d'écoulement (critique et sous-critique), le meilleur modèle théorique adapté a été sélectionné : le modèle de Gilbert adapté a été utilisé pour le régime critique, tandis que l'extension de Gilbert 1 a été choisie pour le régime non critique.

L'objectif de la subdivision en intervalles de  $GOR$  est de prendre en compte les variations des inputs en fonction du rapport gaz/huile  $GOR$ .

Pour chaque intervalle de  $GOR$ , le modèle adapté correspondant a été réadapté afin d'améliorer ses performances. Ce processus vise à ajuster le modèle en fonction des données spécifiques à chaque intervalle de  $GOR$ , dans le but d'optimiser sa précision et sa fiabilité.

#### ❖ Régime critique

Les tableaux (3-15) et (3-16) représentent respectivement les nouveaux coefficients obtenus ainsi que les résultats de training et de test des nouveaux modèles et le diagramme cyclique (3-23) suivant représente la variation de coefficient de détermination en fonction des intervalles de  $GOR$



Tableau 3- 15 les nouveaux coefficients obtenus.

Intervalles de GOR(scf/STB )	Nouveaux coefficients de l'équation de Gilbert adapté pour chaque GOR			
	<i>a</i>	<i>b</i>	<i>c</i>	<i>e</i>
<b>GOR</b> ≤4000	5,87	1,33	0,38	0,76
4000 < <b>GOR</b> ≤ 15000	9,6	1,53	0,7	0,91
15000 < <b>GOR</b> ≤ 35000	14,52	1,77	0,81	0,91
35000 < <b>GOR</b> ≤ 60000	129,38	1,78	0,98	0,84
60000 < <b>GOR</b> ≤ 100000	0,008	1,156	0,17	1,05
<b>GOR</b> >100000	1,01	1,28	0,52	1,01

Tableau 3- 16 les résultats de training et de test des nouveaux modèles.

Intervalles de GOR(scf/STB)	train data				test data			
	R <sup>2</sup>	MAE	RMSE	RAE	R <sup>2</sup>	MAE	RMSE	RAE
<b>GOR</b> ≤4000	0,48	225,48	352,47	0,36	0,42	225,18	351,21	0,42
4000 < <b>GOR</b> ≤ 15000	0,73	106,93	179,44	0,27	0,74	103,67	173,39	0,27
15000 < <b>GOR</b> ≤ 35000	0,8	66,09	119,51	0,26	0,83	67,26	107,081	0,33
35000 < <b>GOR</b> ≤ 60000	0,8	56,53	88,09	0,4	0,9	46,1	65,82	0,26
60000 < <b>GOR</b> ≤ 100000	0,83	22,31	32,81	0,32	0,72	29,61	50,12	0,21
<b>GOR</b> >100000	0,63	22,51	35,3	1,08	0,57	20,48	26,25	0,84

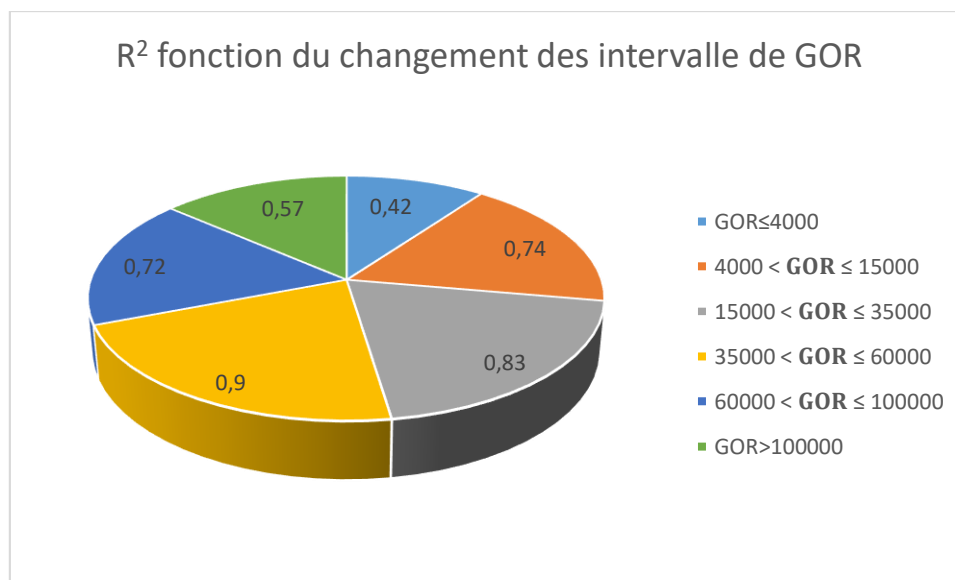


Figure 3- 23 R2 en fonction de variation des intervalle de GOR pour le régime critique

## Interprétation des résultats

- ❖ Selon les résultats, les nouveaux modèles adaptés sur l'ensemble des intervalles fournissent une estimation satisfaisante du débit de production, avec des coefficients de détermination se situant entre 0.72 et 0.9.
- ❖ Sauf les valeurs du *GOR* inférieures à 4000 scf/STB et supérieures de 100000, les erreurs commencent à devenir relativement importantes, comme c'est montré dans le tableau (3-16) où le coefficient de détermination de premier intervalle est 0.42 et l'erreur relatif dans le dernier intervalle 0.84
- ❖ L'analyse des valeurs des coefficients *a*, *b*, *c* et *e*, il est observé que le coefficient *a* et *c* ne présentent pas de tendance claire et qu'il prend différentes valeurs en fonction de l'intervalle de *GOR* étudié. Le coefficient *b* est compris principalement entre 1.156 et 1.78 sur tous les intervalles. Quant au coefficient *e* il demeure relativement stable avec des valeurs comprises entre 0.84 et 1.05 à l'exception du premier intervalle où il prend une valeur de 0.76.

### ❖ Régime sous critique

Les tableaux (3-17) et (3-18) représentent respectivement les nouveaux coefficients obtenus ainsi que les résultats de training et de test des nouveaux modèles et le diagramme cyclique (3-24) suivant représente la variation de coefficient de détermination en fonction des intervalles de *GOR*

**Tableau 3- 17 les nouveaux coefficients obtenus**

Intervalles de <i>GOR</i> (scf/STB)	Nouveaux coefficients de l'extension de Gilbert 1 adapté pour chaque <i>GOR</i>				
	<i>a</i>	<i>b</i>	<i>c</i>	<i>e</i>	<i>f</i>
<b>GOR</b> ≤1000	8,69	0,13	0,32	0,91	0,04
1000 < <b>GOR</b> ≤ 10000	3,33	0,79	0,19	1,48	0,58
10000 < <b>GOR</b> ≤ 30000	17,89	0,63	0,27	1,66	0,76
<b>GOR</b> >30000	40,51	0,69	0,32	1,9	0,97

**Tableau 3- 18 les résultats de training et de test des nouveaux modèles**

Intervalles de <i>GOR</i> (scf/STB )	train data				test data			
	R <sup>2</sup>	MAE	RMSE	RAE	R <sup>2</sup>	MAE	RMSE	RAE
<b>GOR</b> ≤1000	0,28	143,39	215,07	0,35	0,29	119,08	167,32	0,29
1000 < <b>GOR</b> ≤ 10000	0,7	131,93	211,07	0,39	0,67	126,97	223,33	0,36
10000 < <b>GOR</b> ≤ 30000	0,8	65,54	108,03	0,3	0,75	62,28	113,23	0,29
<b>GOR</b> >30000	0,76	22,04	34,46	0,33	0,78	19,21	29	0,32

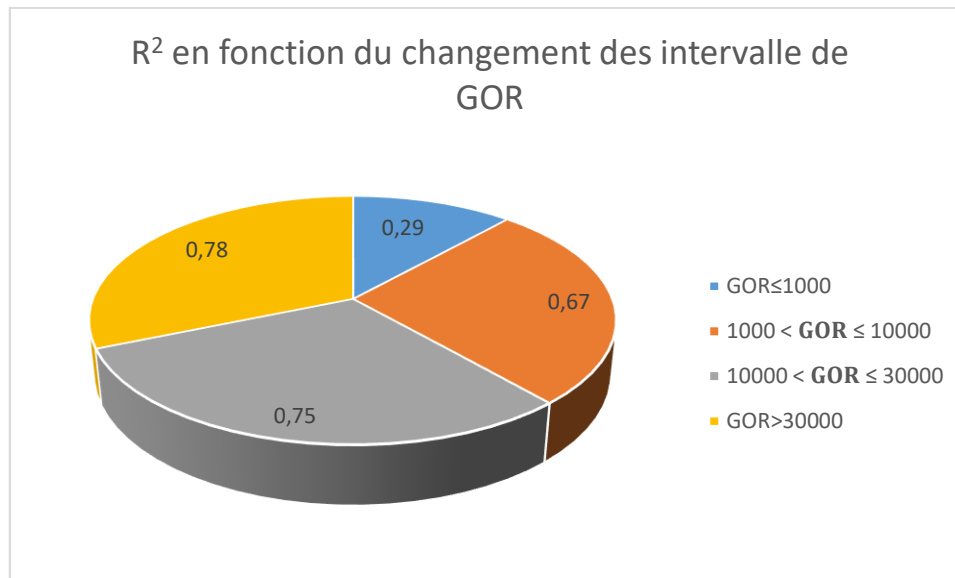


Figure 3- 24  $R^2$  en fonction du changement des intervalle de GOR pour le régime sous critique

### Interprétation des résultats

- ❖ Selon les résultats, les nouveaux modèles adaptés sur l'ensemble des intervalles fournissent une estimation satisfaisante du débit de production, avec des coefficients de détermination se situant entre 0.67 et 0.78
- ❖ À l'exception du premier intervalle où les valeurs du  $GOR$  sont inférieurs à 1000 scf /STB, le coefficient de détermination est 0.29
- ❖ D'après L'analyse des valeurs des coefficients  $a, b, c, e$  et  $f$ , il est observé que le coefficient  $a$  ne présente pas de tendance claire et qu'il prend différentes valeurs en fonction de l'intervalle de  $GOR$  étudié. Les coefficient  $b, e$  et  $f$  sont compris principalement des valeurs stables sur tous les intervalles sauf le premier intervalle.

### Conclusion

Les conclusions suivantes peuvent être tirées des résultats obtenus dans cette partie :

- L'adaptation de la corrélation de Gilbert pour le régime critique et l'extension de Gilbert 1 pour le régime non critique, en fonction des intervalles de  $GOR$ , a permis de développer des modèles plus représentatifs que les modèles théoriques ajustés initialement sur l'ensemble des données. Cela a permis de capturer des relations plus précises entre les variables d'entrée et le débit de production, conduisant ainsi à des prédictions plus fiables ce qui améliore la précision des prédictions du débit de production

- Toutefois, une attention particulière doit être portée aux situations où le *GOR* est inférieur à 4000 et supérieur à 100000 pour le régime critique et inférieure de 1000 pour le régime non critique, où le modèle peut être moins précis.

### 3.5.2.3 Etude de sensibilité de chaque puits

Dans les sections précédentes, on a pris en considération le régime d'écoulement du puits (critique et sous-critique) lors de l'élaboration des modèles pour décrire le débit de production. Cependant, il peut être difficile de choisir le modèle approprié en l'absence d'informations sur le régime d'écoulement. Afin de résoudre ce problème, on a utilisé *LSTM* pour développer différents modèles pour le puits *MD409*.

L'objectif initial était de développer un modèle qui ne prend pas en compte le régime d'écoulement, suivi du développement de modèles plus représentatifs pour décrire la production de puit *MD409*.

Le puits *MD409* a été choisi pour une étude de cas au champ de *HMD*. On dispose de l'historique du puits depuis 2004 jusqu'à 2023 avec 110 points de jaugeage

En utilisant un *LSTM*, on peut mettre en place un modèle de série chronologique pour modéliser et prédire le débit de production. Ce modèle sera capable de capturer les relations temporelles et d'apprendre à partir des données historiques pour fournir des prévisions précises du débit de production à l'avenir.

## Resultats

Le tableau (3-19) représente les résultats de training et de test de modèle *LSTM*. Les Figures (3-25) et (3-26) montrent les cross plot du training et du test du modèle *LSTM* et la figure 4.89 montre le profil de production réel et le débit prédit de ce puits par le modèle *LSTM*.

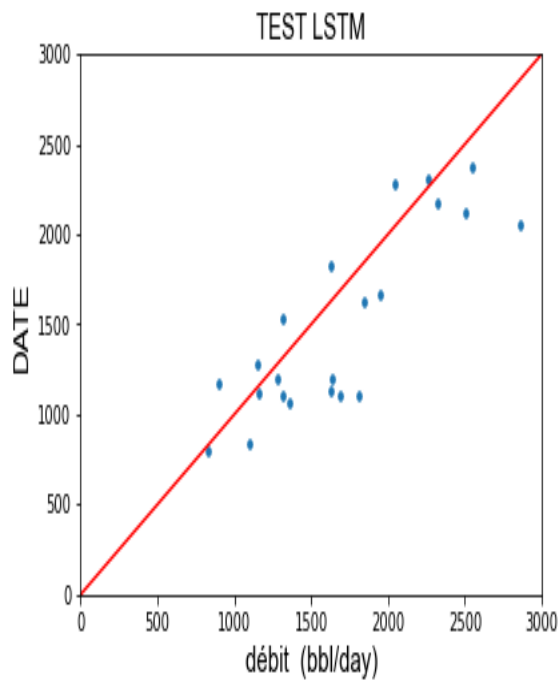
**Tableau 3- 19 les résultats de training et de test de modèle *LSTM***

modèle	train data				test data			
	R <sup>2</sup>	MAE	RMSE	RAE	R <sup>2</sup>	MAE	RMSE	RAE
<b>LSTM</b>	0.9	120.94	167.46	0.07	0.7	219.26	276.43	0.13

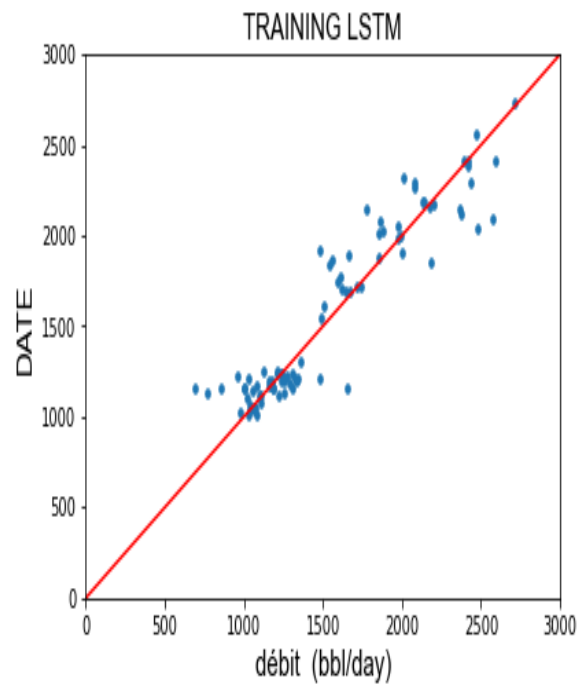
## Interprétation des résultats

- Les résultats obtenus mettent en évidence les performances du modèle *LSTM* utilisé pour prédire le débit de production. Le modèle parvient à estimer le débit réel avec un coefficient de détermination d'environ 0.9 pour training et 0.7 pour le test. De plus une erreur relative absolue de seulement 0.07 pour training et de 0.13 pour test. Ce qui démontre sa capacité à produire des prédictions précises et proches des valeurs réelles.

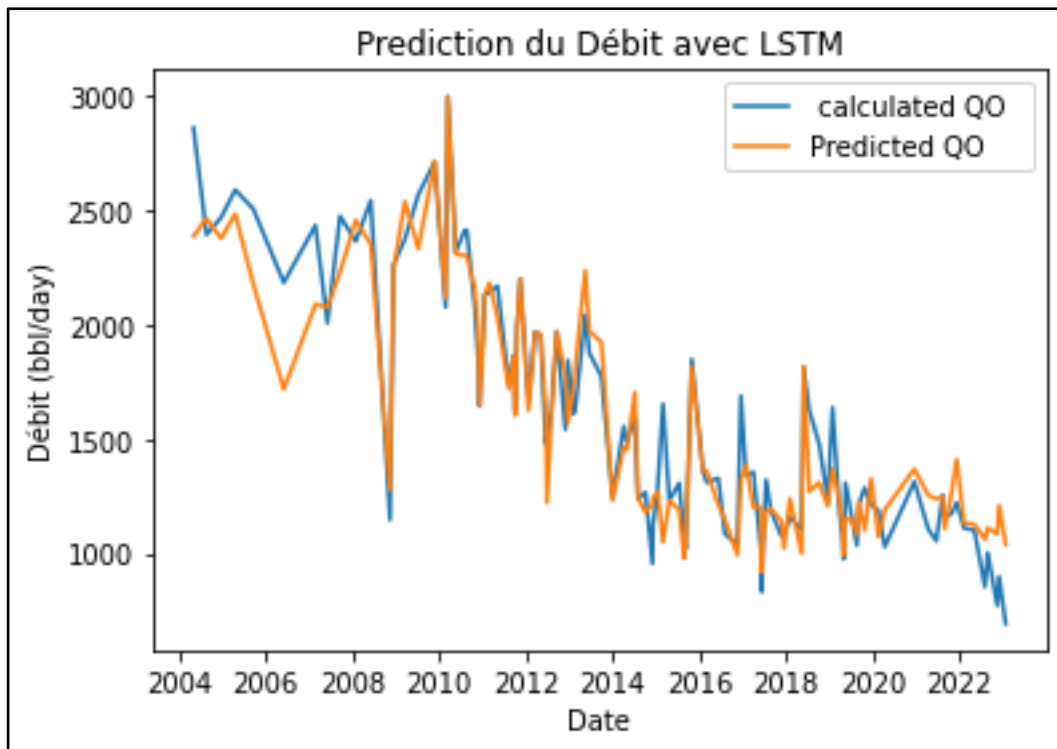
- Les résultats obtenus peuvent être visualisés dans les plots suivants, qui illustrent la comparaison entre les valeurs réelles du débit de production et les prédictions du modèle *LSTM*.



**Figure 3- 25 cross plot de TEST du modèle LSTM**



**Figure 3- 26 cross plot de TRAINING du modèle LSTM**



*Figure 3- 27 prédiction du débit avec LSTM*

**Conclusion**

L'utilisation de modèles LSTM pour étudier chaque puits individuellement a entraîné une augmentation significative de la précision des prédictions. Cette approche a été couronnée de succès, avec des résultats probants.

## **CHAPITRE 4**

### **APPLICATION DUBACK ALLOCATION**

## **4.1 Introduction**

« La répartition de la production des puits est la pierre angulaire de la surveillance et de la bonne gestion des réservoirs. L'apparente simplicité du processus d'allocation conduit souvent à sous-estimer son importance cruciale. Cependant, la précision de l'attribution des taux de production a un impact considérable sur la capacité de l'entreprise à utiliser des données fiables et à effectuer des analyses basées sur des modèles. Par conséquent, la fiabilité des prévisions de production, des estimations de réserves et des efforts d'optimisation du système de production est affectée par l'approche de répartition choisie ». [8]

Il existe différentes approches pour la répartition de la production des puits, qui peuvent varier en fonction des caractéristiques du réservoir, des données disponibles et des objectifs de l'entreprise. Certaines méthodes courantes incluent l'utilisation des données de pression, de débit et de composition des fluides, ainsi que des techniques de modélisation et de simulation avancées.

Il est essentiel de choisir une approche de répartition appropriée et de maintenir sa précision tout au long de la durée de vie du réservoir. Cela permet de garantir la fiabilité des prévisions de production, des estimations de réserves et des efforts d'optimisation du système de production. Une répartition précise des taux de production contribue à une meilleure prise de décision, à une gestion plus efficace des ressources.

Dans notre mémoire, on propose une approche d'allocation de production basée sur les mesures de volume des satellites, ce qui nous permet de corriger les débits calculés des puits en l'absence de leur mesure directe. Cette approche nous offre un moyen de contrôler les puits en détectant les anomalies et les problèmes de performance, tout en améliorant la précision d'allocation de production.

## **4.2 Définition du Back allocation**

Dans l'industrie pétrolière et gazière La back allocation est une méthode d'allocation des hydrocarbures dans laquelle la quantité de production d'un champ est déterminée en analysant les données de production de l'ensemble du champ. Cette méthode est souvent utilisée lorsque les champs produisent des hydrocarbures de manière simultanée et que leur production est mélangée dans un seul flux avant d'être transportée.



Back allocation (La répartition de la production des puits) fait référence au processus de répartition des taux de production entre les différents puits d'un réservoir. C'est un aspect essentiel de la surveillance et de la gestion des réservoirs, car il permet de déterminer la contribution de chaque puits à la production totale.

### 4.3 Les structures back allocation

« Les structures d'allocation sont classées en structures à un niveau et à plusieurs niveaux (Kurian, 2019 ; MPMS 78 Chapitre 20.3, 2011). Le niveau dépend de l'emplacement des stations de mesure et est défini comme le nombre d'étapes nécessaires pour effectuer le processus d'allocation ». [11]

La structure de back allocation fait référence à la méthode utilisée pour répartir la production d'hydrocarbures. Elle dépend de l'emplacement des stations de mesure et correspond au nombre d'étapes nécessaires à l'accomplissement du processus d'attribution. Il existe deux types de structures de back allocation : la structure à un niveau et la structure à plusieurs niveaux.

#### 4.3.1 Structure à un niveau

Elle est effectuée en une seule étape, généralement à partir d'un point central tel qu'une station de mesure fiscale, jusqu'aux champs contributeurs. Comme représente la figure (back au field A et B). Ce type d'allocation est relativement simple et direct et plus faciles à gérer, avec moins de stations de mesure impliquées. Mais elle peut ne pas être suffisamment précise pour répondre à certains besoins d'allocation plus détaillés.

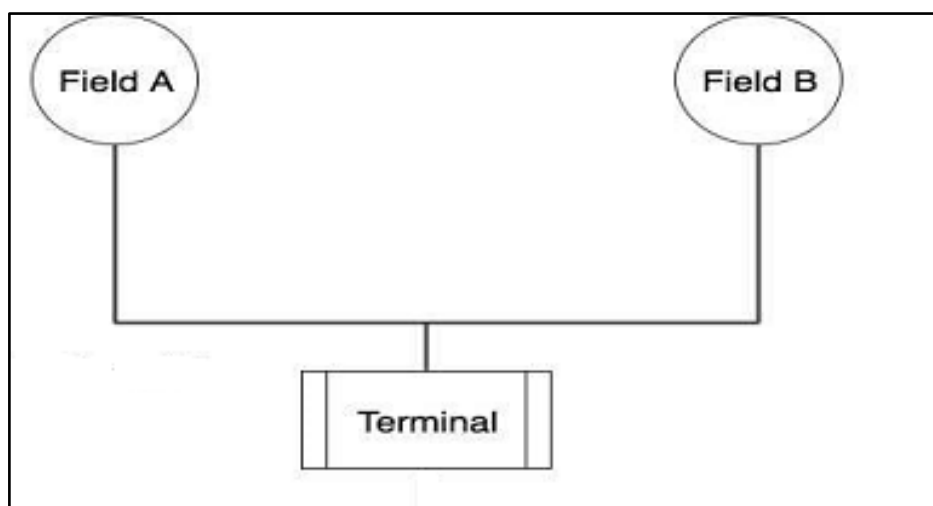


Figure 4- 1 allocation à un niveau (back to Field A et B) [11]

#### 4.3.2 Une structure d'allocation à plusieurs niveaux

Elle implique plusieurs étapes, avec des stations de mesure situées à différents points dans le processus de production et de transport. Cela permet une attribution de la production plus

précise et une allocation plus détaillée des coûts et des revenus aux différentes parties impliquées dans l'opération.

Tel que montré sur la figure : le premier niveau allocation de back au (Field A +Field B) et Field c. et le deuxième niveau de (Field A +Field B) au Field A et B

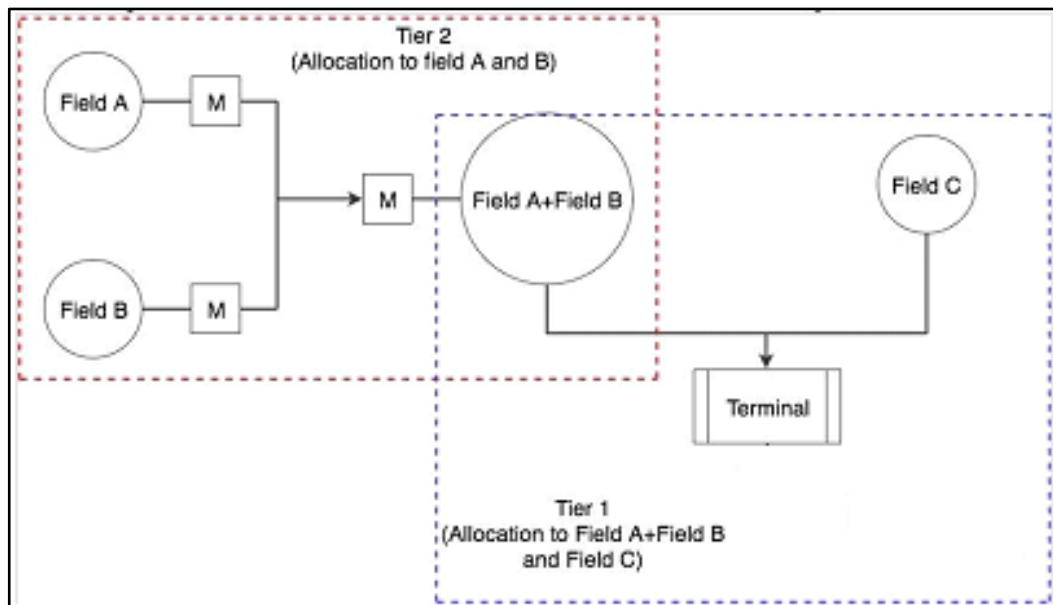


Figure 4- 2 allocation à plusieurs niveaux (back to (field A +field B) et feild c)

#### 4.4 Les types de back allocation :

Il existe plusieurs façons de classer les types d'allocation d'hydrocarbures, mais en général, elles sont classées en fonction de leurs objectifs en allocation de champ et allocation fiscale.

##### Allocation de champ

« Il existe plusieurs façons de classer les types d'allocation des hydrocarbures. Theuveny et Mehdizadeh (2002) ont classé les types d'allocation des hydrocarbures en allocation de vente, allocation de champ, allocation de puits et allocation zonale. Pour plus de simplicité, l'auteur a classé les types d'allocation des hydrocarbures en termes de leurs objectifs, à savoir l'allocation de champ et l'allocation fiscale. La distinction entre ces deux types d'allocation est nécessaire car leurs exigences de mesure de débit et leur méthodologie d'allocation ne sont pas les mêmes. » [11]

Allocation de champ est utilisée pour répartir la production d'hydrocarbures d'un champ entre les différents puits qui y contribuent. Cette méthode d'allocation est utilisée pour déterminer la part de production de chaque puits contribuant à la production globale du champ. Elle est basée sur les mesures de débit et les données de production de chaque puits. Cette allocation est généralement effectuée pour des raisons opérationnelles afin d'optimiser la production du champ, suivre la production de chaque puits individuellement et d'optimiser leur performance et pour prendre des décisions opérationnelles telles que l'arrêt ou la

réduction de la production d'un puits particulier. Parmi les méthodes d'allocation de champ. On peut citer :

- L'allocation proportionnelle : La méthode la plus simple, elle consiste à répartir les volumes produits entre les puits en fonction de leur contribution proportionnelle à la production totale.
- Allocation différentielle : Cette méthode d'allocation est utilisée lorsqu'il y a une différence de qualité entre les fluides produits par différents puits. Dans ce cas, la production de chaque puits est ajustée en fonction de sa qualité relative par rapport à la qualité moyenne du champ.

### **L'allocation fiscale**

L'allocation fiscale est un type d'allocation qui s'applique aux hydrocarbures extraits d'un gisement et qui sont soumis à des obligations fiscales telles que des redevances ou des taxes. Elle est utilisée pour répartir les coûts et les revenus fiscaux provenant de la production d'hydrocarbures entre les différents partenaires d'un contrat de production pétrolière et gazière, tels que l'État, les entreprises pétrolières et les partenaires contractuels, les investisseurs. Elle est basée sur les termes du contrat de partage de production, qui spécifie les pourcentages de revenus à allouer à chaque partie. Les mesures de débit sont effectuées sur les points de mesure fiscaux au point de vente, tels que les stations de mesure fiscale ou les points de transfert de garde. Parmi les méthodes d'allocation de fiscale, on peut citer :

- Allocation de coûts communs : Dans les installations d'exploitation et de production, les coûts varient considérablement entre les champs participants... il y a souvent des coûts qui sont partagés entre plusieurs puits ou champs, tels que coûts de forage, les coûts de traitement ou de transport. La méthode d'allocation de coûts communs est utilisée pour répartir ces coûts entre les différentes parties en fonction de leur utilisation relative des installations.
- Allocation basée sur la valeur : Cette méthode d'allocation est utilisée pour répartir les revenus entre les différents puits ou champs en fonction de la valeur relative de leur production. Cette méthode prend en compte le prix de vente et les coûts associés à la production de chaque puits ou champ.

Ces différents types d'allocation peuvent être combinés ou adaptés en fonction des besoins spécifiques de chaque champ ou installation de production ou projet.

## **4.5 Objectifs de back allocation**

Back allocation, ou la répartition de la production des puits, est un processus essentiel dans l'industrie pétrolière et gazière. Ses objectifs et rôles principaux sont les suivants :

- Contrôle des débits des puits : La back allocation permet de surveiller et de contrôler les débits de chaque puits individuellement. Cela permet d'optimiser la production

en identifiant rapidement les problèmes, les pertes ou les anomalies de production. [23]

- Répartition précise des volumes de production : La back allocation vise à attribuer avec précision les volumes de production à chaque puits ou champ. Cela permet de quantifier la contribution de chaque puits à la production totale. Une répartition précise des volumes de production est essentielle pour une gestion efficace des réservoirs, l'évaluation des performances des puits et la planification de la production future.
- Capacité à répartir correctement la production : Lorsque des mesures précises ne sont pas possibles ou rentables, back allocation permet d'apporter une répartition précise de la production, ce qui garantit une gestion efficace des ressources.
- Gestion des infrastructures partagées : Dans les cas où plusieurs puits ou champs de production utilisent des infrastructures communes telles que des pipelines ou des installations de traitement, la back allocation permet de gérer efficacement l'utilisation de ces infrastructures. Elle assure une répartition équilibrée des volumes de production entre les différentes entités ce qui facilite une meilleure utilisation des infrastructures et contribue à une efficacité opérationnelle accrue. [12]

## **4.6 Mesure de Back allocation**

Back allocation permet de déterminer de manière précise les volumes de production individuels des puits ou champs en utilisant les informations disponibles à partir des mesures effectuées aux différentes étapes du processus amont(upstream), telles que les jauges de puits, les débitmètres aux manifolds et satellites, les séparateurs de production, les systèmes de stockage. etc.

Le processus upstream (processus amont) se réfère à l'ensemble des activités nécessaires à l'exploration, la production et au transport des ressources pétrolières et gazières. Les sections du processus amont peuvent inclure des activités telles que la production de pétrole et de gaz, la séparation, le traitement, le stockage, le transport ainsi que d'autres activités connexes.

Mesure de Back allocation peut être appliquée aux différents composants du processus amont :

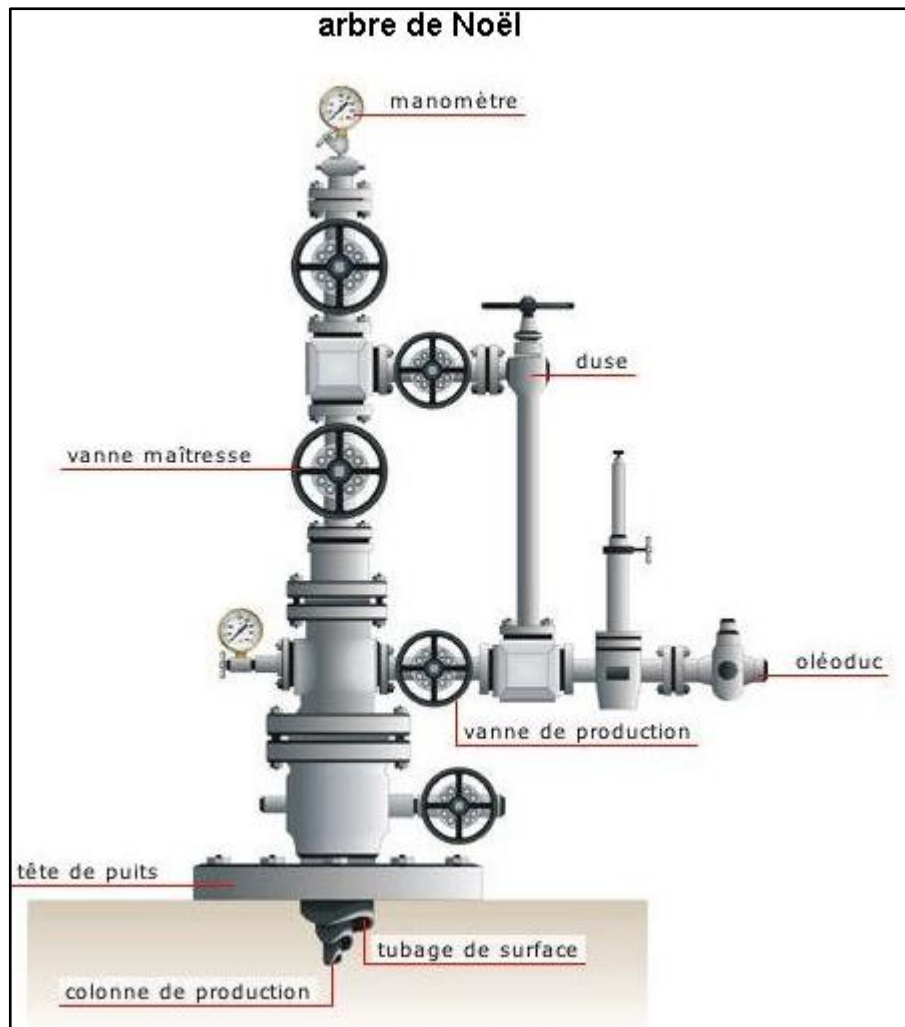
### **4.6.1 La tête de puits de production**

La tête de puits se situe au sommet du puits de pétrole ou de gaz, menant au réservoir. Une tête de puits peut également être un puits d'injection, utilisé pour injecter de l'eau ou du gaz dans le réservoir afin de maintenir la pression et les niveaux pour une production maximale ». [7]

La tête de puits de production, est un équipement installé après la phase de forage et de complétion. Elle est conçue pour permettre la production efficace et contrôlée du pétrole ou du gaz naturel extrait du réservoir souterrain.

La tête de production comprend plusieurs éléments clés qui facilitent le processus de production et de gestion des fluides. Voici quelques composants couramment présents dans une tête de production de puits pétroliers :

- Tubing head (tête de tubage) : C'est la partie supérieure de la tête de production qui reçoit le tubage, un conduit métallique descendant dans le puits. Le tubage est utilisé pour acheminer les fluides de production du puits vers la surface.
- Vannes de production : La tête de production est équipée de vannes de production qui permettent de réguler le débit des fluides provenant du puits. Ces vannes peuvent être ouvertes ou fermées pour contrôler la quantité de pétrole ou de gaz produite.
- Séparateurs et équipements de traitement : Certains puits produisent des fluides mélangés tels que du pétrole, du gaz naturel et de l'eau. La tête de production peut comprendre des équipements tels que des séparateurs, des déshydrateurs et des équipements de traitement pour séparer et traiter ces différents composants.
- Équipement de mesure et de contrôle : La tête de production peut être équipée de capteurs, de jauges de pression et d'autres instruments de mesure pour surveiller et contrôler les paramètres de production tels que la pression, la température et le débit.
- Soupapes de sécurité : Des soupapes de sécurité sont généralement installées sur la tête de production pour éviter toute surpression dangereuse ou tout autre incident pouvant mettre en danger les opérations et la sécurité.



**Figure 4- 3 tête de puits [13]**

La tête de production assure différentes mesures, telles que les mesures de pression et de jaugeage

- Mesure de pression : Des jauges de pression sont installées au niveau de la tête de puits pour mesurer la pression du fluide à l'intérieur du puits. Ces jauges peuvent être des capteurs électroniques ou des manomètres mécaniques. Ces mesures de pression sont souvent enregistrées en continu et surveillées en temps réel à l'aide de systèmes de surveillance et de contrôle tels que le SCADA (Supervisory Control and Data Acquisition). Cela permet aux opérateurs de suivre l'évolution des pressions
- Le jaugeage au niveau des puits : est un processus de mesure et d'analyse utilisé pour évaluer les performances et la productivité d'un puits de pétrole ou de gaz. Il permet de déterminer le débit de gaz, d'huile et d'eau produits par un puits, différentes méthodes de mesure sont utilisées. Dans certains cas, des débitmètres peuvent être installés au niveau de la tête de puits pour mesurer directement le débit ou bien des séparateurs pour séparer les différents composants (gaz, huile et eau) du fluide produit par le puits. Ces méthodes permettent de mesurer et de collecter séparément le débit de chaque composant. Ces méthodes permettent de fournir des mesures précises du débit. Cependant, dans des champs pétroliers ou gaziers de grande

envergure avec un grand nombre de puits, il peut être coûteux d'installer et d'entretenir des débitmètres individuels ou d'effectuer un jaugeage sur chaque puits de manière régulière.

#### **4.6.2 Manifolds et pipelines**

« Les flux provenant de chaque puits sont transportés vers les installations de production principales à travers un réseau de pipelines de collecte et de systèmes de collecteurs. Le but de ces pipelines est de permettre la configuration d'ensembles de puits de production pour un niveau de production donné ». [7]

Les manifolds et les pipelines jouent un rôle crucial dans le transport des flux de chaque puits vers les installations de production principales. Les manifolds et les satellites sont des équipements qui ne servent de points de collecte pour regrouper les flux provenant de plusieurs puits.

Les pipelines de collecte permettent de transporter ces flux de manière efficace et sécurisée vers les installations de production. Les manifolds sont des points de connexion où les flux provenant des puits sont regroupés avant d'être acheminés dans les pipelines. Ces manifolds peuvent être équipés de débitmètres pour mesurer le débit global des flux provenant de tous les puits connectés.

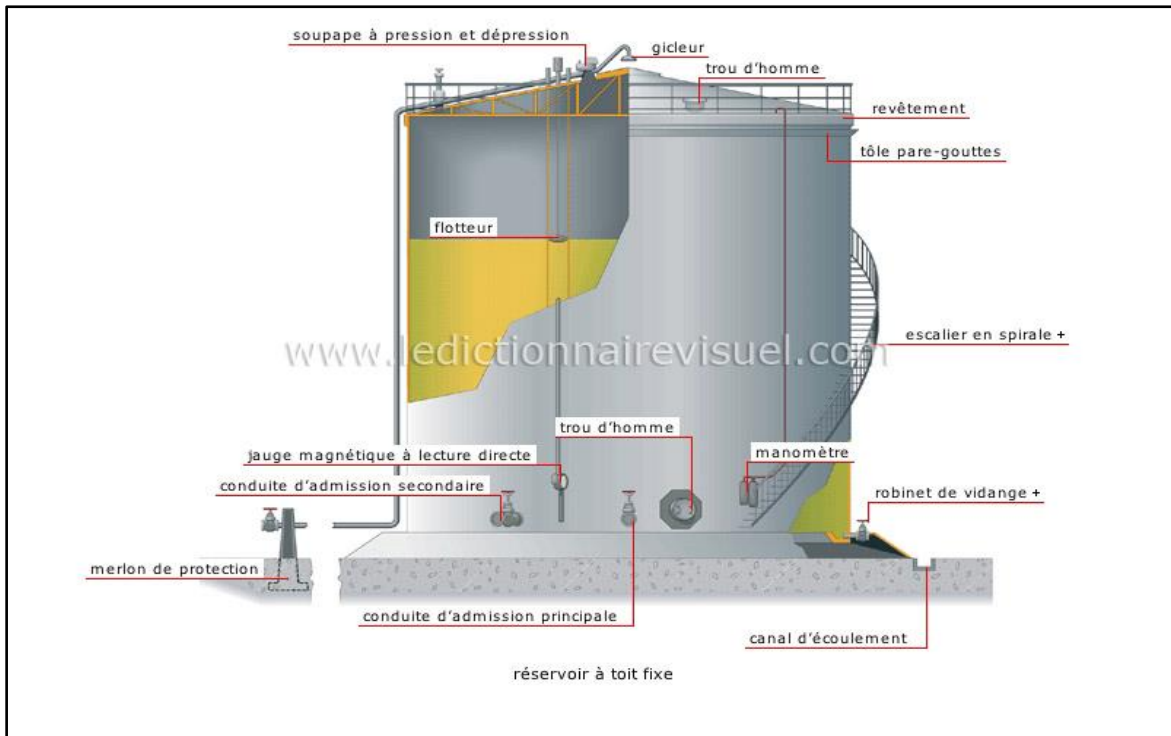
#### **4.6.3 Stockage**

Dans la plupart des sites de production, le pétrole est acheminé directement vers les bacs de stockage qui sont utilisés pour contenir le produit avant sa distribution ultérieure. Il existe deux types courants de bacs de stockage d'huile : les réservoirs à toit fixe et les réservoirs à toit flottant. Les réservoirs à toit flottant sont généralement utilisés lorsque de grands diamètres sont nécessaires et qu'il n'y a pas de restrictions concernant l'évacuation des vapeurs. Ces réservoirs ne fonctionnent qu'à pression atmosphérique, et le toit flotte de haut en bas en fonction du volume de pétrole brut. [10]

Des systèmes spéciaux de jaugeage des bacs, tels que les radars de niveau ou les flotteurs, sont utilisés pour mesurer le niveau d'huile dans bac de stockage pour convertir en volume

Les capteurs de niveau utilisent des ondes pour mesurer précisément la distance entre la surface de l'huile et le capteur installé dans le bac de stockage. Les flotteurs, quant à eux, sont des dispositifs flottants qui montent et descendent en fonction du niveau de liquide.

Des registres précis des volumes et de l'historique sont tenus pour documenter ce que les puits ont produit et ce que le bac a reçu car Les flux produits sont mélangés et peuvent avoir une propriété ou un régime fiscal différent et Entre le point de mélange et le point de mesure (dans l'installation commune), l'état physique des fluides peut changer. En outre, l'équipement utilisé pour mesurer le débit des différents flux peut présenter différents niveaux d'incertitude de mesure. Aussi il peut y avoir des pertes dues à une rupture de pipeline .La figure (4-4) représente un bac de stockage à toit fixe :



**Figure 4- 4 bac de stockage [25]**

#### 4.6.4 Télémétrie/SCADA

« SCADA and telemetry work in tandem to provide a combination of data acquisition, analysis, storage, and reporting. Additionally, SCADA systems allow operators to monitor and control system command functions from a remote location ». [24]

SCADA (Supervisory Control and Data Acquisition) et la télémétrie travaillent en tandem pour fournir une combinaison d'acquisition de données, d'analyse, de stockage et de génération de rapports. De plus, les systèmes SCADA permettent aux opérateurs de surveiller et de contrôler les fonctions de commande du système à distance.

La télémétrie permet la collecte et la transmission de données de mesure à distance à partir de différents points du processus amont, tels que les puits, les pipelines, les stations de compression, les installations de stockage, etc. Les données collectées peuvent inclure des mesures telles que la pression, la température, le débit, les niveaux de réservoir, les conditions de fonctionnement des équipements, etc.

Ces données sont ensuite transmises en temps réel ou à intervalles réguliers via des réseaux de communication à grande échelle vers le système SCADA central.

Le système de contrôle et d'acquisition de données SCADA associés à la télémétrie et aux communications à grande échelle sont utilisés pour collecter les données de mesure telles que la pression, les diamètres, etc...à partir de différents points du processus amont. Ces données sont utilisées pour surveiller et contrôler la production.



## 4.7 Une nouvelle approche de back allocation

Dans un champ pétrolier, le satellite est une installation regroupant plusieurs puits qui peuvent avoir des débits mesurés par jaugeage ou des débits calculés à partir de modèles et de données disponibles. Comme illustre la figure (4-5)

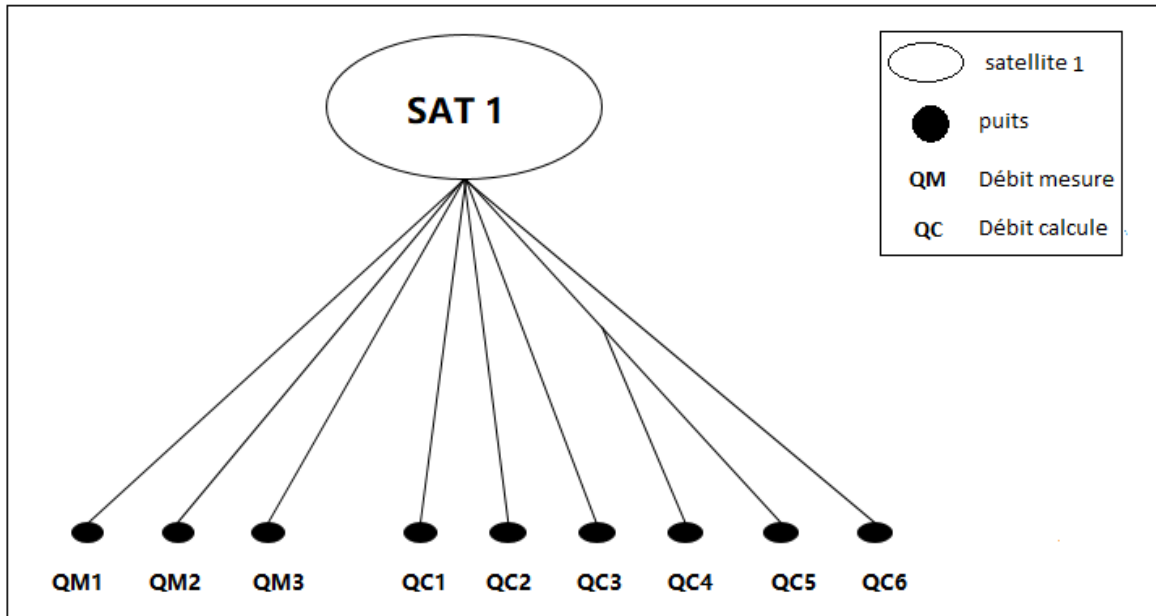


Figure 4- 5 schéma représentatif de la structure de back allocation pour un seul satellite

Les débits mesurés directement sur le terrain sont considérés comme des données fiables. Par contre Les débits calculés pour les autres puits, basés sur des modèles et des estimations, peuvent présenter des écarts par rapport aux débits mesurés.

Pour corriger ces écarts, on propose une méthode de "back allocation" vise à améliorer la précision de l'allocation des débits des puits. Elle repose sur la correction individuelle des débits à l'aide de coefficients spéciaux "k" propres à chaque puits. Ces coefficients sont déterminés en fonction des caractéristiques individuelles de chaque puits, telles que les performances historiques.

Notre approche est basée sur l'équation suivant :

$$Q_{SAT} - Q_m - Q_{corr} = 0 \quad (4.1)$$

Où  $Q_m$  représente la somme des débits des puits mesurées

$$Q_m = \sum_{i=1}^n Q_{M_i} \quad (4.2)$$

$Q_{corr}$  Représente la somme des débits des puits corrigés par le coefficient k

$$Q_{corr} = \sum_{i=1}^n Q_{c_i} / k_i \quad (4.3)$$

$Q_{SAT}, Q_{M_i}, Q_{C_i}, k_i$  Représentent respectivement le volume mesuré au niveau du satellite. Les débits mesurés par jaugeage pour chaque puits individuel, les débits calculés pour chaque puits, le coefficient  $k_i$  pour chaque puits

Il est important de noter que chaque puits peut avoir des coefficients spécifiques, notés  $k_i$ , qui sont définis dans un intervalle donné. noté  $k_{i\ min} < k_i < k_{i\ max}$

## 4.8 Outil mathématique

### ❖ Le score z

Le score z, également connu sous le nom d'écart-type standardisé, est utilisé en statistiques pour mesurer la position d'un point de données par rapport à la moyenne d'une distribution, en termes d'écart-types. Il permet de comparer différentes observations provenant de ensembles de données distincts en les normalisant. Le score z est calculé en soustrayant la moyenne de la valeur donnée, puis en divisant le résultat par l'écart-type. Un score z positif indique que le point de données se situe au-dessus de la moyenne, tandis qu'un score z négatif indique une position en dessous de la moyenne.

La formule pour calculer le score z d'un point de données x dans un ensemble de données avec une moyenne  $\mu$ , et un écart-type,  $\sigma$ , est la suivante :

$$z = \frac{(x - \mu)}{\sigma} \quad (4.4)$$

La déviation standard, également appelée écart type, est une mesure statistique qui quantifie la dispersion ou la variabilité des données par rapport à leur moyenne. Elle indique à quel point les valeurs individuelles d'un ensemble des données s'écartent de la valeur moyenne. Plus la déviation standard est élevée, plus les valeurs sont dispersées autour de la moyenne et inversement.

La déviation standard est calculée à partir de la formule suivante :

$$\sigma = \sqrt{\frac{\Sigma(x - \mu)^2}{N}} \quad (4.5)$$

Où :

$\sigma$  représente la déviation standard.

x est chaque valeur individuelle de l'échantillon.

$\mu$  représente la moyenne de l'échantillon.

N est le nombre total de valeurs dans l'échantillon.

## 4.9 Résultats et discussions

### 4.9.1 Présentation du dataset de production du champ de Hassi Messaoud

#### 4.9.1.1 Description du champ

Le champ pétrolier de *HMD* est le plus grand champ pétrolière d'Algérie .il produisant des quantités importantes de pétrole brut chaque année.il comprend un grand nombre de puits de pétrole. En fait, il est l'un des champs pétroliers les plus vastes et les plus complexes d'Algérie avec environ 1500 puits de pétrole en exploitation. Le champ se compose de deux extensions importantes, le champ de Hassi Messaoud Sud *HMDS* et le champ de Hassi Messaoud Nord *HMDN*, qui sont exploités séparément.

Le champ de *HMDS* comprend plus de 500 puits de pétrole en activité. Ce nombre de puits peut changer en fonction des besoins de production, de l'exploration de nouvelles zones et des activités de maintenance.

Le processus de production dans les puits de pétrole de *HMDS* implique souvent la collecte du pétrole brut de plusieurs puits par des manifolds qui sont diriger vers 10 satellites en général pour permettre une séparation initiale. Le pétrole est ensuite acheminé vers deux bacs de stockage intermédiaires :6 satellites vers le bac UTBS et 4 satellites vers bac CIS comme représente la figure (4-6). Le pétrole brut est ensuite envoyé vers un bac final le BAC SUD où il est stocké avant d'être transporté pour le traitement ultérieur.

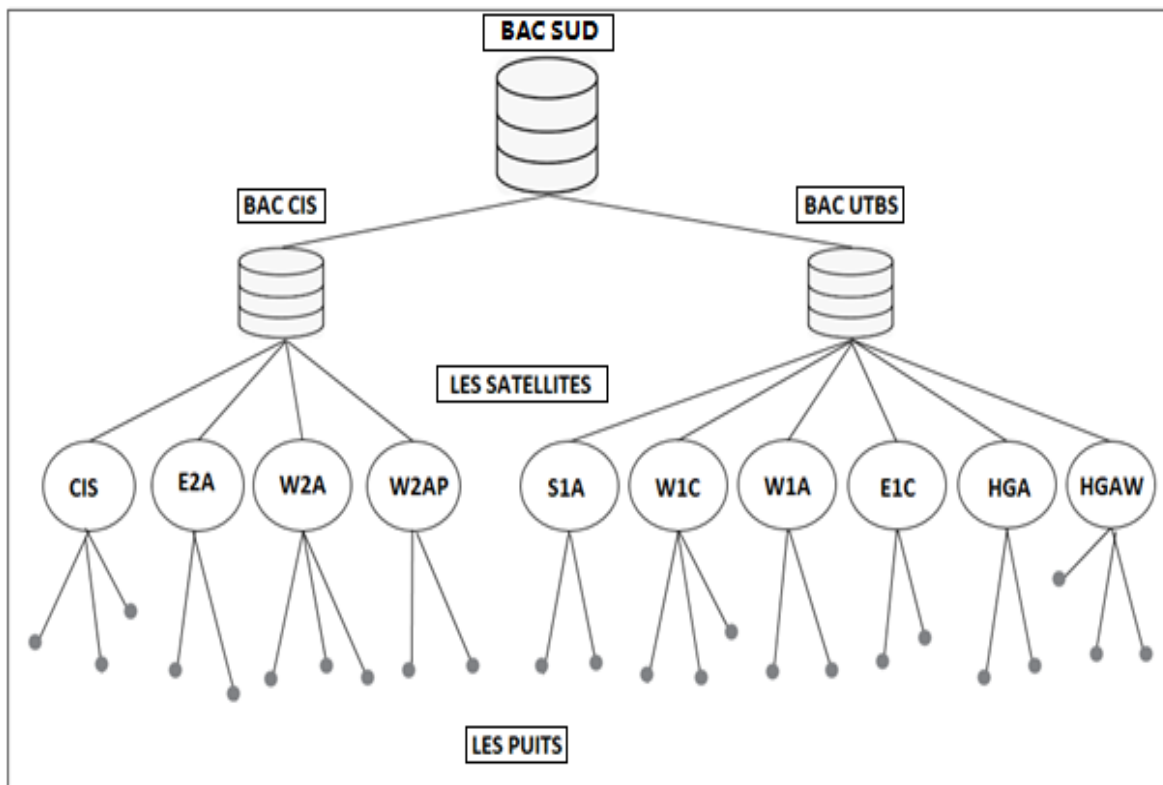


Figure 4- 6 le champs de HMD

#### **4.9.1.2 Description du data**

Les data sont présente sous forme d'un Excel et de base des données dans DB browser sqLite. Il comporte 02 tableaux le premier tableau « DAILY DATA » sous forme de sqLite pour les données brutes de production des puits du champ de *HMD*. Ces données sont comportées les data quotidiennes et les jaugeages data qui s'effectuent via des opérations de jaugeage pour chaque puits.

Le premier tableau contient les variables suivantes : la date, nom du puits, nom du satellite, nom du manifold, le diamètre de la Duse, la pression en tête de puits (pression en amont de la Duse), la pression du pipe (pression en aval de la Duse), les pressions de jaugeage, le débit de gaz, eau et d'huile de jaugeage, le *GOR*, débit de production d'huile , de gaz et d'eau, heur de travaille de puits.

Le deuxième tableau « BAC » sous forme Excel contient les volumes de bac et les volumes de chaque satellite pendant cette durée. Il comporte les variables suivantes : la date, le volume mesure et calcule de bac sud, bac UTBS et bac CIS, le volume mesuré et calcule de 10 satellites, le nombre de puits de chaque bac et chaque satellite.

#### **4.9.1.3 Préparation des data :**

Ces données ne peuvent pas être utilisées tel qu'elles sont car elles contiennent des valeurs manquantes, des satellites mal distribué, des erreurs de mesure, des valeurs négatives. Donc Il est important de traiter ces erreurs de manière appropriée pour éviter des résultats biaisés ou trompeurs qui peuvent rendre difficile l'analyse et la compréhension des résultats.

Pour préparer les données à des fins d'analyse. On a suivi les étapes suivantes :

1-Elimination des lignes ou les satellites sont mal distribués.

2-Correction des erreurs de saisie : Dans certains cas, ces erreurs peuvent se produire lors de la collecte des données. Par exemple, il peut arriver qu'un satellite ait un débit mesuré noté comme nul, alors que les puits sont actifs et que le débit calculé est de 130. Dans ce cas, nous avons corrigé cette erreur en remplaçant le zéro par une valeur de débit mesuré correspondant à un autre jour où le même satellite avait le même débit calculé et le même nombre de puits actifs.

3-la mise en forme de tableau Excel « BAC » sous forme un tableau sqLite.

4-correction les volumes d'huile mesurés des satellites « S1A, W1C, W1A, E1C, HGA, HGAW »de bac UTBS en tenant compte de la quantité d'eau présente dans chaque volume. Pour éliminer ce problème on :

- Calcul de la quantité d'eau : on a calculé la quantité d'eau présente dans le bac UTBS en soustrayant le volume d'arrivée du volume de départ. Cela nous a donné une estimation de la quantité d'eau totale présente dans le bac UTBS.
- Détermination du pourcentage d'erreur : Pour chaque satellite, on a calculé un pourcentage d'erreur basé sur le nombre de puits du SAT par rapport au nombre total de puits du bac UTBS. Ce pourcentage représente la part de chaque SAT dans la quantité d'eau totale.

- Correction du débit mesuré : Pour corriger le débit mesuré de chaque SAT, on a multiplié la quantité d'eau déterminée (eau) par le pourcentage d'erreur correspondant au SAT ( $NB_{SAT} / NB_{UTBS}$ ). Ensuite, nous avons soustrait cette quantité d'eau corrigée du débit mesuré initial du SAT.

Par exemple, la correction du débit mesuré pour le SAT E1C serait réalisée selon la formule suivante :

$$E1C_{corr} = E1C_{mes} - V_{eau} * \frac{NB_{E1C}}{NB_{UTBS}} \quad (4.6)$$

$E1C_{corr}$  .  $E1C_{mes}$  .  $NB_{E1C}$  .  $NB_{UTBS}$  Représentent respectivement le volume de E1C corriger, le volume de E1C mesure, le nombre des puits de E1C. le nombre des puits de bac UTBS

## 4.9.2 Optimisation de débit des puits par la nouvelle méthode de back-allocation

### 4.9.2.1 Méthodologie

La méthode de back-allocation que l'on utilise vise à améliorer la précision en corrigeant individuellement les débits de chaque puits à l'aide de coefficients spéciaux  $k$  propres à chacun. Pour obtenir ces coefficients, on se base sur l'historique des puits, où l'on dispose des débits mesurés et des débits calculés. En divisant le débit calculé par débit mesuré, on obtient un intervalle de coefficients  $ki$  qui décrit l'écart entre ces deux valeurs pour l'ensemble de la durée de vie du puits (historique).

Ce processus de calcul des coefficients  $k$  permet de quantifier l'écart entre les débits mesurés et les débits calculés, et ainsi de déterminer un ajustement spécifique pour chaque puits. Les coefficients  $k$  servent à corriger les débits calculés afin de s'approcher davantage des débits mesurés, en prenant en compte les variations observées dans l'historique des puits.

Une fois que l'on a obtenu ces coefficients  $k$  pour chaque puits, on utilise La régression non linéaire , pour effectuer les calculs de répartition des débits. On dispose effectivement des informations sur le volume total mesure dans le satellite ainsi que des débits mesurés par le jaugeage pour certains puits. Cependant, pour les puits restants qui n'ont pas de débits mesurés, on utilise leurs débits calculés. Ensuite, on ajuste ces débits calculés en utilisant le coefficient  $k$  . Cette correction permet d'obtenir des débits ajustés plus proches de la réalité.

Cette méthode de back-allocation améliorée permet donc d'ajuster les débits des puits individuellement en prenant en compte des caractéristiques spécifiques de chaque puits.

### 4.9.2.2 Résultats de back allocation

Dans cette partie, on a choisi de travailler sur le SAT E1C qui importer plus de 50 puits. Notre choix est basé sur la disponibilité de data et de l'historique de tous leurs puits.

### ❖ Calcul des coefficients $ki$

Pour chaque puits, on calcule le coefficient  $ki$  en divisant le débit calculé par le débit mesuré correspondant, les débits calculés sont estimés à partir de modèle de Random Forest. Cette étape permet d'évaluer l'écart entre les débits mesurés et les débits calculés, et le coefficient  $ki$  représente l'ajustement nécessaire pour corriger cet écart. Cette étape est réalisée pour tous les puits de *E1C*

Le plot suivant représente les valeurs de coefficient  $k$  déterminé à partir de Random Forest pour la période de 2004 au 2023 pour le puit *MD409*

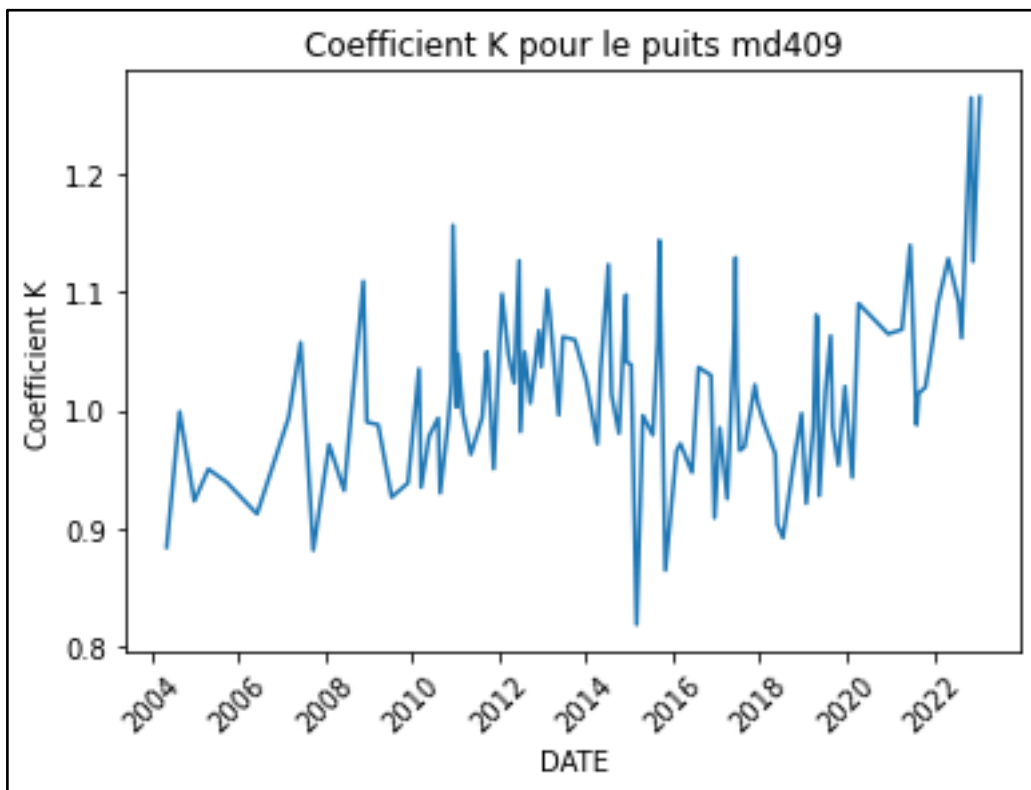


Figure 4- 7 intervalle du coefficient  $K$  pour le puits *md409*

Le plot représente que le coefficient  $K$  prendre des valeurs entre 0.82 et 1.27

### ❖ Création d'un intervalle de coefficients

À partir des valeurs de coefficients  $ki$  obtenues, on crée un intervalle de coefficients pour chaque puits. En ajoutant une incertitude de 20% à chaque intervalle.

Le tableau (4-1) suivant représente les nouvelles valeurs (minimale et maximale) de coefficient  $k$  pour le puit *MD409*

Tableau 4- 1 représente les nouvelles valeurs (minimale et maximale) de coefficient  $k$  pour le puit MD409

	min $k$	max $k$
$k$	0.66	1.52

❖ **Correction des débits des puits**

On utilise la régression non linéaire pour effectuer les calculs de répartition des volumes des puits dans le SAT. Une fois on a trouvé la solution optimale, on applique les coefficients  $k$  ajustés à chaque puits pour corriger individuellement leurs débits. On divise le débit calculé de chaque puits par le coefficient  $k_i$  approprié pour obtenir le débit corrigé.

Les plots suivants représentent respectivement les nouvelles valeurs de coefficient  $k$  déterminées pour corriger le débit calculé pour chaque jour de la période allant du 01-10-2022 au 31-10-2022 pour le puit MD409 et les débits corrigés par back allocation pour le même puit pendant la même période

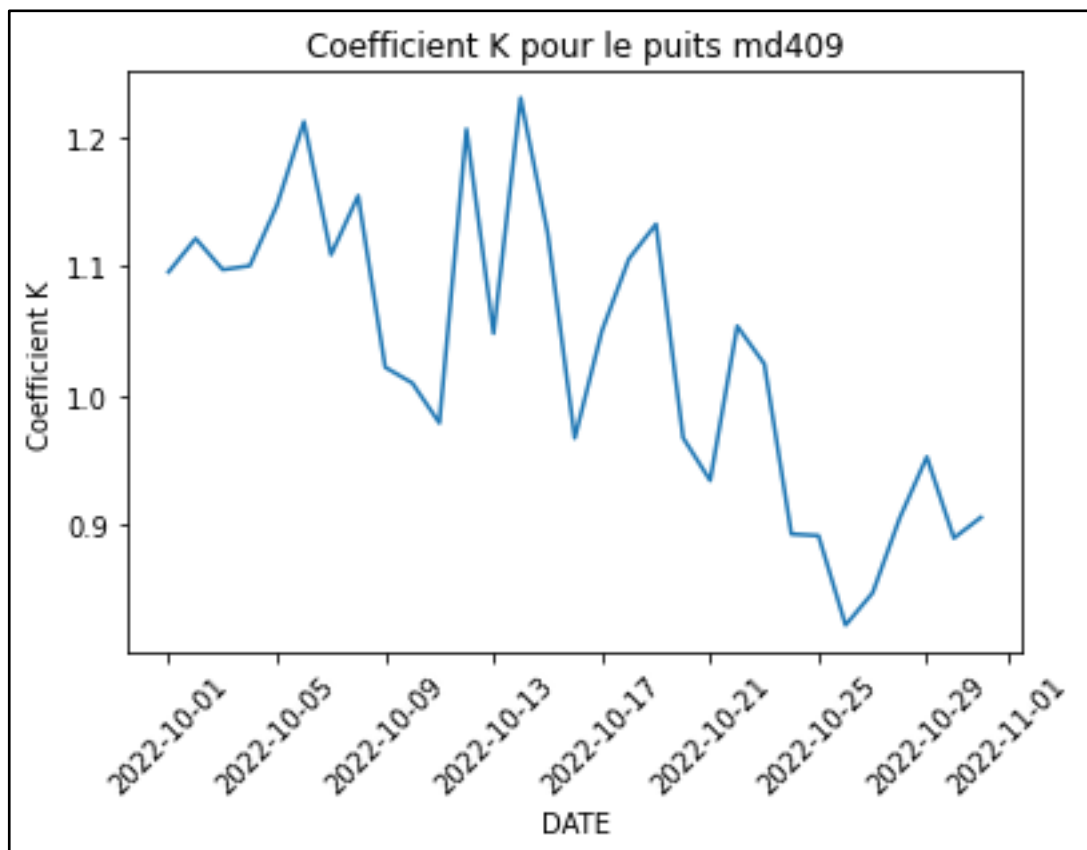


Figure 4- 8 intervalle du coefficient  $K$  pour le puits md409 du 01-10-2022 au 31-10-2022

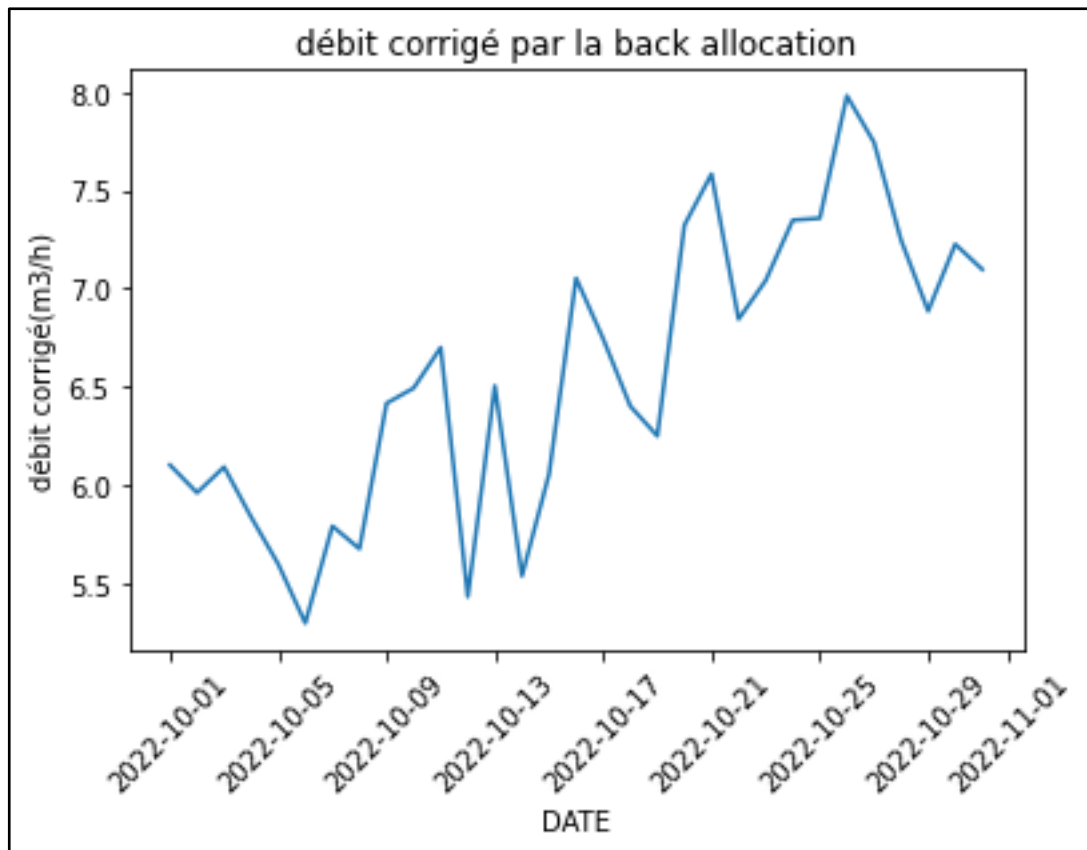


Figure 4- 9 débit corrigé par la back allocation

#### ❖ Détection proactive des anomalies dans les débits des puits

La détection proactive des anomalies dans les débits des puits est un processus important pour assurer une exploitation efficace et fiable des puits. Cela consiste à surveiller en continu les débits des puits et à identifier toute variation significative ou anormale qui peut indiquer un problème opérationnel ou une défaillance potentielle

On peut procéder à la détection d'anomalies en utilisant différentes techniques :

L'évaluation du Z-score est une méthode courante pour détecter les anomalies dans les séries de données. En calculant le Z-score des débits ajustés pour chaque puits, on peut estimer la déviation de chaque valeur par rapport à la moyenne de l'ensemble des valeurs, en termes d'écart-type. En utilisant des seuils d'anomalie définis, indiquant que le débit du puits est très éloigné de la moyenne et peuvent être considérées comme des anomalies.

Une autre approche consiste à utiliser des modèles de séries chronologiques, tels que le *LSTM* (Long Short-Term Memory), pour détecter les anomalies dans les débits des puits. Les modèles *LSTM* sont capables de capturer les tendances et les motifs temporels des données historiques. En entraînant le modèle avec les débits ajustés passés, on peut prédire le débit attendu pour une journée donnée. Si le débit réel (ajustée) diffère considérablement de la prédiction du modèle *LSTM*, cela peut indiquer une anomalie ou un problème.



Le plot représente Analyse des anomalies de débit par le z score

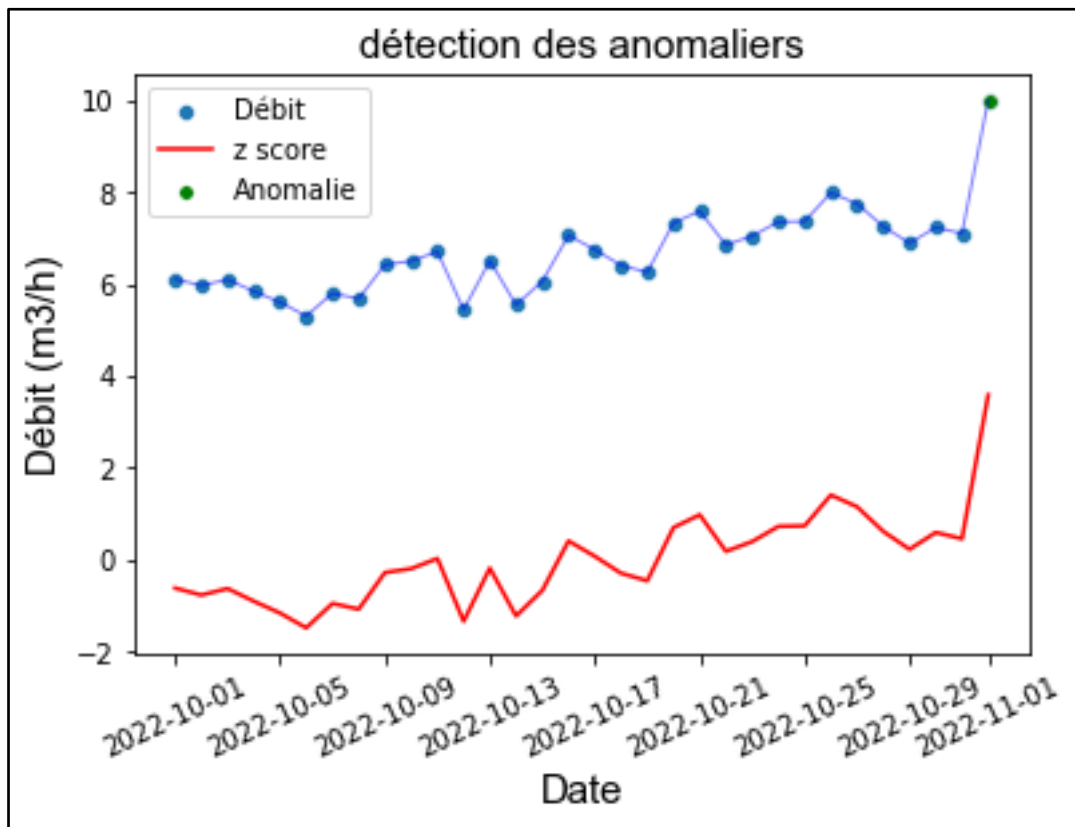


Figure 4- 10 détection des anomalies par z\_score

Le graphe représente le débit du puits corrigé par la back allocation en fonction du temps, et le z-score correspondant à chaque valeur de débit du puits MD409, le débit varie entre 5 et 7 m3/h, avec un z-score entre 2 et -2, sur une période allant du 01-10-2022 au 31-10-2022. Pendant cette période, les débits du puits se situent dans la plage attendue, ce qui signifie qu'ils sont considérés comme normaux car le z-score ne dépasse pas le seuil établi.

Le 01 -11-2022, on observe une anomalie dans le graphe où le débit atteint 10 m3/h. Cette valeur élevée du débit correspond à un z-score qui dépasse la valeur seuil établie. Cette anomalie du débit, avec un z-score élevé, indique une variation importante et inattendue du débit du MD409.

Le plot représente Analyse des anomalies de débit par LSTM :

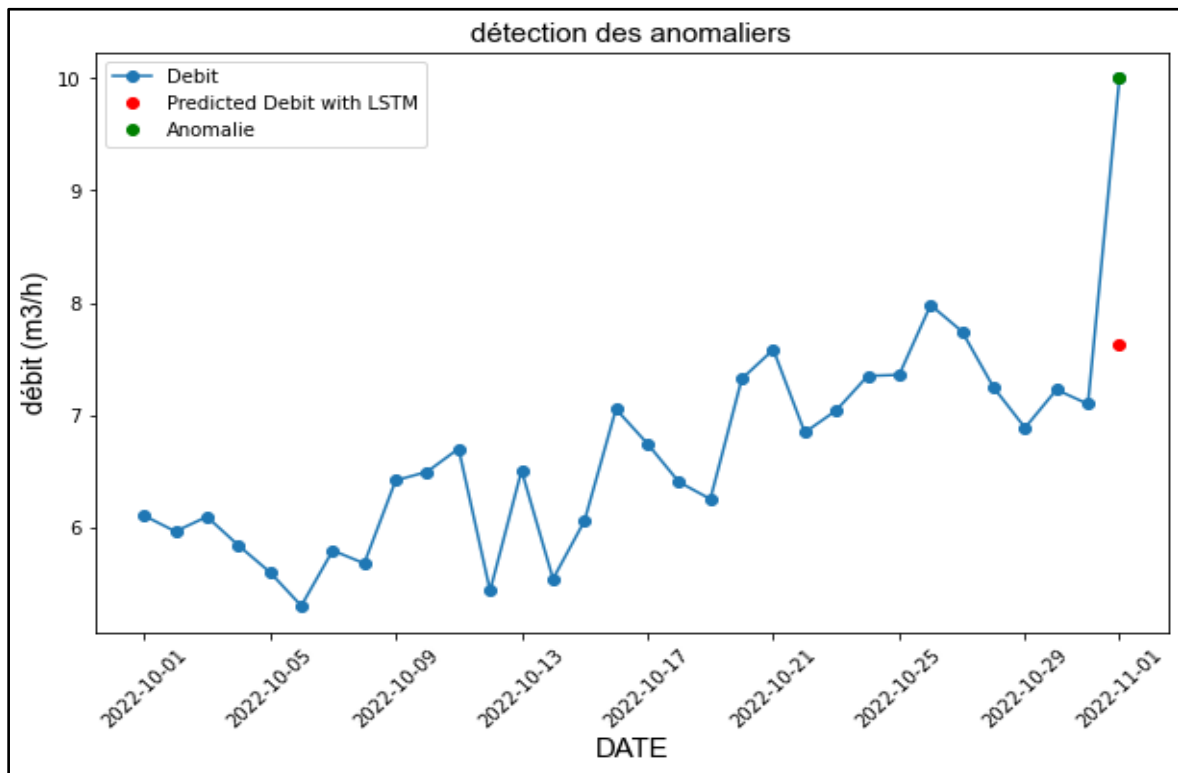


Figure 4- 11 détection des anomalies par LSTM

On utilise le *LSTM* sur le même puits *MD409* et pour la même période précédente, pour prédire le débit du le 01-11-2022 en se basant sur les jours précédents, une prédiction de 7.6 m<sup>3</sup>/h a été obtenue, ce qui était cohérent avec les tendances observées dans les jours précédents. Cependant, la valeur réelle enregistrée pour cette journée était de 10 m<sup>3</sup>/h, ce qui représente une anomalie significative.

#### 4.10 Conclusion

La méthode de back-allocation que l'on utilise permet de corriger individuellement les débits de chaque puits

Une anomalie a été détectée dans le puits *MD409*. Le débit indique une variation importante et inhabituelle par rapport aux jours précédents. Cette anomalie a été identifiée à la fois par l'analyse du débit du puits en fonction du temps et par l'utilisation du z-score et le modèle *LSTM*. Ces résultats convergents confirment la présence d'un problème ou d'une perturbation réelle dans le puits à cette date.

## Conclusion générale

Dans ce mémoire, nous nous attelons sur le développement de modèles prédictifs plus précis pour estimer le débit d'huile pour les puits pétroliers

La corrélation de *Gilbert* adaptée s'est révélée plus performante en termes de prédiction par rapport aux autres corrélations. Cette constatation est clairement évidente lors de l'étude de sensibilité des paramètres, où les variables suivantes ont montré une influence plus significative sur les performances du modèle *ANN* et random forest :  $P_t$ ,  $\phi$  et *GOR*. Ces paramètres jouent un rôle crucial dans la détermination du débit critique

Pour le régime non critique *L'extension Gilbert\_1* en prenant en compte la chute de pression a conduit à une amélioration des prédictions du débit. Cette amélioration est clairement observée dans l'étude de sensibilité des paramètres, où les variables  $P_t$ ,  $\phi$ ,  $dp$  et *GOR* semblent exercer une influence plus importante sur les performances du modèle *ANN* et de random forest

Une autre observation importante est que l'adaptation de ces modèles pour plusieurs intervalles de *GOR* dans les deux régimes d'écoulement a conduit au développement de modèles plus représentatifs par rapport aux modèles théoriques initialement ajustés sur l'ensemble des données.

Cependant, il convient d'être attentif dans les situations où le *GOR* est inférieur à 4000 et supérieur à 100000 pour le régime critique, ainsi que lorsqu'il est inférieure de 1000 pour le régime non critique, car le modèle peut être moins précis dans ces plages de valeurs.

L'étude de chaque puits individuellement en utilisant les modèles *LSTM* a révélé qu'il était possible d'augmenter encore davantage le degré de précision des prédictions. Cette approche a permis de prendre en compte les caractéristiques spécifiques de chaque puits et les dépendances temporelles des données de débit.

La nouvelle approche de back-allocation que nous utilisons permet une correction individualisée des débits pour chaque puits, ce qui améliore la précision des allocations. Cette méthode nous permet d'ajuster les débits de production spécifiquement pour chaque puits, Cette méthode offre une solution efficace en remplacement des mesures individuelles coûteuses qui ne sont pas accessibles

Le processus de calcul des coefficients  $k$  nous permet de quantifier la différence entre les débits mesurés et les débits calculés en utilisant l'historique des puits. Cela nous permet de déterminer un ajustement spécifique pour chaque puits en se basant sur leur historique respectif.

En utilisant soit le z-score, soit les modèles *LSTM*, nous sommes en mesure de détecter les anomalies dans les débits de pétrole par rapport aux jours précédents. Grâce à ces méthodes améliorées, nous sommes en mesure d'identifier les variations significatives dans les données de production, ce qui nous permet de détecter de manière fiable les anomalies de débit de

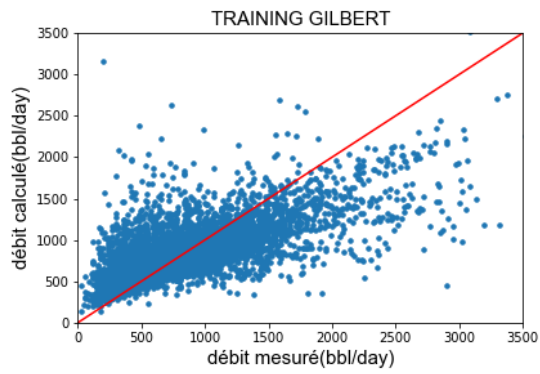
pétrole. L'utilisation de ces techniques nous aide à prendre des mesures appropriées pour résoudre les problèmes et maintenir une production optimale.

## REFERENCE

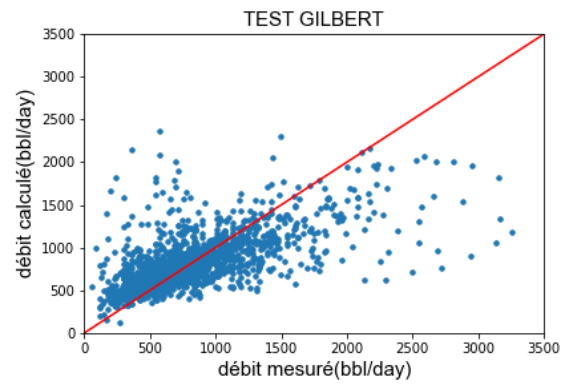
- [1] Dr. James P. Brill and Dr.H.Dale Beggs. « Two-Phase Flow In Pipes ». sixth Edition. January.1991
- [2] Harvey J. Motulsky and Lennart A. Ransnas « Fitting curves to data using nonlinear regression: a practical and nonmathematical review »
- [3] Boyun Guo. Ph.D. University of Louisiana at Lafayette . William C. Lyons. Ph.D. New Mexico Institute of Mining and Technology . Ali Ghalambor. Ph.D. University of Louisiana at Lafayette «Petroleum Production Engineering A Computer-Assisted Approach»
- [4] J. Bouré *et al.* « Two-phase Flow and Heat Transfer in the Power and Process Industries » (1981)
- [5] John R. Fanchi and Richard L. Christiansen « Introduction to Petroleum Engineering »
- [6] Ragnhild Krogsrud Haug « Multiphase Flow Through Chokes»
- [7] Håvard Devold « Oil and gas production handbook An introduction to oil and gas production. transport. refining and petrochemical industry»
- [8] Maxim Zubarev and Denis Zubarev. Amni International Petroleum Resources « Use of Radial Basis Function Networks for Efficient Well Production Allocation» SPE-198860-MS
- [9] panelAnirbid Sircar. Kriti Yadav. Kamakshi Rayavarapu. Namrata Bist. Hemangi Oza « Application of machine learning and artificial intelligence in oil and gas industry »
- [10] Frank Jahn. Mark Cook and Mark Graham TRACS International Consultancy Ltd. Aberdeen. UK « HYDROCARBON EXPLORATION AND PRODUCTION 2ND EDITION »
- [11] Sunday Kanshio « A review of hydrocarbon allocation methods in the upstream oil and gas industry » Journal of Petroleum Science and Engineering (2019)
- [12] *Hwang. R.J.; Baskin. D.K.; Teerman. S.C. (December 2000). « Allocation of commingled pipeline oils to field production ». Organic Geochemistry.*
- [13] <https://vitrinelinguistique.oqlf.gouv.qc.ca/fiche-gdt/fiche/8423681/tete-de-puits>
- [14] Manohar Swamynathan « Mastering Machine Learning with Python in Six Steps »
- [17] <https://rpubs.com/alexvezeau/SCI1035-CH1>
- [18] Aurélien Géron « Hands-On Machine Learning with Scikit-Learn and TensorFlow »
- [19] <https://www.analyticsvidhya.com/blog/2022/01/tutorial-on-rnn-lstm-gru-with-implementation/3>

- [20] <https://www.oracle.com/ch-fr/artificial-intelligence/machine-learning/what-is-machine-learning/>
- [21] <https://builtin.com/data-science/random-forest-algorithm>
- [22] <https://levelup.gitconnected.com/random-forest-regression-209c0f354c84>
- [23] D. Abdelmoula;J. E. Montiel Dunlop «Qatargas Well Production Back Allocation ProjectMethodology & System Requirements . »
- [24] <https://htt.io/resources/scada-basics/>
- [25] <http://www.ikonet.com/fr/ledictionnairevisuel/energies/geothermie-et-energie-fossile/petrole/reservoirs-2.php>
- [26]<https://jlcataj.gob.mx/cid-174/different-types-of-deep-learning-2-29-68-720255219.html>
- [27] <https://www.section.io/engineering-education/introduction-to-random-forest-in-machine-learning/decision-tree-nodes.png>

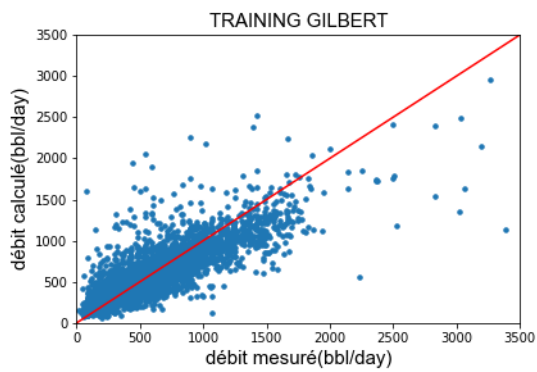
## Annexe



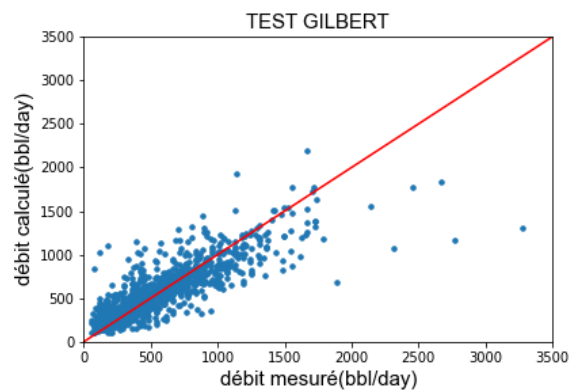
**Figure AN 1 cross plot training Gilbert( $GOR \leq 4000$ ), régime critique**



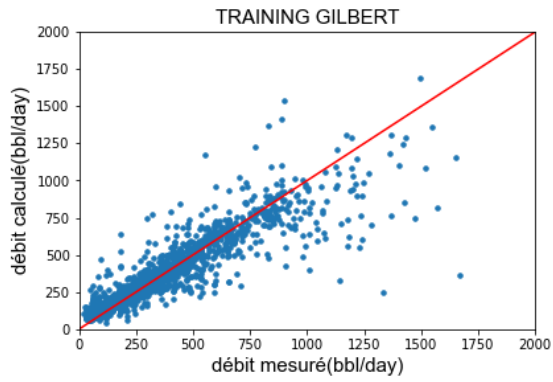
**Figure AN 2 cross plot test Gilbert( $GOR \leq 4000$ ), régime critique**



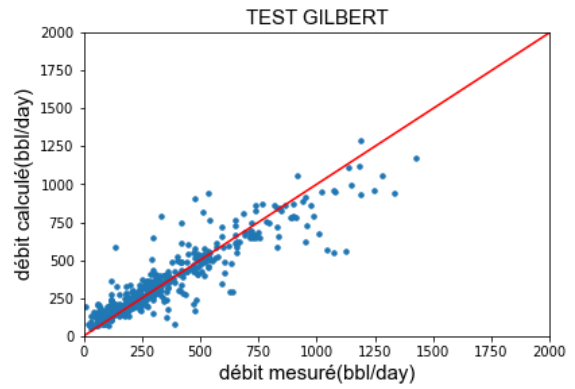
**Figure AN 3 cross plot training Gilbert( $4000 < GOR \leq 15000$ ), régime critique**



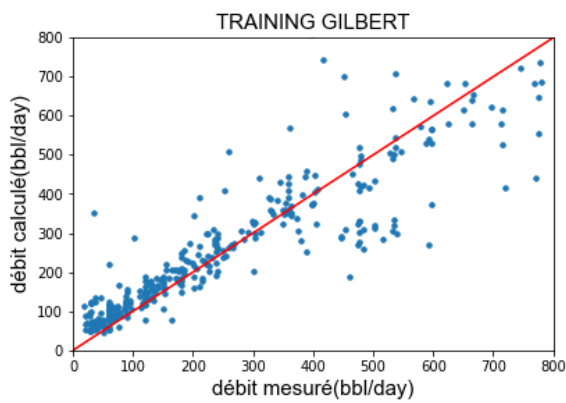
**Figure AN 4 cross plot test Gilbert( $4000 < GOR \leq 15000$ ), régime critique**



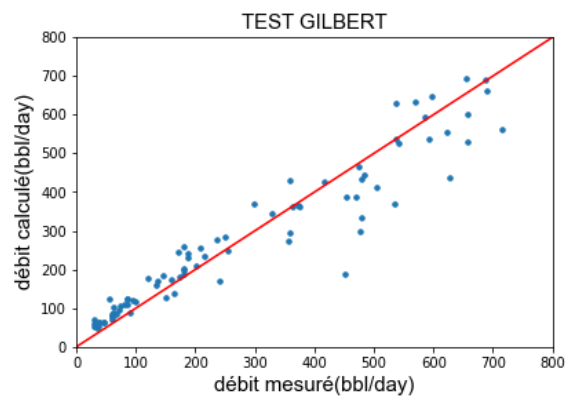
**Figure AN 5 cross plot training Gilbert(15000 < GOR ≤ 35000) ,régime critique**



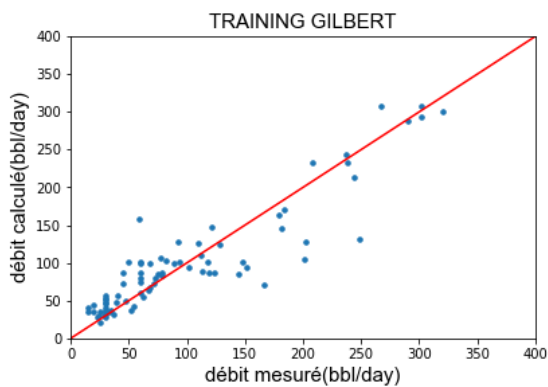
**Figure AN 6 cross plot test Gilbert(15000 < GOR ≤ 35000) ,régime critique**



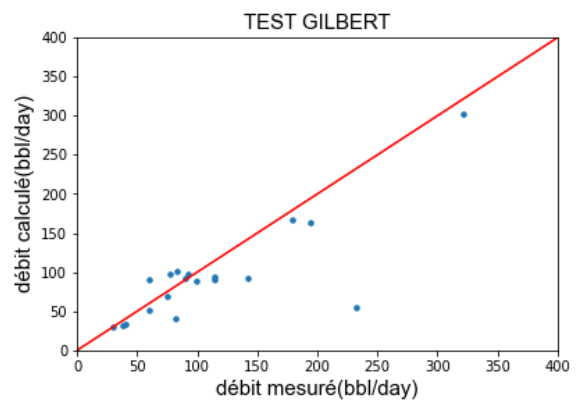
**Figure AN 7 cross plot training Gilbert(35000 < GOR ≤ 60000) ,régime critique**



**Figure AN 8 cross plot test Gilbert(35000 < GOR ≤ 60000) ,régime critique**

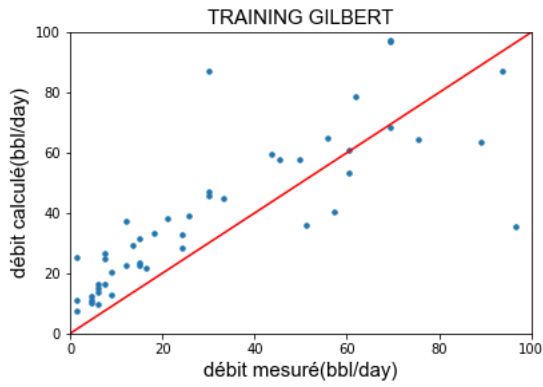


**Figure AN 9 cross plot training Gilbert(60000 < GOR ≤ 100000) ,régime critique**

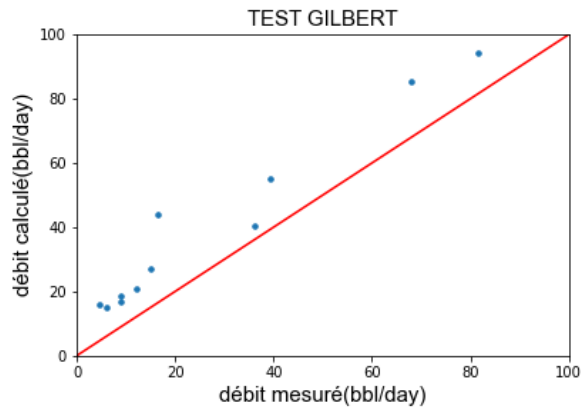


**Figure AN 10 cross plot test Gilbert(60000 < GOR ≤ 100000) ,régime critique**

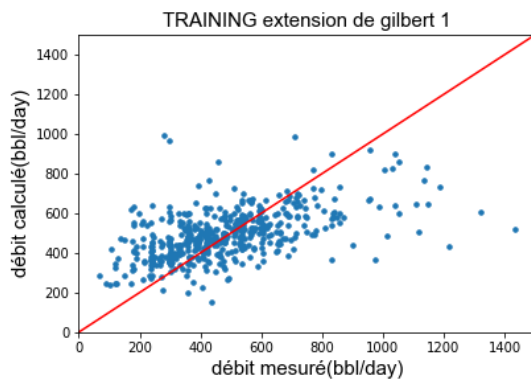




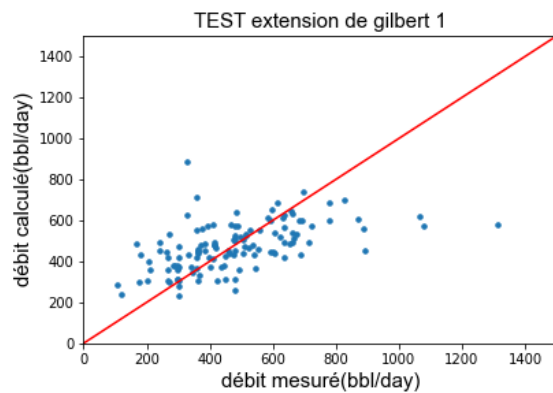
**Figure AN 11 cross plot training Gilbert( $GOR > 100000$ ), régime critique**



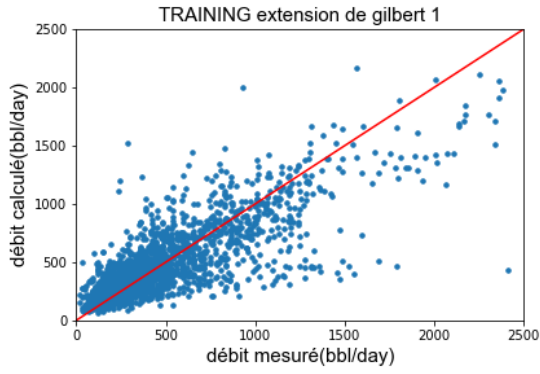
**Figure AN 12 cross plot test Gilbert( $GOR > 100000$ ), régime critique**



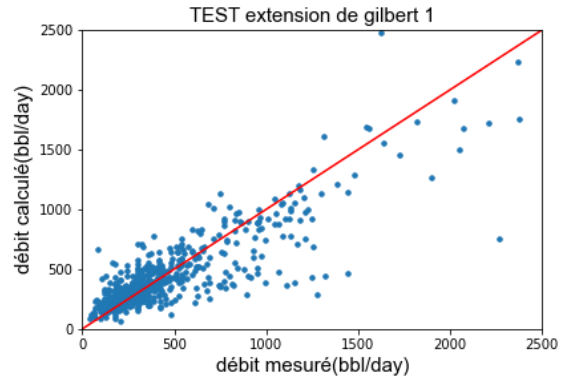
**Figure AN 13 cross plot training Gilbert( $GOR \leq 1000$ ), régime sous critique**



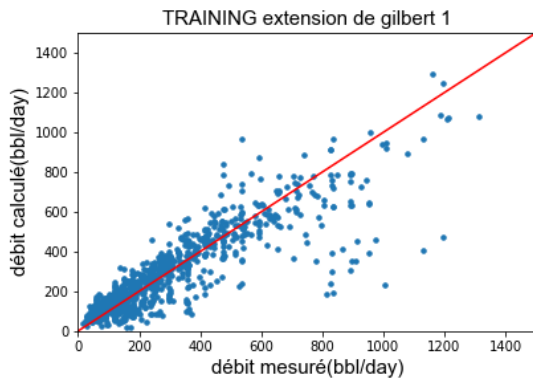
**Figure AN 14 cross plot test Gilbert( $GOR \leq 1000$ ), régime sous critique**



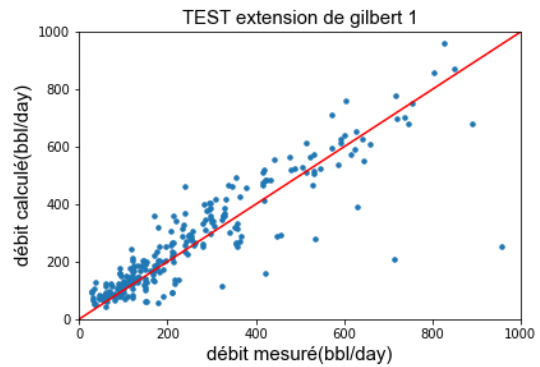
**Figure AN 15 cross plot training Gilbert( $1000 < GOR \leq 10000$ ), régime sous critique**



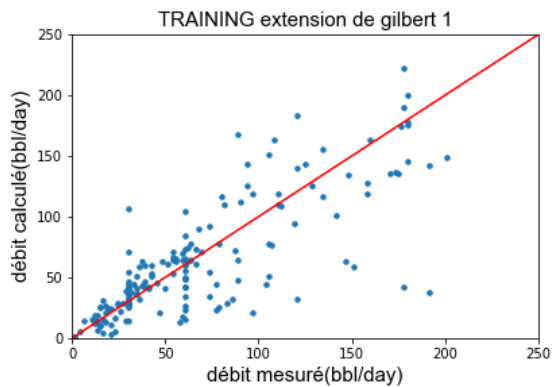
**Figure AN 16 cross plot test Gilbert( $1000 < GOR \leq 10000$ ), régime sous critique**



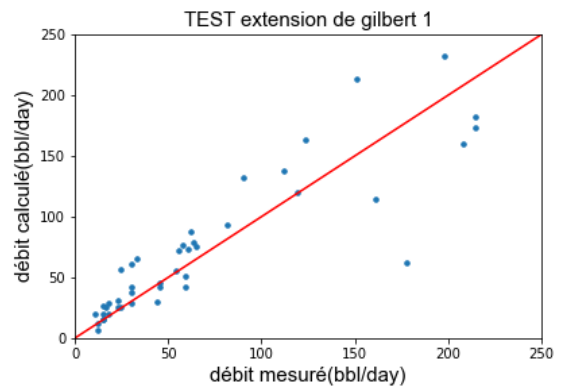
**Figure AN 17 cross plot training Gilbert( $10000 < GOR \leq 30000$ ), régime sous critique**



**Figure AN 18 cross plot test Gilbert( $10000 < GOR \leq 30000$ ), régime sous critique**



**Figure AN 19 cross plot training Gilbert( $GOR > 30000$ ), sous régime critique**



**Figure AN 20 cross plot test Gilbert( $GOR > 30000$ ), régime sous critique**