

N° Ordre ...../Faculté/UMBB/2021

REPUBLIQUE ALGERIENNE DEMOCRATIQUE 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'étude

En vue de l'obtention du diplôme :

**MASTER**

Présenté par

**Thileli HADJARI**

**Amel HADJADJ**

Filière : Hydrocarbures

Option : Génie pétrolier : Production des hydrocarbures.

**Thème**

---

**Monitoring des Puits du Champ de Hassi Messaoud par  
l'Application des Algorithmes de Machine Learning**

---

Devant le jury :

BENYOUNES Khaled	Professeur	UMBB	Président
GARECHE Mourad	Professeur	UMBB	Examineur
SAIFI Redha	MAA	UMBB	Examineur
ZERAIBI Noureddine	Professeur	UMBB	Encadreur
REDOUANE Kheireddine	Docteur	CRD	Co-encadreur

Année universitaire : 2020/2021

## Remerciements

Nous tenons à remercier ALLAH, le clément et miséricordieux de nous avoir donné la force et la patience pour accomplir ce modeste travail.

Nous remercions nos parents et familles pour leurs soutien et encouragement au cours de notre parcours.

Nous voudrions exprimer nos vifs remerciements à notre promoteur Pr. ZERAIBI Noureddine pour sa disponibilité, son sérieux et ses conseils judicieux.

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

Nous tenons aussi à remercier l'ensemble des enseignants de la FHC qui ont contribué à notre formation.

Enfin, tous nos remerciements vont à nos amis de la FHC qui ont été nombreux à nous aider et à nous encourager.

## Résumé

Ce projet de fin d'étude comprend le développement de plusieurs modèles pour l'estimation du débit d'huile à travers les duses au champ de Hassi Messaoud, et ce par l'adaptation de deux approches, à savoir :

- La modélisation semi empirique.
- L'intelligence artificielle.

Lors du recours à ces deux approches, certaines corrélations existantes dans la littérature, sont adaptées à des régimes spécifiques, de même la meilleure équation de chaque régime a été réadaptée sur chaque intervalle de GOR.

Concernant cette problématique, divers modèles intelligents ont été réalisés pour représenter le régime critique et non critique où il a été montré que les réseaux de neurones (ANN) présentent les meilleures performances sur l'ensemble des données de test. En su de cette approche, une autre alternative a été abordée à travers laquelle chaque puits est traité individuellement en faisant appel à :

- La programmation génétique GP.
- Les modèles LSTM.

Effectivement cette dernière alternative a révélé des résultats fortement intéressants et qui reste un domaine à explorer minutieusement.

**Mots clés :** Ecoulements diphasiques, Duse, Estimation de la production, Régime critique et sous critique, Apprentissage automatique, HMD.

## Abstract

This project includes the development of several models for the estimation of oil flow rate through chokes in the Hassi Messaoud field, by adapting two approaches, namely:

- Semi-empirical modeling.
- Artificial intelligence.

When using these approaches, some existing correlations in the literature are adapted to specific regimes, and the best equation of each regime has been readapted on each interval of GOR.

Concerning this problem, various intelligent models have been created to represent the critical and non-critical regime where it has been shown that neural networks (ANN) present the best performances on the whole test data. In addition to this approach, another alternative has been addressed through which each well is treated individually using:

- Genetic Programming GP.
- LSTM models.

Indeed, this last alternative has revealed very interesting results and remains an area to be explored in detail.

**Key words:** Two-phase flows, Choke, Production estimation, Critical and sub-critical regime, Machine learning, HMD

# Table des matières

Remerciements .....	i
Résumé .....	ii
Abstract.....	ii
Introduction .....	1
Revue de littérature .....	3
<b>Chapitre 1: Les écoulements diphasiques .....</b>	<b>7</b>
1.1 Equation de gradient de pression .....	7
1.2 Ecoulements monophasiques .....	8
1.3 Ecoulements diphasiques .....	9
1.3.1 Paramètres descriptifs.....	9
Le Holdup : .....	9
Masse Volumique : .....	10
La vitesse : .....	11
La viscosité : .....	11
Tension superficielle de liquide : .....	12
1.3.2 Configurations d'écoulements diphasiques .....	12
Ecoulement vertical : .....	12
Ecoulements horizontal : .....	13
1.3.3 Modélisation des écoulements diphasiques : .....	14
Catégorie A (modèle homogène).....	15
Catégorie B .....	16
Catégorie C .....	17
<b>Chapitre 2: Modélisation des écoulements à travers les restrictions et calcul de production.....</b>	<b>21</b>
2.1 Modélisation des écoulements à travers les restrictions.....	21
2.1.1 Généralités sur les pertes de charges locales.....	21
2.1.2 Ecoulements à travers les duses.....	25
2.1.2.1 Description de la duse.....	25
2.1.2.2 Régime d'écoulements : critique et sous-critique : .....	26
2.1.2.3 Modélisation des écoulements dans les restrictions en utilisant les corrélations existantes dans la littérature.....	28
2.2 Calcul de production : .....	34
2.2.1 Système de production : .....	34
2.2.2 Principe de calcul de production.....	35

2.2.3	Analyse nodal .....	36
2.2.4	Performances des réservoirs : .....	38
2.2.5	Performances du tubing : .....	39
<b>Chapitre 3 : Outils mathématiques.....</b>		<b>41</b>
3.1	Préparation des données : .....	41
3.2	Régression non linéaire et algorithme de Levenberg Marquardt .....	43
3.3	Introduction au machine learning : .....	44
3.3.1	Définition du machine learning : .....	44
3.3.2	Les types de machine learning : .....	44
3.3.3	Quelques notions de machine learning : .....	45
	Underfitting et overfitting : .....	45
	Fonctions d'activation : .....	46
3.4	Réseaux de neurones : .....	47
3.4.1	Réseau de Neurone à un seul Neurone : Le perceptron .....	47
3.4.2	Multilayer perceptron MLP (réseau à plusieurs neurones) : .....	48
3.4.3	L'entraînement d'un réseau de neurones : .....	49
3.5	Réseau de neurones récurrent (RNN) .....	49
	Types des RNNs.....	50
	Modèles Long Short Term Memory (LSTM) .....	50
3.6	Support Vector Machine Regression SVR : .....	53
3.7	Gaussian process regression GPR (Processus gaussiens).....	54
3.8	Programmation génétique et régression symbolique : .....	55
	Régression symbolique multigénique (Multigene Symbolic Regression) : .....	55
3.9	Métriques d'évaluation d'un modèle.....	56
	Le coefficient de détermination $R^2$ : .....	57
	La racine de l'erreur quadratique moyenne RMSE : .....	57
	L'erreur absolue moyenne MAE : .....	57
	Erreur relative absolue RAE : .....	57
<b>Chapitre 4 : Résultats et discussions.....</b>		<b>58</b>
4.1	Validation de la méthodologie en utilisant les données du champ de Reshadat : .....	58
4.1.1	Description du dataset du champ de Reshadat : .....	58
4.1.2	Evaluation des modèles théoriques : .....	59
4.1.3	Adaptation du modèle de Gilbert et création de modèles intelligents : .....	61
4.1.4	Conclusion I : .....	61
4.2	Présentation du dataset de production du champ de Hassi Messaoud : .....	64
4.2.1	Description du dataset brut : .....	64
4.2.2	Préparation des données : .....	64

Résultats et discussions.....	65
4.3 Evaluation des modèles théoriques : .....	65
4.4 Adaptation des modèles théoriques et création de modèles intelligents : .....	69
4.4.1 Méthodologie : .....	69
4.4.2 Nouveaux modèles : résultats et discussion.....	70
4.4.2.1 Régime critique.....	70
4.4.2.2 Régime non critique.....	77
4.4.3 Conclusion 2.....	90
4.5 Adaptation suivant les intervalles de GOR : .....	91
4.5.1 Régime critique : .....	91
4.5.1.1 Test sur différents intervalles de GOR : .....	91
4.5.1.2 Adaptation sur différents intervalles de GOR : .....	92
4.5.1.3 Comparaison entre le modèle ANN et les modèles adaptés sur différents intervalles de GOR : .....	94
4.5.2 Régime non critique.....	95
4.5.2.1 Test sur différents intervalles de GOR.....	95
4.5.2.2 Adaptation de l'extension de Gilbert sur différents intervalles de GOR.....	95
4.5.2.3 Comparaison entre le modèle ANN et les modèles adaptés sur différents intervalles de GOR.....	97
4.5.3 Conclusion 3.....	98
4.6 Modélisation par puits via la programmation génétique .....	98
4.7 Utilisation des modèles LSTM pour la prédiction du débit : .....	100
4.8 Influence de l'équation de Gilbert adaptée sur le point de fonctionnement d'un puits : .....	102
4.8.1 Conclusion 4:.....	104
Conclusion et recommandations.....	105
Références.....	xiv
Annexe A : graphes de la partie théorique.....	xvi
Annexe B : cross plot après subdivision en intervalles GOR .....	xix
Annexe C : cross plot et équation de la programmation génétique .....	xxxii

# Liste des figures

Figure 1.1 : Diagramme de Moody représentant le facteur de friction en fonction du nombre de Reynolds et de la rugosité relative pour les tuyaux ronds. [15].....	9
Figure 1.2: configurations d'écoulements dans une conduite verticale .....	13
Figure 1.3: configurations d'écoulements dans une conduite horizontale.....	14
Figure 2.1: Illustration des différents paramètres influençant sur la perte de charge à travers une conduite et différence entre une surface lisse (smooth) et rugueuse (rough wall) [15].....	21
Figure 2.2: écoulement à travers une vanne.....	24
Figure 2.1: Illustration des différents paramètres influençant sur la perte de charge à travers une conduite et différence entre une surface lisse (smooth) et rugueuse (rough wall) [15].....	20
Figure 2.2: écoulement à travers une vanne.....	22
Figure 2.3: Quelques expressions du coefficient de pertes de charge singulières K pour des géométries d'écoulements couramment utilisées. [16].....	23
Figure 2.4: schéma d'écoulement à travers une duse.....	24
Figure 2.5: Courbe typique de performance d'une duse [13].....	26
Figure 2.6: un système de production simple avec trois puits de pétrole à écoulement naturel vers séparateurs [23].....	33
Figure 2.7: Système de production d'un puits simple.....	34
Figure 2.8: Exemple de configuration d'analyse nodale avec un nœud au fond du puits [36].....	35
Figure 2.9: Tracé de la pression du nœud en fonction du débit et illustration du point de fonctionnement.....	36
Figure 2.10: IPR en utilisant la méthode de l'indice de productivité.....	37
Figure 2.12: illustration de différentes courbes TPR pour différent diamètres de tubing.....	38
Figure 2. 11: IPR pour un réservoir sur saturé.....	38
Figure 3. 1: schéma représentant les types de machine Learning.....	43
Figure 3. 2: Représenttion des fonctions d'activation.....	45
Figure 3. 3: Exemple d'un perceptron qui reçoit 3 entrées.....	45
Figure 3. 4: représentation d'un réseau de neurones multicouches.....	46
Figure 3. 5: schéma du principe d'entraînement d'un réseau de neurones.....	47
Figure 3. 6: types des RNNs.....	48
Figure 3. 7: Représentation d'une cellule LSTM.....	48
Figure 3. 8: Représentation de la porte d'oubli (forget gate) d'une cellule LSTM.....	49
Figure 3. 9: Représentation de la porte d'entrée (input gate) d'une cellule LSTM.....	50
Figure 3. 10: Représentation de la porte de sortie (output gate) d'une cellule LSTM.....	50
Figure 3. 11: support vector machine regression.....	51
Figure 3. 12: ajustement GPR d'une observation bruyante [35].....	52
Figure 3. 13: modèle multigénique de la sortie y avec des entrées $x_1$ , $x_2$ et $x_3$ .....	54

Figure 4. 1: cross plot de test de la corrélation de Gilbert au champ de Reshadat .....	58
Figure 4. 2: cross plot de test du modèle de Ros au champ de Reshadat.....	59
Figure 4. 3: cross plot de test du modèle de Pilehvari au champ de Reshadat.....	59
Figure 4. 4: cross plot de test du modèle de Baxendell au champ de Reshasat.....	59
Figure 4. 5: cross plot de test du modèle de Achong au champ de Resadat.....	59
Figure 4.6: cross plot de training du modèle de Gilbert adapté au champ de Reshadat.....	61
Figure 4.7: cross plot de test du modèle de Gilbert adapté au champ de Reshadat.....	61
Figure 4. 8: cross plot du trainng du modèle GPR au champ de Reshadat.....	61
Figure 4.9: cross plot du test du modèle GPR au champ de Reshadat.....	61
Figure 4. 10; cross plot de training de modèle ANN au champ de Reshadat.....	62
Figure 4. 11: cross plot de test d modèle ANN au champ de Reshadat.....	62
Figure 4. 12: cross plot de training de modèle SVR au champ de Reshadat.....	62
Figure 4. 13: cross plot de test de modèle SVR au champ de Reshadat.....	62
Figure 4. 14: cross plot du test de la corrélation de Achong.....	66
Figure 4. 15: cross plot du test de la corrélation de Gilbert (régime critique).....	66
Figure 4. 16: cross plot du test de la corrélation de Baxendell.....	66
Figure 4. 17: cross plot du test de la corrélation de Ros.....	66
Figure 4. 18: cross plot du test de la corrélation de Pilehvari.....	67
Figure 4. 19: cross plot du test de la corrélation de Towailib et Marhoun.....	67
Figure 4. 20: cross plot du test de la corrélation de Gilbert au régime non critique.....	67
Figure 4. 21: cross plot du test de la corrélation de Elgibaly et Nashawi.....	67
Figure 4. 22: cross plot du test de l'extension de la corrélation de Gilbert.....	68
Figure 4. 23: cross plot de training du modèle Gilbert_HMD 1.....	70
Figure 4. 24: cross plot de test du modèle Gilbert_HMD 1.....	70
Figure 4. 25: cross plot de training du modèle Gilbert_HMD 2.....	71
Figure 4. 26: cross plot de test du modèle Gilbert_HMD 2.....	71
Figure 4. 27: cross plot de training du modèle SVR.....	71
Figure 4. 28: cross plot de test du modèle SVR.....	71
Figure 4. 29: cross plot de training du modèle GPR.....	72
Figure 4. 30: cross plot de test du modèle GPR.....	72
Figure 4. 31: cross plot de training du modèle ANN.....	72
Figure 4. 32: cross plot de test du modèle ANN.....	72
Figure 4. 33: cross plot de training du modèle de Towailib & M. _HMD.....	74
Figure 4. 34: cross plot de test du modèle de Towailib & M. _HMD.....	74
Figure 4. 35: cross plot de training du modèle SVR.....	74
Figure 4. 36: cross plot de test du modèle SVR.....	74
Figure 4. 37: cross plot de training du modèle GPR.....	75

Figure 4. 38: cross plot de test du modèle GPR.....	75
Figure 4. 39: cross plot de training du modèle ANN.....	75
Figure 4. 40: cross plot de test du modèle ANN.....	75
Figure 4. 41: cross plot de training du modèle Gilbert NC_HMD1.....	77
Figure 4. 42: cross plot de test du modèle Gilbert NC_HMD 1.....	77
Figure 4. 43: cross plot de training du modèle Gilbert NC_HMD 2 au régime non critique.....	78
Figure 4. 44: cross plot de test du modèle Gilbert NC_HMD 2 au régime non critique.....	78
Figure 4. 45: cross plot de training du modèle SVR.....	78
Figure 4. 46: cross plot de test du modèle SVR.....	78
Figure 4. 47: cross plot de training du modèle GPR.....	79
Figure 4. 48: cross plot de test du modèle GPR.....	79
Figure 4. 49: cross plot de training du ANN.....	79
Figure 4. 50: cross plot de test du modèle ANN.....	79
Figure 4. 51: cross plot de training du modèle de Elgibaly & Nashawi . _HMD.....	81
Figure 4. 52: cross plot de test du modèle de Elgibaly & Nashawi . _HMD.....	81
Figure 4. 53: cross plot de training du modèle SVR.....	81
Figure 4. 54: cross plot de test du modèle SVR.....	81
Figure 4. 55: cross plot du training du modèle GPR.....	82
Figure 4. 56: cross plot de test du modèle GPR.....	82
Figure 4. 57: cross plot de training du modèle ANN.....	82
Figure 4. 58: cross plot de test du modèle ANN.....	82
Figure 4. 59: cross plot de training du modèle extension _HMD.....	84
Figure 4. 60: cross plot de test du modèle extension Gilbert _HMD.....	84
Figure 4. 61: cross plot de training du modèle SVR.....	84
Figure 4. 62: cross plot de test du modèle SVR.....	84
Figure 4. 63: cross plot de training du modèle GPR.....	85
Figure 4. 64: cross plot de test du modèle GPR.....	85
Figure 4. 65: cross plot de training du modèle ANN.....	85
Figure 4. 66: cross plot de test du modèle ANN.....	85
Figure 4. 67: cross plot de training du modèle Extension 1_HMD.....	87
Figure 4. 68: cross plot de test du modèle Extension 1_HMD.....	87
Figure 4. 69: cross plot du training de modèle Extension 2_HMD.....	87
Figure 4. 70: cross plot de test du modèle Extension 2_HMD.....	87
Figure 4. 71: cross plot de training du modèle SVR.....	88
Figure 4. 72: cross plot de test du modèle SVR.....	88
Figure 4. 73: cross plot de training du modèle GPR.....	88
Figure 4. 74: cross plot de test du modèle GPR.....	88

Figure 4. 75: cross plot de training de ANN.....	89
Figure 4. 76: cross plot de test de ANN.....	89
Figure 4. 77: Evolution du coefficient a en fonction de l'intervalle de GOR (régime critique).....	92
Figure 4. 78: Evolution du coefficient b en fonction de l'intervalle de GOR (régime critique) .....	92
Figure 4. 79: Evolution du coefficient e en fonction de l'intervalle de GOR (régime critique) .....	92
Figure 4. 80: Evolution du coefficient c en fonction de l'intervalle de GOR (régime critique) .....	92
Figure 4. 81: Evolution du coefficient b en fonction de l'intervalle de GOR (régime non critique) .....	95
Figure 4. 82: Evolution du coefficient a en fonction de l'intervalle de GOR (régime non critique).....	95
Figure 4. 83: Evolution du coefficient c en fonction de l'intervalle de GOR (régime non critique) .....	95
Figure 4.84: Evolution du coefficient e en fonction de l'intervalle de GOR (régime non critique) .....	95
Figure 4.85: Evolution du coefficient f en fonction de l'intervalle.....	95
Figure 4. 86: Tracé du profil de production réel et prédit du puits Well1289.....	100
Figure 4. 87: cross plot du training de LSTM.....	100
Figure 4. 88: cross plot du training de LSTM.....	100
Figure 4. 89: modèle puits en PIPESIM.....	101
Figure 4.90: points de fonctionnement où le diamètre de la duse est 1.019 dans le cas de Gilbert_HMD 2.....	103
Figure 4. 91: points de fonctionnement où le diamètre de la duse est 1.019 dans le cas de Gilbert.....	103

## Liste des tableaux

Tableau 2.1 : constantes de la corrélation de Gilbert suivant différents auteurs [19],[20] .....	30
Tableau 2.2 : les coefficients de la corrélation de Omana .....	30
Tableau 2.3 : Les coefficients de la corrélation de Towailib et Marhoun .....	31
Tableau 2.4 : les coefficients de la corrélation de Attar et Abdul Majeed .....	32
Tableau 2.5 : les coefficients de Elgibaly et Nashawi .....	32
Tableau 2.6 : les coefficients de l'extension de l'équation de Gilbert .....	33
Tableau 4.1 : Différente plage des paramètres des données de champ Reshadat.....	58
Tableau 4.2 : statistiques de test des modèles de régime critique.....	59
Tableau 4.3 : les nouveaux coefficients de l'équation de Gilbert adaptée au champ de Reshadat.....	61
Tableau 4.4 : statistiques de training et de test du modèle théorique de Gilbert adapté au champ Reshadat et des modèles intelligents.....	61
Tableau 4.5 : Différentes plages des paramètres de données du champ de HMD.....	65
Tableau 4.6 : statistiques de test des modèles théoriques au régime critique (HMD).....	66
Tableau 4.7 : statistiques de test des modèles théoriques au régime non critique (HMD).....	66
Tableau 4.8 : les nouveaux coefficients de l'équation de Gilbert adaptée au champ de HMD (Régime critique).....	70
Tableau 4.9 : statistiques de training et de test du modèle théorique de Gilbert adapté et des modèles intelligents (Régime critique).....	70
Tableau 4.10 : les nouveaux coefficients de l'équation de Towailib et Marhoun adaptée au champ de HMD.....	74
Tableau 4.11 : statistiques de training et de test du modèle théorique de Towailib et Marhoun adapté et des modèles intelligents.....	74
Tableau 4.12 : les nouveaux coefficients de l'équation de Gilbert adaptée au champ de HMD (Régime non critique).....	77
Tableau 4.13 : statistiques de training et de test du modèle théorique de Gilbert adapté au champ de HMD et des modèles intelligents (Régime non critique).....	77
Tableau 4.14 : les nouveaux coefficients de l'équation de Elgibaly et Nashawi adaptée au champ de HMD.....	81
Tableau 4.15 : statistiques de training et de test du modèle de Elgibaly et Nashawi adapté au champ de HMD et des modèles intelligents.....	81
Tableau 4.16 : les nouveaux coefficients de l'équation de l'extension de Gilbert adaptée au champ de HMD.....	84
Tableau 4.17 : statistiques de training et de test du modèle extension de Gilbert adapté et des modèles intelligents.....	84
Tableau 4.18 : les nouveaux coefficients des extensions utilisées à HMD.....	87
Tableau 4.19 : statistiques de training et de test des extensions utilisées à HMD adaptées et des modèles intelligents.....	87
Tableau 4.20 : résultats de test du modèle Gilbert_HMD 2 sur différents intervalles de GOR.....	91

Tableau 4. 21 : nouveaux coefficients du modèle de Gilbert adapté sur chaque intervalle de GOR (régime critique).....	92
Tableau 4. 22 : résultats de training et de test des modèles adaptés pour chaque intervalle de GOR (régime critique).....	92
Tableau 4. 23 : résultats de test du modèle de Gilbert adapté sur chaque intervalle GOR et le test du modèle ANN.....	94
Tableau 4. 24 : résultats de test du modèle Extension Gilbert _HMD sur différents intervalles de GOR régime non critique.....	95
Tableau 4. 25 : nouveaux coefficients de l'extension de Gilbert adaptée sur chaque intervalle de GOR (régime non critique).....	95
Tableau 4. 26 : résultats de training et de test des modèles adaptés pour chaque intervalle de GOR (régime non critique).....	97
Tableau 4. 27: résultats de test du modèle de Gilbert adapté sur chaque GOR et le test de test du modèle ANN.....	98
Tableau 4. 28 : statistiques de training et test de GP pour chaque puits avec différentes variables.....	99
Tableau 4. 29 : statistiques de training et de test du modèle LSTM.....	100
Tableau 4. 30 : le point de fonctionnement en utilisant l'équation de Gilbert originale et celle adaptée à HMD.....	103



# Nomenclature

$A_o$	Section d'écoulement dans la duse , ft <sup>2</sup>
$B_o$	Facteur volumétrique d'huile, bbl/STB
$B_w$	Facteur volumétrique d'eau, bbl/STB
$C$	Coefficient d'écoulement
$C_d$	Coefficient de décharge
$D$	Diamètre intérieur du tube, ft
$d$	Diamètre de l'orifice, ft
$f_{2F}$	Facteur de friction de Fanning pour l'écoulement diphasique.
$GOR$	Rapport gaz/huile de production, scf /STB
$GLR$	Rapport gaz/ liquide produit scf/STB
$HMD$	Hassi Messaoud
$M$	Masse totale associée à 1 STB d'huile
$MPFM$	Multiphase flow meter
$N_d$	Nombre de diamètre de pipe (Pipe Diameter Number)
$N_{gv}$	Nombre de vitesse du gaz (Gas Velocity Number),
$N_L$	Nombre de viscosité du liquide (Liquid Viscosity Number)
$N_{Lv}$	Nombre de vitesse du liquide (liquid velocity number)
$OPR$	Outflow performance relationship
$P$	Pression in situ, psia
$P_{down}, P_{outlet}$	Pression en aval de la duse (downstream pressure), psia
$P_u$	Pression en amont de la duse (upstream pressure), psia
$q_g$	Débit de gaz,
$Q, q_o$	Débit d'huile, bbl/day
$q$	Débit du fluide, ft <sup>3</sup> /ses
$R$	Rapport gaz/ huile aux conditions standards, scf/STB

$R_L$	Rapport gaz/liquide en amont de la duse
$R_p$	Rapport gaz/liquid produit, scf/STB
$R_s$	Rapport de solubilité, scf/STB
$T$	Température in situ, R
$t$	Température in situ, °F
$T_{up}$	Température en amont, R
$TPC$	La courbe de performance de tubing (Tubing performance curve)
$V_m$	Volume du mélange associé à 1 STB d'huile, ft <sup>3</sup>
$WOR$	Rapport eau/huile de production, bbl/STB
$Z$	Facteur de compressibilité du gaz à P et T définie
$Z_{up}$	Facteur de compressibilité au condition amont
$\Delta h$	Incrément de profondeur, ft
$\Delta P$	Différence de pression entre $P_{up}$ et $P_{down}$ , psia
$\rho$	Masse volumique du fluide, lbm/ft <sup>3</sup>
$\bar{\rho}$	La masse volumique moyenne du mélange, lb/ft <sup>3</sup>
$\gamma_g$	Densité de gaz
$\gamma_o$	Densité d'huile
$\gamma_w$	Densité de l'eau
$\gamma_m$	Densité de mélange
$\phi, d_{bean}$	Diamètre de la duse (1/64inch)
$\sigma_L$	Tension de surface (conditions en amont), dynes/cm
ANN	Artificial neural network
SVR	Support Vector Regresion
GPR	Gaussian Process Regression
GP	Genetic Programming
LSTM	Long Short Term Memory

# Introduction

L'estimation de la production d'huile provenant de chaque puits séparément est l'une des préoccupations majeures des ingénieurs de production pétrolière. Elle permet de comprendre d'où vient la production, et quelle fraction de pétrole exporté provient de quel puits. Avoir une bonne estimation des débits produits est crucial pour différentes raisons techniques et économiques, les plus importantes étant :

- Détermination de revenu financier de chaque partenaire dans le champ, qui n'est pas toujours le même pour tous les puits traités au même endroit.
- Une bonne estimation du débit de production combinée avec l'estimation des prix de pétrole est un outil très puissant d'aide à la décision lors de l'élaboration des contrats de vente.
- Le contrôle de la production (production monitoring) de chaque puits individuellement permet la détection de tout changement de profil de production, de composition du fluide ainsi que les profils de pression et température, ce qui est reconnu comme la meilleure façon d'optimiser les performances d'un champ pétrolier.
- La connaissance du débit de chaque puits séparément est l'un des éléments clés pour une gestion efficace des réservoirs afin de maximiser la récupération.

Il existe principalement trois approches utilisées pour la quantification de production : la mesure directe, la modélisation physique et l'approche data driven

Les mesures directes sont considérées comme les plus exactes, mais elles ont l'inconvénient d'être coûteuses, en plus ils perturbent la production d'un champ suite à la fermeture des puits ; de ce fait, généralement elles ne sont envisagées que périodiquement pour les opérations onshore.

La modélisation physique consiste principalement à faire appel aux principes fondamentaux de la physique du procès et à des expériences passives ou actives afin de fermer le système pour obtenir des corrélations hybrides pouvant décrire les écoulements à travers les duses. L'avantage majeur de ces modèles, c'est qu'ils sont simples à utiliser et à implémenter dans d'autres logiciels de simulation, ils permettent de donner une bonne estimation du débit d'une façon très économique. La complexité de cette méthode réside principalement dans la nature des écoulements multiphasiques des effluents produits (huile, gaz, eau, sable...) qui fait qu'aucune corrélation développée jusque-là ne peut être considérée comme ayant une application universelle. Des adaptations doivent être considérées à chaque fois en fonction de champ d'étude.

Enfin, l'approche data driven qui est relativement récente, se base sur l'utilisation combinée de données et des algorithmes de machine learning. Cette dernière se présente comme une sorte de débitmètre virtuel qui produit des résultats très intéressants lorsque les données sont à la fois suffisantes et représentatives.

Le problème de quantification de production est universel, et plusieurs auteurs de différents pays ont travaillé pour faire adapter les corrélations présentes dans la littérature à leurs zones d'études. Et puisqu'en Algérie le problème subsiste toujours, on va tenter d'apporter notre

contribution à travers ce projet de fin d'études. Dans le cadre de ce travail, on va s'intéresser à l'étude de l'un des plus grands champs pétrolier algériens, il s'agit du champ de Hassi Messaoud (HMD). Deux approches seront principalement utilisées, à savoir, la modélisation physique de la duse via des modèles semi-empiriques et l'approche data driven.

Notre travail va être structuré de la façon suivante :

- ❖ Une partie théorique qui comprend 3 chapitres. A travers le premier chapitre, nous allons essayer de mettre l'accent sur les concepts fondamentaux relatifs à la description des écoulements multiphasiques. Puis, dans le deuxième chapitre, nous allons aborder la modélisation des écoulements à travers les duses qui va nous servir de base dans la partie pratique. Finalement, les différents outils mathématiques utilisés dans la présente étude seront exposés dans le chapitre 3.
- ❖ Une partie pratique comprenant l'essentiel de notre travail et qu'on peut résumer dans les points suivant :
  - Validation de la méthodologie appliquée dans la première section spécifique à l'étude du champ de Hassi Messaoud en se basant sur les données d'un travail publié dans un article qui concerne l'étude du champ de Reshadat
  - Evaluation de quelques modèles théoriques décrivant les écoulements à travers les duses et étude de leur applicabilité au champ de HMD
  - Adaptation des modèles théorique précédemment évalués suivant les données du champ de HMD pour les deux régimes d'écoulement (critique et sous critique) et ce en utilisant l'algorithme de régression non linéaire de Levenberg-Marquardt, puis création de quelques modèles intelligents à savoir : support vector regression SVR, gaussian process regression GPR, et artificial neural network ANN
  - Effectuer une subdivision en intervalles de 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.
  - Tenter une approche à travers laquelle chaque puits sera traité individuellement et ce en appliquant la programmation génétique et les modèles LSTM.
  - Etude de sensibilité de choix de l'équation de la duse sur le point de fonctionnement d'un puits.

Ce travail est réalisé en utilisant les langages de programmation MATLAB, Python ainsi que le logiciel PIPESIM de Schlumberger.

## Revue de littérature

La duse est l'un des équipements les plus importants dans les systèmes de production de pétrole et de gaz. Une duse est une restriction essentiellement conçue pour contrôler les débits de production à travers le puits, imposer une contre-pression à la formation afin de la protéger des endommagements, empêcher les perturbations en aval de se propager en amont (écoulement critique) ainsi que pour protéger les installations de surface contre les pressions très élevées. [10]

Au cours de ces dernières décennies, de nombreux travaux et articles ont traité le problème de quantification de production à travers les duses lors des écoulements multiphasiques, soit en proposant de nouveaux modèles, soit en réalisant une étude comparative pour tester l'efficacité des modèles existants sur un champ donné. Les paragraphes ci-dessous résument l'histoire de ces différents modèles.

Le premier développement théorique important dans le domaine de l'écoulement diphasique à travers des restrictions a peut-être été publié par Tangren et al. en 1949 [2]. En supposant un liquide incompressible, un gaz idéal, un mélange homogène, aucun transfert de masse entre les phases, un écoulement isotherme, adiabatique, unidimensionnel et laminaire, ces deux auteurs ont développé des équations basées sur les principes de base de la mécanique des fluides. Bien que l'étude ait porté sur tous les aspects des effets de compressibilité dans l'écoulement diphasique, une partie de l'étude concernait l'écoulement dans une conduite convergent-divergent. Toutes les expériences ont été menées dans des conditions d'écoulement critiques. [1]

Gilbert [3] (1954) a développé un modèle d'écoulement multiphasique pour calculer le débit de liquide à travers une duse. Il a utilisé 268 données d'essai de production du champ de Ten Section en Californie pour des diamètres de duse compris entre 6/64 et 18/64 pouces. Gilbert a cependant noté que son modèle est bon lorsque la pression amont de l'étrangleur est au moins 70% supérieure à la pression aval.

Ensuite, plusieurs auteurs ont essayé d'adapter l'équation de Gilbert à d'autres champs d'étude, et ce en introduisant de nouvelles constantes et exposants sur l'équation de Gilbert, parmi lesquels : Baxendell, Ros, Achong, Pilehvari, Safar Beiranvand et al.

Ros [4] (1960) a étendu les travaux de Tangren et al. à des rapports gaz-liquide plus élevés où le gaz est la phase continue.

Fondamentalement, Tangren et al ont supposé un mélange "homogène", c'est-à-dire que les bulles de gaz sont si petites et uniformément distribuées dans le liquide que les vitesses des deux phases peuvent être considérées comme égales pendant l'expansion. Cette hypothèse semble raisonnable pour les systèmes dans lesquels la phase liquide est continue, ce qui se produit lorsque le rapport volumétrique gaz/liquide est inférieur à environ 1. Cependant, lorsque ce rapport dépasse 1, la situation inverse est susceptible de se produire. Alors, la phase liquide étant la phase dispersée, il faut s'attendre à une différence de vitesse entre les phases, ce qui entraîne des pertes d'énergie dues au glissement qui ont été ignorées dans l'étude de Tangren.

Le développement de Ros supposait une expansion polytropique du gaz, et une énergie potentielle et des pertes irréversibles négligeables, à l'exception de celles dues au glissement. Il a postulé qu'un écoulement à brouillard se produit dans la restriction, que tout le film liquide sur la paroi serait négligeable, que la dispersion du liquide se produit à l'entrée de la restriction, que les gouttelettes de liquide sont accélérées par la vitesse plus élevée du gaz, que le glissement à l'extrémité de la gorge peut être négligé et que la friction de la paroi peut être ignorée si la restriction est courte.

Poettmann et Beck [5] (1963) ont converti l'équation de Ros en unités de champ pétrolier et ils l'ont présenté sous forme graphique pour trois gravités API différentes. Ils ont comparé leurs courbes à 108 essais sur le terrain et ont prédit la production.

Omana [6] (1969) a introduit les propriétés des fluides dans son modèle. Il a réalisé des expériences avec de l'eau et du gaz naturel pour établir sa corrélation. Les essais comprenaient 35 essais d'écoulement en phase liquide, 24 en phase gazeuse et 47 en écoulement diphasique et couvraient de larges plages de débit, de taille de duses et de pression. Les schémas d'écoulement ont été photographiés au-dessus et au-dessous de l'étrangleur. Une analyse dimensionnelle a été effectuée et a donné lieu à huit groupes sans dimension

Ashford et Pierce [7] (1975) ont étendu la théorie de Ros et ont développé une équation pour l'écoulement sous-critique à travers des restrictions. Il a été énoncé dans leur article qu'une expansion polytropique du gaz a été utilisée ; le gaz libre et le gaz en solution ont été pris en compte ; une expression améliorée pour le débit de liquide qui tient compte du gaz en solution ; les conditions d'écoulement critiques peuvent être déterminées ; et l'écoulement sous-critique peut être traité de sorte que la chute de pression à l'orifice soit liée aux débits et aux propriétés du fluide. Les expériences ont été menées sur un puits à écoulement de 12 000 pieds de profondeur avec une vanne souterraine Otis J type 22J037 située à 3500 pieds pour servir d'orifice. Trois diamètres de duse ont été considérés : 14/64, 16/64 et 20/64.

Al-Attar et Abdul-Majeed [8] (1988) ont mené une étude comparative sur trois modèles ; Gilbert, Poettmann and Beck et Ashford, et ce dans le but de définir la gamme d'applicabilité de ces techniques et de trouver une technique qui s'adapterait le mieux aux données de production observées sur le champ pétrolifère de East Bagdad. Il a donc été nécessaire d'utiliser certaines données de production publiées et d'autres données de production provenant de cinq puits de pétrole dans le champ pétrolifère d'East Bagdad. Il a été conclu que la corrélation d'Ashford surestime grossièrement les taux de production. D'autre part, les corrélations de Poettmann et Beck et de Gilbert sous-estiment largement les taux de production. La corrélation de Gilbert, cependant, semble être la meilleure. Comme il a étudié l'influence de la densité API de l'huile, GOR, diamètre de la duse et la pression en tête de puits sur la performance de chaque modèle. Il a finalement proposé deux formes révisées du modèle de Gilbert en introduisant la gravité API dans l'une d'entre elles, et ce pour trouver les corrélations qui s'adaptent le mieux aux données mesurées du champ pétrolifère East Bagdad .

Abdul-Majeed et Maha [9] (1991) ont comparé les corrélations de Gilbert, Ros, Baxendall, Achong, Poettmann-Beck, Omana, Ashford, Pilehvari, Sachdeva et al. Et Hazim-Ghassan. Certaines de ces corrélations nécessitent des valeurs des propriétés physiques du fluide, des combinaisons de corrélations PVT ont été utilisées pour les estimer.

Il a été observé que le choix des corrélations PVT semble n'avoir qu'un effet mineur sur les résultats statistiques finaux calculés. Les deux auteurs ont également constaté que la plupart des corrélations comparées ont donné des résultats insatisfaisants, ils ont donc essayé de trouver les corrélations qui correspondent le mieux aux données mesurées. Par conséquent, douze nouvelles corrélations révisées sont obtenues par une analyse de régression et ce en utilisant les corrélations d'Omana, Gilbert et Hazim-Ghassan, dont quatre d'entre elles se sont avérées avoir de meilleures performances sur les données de test pour la prédiction du débit.

Towailib A.I. and Al-Marhoun [11] (1992) ont développé une corrélation empirique qui met en relation le débit d'huile à d'autres paramètres. Cette corrélation couvre une large gamme de débits et de tailles de duse. Elle était basée sur 3554 données de tests de production provenant de dix champs du Moyen-Orient. La corrélation s'applique aux conditions d'écoulement critiques. Towailib et Al-Marhoun ont utilisé une régression multiple non linéaire et ont trouvé que le débit d'huile est en fonction du diamètre de la duse, de la pression en amont de la tête de puits et de la densité du mélange.

Elgibaly and Nashawi [10] (1998) ont mené une recherche qui avait comme objectif le développement de nouvelles corrélations empiriques simples qui s'adaptent le mieux aux données d'essai du Moyen-Orient, l'étude de l'effet du couplage de l'approche Ashford-Pierce pour l'écoulement multiphasique à travers les duses avec les corrélations PVT développées pour le pétrole brut et le gaz du Moyen-Orient, et l'évaluation de performance prédictive des nouvelles corrélations développées par rapport aux corrélations empiriques disponibles sur la base d'une analyse statistique des erreurs. A la fin de leur travail, une corrélation empirique de la forme initialement proposée par Gilbert a été développée dans cette étude pour le régime d'écoulement critique. Les résultats ont montré que la nouvelle corrélation a fourni la meilleure précision, suivie de Gilbert, Ros et Achong ; tandis que les corrélations Pilehvari, Osman et Dokla ont donné le plus faible degré de précision. L'incorporation des nouvelles corrélations PVT développées pour les mélanges de pétrole et de gaz du Moyen-Orient, dans la corrélation Ashford-Pierce utilisée pour le régime d'écoulement critique, a amélioré la précision prédictive de cette corrélation. Cependant, la corrélation Ashford-Pierce modifiée était moins précise que la nouvelle corrélation empirique. Une autre corrélation empirique a été développée pour le régime d'écoulement sous-critique.

Récemment, l'IA a été appliquée avec succès dans le secteur pétrolier. Le domaine de l'IA comprend un nombre important de techniques dont certaines sont simples à utiliser, alors que beaucoup d'autres nécessitent une compréhension plus complexe des processus d'apprentissage mathématique [26].

De nombreuses analyses ont été réalisées sur l'utilisation de différents schémas d'intelligence artificielle (IA) pour estimer le flux diphasique dans les réservoirs et les conduites [31]. Parmi les schémas utilisés : la régression logistique (LR), les perceptrons multicouches (multilayer perceptrons (MLP)), les machines à vecteurs de support (support vector machines (SVM)), les réseaux neuronaux artificiels (artificial neural networks (ANN)), mémoire à long et court terme (Long short term memory, LSTM), régression par processus gaussien (Gaussian process GPR) et programmation génétique (genetic programming, (GP)).

Pejman Shoeibi omrani, Iulian Dobrovolschi et Stephane Belfroid, Peter Kronberger et Esteban Munoz [29] ont exploré l'approche data driven en utilisant les réseaux de neurones artificiels ANN (artificial neural network) pour la prédiction du débit d'huile et de gaz produit par

chaque puits. Les approches développées ont été testées pour évaluer leurs performances et ce en comparant les résultats obtenus avec des données de terrain. Les résultats ont montré que les modèles ANN sont capables d'estimer avec une grande certitude les données de terrain.

Seidi et Sayahi [27] ont établi, pour le régime d'écoulement diphasique sous-critique, une nouvelle corrélation de type Gilbert sur la base de la corrélation de Nasriani et Kalantariasl [28] en ajoutant à la chute de pression un exposant. Ils ont utilisé un algorithme génétique et une technique de régression non linéaire. Les nouveaux modèles, ainsi que d'autres modèles de la littérature ont été testés sur 67 mesures de 10 champs différents dans des réservoirs de condensat de gaz à haut débit dans le sud de l'Iran. Ils ont constaté que l'analyse de régression non linéaire est plus efficace pour ajuster les données que l'algorithme génétique.

Pour la prédiction de l'écoulement diphasique à travers la duse en utilisant les réseaux de neurones, une recherche a été faite par Al-Khalifa et Al-Marboun(2013). Il a étudié l'application du réseau neuronal pour l'écoulement diphasique à travers la duse. L'étude a utilisé 4031 points de données collectés dans différents puits du Moyen-Orient. Dans sa recherche, il a développé un modèle ANN pour prédire la taille de la duse et le débit de production en utilisant un ratio de 80% pour le training, 10% pour la validation et 10% pour le test pour les données étudiées. Une comparaison a été faite entre les résultats obtenus par ANN et les corrélations existantes. Il a été démontré que le modèle ANN donne de meilleurs résultats parmi toutes les corrélations et tous les modèles. [31]

Pour conclure, même si de nombreux chercheurs ont tenté d'établir des corrélations pour décrire les écoulements à travers la duse, aucune équation ne peut être considérée comme ayant une application universelle. Néanmoins, il y a des points clés à prendre en compte dans chaque étude qui permettent une compréhension globale de l'écoulement multiphasique à travers les duses. De ce fait l'approche par l'apprentissage automatique, se présente comme une alternative très intéressante pour combler aux différents inconvénients des approches dites classiques, ce que nous allons essayer d'étudier, d'analyser et de décortiquer le long de notre mémoire de fin d'études.

# Chapitre 1 :

## Les écoulements diphasiques

Pour évaluer la performance d'écoulement, il est nécessaire de pouvoir calculer les pertes de pression dans tous les composants du système. Dans le cas d'un écoulement monophasique, liquide ou gazeux, les pertes de charge peuvent être calculées facilement, à condition que les caractéristiques intrinsèques du système, telles que la taille et la rugosité soient connues. En réalité, la plupart des puits de pétrole ou de gaz fonctionnent sous des conditions d'écoulement multiphasiques. Lorsque les conditions de pression changent, un changement de phase dans les fluides produits peut avoir lieu. Cela entraîne des changements dans les densités, les vitesses, les volumes de chaque phase et les propriétés des fluides. La présence simultanée de liquide et de gaz rend le calcul de perte de pression plus compliquée.

### 1.1 Equation de gradient de pression

La base théorique de la plupart des équations d'écoulement des fluides est l'équation générale de l'énergie. Elle représente la conservation de l'énergie entre deux points d'un système (elle stipule simplement que l'énergie d'un fluide entrant dans un volume de contrôle, plus tout travail effectué sur ou par le fluide, plus toute énergie thermique ajoutée ou retirée du fluide, doit être égale à l'énergie du fluide quittant le volume de contrôle). [17]

L'équation d'énergie est développée en premier lieu en utilisant les principes thermodynamiques, puis modifiée en une forme d'équation de gradient de pression pour prendre la forme générale ci-dessous.

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

Avec :

$$\diamond \left(\frac{dP}{dL}\right)_{ele} = \frac{g}{g_c} \rho \sin \theta \quad (\text{Terme de perte par élévation})$$

$$\diamond \left(\frac{dP}{dL}\right)_f = \frac{f_{tp} \cdot \rho \cdot v^2}{2 \cdot g_c \cdot d} \quad (\text{Terme de perte avec frottement})$$

$$\diamond \left(\frac{dP}{dL}\right)_{acc} = \frac{\rho v dv}{g_c dL} \quad (\text{Terme de perte par accélération})$$

L'équation (1.1) s'applique à tout fluide en régime permanent, en écoulement unidimensionnel. Elle sera ensuite adaptée à diverses conditions de systèmes de tuyauterie et de fluides.

## 1.2 Ecoulements monophasiques

Certains aspects de l'équation du gradient de pression telle qu'elle s'applique à l'écoulement monophasique sont discutés afin de développer une compréhension approfondie de chaque composante avant de la modifier pour l'écoulement diphasique.

**Perte de charge par élévation** : elle est causée par le poids de la colonne de fluide dans le puits. Cette composante est nulle pour un écoulement horizontal. Elle s'applique pour un écoulement compressible ou incompressible, en régime permanent ou transitoire, dans un écoulement vertical ou incliné.

**Perte de charge par accélération** : cette composante s'applique à toutes les conditions d'écoulement transitoires, mais elle est nulle pour un écoulement incompressible. Pour toute condition d'écoulement dans laquelle un changement de vitesse se produit, comme un écoulement compressible, une chute de pression se produira dans la direction de l'augmentation de la vitesse.

**Perte de charge par frottement** : elle provoque souvent une chute de pression considérable, que ce soit pour les puits à l'huile ou à gaz. Elle est liée aux forces visqueuses dans le fluide ainsi qu'aux propriétés de ce dernier. Cette composante devient de plus en plus importante lorsque la vitesse du fluide augmente, ce qui revient à dire qu'elle dépend fortement du régime d'écoulement (laminaire ou turbulent).

**Le facteur de friction** : nécessaire pour le calcul de la perte de charge par friction, est une fonction du nombre de Reynolds  $Re$  et de la rugosité relative de la conduite  $\frac{\varepsilon}{D}$

$$f = \varphi'(Re, \frac{\varepsilon}{D}) \quad (1.2)$$

Pour un écoulement laminaire, Le facteur de friction est simplement  $= 64/Re$  , il est indépendant de  $\varepsilon/D$  . Pour un écoulement turbulent ; qui est souvent le cas pour les écoulements de pétrole et de gaz, la dépendance fonctionnelle du facteur de friction par rapport au nombre de Reynolds et à la rugosité relative est complexe, elle peut être obtenue à partir du diagramme de moody (figure 1.1).

L'équation suivante de Colebrook est valide pour le calcul du facteur de friction pour tout l'intervalle de l'écoulement non laminaire du diagramme de moody, en effet le diagramme de moody est une représentation graphique de cette équation.

$$\frac{1}{\sqrt{f}} = -2 \log \left( \frac{\varepsilon/D}{3.7} + \frac{2.51}{Re\sqrt{f}} \right) \quad (1.3)$$

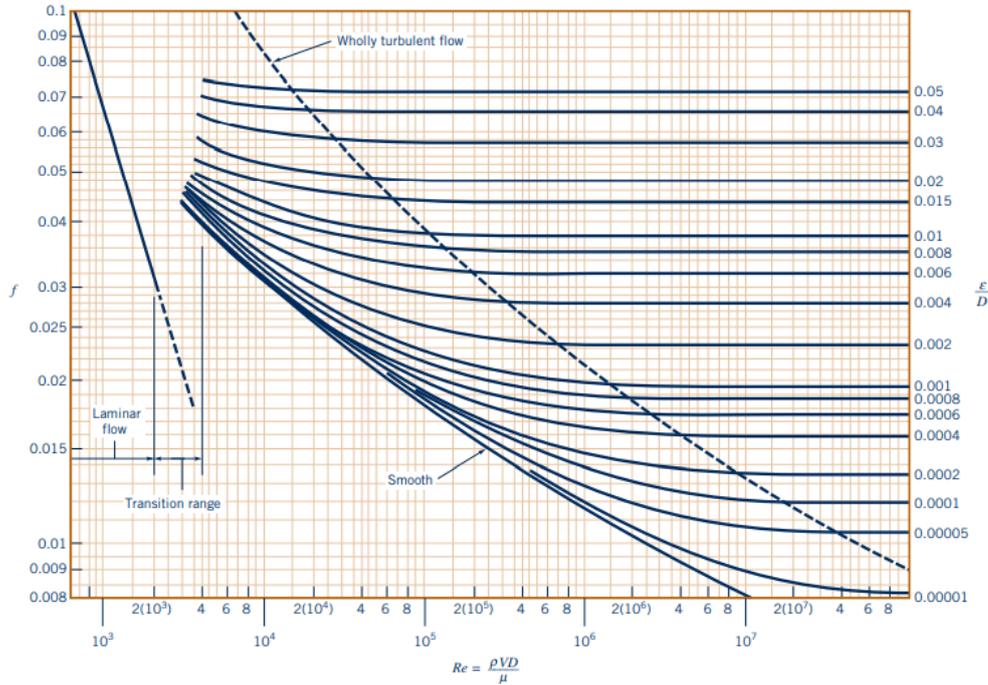


Figure 1.1 : Diagramme de Moody représentant le facteur de friction en fonction du nombre de Reynolds et de la rugosité relative pour les tuyaux ronds. [15]

### 1.3 Écoulements diphasiques

L'introduction d'une seconde phase dans un flux d'écoulement complique l'analyse de l'équation du gradient de pression. Les fluides peuvent se séparer en raison des différences de densité et s'écouler à des vitesses différentes dans le tuyau. Les propriétés telles que les densités, la vitesse et la viscosité, qui sont relativement simples pour les fluides individuels, peuvent être modifiées et deviennent parfois difficiles à déterminer.

#### 1.3.1 Paramètres descriptifs

Pour décrire les problèmes des écoulements diphasiques, il est important d'introduire quelques paramètres et propriétés afin de pouvoir adapter l'équation du gradient de pression pour ce type d'écoulements.

##### Le Holdup :

Lors des écoulements multiphasiques dans les conduites, le Holdup est défini comme la fraction d'un fluide particulier présent dans un intervalle de la conduite.

Le Holdup liquide  $H_L$ , est défini comme la fraction d'un élément de tube qui est occupée par un liquide à un moment donné.

$$H_L = \frac{\text{volume de liquide dans un segment de pipe}}{\text{le volume du segment de pipe}} \quad (1.4)$$

Et le reste du segment est occupé par du gaz ce qui est connu comme le gaz Holdup  $H_g$

$$H_g = 1 - H_L \quad (1.5)$$

Le Holdup liquide peut être mesuré expérimentalement par plusieurs méthodes, telles que les sondes de résistivité, les densitomètres nucléaires, ou en piégeant un segment du flux d'écoulement entre des vannes à fermeture rapide et en mesurant la quantité de liquide retenue. Comme il peut être déterminé à partir de corrélations empiriques, il est en fonction des propriétés du gaz et du liquide, du schéma d'écoulement, le diamètre de la conduite, etc.

**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 :

$$\lambda_L = \frac{q_L}{q_L + q_g} \quad (1.6)$$

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 :

$$\lambda_g = 1 - \lambda_L = \frac{q_g}{q_L + q_g} \quad (1.7)$$

**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.

❖ **La masse volumique de liquide  $\rho_L$**  peut être calculée à partir des masses volumiques d'huile et d'eau en supposant l'absence de glissement entre les phases d'huile et d'eau :

$$\rho_L = \rho_o f_o + \rho_w f_w \quad (1.8)$$

Avec :

$$f_o = \frac{q_o}{q_o + q_w} \quad (1.9)$$

Et

$$f_w = 1 - f_o = \frac{q_w}{q_o + q_w} \quad (1.10)$$

❖ **La masse volumique de mélange liquide/gaz** Ce calcul nécessite la connaissance de la valeur de holdup liquide. Trois équations ont été utilisées par divers chercheurs pour le calcul de la masse volumique liquide/gaz pour un écoulement diphasique.

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

$$\rho_n = \rho_L \cdot \lambda_L + \rho_g \cdot \lambda_g \quad (1.12)$$

$$\rho_k = \frac{\rho_L \cdot \lambda_L^2}{H_L} + \frac{\rho_g \cdot \lambda_g^2}{H_g} \quad (1.13)$$

La formule (1.11) est généralement utilisée pour le calcul de la perte de charge par élévation. Certaines corrélations sont basées sur l'hypothèse de non glissement donc l'équation (1.12) est utilisée pour la densité diphasique. L'équation (1.13) est utilisée par certains chercheurs pour définir la densité du mélange pour le calcul du terme de perte de charge par frottement et le nombre de Reynolds.

### La vitesse :

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. La vitesse superficielle du gaz est calculée à partir de :

$$v_{sg} = \frac{q_g}{A} \quad (1.14)$$

La vitesse superficielle du liquide est calculée à partir de :

$$v_{sL} = \frac{q_L}{A} \quad (1.15)$$

La section réelle à travers laquelle le gaz circule est réduite par la présence du liquide à  $AH_g$ . Par conséquent, la vitesse réelle du gaz est calculée à partir de :

$$v_g = \frac{q_g}{AH_g} \quad (1.16)$$

Pour la vitesse réelle du liquide :

$$v_L = \frac{q_L}{AH_L} \quad (1.17)$$

Vitesse du mélange :

$$v_m = v_{sL} + v_{sg} \quad (1.18)$$

Vitesse de glissement :

$$v_s = v_g - v_L \quad (1.19)$$

### La 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.

$$\mu_s = \mu_L^{H_L} + \mu_g^{H_g} \quad (1.20)$$

$$\mu_{ns} = \mu_L \lambda_L + \mu_g \lambda_g \quad (1.21)$$

$$\mu_s = \mu_L H_L + \mu_g H_g \quad (1.22)$$

Dont la viscosité d'un mélange huile et eau est calculé par :

$$\mu_L = \mu_o f_o + \mu_w f_w \quad (1.23)$$

Les viscosités du gaz naturel, du pétrole brut et de l'eau peuvent être estimées à partir de corrélations empiriques si les viscosités mesurées ne sont pas disponibles.

### Tension superficielle de liquide :

Si la phase liquide contient à la fois de l'eau et de l'huile, la tension superficielle est calculée par :

$$\sigma_L = \sigma_o f_o + \sigma_w f_w \quad (1.24)$$

## 1.3.2 Configurations d'écoulements diphasiques

Lors d'un écoulement diphasique (gaz, liquide), l'apparition d'un régime d'écoulement dépend de plusieurs facteurs tels que les propriétés du fluide, les conditions d'opération, le débit de chaque phase, la géométrie de système, la dispersion, la coalescence, ainsi que les effets du flux thermique. Quelques régimes d'écoulements diphasiques sont présentés dans cette partie pour le cas de conduites verticales et horizontale.

### Écoulement vertical :

Pour un écoulement vertical, les principales configurations suivantes sont observées, comme c'est illustré dans la figure (1.2)

➤ **Écoulement 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.

➤ **Écoulement 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.

➤ **Écoulement 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.

➤ **Écoulement 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.

➤ **Écoulement 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.

➤ **Écoulement à brouillard (Mist flow) :**

À des pressions très basses, le taux de gaz est très élevé, il surmonte la force de tension liquide-tuyau 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.

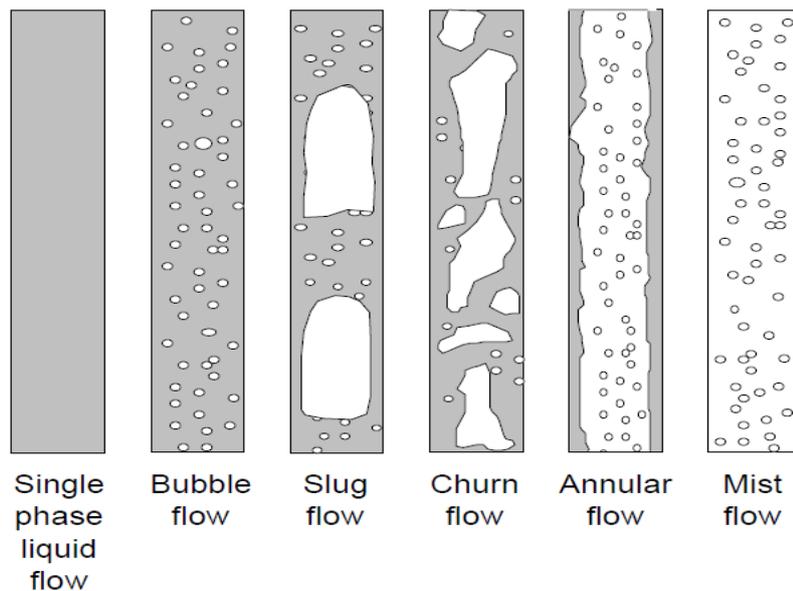


Figure 1.2 : configurations d'écoulements dans une conduite verticale

**Écoulements 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 suivants sont généralement observés (voir la figure 1.3).

➤ **Écoulement en bulles (Bubble flow) :**

Comme dans le cas d'un écoulement vertical, la phase est composée de bulles de gaz dispersées dans la phase liquide. Cependant, en raison de l'effet des forces de flottabilité, les bulles ont tendance à s'accumuler dans la partie supérieure du tuyau.

➤ **Écoulement par bouchon (Slug flow) :**

Ce régime est caractérisé par le passage le long du canal de "bouchons" mousseux qui remplissent complètement le tube. Les bouchons occupant plus la partie supérieure du tube peuvent souvent être très larges et une grande source de problèmes lors des opérations sur des pipelines horizontaux.

➤ **Écoulement stratifié (Stratified flow) :**

Ce régime se produit lorsque la séparation gravitationnelle est complète. Le liquide s'écoule le long du fond du tube et le gaz le long de la partie supérieure du tube, comme le montre la figure (1.3). Ce régime se produit pour des faibles vitesses de liquide et de vapeur en écoulements laminaires.

➤ **Écoulement annulaire (Annular flow) :**

Cet écoulement s'observe lorsque le débit de gaz augmente, le gaz occupe le milieu de la conduite tandis que le liquide occupe les parois.

➤ **Écoulement à brouillard (Mist flow) :**

Comme pour le cas d'un écoulement vertical, ce régime se produit à des pressions très basses, donc le taux de gaz est très élevé.

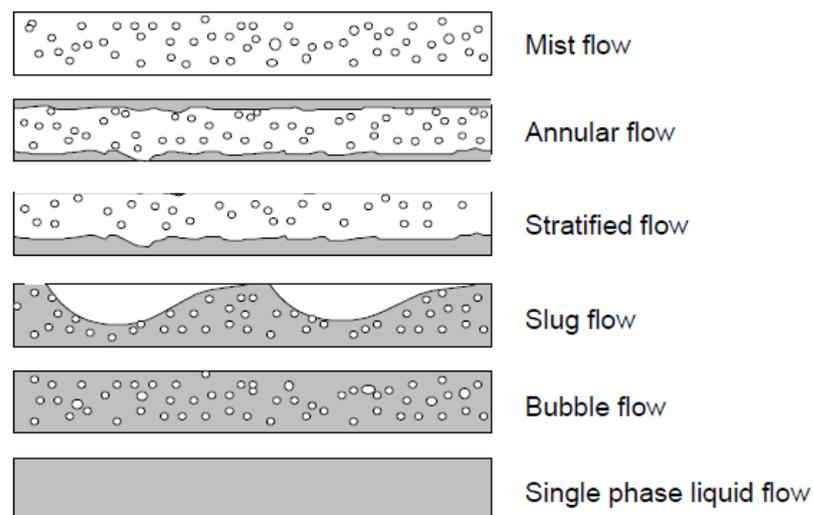


Figure 1.3 : configurations d'écoulements dans une conduite horizontale

### 1.3.3 Modélisation des écoulements diphasiques

L'équation de gradient de pression prend la même forme générale que l'équation (1.1), bien sûr avec quelques adaptations pour tenir compte de l'écoulement diphasique.

En remplaçant chaque terme dans l'équation générale, l'équation de gradient de pression pour un mélange diphasique prend la forme suivante tel que  $m$  représente l'indice du mélange.

$$\left(\frac{dP}{dL}\right) = \frac{g}{g_c} \rho_m \sin \theta + \frac{(f \rho v^2)_m}{2 \cdot g_c \cdot d} + \frac{(\rho v dv)_m}{g_c dL} \quad (1.25)$$

Suite à l'étude des différentes composantes contribuant au gradient de pression total, il s'ensuit que les principales considérations pour le développement du gradient de pression sont le développement de méthodes pour prédire le holdup et le facteur de friction diphasique où différents modèles d'écoulement ont été développés à travers les années.

## Modèles d'écoulements

L'un des composants les plus importants de l'ensemble du système de production est la colonne de production. Jusqu'à 80 pour cent de la perte de pression totale se produit à ce niveau pour soulever les fluides depuis le fond jusqu'à la surface. Dans cette partie, quelques modèles d'écoulements verticaux seront développés.

On distingue 3 catégories de modèles en fonction des considérations pour le régime d'écoulement et la prise en compte du glissement entre les phases.

### ❖ Catégorie A (modèle homogène)

Pour cette catégorie la configuration est non considérée et la vitesse de glissement est nulle.

De nombreux modèles d'écoulement homogène ont été développés pour analyser les écoulements diphasiques depuis les travaux pionniers de Poettmann et Carpenter (1952). Le modèle de Poettmann Carpenter utilise un facteur de friction diphasique empirique pour les calculs de perte de pression par friction sans tenir compte de l'effet de la viscosité du liquide. L'effet de la viscosité du liquide a été pris en compte par des chercheurs ultérieurs, notamment Cicchitti (1960) et Dukler et al. (1964). Une revue complète de ces modèles a été donnée par Hasan et Kabir (2002).

Poettmann et Carpenter (1952) ont présenté un modèle simplifié d'écoulement liquide-gaz pour calculer les pertes de pression dans le puits en estimant la densité du mélange et le facteur de friction. Selon Poettmann et Carpenter, l'équation suivante peut être utilisée pour calculer la traversée de pression dans un tubing vertical lorsque le terme d'accélération est négligé [13]

$$\Delta P = (\bar{\rho} + \bar{k}/\bar{\rho}) \frac{\Delta h}{144} \quad (1.26)$$

Et :

$$\bar{k} = \frac{f_{2F} q_o^2 M^2}{7.4137 \times 10^{10} D^5} \quad (1.27)$$

La masse volumique moyenne du mélange  $\bar{\rho}$  peut être calculée par :

$$\bar{\rho} = (\rho_1 + \rho_2)/2 \quad (1.28)$$

$\rho_1$  et  $\rho_2$  représentent respectivement les densités du mélange au sommet et à au bas du segment de tubing exprimées en lb/ft<sup>3</sup>

La masse volumique du mélange en un point donné peut être calculée par :

$$\rho = \frac{M}{V_m} \quad (1.29)$$

Avec :

$$M = 350.17(\gamma_o + WOR \gamma_w) + GOR \gamma_g \quad (1.30)$$

$$V_m = 5.615(B_o + WOR B_w) + (GOR - R_s) \left( \frac{14.7}{P} \right) \left( \frac{T}{520} \right) \left( \frac{Z}{1.0} \right) \quad (1.31)$$

Le facteur volumétrique d'huile et le rapport de solubilité peut être calculés en utilisant ces corrélations :

$$R_s = \gamma_g \left[ \frac{P \ 10^{0.0125 \ API}}{18 \ 10^{0.00091 \ t}} \right]^{1.2048} \quad (1.32)$$

$$B_o = 0.9759 + 0.00012 \left[ R_s \left( \frac{\gamma_g}{\gamma_o} \right)^{0.5} + 1.25t \right]^{1.2} \quad (1.33)$$

Concernant le calcul de  $f_{2F}$ , Guo et Ghalambor (2002) ont développé la corrélation suivante :

$$f_{2F} = 10^{1.444 - 2.5 \log(D\rho v)} \quad (1.34)$$

Où  $(D\rho v)$  est le numérateur du nombre de Reynolds représentant la force d'inertie et peut être formulé comme suit :

$$(D\rho v) = \frac{1.4737 \times 10^{-5} M q_o}{D} \quad (1.35)$$

## ❖ Catégorie B

Pour cette catégorie, la configuration est non considérée et la vitesse de glissement est non nulle. Une corrélation est nécessaire pour le calcul à la fois de la rétention de liquide (liquid holdup) et le facteur de friction.

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. Calcul des nombres adimensionnels :

$$N_{LV} = 1.938 V_{sl} \sqrt[4]{\frac{\rho_L}{\sigma_L}} \quad (1.36)$$

$$N_{gV} = 1.938 V_{sg} \sqrt[4]{\frac{\rho_L}{\sigma_L}} \quad (1.37)$$

$$N_d = 120.872d \sqrt{\frac{\rho_L}{\sigma_L}} \quad (1.38)$$

$$N_L = 0.15726 \mu_L^4 \sqrt{\frac{1}{\rho_L \sigma_L^3}} \quad (1.39)$$

2. Détermination la valeur de  $CN_L$  à partir du graphe de la figure (An.A.1) ( $N_L$  vs  $CN_L$ )
3. Détermination de la valeur de  $\frac{H_L}{\psi}$  à partir du graphe (An.A.2)

$$\left( \frac{N_{LV}}{N_{gV}^{0.575}} \right) \left( \frac{p}{p_a} \right)^{0.1} \left( \frac{CN_L}{N_d} \right) \text{ vs } \frac{H_L}{\psi}$$

4. Déterminer la valeur de  $\psi$  à partir du graphe (An.A.3)  $\left( \frac{N_{gV} N_L^{0.380}}{N_d^{2.14}} \right)$  vs  $\psi$

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

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

$$N_{Re} = \frac{\rho_s v_m d}{\mu_s} \quad (1.40)$$

Où :

$$v_m = v_{sL} + v_{sg} \quad (1.41)$$

$$\mu_s = \mu_L^{H_L} + \mu_g^{H_g} \quad (1.42)$$

6. Une fois les valeurs du 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

Cette catégorie est la plus efficace car elle intègre le régime d'écoulement et considère la vitesse de glissement dans le calcul de perte de pression. Plusieurs auteurs ont développé des corrélations pour quelques paramètres utilisées dans le calcul de perte de pression, parmi eux on cite : Duns et Ros, Orkiszewski. Aziz, Govier et Fogarasi. Chierici, Ciucci et Sclocchi. Biggs et Brill.

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. [3]

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

1. calculer les nombres adimensionnels  $N_{LV}$ ,  $N_{gv}$ ,  $N_d$ ,  $N_L$ ,  $L_s$ ,  $L_m$ ,  $L_1$  et  $L_2$

Tel que :

$$L_s = 50 + 36N_{Lv} \quad (1.43)$$

$$L_m = 75 + 84N_{Lv}^{0.75} \quad (1.44)$$

Et  $L_1$  et  $L_2$  sont déterminés à partir du graphe (An.A.4) ( $N_d$  vs  $L_1, L_2$ )

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 :

$$S = F_1 + F_2 N_{Lv} + F_3' \left( \frac{N_{gv}}{1 + N_{Lv}} \right)^2 \quad (1.45)$$

$$F_3' = F_3 - \frac{F_4}{N_d} \quad (1.46)$$

Où  $F_2$  et  $F_3$  sont déterminées à partir du graphe (An.A.5).

○ Slug flow

$$S = (1 + F_5) \frac{N_{gv}^{0.982} + F_6'}{(1 + F_7 N_{Lv})^2} \quad (1.47)$$

$$F_6' = 0.029 N_d + F_6 \quad (1.48)$$

$F_5, F_6$  et  $F_7$  sont déterminées à partir des graphes de la figure (An.A.6)

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

○ Bubble flow et slug flow

$$v_s = \frac{S}{\left( \frac{\rho_L}{\sigma_L g} \right)^{1/4}} \quad (1.49)$$

$$\left(\frac{dp}{dL}\right)_t = \left(\frac{dp}{dL}\right)_{elev} + \left(\frac{dp}{dL}\right)_f \quad (1.50)$$

$$H_L = \frac{v_s - v_m + \left[ (v_s - v_m)^2 + 4v_m v_{sL} \right]^{1/2}}{2v_s} \quad (1.51)$$

- Mist flow

$$H_L = \lambda_L \quad (1.52)$$

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

$$\frac{dp}{dL} = \frac{g}{g_c} [\rho_L H_L + \rho_g (1 - H_L)] \quad (1.53)$$

6. Calculer la perte de charge due au frottement

- Bubble flow et slug flow

$$\left(\frac{dp}{dL}\right)_f = \frac{f_m \rho_L v_{sL} v_m}{2g_c d} \quad (1.54)$$

$$f_m = f_1 f_2 / f_3 \quad (1.55)$$

Avec :

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

$f_2$  : Fcteur à tirer du graphe de la figure (An. A 7)

Et :

$$f_3 = 1 + f_1 \sqrt{\frac{v_{sg}}{50v_{sL}}} \quad (1.56)$$

- Mist flow

$$\frac{dp}{dL} = \frac{f \rho_g v_{sg}^2}{2g_c d} \quad (1.57)$$

Où  $f$  est facteur de frottement à partir du diagramme de Moody et  $Re = \rho_g v_{sg} d / \mu_g$

7. Calcul de la perte de charge total

- Bubble flow, slug flow, mist flow

○ Transition flow

$$\left(\frac{dp}{dL}\right)_t = A\left(\frac{dp}{dL}\right)_{t,slug} + B\left(\frac{dp}{dL}\right)_{t,mist} \quad (1.58)$$

Où :

$$A = \frac{L_m - N_{gv}}{L_m - L_s} \quad B = \frac{N_{gv} - L_s}{L_m - L_s} = 1 - A, \quad (1.59)$$

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

$$\rho_g' = \frac{\rho_g N_{gv}}{L_m} \quad (1.60)$$

## Chapitre 2 :

# Modélisation des écoulements à travers les restrictions et calcul de production

L'écoulement des fluides est un phénomène physique très complexe. Et pour comprendre le comportement des fluides dans leur chemin d'écoulement, la modélisation prend place. La modélisation consiste à transformer le phénomène physique en une équation ou un modèle généralement développé à travers l'expérience.

## 2.1 Modélisation des écoulements à travers les restrictions

### 2.1.1 Généralités sur les pertes de charges locales

Pour résoudre les problèmes d'écoulement de pétrole et de gaz dans les conduites, il est important de connaître les pertes de charges à travers celles-ci.

Un système de conduites typique se compose généralement de différentes longueurs de tuyaux droits connectés avec différents types de composants (restrictions, vannes, coudes, etc.). La perte de charge globale du système notée  $h$  est constituée de la perte de charge due aux effets visqueux dans les canalisations droites, appelée perte de charge linéaire (ou régulière) notée  $h_L$  et de la perte de charge dans les différents composants de la conduite, appelée perte singulières (ou locales) notée  $h_s$ .

$$h = h_L + h_s \quad (2.1)$$

Dans cette section, on va apporter plus de détails sur le calcul de ces dernières et les différents paramètres qui les affectent. Les pertes de charges linéaires seront abordées en premier lieu pour développer une meilleure compréhension des pertes de charges locales.

#### ❖ Pertes de charges linéaires (régulières)

Les écoulements de pétrole et de gaz sont considérés souvent turbulents. L'analyse dimensionnelle fournit la base la plus commode pour examiner ce type d'écoulements.

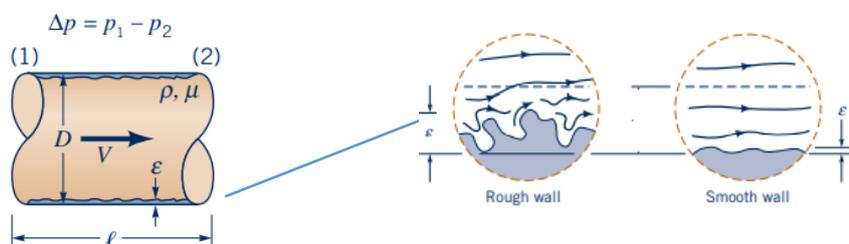


Figure 2.1 Illustration des différents paramètres influençant sur la perte de charge à travers une conduite et différence entre une surface lisse (smooth) et rugueuse (rough wall) [15]

La chute de pression, pour un écoulement permanent, d'un fluide incompressible dans une conduite ronde horizontal de diamètre  $D$ , peut être écrite sous la forme fonctionnelle suivante :

$$\Delta P = f(V, D, l, \varepsilon, \mu, \rho) \quad (2.2)$$

Où  $V$  est la vitesse moyenne du fluide,  $l$  : la longueur du tuyau,  $\varepsilon$  : la rugosité de la paroi du tuyau,  $\mu$  : viscosité du fluide et  $\rho$  : masse volumique du fluide.

Vu qu'il y a sept variables ( $k = 7$ ) qui peuvent être écrites en termes des trois dimensions de référence MLT ( $r = 3$ ), l'équation peut être écrite sous forme adimensionnelle en termes de  $k - r = 3$  groupes sans dimension, donc l'équation (2.2) devient :

$$\frac{\Delta P}{\frac{1}{2}\rho V^2} = f'\left(\frac{\rho V D}{\mu}, \frac{l}{D}, \frac{\varepsilon}{D}\right) \quad (2.3)$$

$\frac{1}{2}\rho V^2$  étant le terme de la pression dynamique. En introduisant la définition du nombre de Reynolds  $Re = \rho V D / \mu$ , et en supposant que la chute de pression est proportionnelle à la longueur de la conduite  $l$  l'équation (2.3) peut s'écrire sous la forme :

$$\frac{\Delta P}{\frac{1}{2}\rho V^2} = \frac{l}{D} \varphi\left(Re, \frac{\varepsilon}{D}\right) \quad (2.4)$$

En d'autre terme, la chute de pression dans une conduite peut se mettre sous la forme suivante :

$$\Delta P = h_L = f \frac{l}{D} \frac{\rho V^2}{2} \quad (2.5)$$

Où  $f$  est le facteur de friction dont le calcul est abordé dans le chapitre précédent tel que :

$$f = \varphi'\left(Re, \frac{\varepsilon}{D}\right) \quad (2.6)$$

#### ❖ Pertes de charges singulières

Les pertes de charge singulière apparaissent lors de changement de direction d'écoulement, de changement de diamètre, et lorsque des dispositifs (vannes, raccords et manchons, coudes, tés, etc.) sont disposés sur la ligne d'écoulement [16]. Ces pertes sont typiquement exprimées par la relation :

$$h_s = \sum_i K_i \frac{V_i^2}{2g} \quad (2.7)$$

Où  $h_s$  est la perte de charge singulière  $V_i$  est la vitesse moyenne pour le dispositif et  $K_i$  le coefficient de perte de charge pour le dispositif  $i$  défini par :

$$K_i = \frac{h_{si} 2g}{V_i^2} = \frac{\Delta p_i}{\rho V_i^2 / 2} \quad (2.8)$$

La valeur du coefficient de perte de charge  $K_i$  dépend fortement de la géométrie de l'élément considéré. Elle peut également dépendre des propriétés du fluide. C'est-à-dire que,

$$K_i = \varphi(\text{géométrie}, Re) \quad (2.9)$$

Où  $Re = \rho V D / \mu$  est le nombre de Reynolds. Pour de nombreuses applications pratiques, le nombre de Reynolds est suffisamment grand pour que l'écoulement à travers le composant soit dominé par les effets d'inertie, les effets visqueux étant d'importance secondaire. Ceci est vrai en raison des accélérations et décélérations relativement importantes subies par le fluide lorsqu'il s'écoule le long d'une trajectoire plutôt incurvée, de surface variable, peut-être même tortueuse, à travers le composant (voir la figure 2.2)

Dans un écoulement dominé par les effets d'inertie plutôt que par les effets visqueux, on constate généralement que les chutes de pression et les pertes de charge sont en corrélation directe avec la pression dynamique. C'est la raison pour laquelle le facteur de friction pour un très grand nombre de Reynolds dans une conduite est indépendant du nombre de Reynolds.

La même condition s'applique à l'écoulement à travers les composants et les restrictions d'un tuyau. Ainsi, dans la plupart des cas les coefficients de perte de charge des composants sont uniquement fonction de la géométrie [15].

$$K_i = \varphi(\text{géométrie}) \quad (2.10)$$

On note que le coefficient de perte de charge singulière est plus grand lorsque l'on considère des changements brusques de section de conduites que des changements progressifs. [16] Quelques valeurs de ce coefficient sont reportées dans le tableau de la figure (2.3)

Les pertes de charges locales sont parfois données en termes de longueur équivalente,  $l_{eq}$ . Dans cette terminologie, la perte de charge à travers un composant est donnée en termes de longueur équivalente de tuyau, qui produirait la même perte de charge que le composant. C'est-à-dire,

$$h_{is} = K_i \frac{V_i^2}{2g} = f \frac{l_{eq}}{D} \frac{V^2}{2g} \quad (2.11)$$

$$l_{eq} = K_i \frac{D}{f} \quad (2.12)$$

Où  $f$  et  $D$  sont respectivement le facteur de friction et le diamètre de la conduite contenant le composant  $i$ .

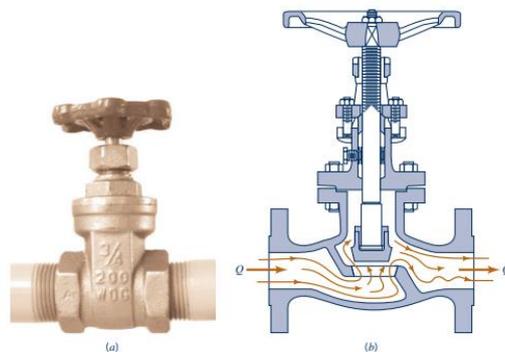


Figure 2.2 : écoulement à travers une vanne

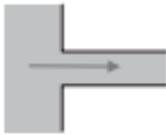
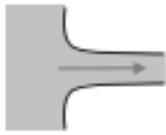
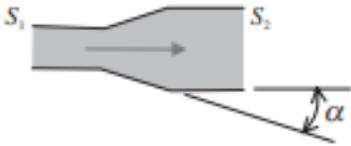
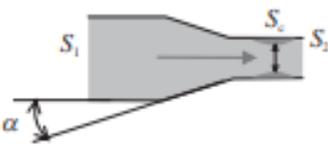
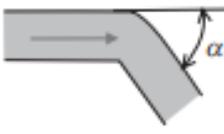
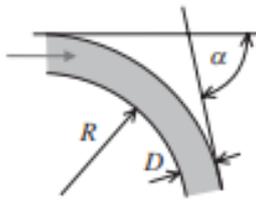
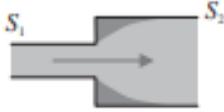
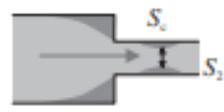
<p>Entrée brusque de canalisation <math>K = 0,5</math></p> 	<p>Entrée progressive de canalisation <math>K = 0,04</math></p> 
<p>Divergent <math>\left(1 - \frac{S_1}{S_2}\right)^2 \sin \alpha</math></p> 	<p>Convergent <math>\left(\frac{1}{\xi} - 1\right)^2 \sin \alpha, \xi = \frac{S_c}{S_2}</math></p> 
<p>Coude brusque <math>\sin^2 \alpha + 2 \sin^4 \frac{\alpha}{2}</math></p> 	<p>Coude progressif <math>\frac{\pi}{\alpha} \left(0,131 + 1,847 \left(\frac{D}{R}\right)^{7/2}\right)</math></p> 
<p>Élargissement brusque <math>\left(1 - \frac{S_1}{S_2}\right)^2</math></p> 	<p>Rétrécissement brusque <math>\left(\frac{1}{\xi} - 1\right)^2 ; \xi = \frac{S_c}{S_2}</math></p> 

Figure 2.3 : Quelques expressions du coefficient de pertes de charge singulières  $K$  pour des géométries d'écoulements couramment utilisées. [16]

Dans ce qui suit, les écoulements à travers un type particulier de restrictions seront abordés, plus précisément, les écoulements à travers les duses.

## 2.1.2 Écoulements à travers les duses

L'écoulement de fluides à travers des restrictions est un phénomène courant dans l'industrie du pétrole et du gaz naturel. Il se produit lorsque des liquides et/ou des gaz sont mesurés avec des compteurs à orifice, lorsque des fluides s'écoulent à travers des étranglements de surface ou de fond, lorsque des fluides s'écoulent à travers des soupapes de sécurité souterraines et lorsqu'ils s'écoulent à travers tout type de soupape de surface. Un débit restreint peut également être causé par des éléments tels que des formations de paraffine ou d'hydrate, des étampes et d'autres types de raccords de tuyauterie. [1]

Dans le cadre de ce mémoire, les écoulements à travers un type spécifique de restrictions seront détaillés, il s'agit des duses.

### 2.1.2.1 Description de la duse

La duse qui est un orifice calibré placé à la sortie de la tête de puits pour étrangler l'écoulement et de contrôler le débit afin d'atteindre un taux de production soigneusement calculé. Les duses ayant un diamètre réglable sont souvent utilisés pendant les opérations de complétion pour permettre à l'opérateur de nettoyer et de tester le débit du puits. Une fois le débit optimal déterminé, la duse réglable est généralement remplacé par une duse positive (qui a un diamètre fixe) pour la production

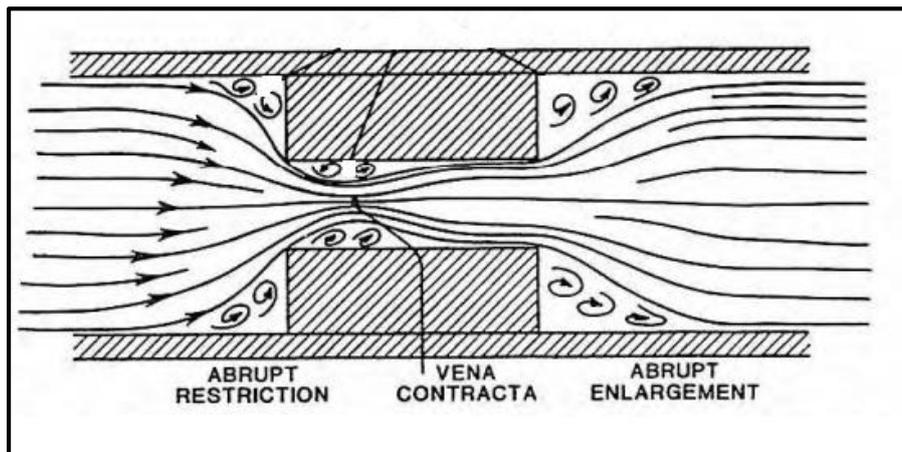


Figure 2.4 : schéma d'écoulement à travers une duse

La fonction principale d'une duse est de dissiper de grandes quantités d'énergie potentielle (c'est-à-dire des pertes de pression) sur une très courte distance. La conception d'une duse prend l'avantage du régime d'écoulement résultant d'une perturbation soudaine de l'écoulement continu dans une conduite circulaire. Il existe un effet combiné d'une restriction soudaine à l'entrée (abrupt restriction), d'un tube d'écoulement de petit diamètre et d'un élargissement brusque à la sortie (abrupt enlargement).

Comme c'est illustré dans la figure (2.4), à l'approche de l'orifice, le fluide quitte la paroi du tube et se contracte pour former un jet à grande vitesse. Le jet converge vers un minimum appelé gorge ou « vena contracta », puis il s'élargit vers la paroi de l'alésage d'étranglement.

Une zone de turbulence à l'entrée et à la sortie de l'étrangleur contribue également à la perte de pression. Les pertes irréversibles totales sont résumées dans ce qui suit :

- La friction dans l'ensemble de la duse et des zones qui lui sont proches.
- La turbulence près de l'entrée et de la sortie de la duse.
- Mouvements tourbillonnaires lents entre le jet contracté et les parois de la conduite.
- Expansion brusque à la sortie du starter.

Une observation importante sur la nature générale de l'écoulement compressible à travers un étrangleur est qu'il existe un débit maximum qui peut s'écouler à travers l'orifice pour des conditions amont données.

### Fonctions principales d'une duse :

Les duses sont largement utilisés dans l'industrie pétrolière et gazière pour :

- Contrôler le flux du puits ;
- Produire le réservoir au taux le plus optimal ;
- Empêcher la formation de cônes d'eau ou de gaz ;
- Maintenir la contre-pression ;
- Protéger la formation et les équipements de surface contre les fluctuations anormales de pression.

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

La chute de pression à travers les duses est généralement très importante. Il n'existe pas d'équation universelle pour prédire cette chute de pression pour tous les types de fluides de production. Différents modèles d'écoulement sont disponibles dans la littérature, et ils doivent être choisis en fonction de la fraction de gaz dans le fluide et des régimes d'écoulement, c'est-à-dire critique ou sous-critique.[13]

Le régime critique apparaît lorsque la vitesse d'écoulement du fluide à travers une duse atteint la vitesse de déplacement du son dans le fluide dans les conditions in situ, l'écoulement est donc appelé écoulement critique (ou sonique). Dans ces conditions, l'onde de pression en aval de la duse ne peut pas aller en amont. Ainsi, tout changement de la pression aval ne peut être détectée par le manomètre amont, de même, tout changement de la de la pression amont ne peut pas non plus être détectée par le manomètre aval.

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.

Une équation généralisée pour décrire le type du régime d'écoulement a été développée pour un écoulement isentropique de gaz à travers une restriction, elle est présentée sous la forme suivante :

$$\left(\frac{P_{up}}{P_{outlet}}\right) = \left[1 + \left(\frac{k-1}{2}\right) M^2\right]^{k/k-1} \quad (2.13)$$

Où :  $k = C_p/C_v$ ,  $P_{up}$  est la pression en amont de la duse,  $P_{outlet}$  : est la pression en aval de la duse,  $k$ : le rapport de chaleur spécifique et  $M$  : est le nombre de Mach. Si un écoulement

sonique existe ( $M = 1$ ) et qu'un écoulement de gaz naturel est considéré,  $k = 1.28$  l'équation (2.13) devient :

$$\left(\frac{P_{outlet}}{P_{up}}\right) = \left(\frac{2}{k+1}\right)^{k/k-1} = 0.55 \quad (2.14)$$

L'équation (2.14) est utilisée pour identifier si le régime d'écoulement est critique ou sous-critique, le terme gauche est appelé rapport de pression critique

Ainsi, si le rapport des pressions est inférieur au rapport de pression critique, un écoulement critique (sonique) existe. Si le rapport est supérieur ou égal au rapport de pression critique, un écoulement sous-critique (subsonique) existe.

Bien que la théorie ci-dessus ait été développée pour l'écoulement des gaz, des observations similaires pour l'écoulement critique des liquides à travers des restrictions s'appliquent également. Puisque l'écoulement sonique pour les gaz et les liquides apparaît à des vitesses différentes, on doit se poser la question - Quel est le flux critique pour un mélange biphasé ? Cette question n'a pas encore trouvé de réponse adéquate. Le critère normalement appliqué est logiquement que le débit critique existe si la réduction de la pression en aval n'augmente pas le débit. [14]

La figure ci-dessous montre une courbe typique de la variation du débit en fonction du rapport  $P_{outlet}/P_{up}$

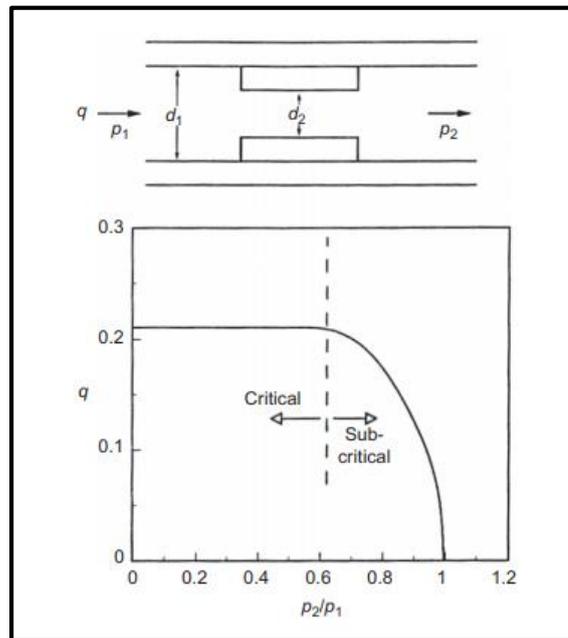


Figure 2.5 : Courbe typique de performance d'une duse [13]

### 2.1.2 .3 Modélisation des écoulements dans les restrictions en utilisant les corrélations existantes dans la littérature

#### Écoulements monophasiques : Equation de Bernoulli

Bernoulli a développé une équation à partir de principe de conservation d'énergie pour modéliser un fluide monophasique dans une conduite. Sa formule est décrite comme suit :

$$\frac{v_1^2}{2g} + \frac{P_1 g_c}{\rho_1 g} + z_1 = \frac{v_o^2}{2g} + \frac{P_o g_c}{\rho_o g} + z_o \quad (2.15)$$

L'équation de continuité qui décrit le principe de conservation de masse s'écrit :

$$\rho_1 v_1 A_1 = \rho_o v_o A_o \quad (2.16)$$

Après la combinaison de l'équation (1.15) et (1.16), en supposant un écoulement incompressible et que  $z_1 = z_o$ , où les conditions correspondent à l'écoulement dans la restriction on obtient :

$$q = v_o A_o = \frac{C_d}{\sqrt{1 - \left(\frac{A_o}{A_1}\right)^2}} \cdot A_o \sqrt{\frac{2 g_c \Delta P}{\rho}} \quad (2.17)$$

L'équation précédente est fréquemment écrite :

$$q = C A_o \sqrt{\frac{2 g_c \Delta P (144)}{\rho}} \quad (2.18)$$

$C$  et  $C_d$  représentent le coefficient d'écoulement et le Coefficient de décharge respectivement.  $\rho$  est la masse volumique du fluide donné en lbm /ft<sup>3</sup> et le débit en ft<sup>3</sup>/ sec.

Le coefficient d'écoulement  $C$  se détermine expérimentalement. Il dépend du type de restriction (orifice, venturi...) et du rapport entre le diamètre de la restriction et celle de la conduite.

Pour des écoulements de gaz, l'équation d'énergie, l'équation d'état de gaz et l'équation de continuité sont combinées pour avoir un résultat similaire à l'équation dérivée pour les liquides, avec l'introduction en plus, du facteur d'expansion du gaz  $Y$ . L'équation finale pour les gaz s'écrit :

$$q_g = Y C A_o \sqrt{\frac{2 g_c \Delta P (144)}{\rho}} \quad (2.19)$$

$Y$  décrit l'effet de compressibilité du gaz, il est donné pour le cas d'un orifice comme :

$$Y = 1 - [0.41 + 0.35 \left(\frac{d}{D}\right)^4] \left(\frac{1}{k}\right) \left(\frac{\Delta P}{P_1}\right) \quad (2.20)$$

Où  $d$  représente le diamètre de l'orifice,  $D$  est le diamètre de la conduite,  $k$  = rapport de chaleur spécifique  $C_p/C_v$

### Écoulements diphasiques :

Plusieurs approches ont été suggérées pour la prédiction des débits d'écoulement multiphasique à travers les duses. Ces approches peuvent être classées en quatre groupes. Le premier groupe consiste en des corrélations similaires à celle proposée à l'origine par Gilbert, où les propriétés PVT des fluides produits ont été négligées. Le deuxième groupe combine les corrélations empiriques du calcul du débit avec les propriétés des fluides. Un exemple de corrélations appartenant à ce groupe sont celles de Al-Attar et Abdul-Majeed, et Al-Towailib et Al-Marhoun. Le troisième groupe concerne les corrélations basées sur l'analyse dimensionnelle. Une corrélation typique a été développée par Omana et al. Le quatrième groupe comprend les approches théoriques essentiellement dérivées du bilan énergétique de l'écoulement des fluides à travers une restriction tel que la corrélation de Ashford et Pierce.

Dans cette section quelques corrélations pour le calcul du débit d'écoulement à travers les duses seront présentées pour le régime d'écoulement critique et sous critique.

### Régime critique

#### ❖ La corrélation de Gilbert

Gilbert a développé une corrélation en utilisant 268 points de données provenant du champ Ten Section de Californie. Il a mis en relation la pression en amont de la duse, diamètre de la duse, le GOR et le débit de liquide produit. La corrélation de Gilbert est décrite sous la forme :

$$Q = a \frac{P_u \cdot \Phi^b}{GLR^c} \quad (2.21)$$

Il existe également une autre forme de l'équation de Gilbert ; décrite dans le simulateur commercial PIPESIM de Schlumberger, en introduisant une puissance à l'équation originale. Cette deuxième forme peut être reformulée pour s'écrire comme suit :

$$Q = \left( a \frac{P_u \cdot \Phi^b}{GLR^c} \right)^e \quad (2.22)$$

L'analyse de l'équation de Gilbert montre clairement que le taux de production est indépendant de la pression en aval de la duse ce qui est une caractéristique de l'écoulement critique. Gilbert a déclaré que sa formule (ou ses coefficients) ne s'applique que lorsque la pression en amont de la duse est supérieure d'au moins 70 % à la pression de ligne (pression en aval). Il a également souligné que son équation est très sensible à la taille de la duse et qu'une erreur de 1/128 de pouce dans sa taille peut entraîner des erreurs considérables dans l'estimation du débit.

Cette corrélation a été ensuite ajustée par d'autres auteurs pour l'adapter à d'autres champs pétroliers. Les coefficients a, b, c et e des différents auteurs sont cités dans le tableau ci-dessous :

Tableau 2.1 : constantes de la corrélation de Gilbert suivant différents auteurs [19],[20]

Corrélation	a	b	c	e
Gilbert	0.1	1.89	0.546	1
Achong	0.26	1.88	0.650	1
Baxendall	0.104	1.93	0.546	1
Ros	0.057	2	0.5	1
Pilehvari	0.021	2.11	0.313	1

#### ❖ La corrélation de Omana

Omana a utilisé des données de terrain prises dans les installations de Union Company of California's Tiger Lagon Field en Louisiane pour développer sa corrélation. Des expériences sur le terrain ont été menées avec de l'eau et du gaz naturel. Il a ensuite effectué une analyse adimensionnelle pour résoudre ce problème d'écoulement diphasique à travers les duses. Omana a introduit de nouveaux paramètres dans son équation tel que la tension de surface et la masse volumique. L'équation de Omana est définie comme suit :

$$Q_l = a \sigma_L^b \cdot \rho_L^c \cdot (1 + R_L)^e \cdot d_{bean}^f \cdot \rho_G^m \cdot P_{up}^n \quad (2.23)$$

Avec :

$$R_L = \frac{T_{up} Z_{up} (R - R_s)}{198.6 P_{up}} \quad (2.24)$$

Tableau 2.2 : les coefficients de la corrélation de Omana

a	b	c	e	f	m	n
$1.95 \times 10^{-3}$	-1.245	1.545	-0.657	1.8	-3.49	3.19

Omana a déclaré que sa corrélation était valable pour le régime critique et que la fiabilité des résultats sera probablement moindre à des débits plus importants (>800 bbl/d) et pour des tailles de duses supérieures à 14/64 pouces.

#### ❖ La corrélation d'Ashford

La formule d'Ashford et Pierce est valable pour le régime critique et non critique. Elle est décrite comme suite :

$$q_o = \frac{0.858 C \beta (d_{bean})^2}{A \sqrt{P_u + 0.56}} \cdot \sqrt{\frac{A + 0.76 P_u}{[B + 0.01353 \gamma_g R_p]^2 [B + 0.01353 \gamma_g R_s]^{-1}}} \quad (2.25)$$

Avec  $A = 5.04 \cdot 10^{-3} T_1 Z_1 (R_p - R_s)$  (2.26)

$$B = 62.4 [\gamma_{o\ sc} + \gamma_{w\ sc} WOR] \quad (2.27)$$

$$\beta = 1/\sqrt{B_o WOR} \quad (2.28)$$

$R_p$  est le rapport gaz / liquide produit (scf/stb),  $R_s$  est le rapport gaz/ liquide au conditions de pression  $P_u$ , et  $C$  est le coefficient d'écoulement.

Ashford a vérifié à nouveau sa corrélation sur 14 puits et a constaté que le coefficient de décharge nécessaire pour prédire les taux de production mesurés allait de 0,642 à 1,218.

#### ❖ La corrélation de Towailib et Marhoun

Ces deux auteurs ont développé une corrélation qui semble à celle de Gilbert en remplaçant le GLR avec la densité du mélange. Elle était basée sur 3554 données de tests de production provenant de dix champs du Moyen-Orient. La corrélation s'applique aux conditions d'écoulement critique. La corrélation prend la forme suivante :

$$Q = \frac{P_u^b \cdot \phi^c}{a * \gamma_m^d} \quad (2.29)$$

Où :  $\gamma_m = \gamma_o + 2.18 * 10^{-4} R_p \gamma_g$  (2.30)

Tableau 2. 3 : Les coefficients de la corrélation de Towailib et Marhoun

a	b	c	d
545.5	0.9813	2.0704	1.4637

#### ❖ La corrélation de Attar et Abdul Majeed

Ces deux auteurs ont développé une corrélation sur la base de données des champs de l'est du Bagdad. Cette corrélation ressemble à celle de Gilbert, avec l'introduction de la densité de l'huile au dénominateur de la formule pour qu'elle prenne la forme suivante :

$$Q = a \frac{P_u^b \cdot \phi^c}{g^d GLR^e} \quad (2.31)$$

Où  $g$  est la densité API, elle est calculée par :

$$^{\circ}API = \frac{141}{\gamma_o} - 131.5 \quad (2.32)$$

Tableau 2.4 : les coefficients de la corrélation de Attar et Abdul Majeed

a	b	c	d	e
0.033567	0.8756	1.796	0.43957	0.2693

### Régime sous-critique

#### ❖ La corrélation de Elgibaly et Nashawi

Nashawi et Elgibaly ont développé une formule pour le régime non critique où ils ont remplacé la pression en amont dans la formule de Gilbert avec la différence de pression entre l'amont et l'aval de la duse.

$$Q = \left[ \frac{a \Delta P \phi^b}{GOR^c} \right]^{\frac{1}{2}} \quad (2.33)$$

Tableau 2.5 : les coefficients de Elgibaly et Nashawi

a	b	c
1204.8193	2.747	1.13501

#### ❖ La corrélation mécaniste

L'équation mécaniste est développée pour le régime non critique. C'est une combinaison de l'équation de Bernoulli (1) avec l'équation de continuité ( $\rho A_1 V_1 = \rho A_2 V_2$ )

La chute de pression à travers la duse est calculée par la somme pondérée de la chute de pression dans la phase liquide et la phase gazeuse :

$$\Delta P = \lambda_L \Delta P_L + \lambda_G \Delta P_G \quad (2.34)$$

Où  $\lambda_L$  et  $\lambda_G$  représentent les fractions s'écoulant des phases liquide et gazeuse

La chute de pression en phase liquide est :

$$\Delta P_L = \frac{\rho_n}{2C} \left( \frac{v}{c_{vL} \cdot Z_L} \right)^2 \quad (2.35)$$

La chute de pression en phase gazeuse est :

$$\Delta P_G = \frac{\rho_n}{2C} \left( \frac{v}{c_{vG} \cdot Z_G} \right)^2 \quad (2.36)$$

La vitesse de mélange à travers la duse est calculée en fonction du débit massique  $q$  (lb/s), la masse volumique de non glissement (en lb/ft<sup>3</sup>) et la surface de la section de la duse  $A_{bean}$  (en ft<sup>2</sup>)

$$v = \frac{q}{A_{bean} \cdot \rho_n} \quad (2.37)$$

$$A_{bean} = \pi \cdot \frac{(d_{bean})^2}{4} \quad (2.38)$$

$$\rho_n = \lambda_L \rho_L + \lambda_G \rho_G \quad (2.39)$$

La compressibilité de liquide  $Z_L$  et de gaz  $Z_G$  sont sonnées par :

$$Z_L = 1 \text{ et } Z_G = 1 - \frac{0.41 + 0.35\delta^4}{\gamma} \left( \frac{\Delta P}{P_{up}} \right) \quad (2.40)$$

c : facteur de conversion pour des unités d'ingénierie (c=144.g) en lb/ ( ft.s<sup>2</sup>)/psi

#### ❖ Extensions de l'équation de Gilbert

Il existe quelques extensions de la forme de Gilbert pour décrire le débit à travers une duse pour le régime non critique. Parmi lesquelles les extensions suivantes et qui seront utilisées dans ce qui suit.

##### ✓ 1<sup>ère</sup> forme :

Cette forme a été proposée par Mohammad Ali Kargarpour dans son article intitulé « Oil and gas well rate estimation by choke formula: semi-analytical approach ». Il s'est basé sur concepts de base de la mécanique des fluides, il a démontré que le débit à travers une duse dépend généralement de la chute de pression à travers l'étrangleur, en plus d'autres facteurs. Cette nouvelle formule a été validée en utilisant des données de terrain comprenant 399 points de données. L'équation s'écrit comme suit :

Le débit  $Q$  est exprimé en STB/day, le  $GOR$  en scf/STB, la pression en psi et le diamètre en 1/64 inch

$$Q = a \frac{P_u^b \Delta P^c d^e}{GOR^f} \quad (2.41)$$

Tableau 2. 6 : les coefficients de l'extension de l'équation de Gilbert

a	b	c	e	f
0.098	0.41	0.44	2	0.42

##### ✓ 2<sup>ème</sup> forme :

Dans ce cas, une fonction qui dépend de la pression amont  $P_u$  et aval  $P_d$  de la duse respectivement. Ces extensions sont utilisées au champ de Hassi Messaoud (selon [19])

##### ❖ 1<sup>ère</sup> extension utilisée à HMD :

$$Q = \frac{a P_u \cdot \phi^b}{GLR^c} \cdot f(P_u, P_d); \quad f = \frac{P_d}{P_u} \cdot \sqrt{\left(\frac{P_u}{P_d}\right) - 1} \quad (2.42)$$

##### ❖ 2<sup>ème</sup> extension utilisée à HMD :

$$Q = \frac{a P_u \cdot \phi^b}{GLR^c} \cdot f(P_u, P_d); \quad f = \left( \frac{P_d}{P_u} \cdot \sqrt{\left(\frac{P_u}{P_d}\right) - 1} \right)^e \quad (2.43)$$

Le débit  $Q$  est exprimé en STB/day, le  $GOR$  en scf/STB, la pression en psi et le diamètre en 1/64 inch

## 2.2 Calcul de production :

### 2.2.1 Système de production :

Un système de production est essentiellement un système qui transporte le fluide du réservoir vers la surface et le sépare en pétrole, gaz et eau. Les flux de pétrole et de gaz sont traités si nécessaire et préparés pour être vendus ou transportés hors du champ. L'eau produite est également traitée et préparée pour être éliminée ou réinjectée dans le réservoir. Les éléments mécaniques de base d'un système de production sont les suivants [18] :

1. les puits
2. les conduites d'écoulement
3. le manifold de production
4. les séparateurs et les équipements de traitement
5. les instruments de mesure
6. les cuves de stockage

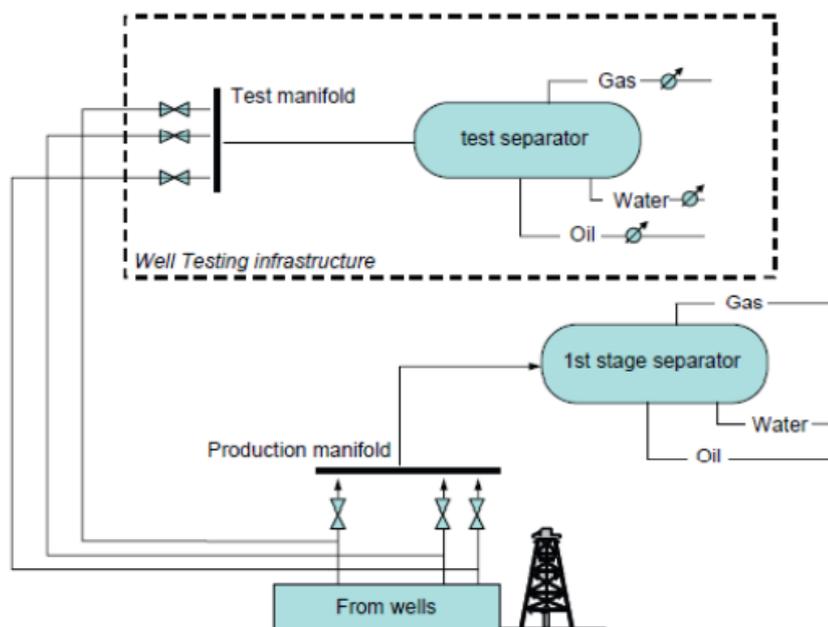


Figure 2.6 : un système de production simple avec trois puits de pétrole à écoulement naturel vers séparateurs [23]

La production des trois puits de la figure 2.6 passe par les conduites d'écoulement jusqu'au collecteur de production. Grâce à un dispositif de vanne approprié dans le collecteur, il est possible de dévier le flux vers l'un des deux séparateurs illustrés. Dans les séparateurs, le fluide produit est divisé en flux liquides et gazeux, qui sont mesurés et transférés vers un traitement ultérieur ou vers la vente. Dans ce système de production simple, un séparateur sert au traitement continu de la production (séparateur de production) tandis que le second est utilisé pour les tests de routine des puits individuels (séparateur de test).

Pour voir de plus près les éléments d'un système de production simple, la figure ci-dessous montre un puits de pétrole à écoulement naturel. Les principaux éléments d'un puits de production sont : le tubage de production (production casing), le tubing, le packer de production, la tête de puits et le tubing hanger, l'arbre de Noël qui comporte plusieurs vannes (master valve, wing valve et swab valve) et la duse (choke).

La duse est placée juste à la connexion entre la sortie de la tête de puits et le flowline. Comme c'est illustré dans la figure (2.7)

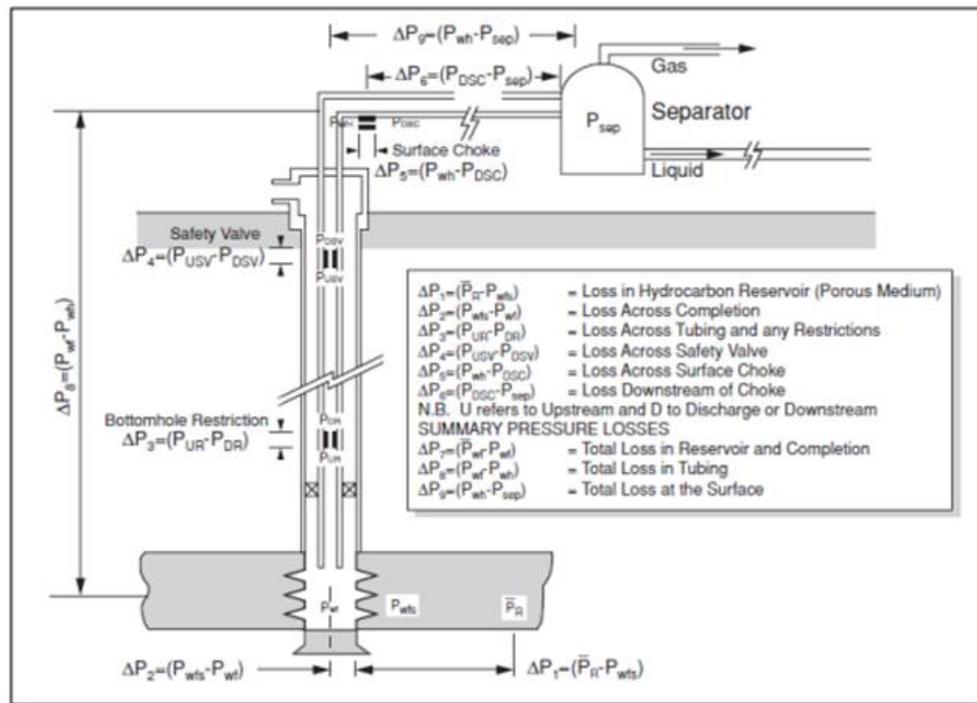


Figure 2.7 : Système de production d'un puits simple

## 2.2.2 Principe de calcul de production

Il existe principalement 3 approches pour le calcul du débit produit par chaque puits à savoir : la mesure directe, la modélisation physique et l'approche data driven.

### ❖ Modélisation physique de la duse :

La modélisation physique consiste principalement à faire appel aux principes fondamentaux de la physique du procès et à des expériences passives ou actives afin de fermer le système pour obtenir des corrélations hybrides pouvant décrire les écoulements à travers les duses.

Plusieurs auteurs ont essayé de quantifier le débit d'huile sortant à travers les duses en utilisant différentes corrélations à savoir : Gilbert, Achong, Ros, Ashford et pierce et Omana ainsi que différents modèles élaborés dans la section précédente. Ce type de modélisation est considéré comme très pratique, simple et une solution très économique pour l'estimation du débit de production avec un degré acceptable de certitude.

#### ❖ Mesures directs :

Ces mesures sont considérées les plus exactes. Dans ce cas, un séparateur de test est connecté au puits et une opération de jaugeage est envisagée. Le fluide à sa sortie du puits est dirigé directement vers un séparateur pour le séparer en huile, eau et gaz, le débit de chaque phase est ainsi évalué. Cette opération nous permet également de mesurer d'autres paramètres tel que la pression en tête du puits, la pression pipe, le diamètre de la duse, puis le calcul de certains paramètres comme GOR, masse volumique, facteur volumétrique de formation ... Ces opérations sont considérées souvent coûteuses donc elles ne sont pas envisagées tout le temps.

Au début des années 1980, l'industrie pétrolière et gazière a commencé à s'intéresser au développement des débitmètres multiphasiques (MPFM). Les MPFM sont capables de mesurer avec précision les débits de pétrole, de gaz et d'eau de chaque puits et également d'un groupe de puits sans séparation des phases. Cependant, l'installation et la maintenance de ces MPFM sont généralement coûteuses et prennent du temps. Ils sont souvent installés dans des systèmes de production sous-marins. [24]

#### ❖ Méthodes Data driven :

En de nombreuses occasions, les mesures physiques directes sont indisponibles puisque ce type de mesures ne se font que périodiquement, de plus, chaque puits dispose d'un historique de données assez intéressant qu'on peut exploiter via les modèles data driven. L'approche Data Driven se base sur l'utilisation combinée des données et des modèles de machine et de deep learning qui ont connu un développement remarquable dernièrement. Cette approche produit des résultats encourageants lorsque les données mesurées sont suffisantes, y compris des tests de puits fréquents (par exemple, 8 à 12 par an), des capteurs à la tête de puits et les capteurs de la conduite d'écoulement (pression, température).

Plusieurs modèles de machine learning peuvent être utilisés dans ce contexte à savoir : les réseaux de neurones artificiels ANN, genetic programming, gaussian process, les réseaux de neurones récurrents RNN comme les modèles long short term memory LSTM ...

### 2.2.3 Analyse nodal

L'approche de l'analyse des systèmes, souvent appelée Analyse Nodal (NODAL Analyse), est appliquée depuis de nombreuses années pour analyser les performances des systèmes comprenant plusieurs composants en interaction à titre. Les circuits électriques, les réseaux complexes de pipelines ...

L'application de l'Analyse Nodal aux systèmes de production de pétrole et de gaz a été proposée pour la première fois par Gilbert en 1954, puis discutée par Nind en 1964 et Brown en 1978.

La procédure consiste à sélectionner un point dans le système appelé nœud, puis à diviser le système en ce point. [17]

Tous les composants en amont du nœud représentent la partie Inflow, alors que la partie Outflow est constituée par tous les éléments en aval du nœud. Une relation entre le débit et la chute de pression doit être établie pour chaque élément du système. Le débit à travers le système est déterminé une fois les conditions suivantes sont satisfaites :

- Le débit entrant dans le nœud est égal au débit sortant à travers ce dernier.
- Au niveau du nœud, une seule pression peut exister.

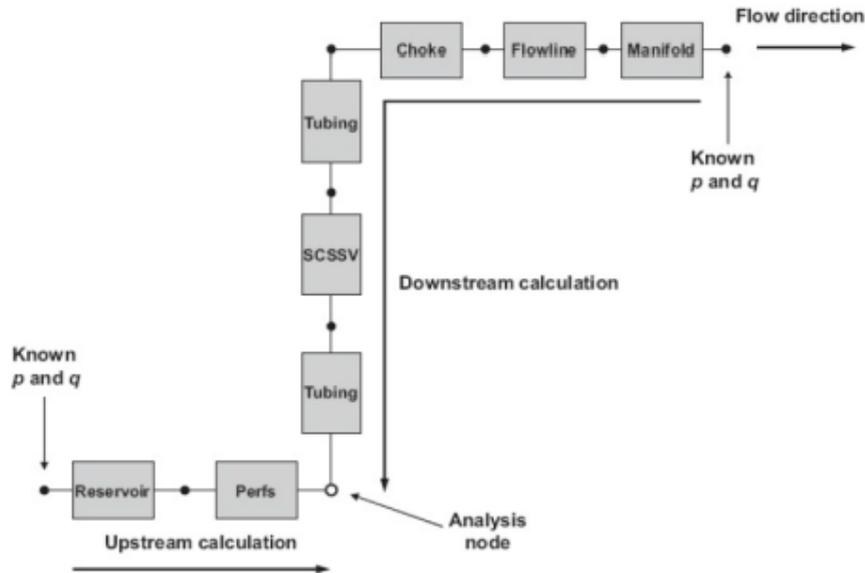


Figure 2.8 : Exemple de configuration d'analyse nodale avec un nœud au fond du puits [36]

A un temps donné de la durée de vie d'un puits, il existe toujours deux pressions fixes, qui ne sont pas en fonction du débit : la pression du réservoir  $P_R$  et la pression à la sortie du système qui est généralement la pression du séparateur  $P_{sep}$ . Une fois le nœud est sélectionné, la pression à ce dernier  $P_{noeud}$  est déterminée comme suit :

- Inflow : 
$$P_{noeud} = P_R - \Delta P(\text{composants en amont du noeud})$$
- Outflow : 
$$P_{noeud} = P_{sep} + \Delta P(\text{composants en aval du noeud})$$

Puisque la chute de pression  $\Delta P$  à travers chaque composant varie en fonction de débit le débit  $q$ , donc le tracé de la pression du nœud  $P_{noeud}$  en fonction du débit va produire deux courbes (courbe inflow et courbe outflow) dont l'intersection donne le point de fonctionnement du système (figure 2.9)

L'analyse nodal est une méthode très flexible qui peut être utilisée pour améliorer la performance des systèmes de production et optimiser les dimensions des conduites d'écoulement, des duses ainsi que d'autres équipements.

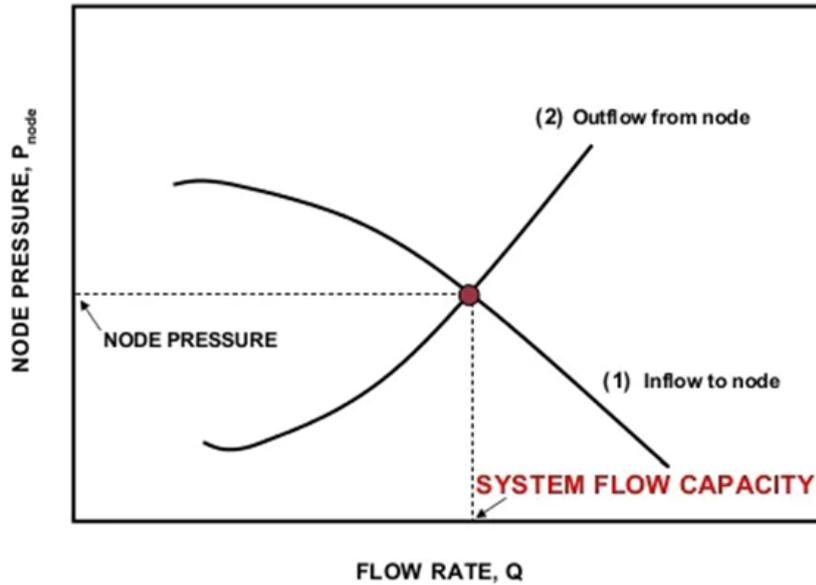


Figure 2. 9 : Tracé de la pression du nœud en fonction du débit et illustration du point de fonctionnement

#### 2.2.4 Performances des réservoirs :

"Inflow performance relationship" (IPR) est habituellement utilisée pour définir la relation entre le débit d'hydrocarbure en surface et la pression de fond en écoulement

L'équation (IPR) la plus simple et la plus utilisée est la ligne droite, qui indique que le débit est directement proportionnel à la chute de pression dans le réservoir. La constante de proportionnalité ( $J$ ) est appelée indice de productivité définie comme étant le nombre de barils de liquide qui peut être produit pour une valeur donnée de draw-down sur le réservoir. Cette méthode s'applique seulement pour des réservoirs sous saturé ; où la pression est supérieure à la pression de bulle. L'équation peut être écrite comme suit :

$$q_o = J(\bar{P}_r - P_{wf}) \quad (2.44)$$

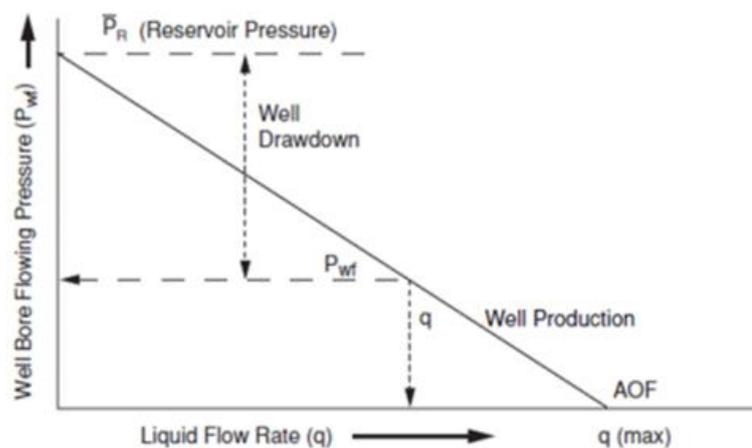


Figure 2.10 : IPR en utilisant la méthode de l'indice de productivité

Pour des réservoirs sur-saturé (figure 2.11) où la pression est inférieure à la pression de bulle ( $\bar{P}_r < P_b$ ), Vogel a introduit un modèle mathématique généralisé pour estimer la productivité d'un réservoir. Il s'écrit sous la forme :

$$\frac{q_0}{q_{0(\max)}} = 1 - 0,2 \cdot \left( \frac{P_{wf}}{P_R} \right) - 0,8 \cdot \left( \frac{P_{wf}}{P_R} \right)^2 \quad (2.45)$$

Où  $q_{0\max}$  est le débit maximum avec une pression fond nulle  $P_{wf} = 0$

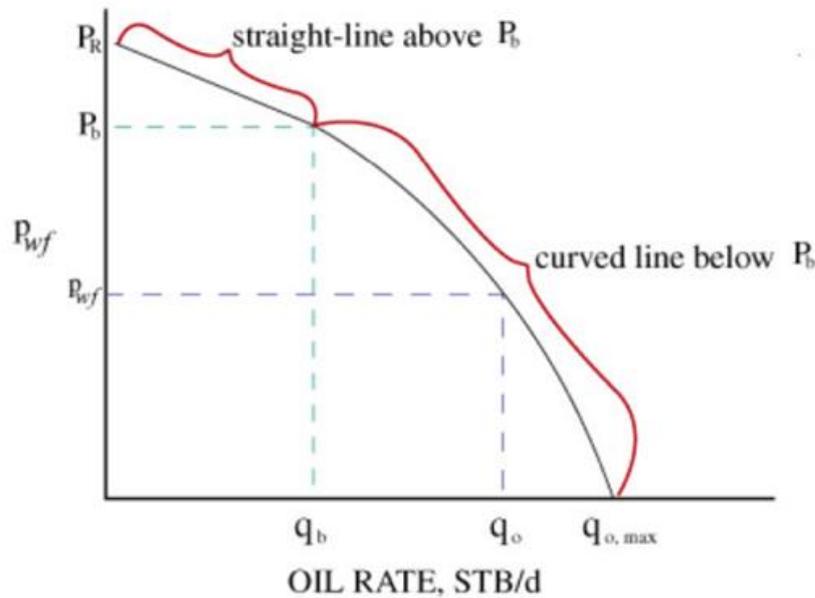


Figure 2. 11 : IPR pour un réservoir sur saturé

### 2.2.5 Performances du tubing :

Près de 80 % de la perte de pression totale dans un puits de pétrole peut se produire lors du déplacement des fluides dans le tubing du fond du trou vers la surface.

La chute de pression nécessaire pour soulever un fluide dans le tubing à un débit donné est l'un des principaux facteurs déterminant la productivité d'un puits. La chute de pression le long du tubing peut être calculée à l'aide de corrélations (développées au chapitre 1). La relation qui résulte entre la pression d'écoulement du fond du puits et le débit de pétrole est appelée relation de performance du tubage (TPC) ou outflow performance relationship (OPR) Il existe une taille de tubing optimale pour tout système de puits. Un tubing trop petit limitera le taux de production en raison d'une perte de friction excessive, tandis qu'un tubing trop grand entraînera le puits à se charger de liquides ce qui va tuer le réservoir.

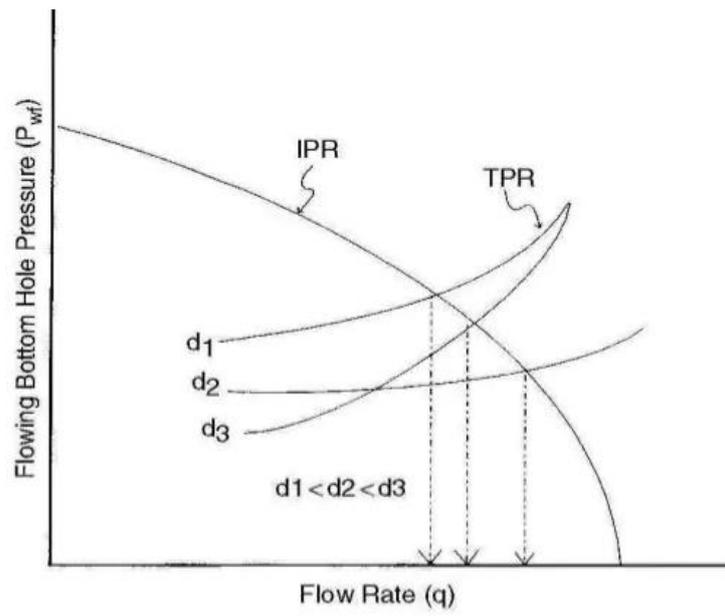


Figure 2.12 : illustration de différentes courbes TPR pour différent diamètres de tubing

## Chapitre 3 :

### Outils mathématiques

#### 3.1 Préparation des données :

Avec l'arrivée de l'Internet des objets, et le développement de la puissance de calcul, les données se multiplient et deviennent plus accessibles. Elles sont désormais une nouvelle arme dans notre ère. Elles constituent le centre des algorithmes de machine learning.

Les données du monde réel contiennent généralement des bruits, des valeurs manquantes, et peuvent être dans un format inutilisable. Par conséquent, préparer au mieux ces données, permettra d'en tirer le maximum de valeur. Il faut noter que l'étape de préparation des données (Data pre-processing) représente le plus grand temps passé sur un projet basé sur celles-ci comme ceux de Machine Learning.

La performance des modèles créés à partir des données dépend non seulement de **quantité** de ces dernières mais aussi de leurs qualités qui représente une importance majeure. Voici la liste de quelques étapes à suivre pour bien préparer nos données :

- ❖ **Obtention du dataset** : pour créer un modèle, la première chose dont nous avons besoin est un ensemble de données. Le dataset peut être sous différents formats, pour l'utiliser dans notre code, nous le plaçons généralement dans un fichier csv, xlsx.
- ❖ **Traitement des données manquantes** : l'étape suivante de préparation des données consiste à traiter les données manquantes. Il existe principalement deux façons de le faire :
  - ✓ Supprimer la ligne en question : elle est utilisée pour traiter couramment les valeurs nulles. Dans ce cas, nous supprimons simplement la ligne ou la colonne spécifique qui contient des valeurs nulles. Mais cette méthode n'est pas si efficace et la suppression des données peut entraîner une perte d'informations.
  - ✓ Calculer la moyenne : de cette façon, nous calculons la moyenne de la colonne ou de la ligne qui contient une valeur manquante et nous la plaçons à la place de la valeur manquante. Cette stratégie est utile pour les caractéristiques qui ont des données numériques telles que l'âge, le salaire, l'année, etc.
- ❖ **Encodage des données catégorielles** : les données catégorielles sont des données qui comportent certaines catégories (exemple : huile/ gaz). Puisque le modèle d'apprentissage automatique par exemple fonctionne entièrement sur les mathématiques et les nombres, donc il est nécessaire d'encoder les variables catégorielles en nombres. (exemple : homme=0, femme=1)
- ❖ **Data splitting** (Fractionnement des données) : l'overfitting est l'un des principaux défis rencontrés pendant l'entraînement des modèles de machine learning. C'est le résultat d'entraînement du modèle sur des détails infimes ou des bruits présents dans les données d'apprentissage. Cela peut conduire à de bonnes performances sur les données d'apprentissage, mais la précision du modèle chute de manière significative pour les

données non vues. Pour aider les modèles à devenir plus généralisés, le fractionnement de l'ensemble de données est essentiel. Un ensemble de données est souvent décomposé en trois ensembles d'observations :

- ✓ **Training data (données d'entraînement)** : cet ensemble de données est utilisé par l'algorithme pour apprendre les paramètres du modèle.
  - ✓ **Validation data (données de validation)** : Cet ensemble de données est utilisé pour évaluer la performance de prédiction du modèle pendant le processus de formation. La tendance observée dans les erreurs de prédiction en utilisant les données de validation peut aider à identifier le nombre d'itérations pour un entraînement adéquat. Lors de l'entraînement d'un modèle typique, l'erreur de prédiction doit continuer à diminuer pour les l'ensemble de données d'apprentissage et de validation tout au long du processus d'entraînement. Cependant, si nous observons un point où l'erreur de prédiction continue à diminuer pour les données de test mais commence à augmenter pour les données de validation, entraînement doit être arrêté
  - ✓ **Test data (données de test)** : Il s'agit de l'ensemble de données qui est utilisé uniquement pour évaluer les performances d'un modèle entièrement entraîné. La précision de la prédiction sur les données de test est un indicateur de la performance du modèle sur les données inconnues rencontrées.
- ❖ **Normalisation des données** : une étape importante du prétraitement des données consiste à remettre à l'échelle les caractéristiques d'entrée et la variable de sortie (pour les problèmes de régression), et ceci afin de les rendre cohérentes avant de les introduire dans l'algorithme d'apprentissage automatique. Pour comprendre ce concept, considérons un exemple avec différentes plages de donnée, telle que la production cumulée (plage de 1000 barils à 10 000 barils). Alors qu'une autre caractéristique peut avoir une très petite plage, comme la porosité (de 0,05 à 0,60).
- ❖ **Normalisation** : la normalisation est un processus de remise à l'échelle des données à partir de la gamme originale afin que les valeurs se situent dans l'intervalle de 0 à 1 ou de -1 à 1. Cette méthode peut être appliquée lorsque les limites supérieures  $X_{max}$  et inférieures  $X_{min}$  approximatives des données sont connues, le calcul se fait comme suit :

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (3.1)$$

- ❖ **Standardisation (Z-Score Normalization)** : la standardisation (également appelée Z-Score Normalization en anglais) est une technique de mise à l'échelle telle que, lorsqu'elle est appliquée, les caractéristiques seront remises à l'échelle de manière à ce qu'elles aient les propriétés d'une distribution normale standard avec une moyenne,  $\mu=0$  et un écart type,  $\sigma=1$ . Les scores standard (également appelés z-scores) des échantillons sont calculés comme suit :

$$z = \frac{x - \mu}{\sigma} \quad (3.2)$$

Où :  $\mu$  est la moyenne et  $\sigma$  est l'écart type de la moyenne,  $z$  est le score standard et  $x$  est la donnée d'entrée.

- ❖ **Feature engineering** : Il consiste à la création de nouvelles caractéristiques ou variables (factures). Exemple : si on prend un dataset de production de pétrole qui contient les features ;  $x_1$  = débit d'huile et  $x_2$  = débit de gaz, alors il est possible de créer  $x_3 = x_2 / x_1$  qui équivaut au GOR.

## 3.2 Régression non linéaire et algorithme de Levenberg Marquardt

Lors de l'ajustement d'une fonction modèle  $f(x_i, \beta)$  d'une variable indépendante  $x$  et d'un vecteur de  $n$  paramètres  $\beta$  à un ensemble de  $m$  points de données  $(x_i, y_i)$ , il est habituel de minimiser la somme des carrés des erreurs (ou de résidus) entre les données  $y_i$  et la fonction d'ajustement de la courbe  $f(x_i, \beta)$ .

$$E(\beta) = \sum_{i=1}^m [y_i - f(x_i, \beta)]^2 \quad (3.3)$$

Si la fonction d'ajustement n'est pas linéaire dans ses paramètres, le problème des moindres carrés nécessite un algorithme de solution itératif. L'algorithme de Levenberg-Marquardt a été développé au début des années 1960 pour résoudre problèmes non linéaires de moindres carrés.

L'algorithme de Levenberg-Marquardt combine deux algorithmes de minimisation numérique : la méthode de descente du gradient et la méthode de Gauss-Newton. Dans la méthode de descente du gradient, la somme des erreurs au carré est réduite en actualisant les paramètres dans la direction de la descente la plus raide. Dans la méthode de Gauss-Newton, la somme des erreurs au carré est réduite en supposant que la fonction des moindres carrés est localement quadratique dans ses paramètres, et en trouvant le minimum de cette quadratique.

La méthode de Levenberg-Marquardt agit plus comme une méthode de descente de gradient lorsque les paramètres sont éloignés de leurs valeurs optimales, et se rapproche de la méthode de Gauss-Newton lorsque les paramètres se rapprochent de leurs valeurs optimales. [32]

La base de l'algorithme consiste à une approximation linéaire de la fonction  $f$  au voisinage de  $\beta$ , le développement en série Taylor s'écrit sous la forme :

$$f(x_i; \beta + \delta_\beta) = f(x_i; \beta) + J_i \delta_\beta \quad (3.4)$$

Avec :

$$J_i = \frac{\partial f(x_i; \beta)}{\partial \beta} \quad (3.5)$$

La somme des erreurs au carrées  $E(\beta)$  peut-être donc écrite sous la forme :

$$E(\beta + \delta_\beta) = [y_i - f(x_i; \beta) - J_i \delta_\beta]^2 \quad (3.6)$$

En partant d'une valeur initiale  $\beta_0$ , l'algorithme produit une série de vecteurs  $\beta_1, \beta_2, \dots$  qui convergent vers le minimum de  $E$ , donc à chaque itération, il faut trouver  $\delta_\beta$  qui minimise  $E(\beta + \delta_\beta)$ . En dérivant  $E(\beta + \delta_\beta)$  par rapport à  $\delta_\beta$  et annulant le deuxième terme, le résultat peut être écrit sous une notation vectorielle comme suit :

$$(\mathbf{J}^T \mathbf{J}) \delta_{\beta} = \mathbf{J}^T [\mathbf{Y} - \mathbf{F}(\beta)] \quad (3.7)$$

Où  $\mathbf{J}$  est la matrice jacobienne dont la  $i^{\text{ème}}$  ligne est égale à  $J_i \cdot \mathbf{F}(\beta)$ ,  $\mathbf{Y}$  sont des vecteurs tel que le  $i^{\text{ème}}$  composant est  $f(x_i; \beta)$  et  $y_i$  respectivement.

La contribution de Levenberg consiste à remplacer l'équation précédente par une "version amortie" tel que :

$$(\mathbf{J}^T \mathbf{J} + \lambda \mathbf{I}) \delta_{\beta} = \mathbf{J}^T [\mathbf{Y} - \mathbf{F}(\beta)] \quad (3.8)$$

Où  $\mathbf{I}$  est la matrice identité et  $\lambda$  est le facteur d'amortissement.  $\lambda$  est ajusté à chaque itération, si la réduction de  $E$  est rapide, une valeur plus petite sera utilisé se rapprochant ainsi de l'algorithme de Gauss-Newton, par contre si l'itération donne une réduction insuffisante de  $E$ , la valeur de  $\lambda$  peut être augmentée se rapprochant de l'algorithme de descente de gradient. Marquardt a suggéré de remplacer  $\mathbf{I}$  avec la diagonale de  $\mathbf{J}^T \mathbf{J}$  ce qui donne l'algorithme de Levenberg-Marquardt :

$$(\mathbf{J}^T \mathbf{J} + \lambda \mathbf{diag}(\mathbf{J}^T \mathbf{J})) \delta_{\beta} = \mathbf{J}^T [\mathbf{Y} - \mathbf{F}(\beta)] \quad (3.9)$$

### 3.3 Introduction au machine learning :

Aujourd'hui, et avec l'augmentation de la puissance de calcul et la disponibilité des données, le machine learning est devenu un outil très puissant dans l'exécution de tâches très complexes comme la reconnaissance faciale, la détection d'objet, les voitures autonomes ainsi que divers problèmes d'engineering, et bien sûr, le domaine pétrolier n'est pas épargné.

#### 3.3.1 Définition du machine learning :

Avant d'en arriver à ce qu'est le machine learning ou l'apprentissage automatique, il est nécessaire premièrement introduire le concept d'intelligence artificielle. L'intelligence artificielle (IA) est un vaste domaine dans lequel nous essayons de simuler le comportement humain dans le but de rendre les machines très puissantes pour effectuer des tâches très complexes.

Selon son inventeur, Arthur Samuel, un mathématicien américain, le machine learning est définit comme suit « **Machine Learning is the science of getting computers to learn without being explicitly programmed** ». Le machine learning est donc un domaine de l'intelligence artificielle (IA) qui donne aux systèmes la capacité d'apprendre automatiquement et de s'améliorer à partir de l'expérience sans être explicitement programmés. Ce domaine se concentre sur le développement de programmes informatiques capables d'accéder aux données et de les utiliser pour apprendre par eux-mêmes.

#### 3.3.2 Les types de machine learning :

Pour donner à un ordinateur la capacité d'apprendre, on utilise des méthodes d'apprentissage. Parmi lesquelles on compte :

- ❖ **Apprentissage supervisé** : Il consiste à relier des variables ( $x$ ) en entrée (input) à des variables ( $y$ ) en sorties (output) en laissant l'algorithme apprendre à cartographier le

chemin entre les deux. L'objectif de l'apprentissage supervisé est de produire des modèles prédictifs pertinents en entraînant des algorithmes aux travers d'exemples que nous leur fournissons. L'apprentissage supervisé peut résoudre des problèmes de classification et de régression.

- **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** : on cherche à prédire la valeur d'une variable continue, c'est-à-dire une variable qui peut prendre une infinité de valeurs. Par exemple ; le calcul du débit en fonction de quelques paramètres tel que le diamètre, la pression, GOR ...
- ❖ **Apprentissage non supervisé** : Il est utilisé lorsqu'on dispose de données en input ( $x$ ) mais pas de données correspondantes en output ( $y$ ). Le rôle de l'algorithme est alors de trouver lui-même des structures ou des associations pertinentes à l'intérieur des données.
- ❖ **Apprentissage par renforcement** : dans ce cas, un programme informatique interagit avec un environnement dynamique dans lequel il doit atteindre un certain but, par exemple conduire un véhicule ou affronter un adversaire dans un jeu.

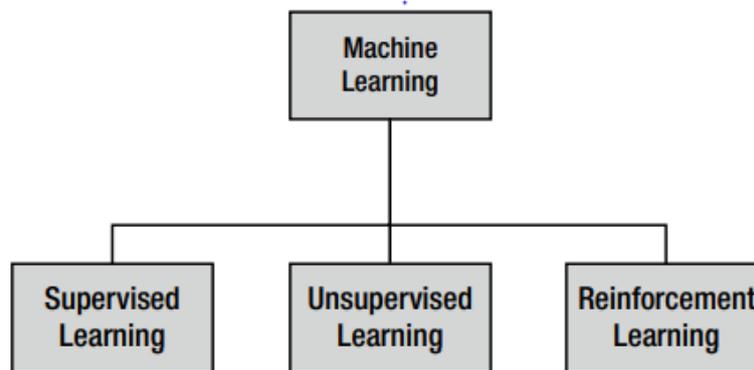


Figure 3. 1 : schéma représentant les types de machine Learning

### 3.3.3 Quelques notions de machine learning :

#### Underfitting et overfitting :

- ❖ **Underfitting** : lorsque le processus d'entraînement est arrêté trop tôt ou que le modèle est insuffisamment exposé aux données d'entraînement. Dans un tel scénario, le modèle n'est pas capable d'apprendre les modèles généraux présents dans les données d'apprentissage. Une faible précision de prédiction à la fois sur les données d'entraînement et les données non vues est un indicateur d'un underfitting (ou sous-apprentissage).
- ❖ **Overfitting** : on peut choisir d'entraîner un modèle en utilisant toutes les données disponibles dans un ensemble de données d'entraînement pour un plus grand nombre d'itérations afin qu'il apprenne toutes les tendances présentes dans les données d'entraînement. En conséquence, le modèle peut faire preuve d'une très grande précision sur les données d'apprentissage. Cependant, le modèle est tellement adapté aux données d'apprentissage qu'il perd sa capacité à généraliser au-delà des données d'apprentissage

qu'il a vue. Dans de tels scénarios, le modèle présente une faible précision de prédiction pour les données non vues, même s'il a pu générer des prédictions pour les données d'apprentissage avec une grande précision. Ce type de scénario indique un overfitting.

### Fonctions d'activation :

Les fonctions d'activation sont des fonctions mathématiques qui déterminent la sortie d'un réseau neuronal. La fonction est attachée à chaque neurone du réseau et détermine s'il doit être activé ou non. Les fonctions d'activation déterminent la précision du modèle d'apprentissage profond et ont un effet majeur sur la capacité du réseau neuronal à converger et sur la vitesse de convergence.

#### ❖ Fonction linéaire :

Il s'agit d'une fonction linéaire où l'activation est proportionnelle à l'entrée (qui n'est autre que la somme pondérée des neurones). Elle produit une valeur réelle entre les plages  $(-\infty, +\infty)$ . En général, elle est utilisée comme nœud de sortie.

$$z = w * x + b$$

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

#### ❖ La fonction Sigmoidé :

La fonction d'activation Sigmoidé, ou fonction logistique, prend une ou plusieurs valeurs réelles en entrée et produit un nombre compris entre 0 et 1. Elle convient à la classification binaire. Mais cette fonction pose un problème pour certains cas d'utilisation car la dérivée de la fonction d'activation  $\sigma(z)$  est nulle lorsque les valeurs ( $z$ ) sont des nombres positifs/négatifs très importants. Cela conduit à un problème de vanishing gradients.

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (3.10)$$

#### ❖ La fonction Tanh :

Tanh (tangente hyperbolique) est une fonction d'activation non linéaire qui prend un ou plusieurs entrés et retourne des valeurs entre  $[-1, 1]$ . La dérivée de cette fonction est plus importante que la fonction sigmoïde, mais elle a toujours le problème de Vanishing gradient.

$$g(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \quad (3.11)$$

Avec 
$$z = w_0 + w_1x_1 + w_2x_2 + \dots + w_nx_n \quad (3.12)$$

Le vecteur  $x$  est l'entrée de la couche

#### ❖ La fonction ReLU:

La ReLU est la fonction d'activation la plus utilisée, elle est utilisée dans presque tous les réseaux de neurones d'apprentissage profond. La ReLU est à moitié rectifiée (à partir du bas),  $f(z)$  est nulle lorsque  $z$  est inférieur à zéro et  $f(z)$  est égale à  $z$  lorsque  $z$  est supérieur ou égal à zéro. Mais le problème est que toutes les valeurs négatives deviennent immédiatement nulles, ce qui réduit la capacité du modèle à s'adapter ou à s'entraîner correctement à partir des données.

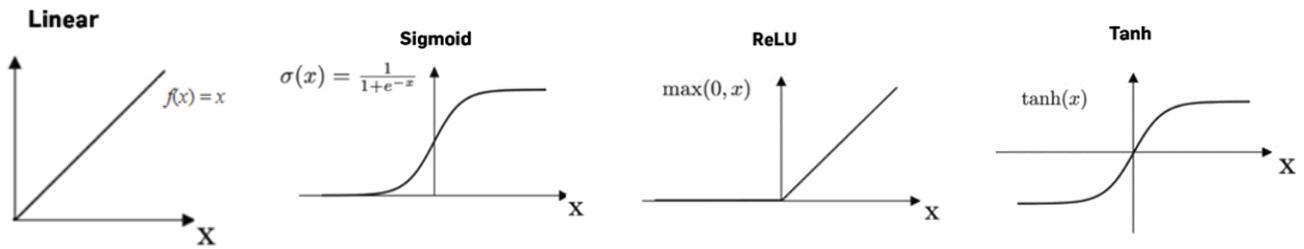


Figure 3. 2 : Représentation des fonctions d'activation

### 3.4 Réseaux de neurones :

Les Réseaux de Neurones sont des modèles bien plus complexes que tous les autres modèles de Machine Learning dans le sens où ils représentent des fonctions mathématiques avec beaucoup de paramètres à apprendre.

#### 3.4.1 Réseau de Neurone à un seul Neurone : Le perceptron

Le réseau de Neurones le plus simple qui existe porte le nom de perceptron. Les entrées du neurone sont les features  $x$  multipliées par des paramètres  $w$  à apprendre. Le calcul effectué par le neurone (appelé également nœud) peut être divisé en deux étapes :

1. Le neurone calcule la **somme**  $z$  de toutes les entrées (features)  $x$  multipliés par leurs poids  $w$  respectif, et ajoute une valeur  $b$  appelée biais . C'est un calcul linéaire. Comme dans l'exemple montré dans la figure ci-dessous, la somme est calculée comme suit :

$$z = w_1 \times x_1 + w_2 \times x_2 + w_3 \times x_3 + b \quad (3.13)$$

2. Le neurone passe la somme dans sa fonction d'activation. C'est un calcul non-linéaire :

$$y = \varphi(z) \quad (3.14)$$

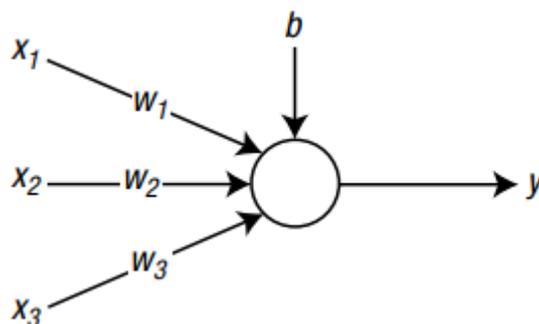


Figure 3. 3 : Exemple d'un perceptron qui reçoit 3 entrées

### 3.4.2 Multilayer perceptron MLP (réseau à plusieurs neurones) :

Pour créer un Réseaux de Neurones, il suffit de développer plusieurs perceptrons et de les connecter les uns aux autres d'une certaine façon.

On réunit les neurones en colonne (on dit qu'on les réunit en couche, en layer). Au sein de leur colonne, les neurones ne sont pas connectés entre eux.

On connecte toutes les sorties des neurones d'une colonne à gauche aux entrées de tous les neurones de la colonne de droite qui suit. On peut ainsi construire un réseau avec autant de couches et de neurones que l'on veut. Plus il y a de couches, plus on dit que le réseau est profond (deep) et plus le modèle devient riche, mais aussi difficile à entraîner. C'est ça, le Deep Learning.

Le perceptron multicouche (MLP) est le type d'algorithme de deep learning le plus basique dont les éléments de base sont les neurones. Les neurones calculent une somme pondérée des valeurs d'entrée en utilisant les poids associés à l'entrée, puis appliquent une fonction d'activation non linéaire ( $f$ ) à la somme pondérée. Les transformations non linéaires effectuées par les neurones permettent aux MLP et autres réseaux neuronaux profonds de modéliser des processus hautement non linéaires, qui se produisent dans un grand nombre de phénomènes naturels et de problèmes industriels.

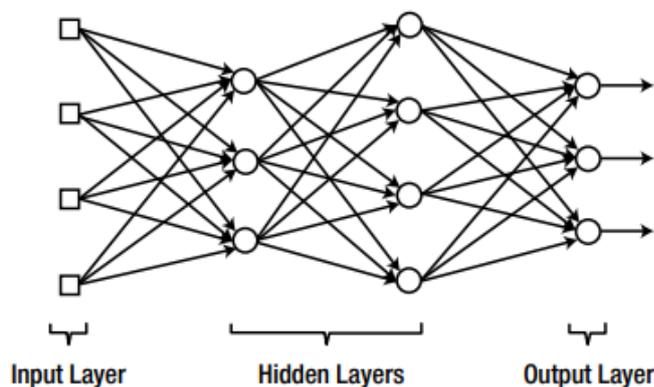


Figure 3. 4 : représentation d'un réseau de neurones multicouches

Ainsi, un réseau de neurones comme c'est schématisé dans la figure (3.4) constitué comme suit :

- **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.
- **Couches cachées (hidden layers):** ces couches contiennent des neurones ou des nœuds, avec une fonction d'activation non linéaire. La couche cachée est représentée par une matrice de poids ( $w$ ), et un vecteur de biais ( $b$ ). Les dimensions de la matrice  $w$  et du vecteur  $b$  dépendent du nombre de nœuds dans les couches. Par exemple la couche d'entrée, qui se connecte à  $m$  entrées et diffuse des valeurs à  $n$  nœuds dans la première couche, possède une matrice de poids de dimension  $m \times n$ . De même, le vecteur de biais  $b$  comporte  $n$  éléments. La sortie de cette couche est  $f(wX + b)$ , où  $f$  est la fonction d'activation.
- **Couche de sortie (output layer) :** La couche de sortie se connecte à la dernière couche cachée, et garantit que le réseau neuronal fournit une sortie cohérente avec la formulation du problème.

### 3.4.3 L'entraînement d'un réseau de neurones :

Afin que la machine puisse construire un modèle qui soit le plus proche de la réalité, il faut que l'erreur entre la valeur de la sortie calculée par le modèle et la valeur réelle soit raisonnable. Ainsi, pour avoir un bon modèle, les étapes suivantes sont suivies :

- Étape 1 : initialiser le network : les valeurs des poids (weights) et des biais sont initialisés.
- Étape 2 : Feed Forward : L'information est transmise à travers le réseau de l'entrée à la couche cachée jusqu' à la couche de sortie via les fonctions d'activation des nœuds et les poids.
- Étape 3 : évaluation de l'erreur : La sortie du réseau est évaluée par rapport à la sortie connue. Si l'erreur est inférieure à un seuil prédéfini, le réseau est formé et l'algorithme se termine.
- Étape 4 : back propagation, l'erreur de la couche de sortie est utilisée pour remodifier les poids. L'algorithme propage l'erreur en arrière dans le réseau et calcule le gradient de la modification de l'erreur par rapport aux changements dans les valeurs de poids.
- Étape 5 : Ajustement, les poids sont ajustés en utilisant les gradients de changement dans le but de réduire l'erreur. Les poids et les biais de chaque neurone sont ajustés par un facteur basé sur la dérivée de l'erreur.

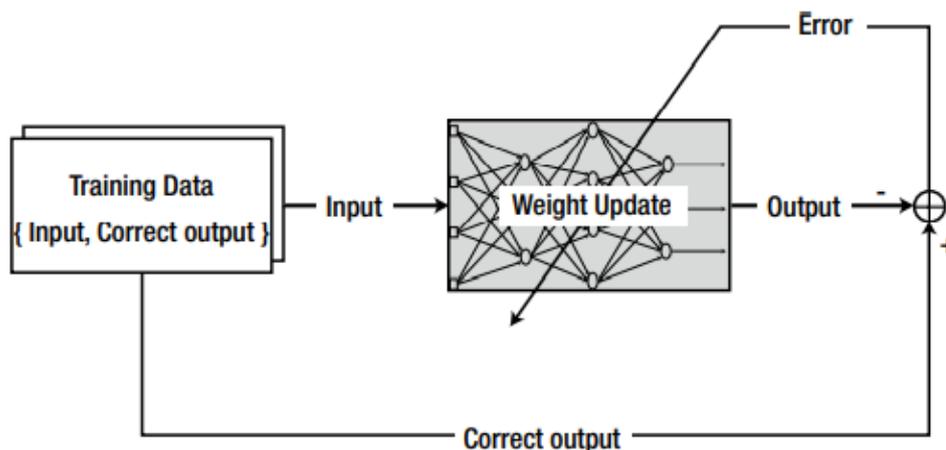


Figure 3. 5 : schéma du principe d'entraînement d'un réseau de neurones

### 3.5 Réseau de neurones récurrent (RNN)

Un réseau de neurones récurrent (RNN) est un type de réseau de neurones artificiel qui utilise des données séquentielles ou des données de séries temporelles. Ces algorithmes d'apprentissage profond sont couramment utilisés pour des problèmes ordinaux ou temporels, tels que la traduction de langues, la reconnaissance vocale

Les réseaux de neurones récurrents se distinguent par leur "mémoire", car ils utilisent les informations des entrées précédentes pour influencer l'entrée et la sortie actuelles. Alors que les réseaux neuronaux profonds traditionnels supposent que les entrées et les sorties sont indépendantes les unes des autres, la sortie des réseaux neuronaux récurrents dépend des éléments antérieurs de la séquence.

## Types des RNNs

Comme les RNN ne sont pas limités au traitement d'entrées de taille fixe, ils étendent réellement les possibilités de calcul des réseaux de neurones. La figure suivante illustre visuellement certaines combinaisons de séquences que nous pouvons réaliser.

- ❖ One-to-one (un à un): Il s'agit du traitement non séquentiel, tel que les feedforward neural network.
- ❖ One-to-many (un à plusieurs) : Cette méthode génère une séquence basée sur une seule entrée, par exemple la génération de légendes à partir d'une image.
- ❖ Many-to-one (Plusieurs à un) : ceci produit un résultat unique basé sur une séquence, par exemple, classification des sentiments à partir d'un texte.
- ❖ Many-to-many (Plusieurs à plusieurs) : cette architecture prend plusieurs entrées et donne plusieurs sorties, par exemple, l'étiquetage des phonèmes dans la reconnaissance vocale.

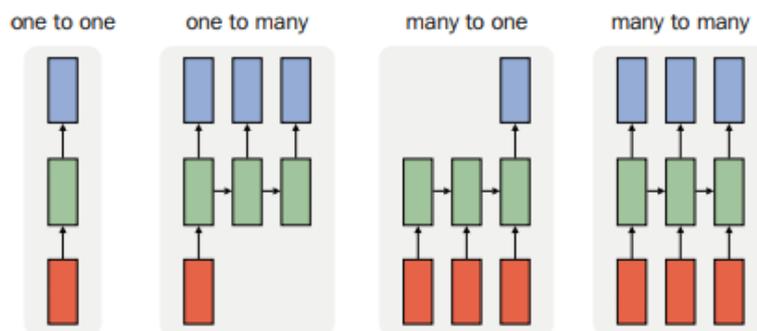


Figure 3. 6 : types des RNNs

Rouge : entrées (inputs), vert : les états (states), bleu : les sorties (outputs)

## Modèles Long Short Term Memory (LSTM)

Long Short Term Memory (LSTM) est une architecture spéciale de réseaux de neurones récurrents. Les RNN ont un mauvais caractère qui est le problème de la distribution de gradient (d'oublier le passé), et c'était la majeure cause de l'invention du LSTM.

La cellule LSTM est une adaptation de la couche récurrente qui permet aux signaux plus anciens des couches profondes de se déplacer vers la cellule du présent. Les trois cellules de LSTM sont présentées dans la figure ci-dessous :

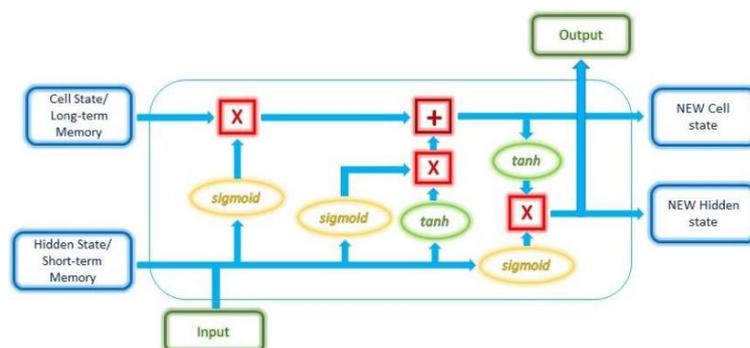


Figure 3. 7 : Représentation d'une cellule LSTM

❖ La porte d'oubli (forget gate) :

La porte d'oubli décide quelles informations de la mémoire à long terme doivent être conservées ou supprimées. Pour obtenir le vecteur d'oubli, la mémoire à court terme et l'entrée actuelle sont passés par une fonction sigmoïde.

$$f = \sigma(w_{forget}(h_{t-1}, x_t) + b_{forget}) \quad (3.15)$$

Les sorties de la porte d'entrée et de la porte d'oubli seront utilisées pour donner une nouvelle version de la mémoire à long terme, qui sera transmise à la cellule suivante. Cette nouvelle mémoire à long terme sera également utilisée dans la porte de sortie.

$$C_t = C_{t-1} * f + I_{input} \quad (3.16)$$

Avec :  $C_t$  : état interne

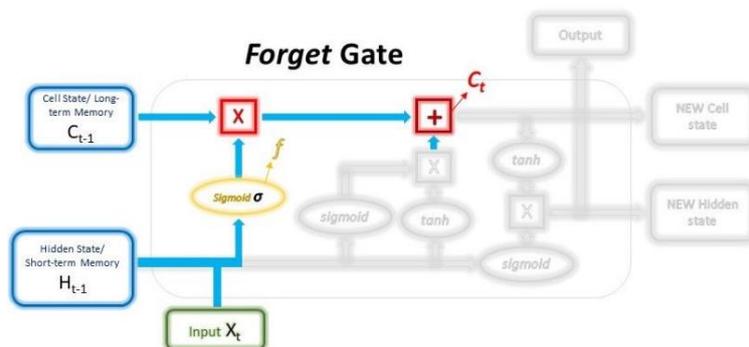


Figure 3. 8 : Représentation de la porte d'oubli (forget gate) d'une cellule LSTM

❖ La porte d'entrée (Input gate) :

Cette porte décide quelles nouvelles informations seront stockées dans la mémoire à long terme. Ceci peut être réalisé essentiellement suivant les étapes suivantes :

- Régulation des valeurs qui doivent être ajoutées à l'état de la cellule en faisant intervenir une fonction sigmoïde. Celle-ci est fondamentalement très similaire à la porte d'oubli et agit comme un filtre pour toutes les informations provenant de  $h_{t-1}$  et  $x_t$ .

$$I_1 = \sigma(w_{i1}(h_{t-1}, x_t) + b_i) \quad (3.17)$$

- Création d'un vecteur contenant toutes les valeurs possibles qui peuvent être ajoutées (à l'état de la cellule. Pour ce faire, la fonction tanh est utilisée.

$$I_2 = \tanh(w_{i2}(h_{t-1}, x_t) + b_i) \quad (3.18)$$

- Les sorties de ces 2 couches sont ensuite multipliées, et le résultat final représente les informations à conserver dans la mémoire à long terme et à utiliser comme sortie.

$$I_{input} = I_1 * I_2 \quad (3.19)$$

Avec :  $h_{t-1}$ ,  $x_t$ ,  $b_i$ ,  $w_i$  représentent respectivement: la sortie à l'instant t-1, entrée courant à l'instant t, un terme biais et le poids

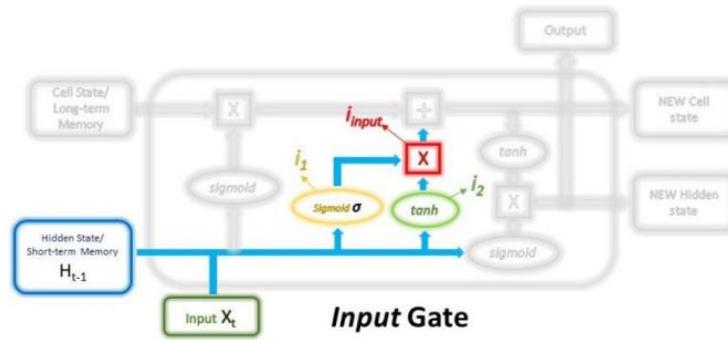


Figure 3. 9 : Représentation de la porte d'entrée (input gate) d'une cellule LSTM

### La porte de sortie (output gate) :

Cette porte prend l'entrée actuelle, la mémoire à court terme précédente et la mémoire à long terme nouvellement calculée pour produire la nouvelle mémoire à court terme qui sera transmise à la cellule au prochain pas de temps.

Tout d'abord, la mémoire à court terme précédente et l'entrée actuelle seront passées dans une fonction sigmoïde pour créer le dernier filtre. Ensuite, nous mettons la nouvelle mémoire à long terme via une fonction d'activation tanh. La sortie de ces 2 processus sera multipliée pour produire la nouvelle mémoire à court terme.

$$O_1 = \sigma(w_{output1} - (h_{t-1}, x_t) + b_{output1}) \quad (3.20)$$

$$O_2 = \tanh(w_{output2} - C_t + b_{output2}) \quad (3.21)$$

$$O_t = O_1 * O_2 \quad (3.22)$$

Avec :  $O_t$  : la sortie

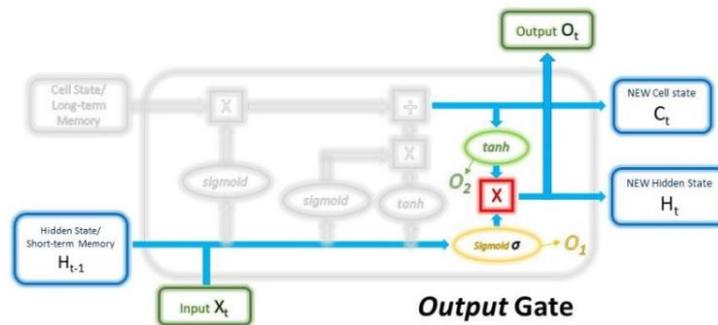


Figure 3. 10 : Représentation de la porte de sortie (output gate) d'une cellule LSTM

### 3.6 Support Vector Machine Regression SVR :

Support Vector Machine (SVM) sont bien connues dans les problèmes de classification. L'objectif de l'algorithme SVM est de créer la meilleure ligne ou frontière de décision qui peut séparer l'espace à n dimensions en classes, afin que nous puissions facilement placer le nouveau point de données dans la bonne catégorie à l'avenir. Cette meilleure frontière de décision est appelée un hyperplan. Le SVM choisit les points extrêmes qui aident à créer l'hyperplan. Ces cas extrêmes sont appelés vecteurs de soutien (support vectors), et l'algorithme est donc appelé support vector machine.

Support Vector Regression utilise le même principe que les SVM pour prédire une variable continue. L'idée de base derrière SVR est de trouver la meilleure ligne d'ajustement. Dans SVR, la meilleure ligne d'ajustement est l'hyperplan qui a le nombre maximum de points. Contrairement à d'autres modèles de régression qui tentent de minimiser l'erreur entre la valeur réelle et la valeur prédite, le SVR essaie d'ajuster la meilleure ligne dans une valeur seuil. La valeur seuil est la distance entre l'hyperplan et la ligne frontière (support vector)

Pour chaque point de données  $y$ , il est acceptable d'obtenir une valeur prédite dans l'intervalle entre  $(y + \epsilon)$  et  $(y - \epsilon)$ . Dans la figure (3-11), la marge d'erreur  $\epsilon$  est représentée par deux vecteurs de soutien.

Certains paramètres clés utilisés dans SVR sont mentionnés ci-dessous :

- ❖ **Hyperplan** : les hyperplans sont des limites de décision utilisées pour prédire la sortie continue. Les points de données de chaque côté de l'hyperplan qui sont les plus proches sont appelés support vector. Ceux-ci sont utilisés pour tracer la ligne requise qui montre la sortie prévue de l'algorithme.
- ❖ **Noyau (kernel)** est un ensemble de fonctions mathématiques qui prennent les données en entrée et les transforment sous la forme requise. Les noyaux les plus largement utilisés incluent linéaire, non linéaire, polynomial, fonction de base radiale (RBF) et sigmoïde.
- ❖ **Lignes de délimitation** : ceux sont les deux lignes qui sont tracées autour de l'hyperplan à une distance d'épsilon. Elles sont utilisées pour créer une marge entre les points de données.

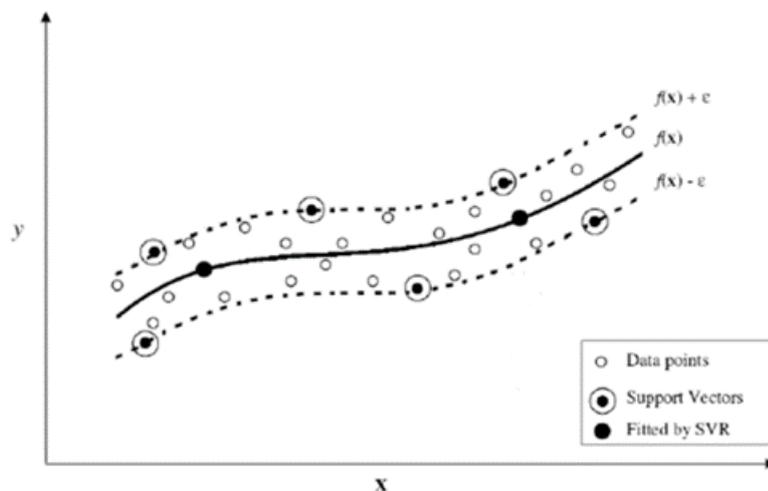


Figure 3. 11 : support vector machine regression

### 3.7 Gaussian process regression GPR (Processus gaussiens)

Processus gaussiens (GP) sont des méthodes d'apprentissage supervisés génériques conçue pour résoudre des problèmes de régression et de classification probabiliste. Le processus gaussien est basé sur le calcul d'une distribution de probabilité et peut être appliqué de manière flexible grâce à l'approche bayésienne.

La régression par processus gaussien utilise l'astuce du noyau. Les noyaux projettent les données d'entrée du problème sur un espace dimensionnel, de telle sorte qu'un problème de régression hautement non linéaire devient linéaire, ou approximativement linéaire. Cela permet une solution plus efficace à un problème de régression non linéaire.

Le processus gaussien (GP) est une collection de variables aléatoires, où toute collection finie de ces variables a une distribution normale multivariée.

Disons qu'il existe des points  $\mathbf{x}$ , où les valeurs observées de la propriété observées sont données par  $f(\mathbf{x})$  ou simplement  $f$ . De plus, sur la base de cette information, pour un ensemble de points  $\mathbf{x}^*$  nous essayons d'estimer  $f(\mathbf{x}^*)$  ou  $f^*$ . Cela signifie que nous essayons de calculer la distribution de probabilité ( $P(f^*|f)$ ).

Sous l'hypothèse d'une moyenne nulle, nous pouvons approximer  $f(\mathbf{x})$  de la manière suivante :

$$f(\mathbf{x}) \sim N(\mathbf{0}, K(\theta, \mathbf{x}, \mathbf{x}')) \quad (3.23)$$

Où  $N$  désigne une distribution normale avec une moyenne nulle, et  $K(\theta, \mathbf{x}, \mathbf{x}')$  représente la matrice de covariance entre toutes les paires de points de données observées  $(\mathbf{x}, \mathbf{x}')$ .

La matrice de covariance garantit que les valeurs proches de l'entrée génèrent des valeurs de sortie, qui sont également proches les unes des autres. Les valeurs moyennes estimées aux points  $\mathbf{x}^*$  sont données par :

$$\mathbf{y} = \mathbf{k}(K(\theta, \mathbf{x}^*, \mathbf{x})) K(\theta, \mathbf{x}, \mathbf{x}')^{-1} \mathbf{f}(\mathbf{x}) \quad (3.24)$$

Dans l'équation précédente un nouveau terme  $K(\theta, \mathbf{x}^*, \mathbf{x})$ . Ce terme tient compte covariance entre les points de données observés et les points où nous essayons d'obtenir des estimations de propriétés.

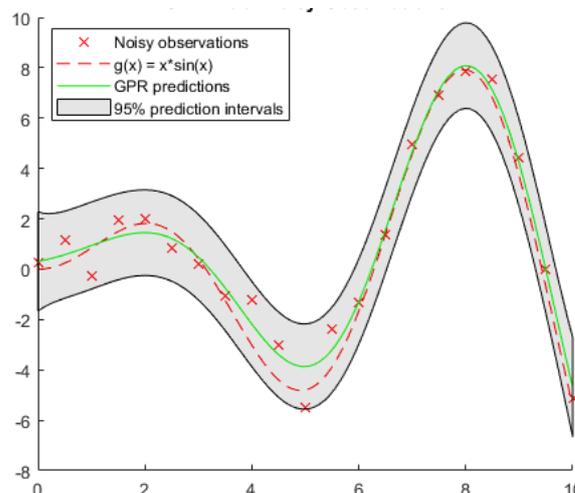


Figure 3.12 : ajustement GPR d'une observation bruyante [35]

### 3.8 Programmation génétique et régression symbolique :

La Programmation Génétique PG est une méthode d'apprentissage automatique inspiré de l'évolution biologique, elle fait évoluer des programmes informatiques pour qu'ils accomplissent une tâche spécifique.

Cela se fait en générant aléatoirement une population de programmes informatiques (représentés par des structures arborescentes), puis en mutant et en croisant les arbres les plus performants pour créer une nouvelle population. Ce processus est itéré jusqu'à ce que la population contienne des programmes qui (avec un peu de chance) résolvent bien la tâche.

Lorsque la tâche consiste à construire un modèle mathématique avec des données acquises à partir d'un processus ou d'un système, la PG est souvent connue sous le nom de **régression symbolique**. Contrairement à l'analyse de régression traditionnelle (dans laquelle l'utilisateur doit spécifier la structure du modèle), la PG fait évoluer automatiquement à la fois la structure et les paramètres du modèle mathématique.

Le programme GP est constitué de cinq éléments essentiels :

- ❖ **Population** : la population est un ensemble d'individus et chaque individu est une solution au problème qu'on veut résoudre. Ils sont caractérisés par un ensemble de paramètres appelés gènes. Ces derniers sont réunis en une chaîne pour former une solution.
- ❖ **Fonction de fitness** : la fonction de fitness évalue l'individu et lui affecte un score. La probabilité qu'un individu soit sélectionné pour la reproduction est basée sur son score de fitness.
- ❖ **Sélection** : la sélection est un processus par lequel les individus qui ont eu un score élevé sont sélectionnés parmi la génération actuelle et serviront de parents à la génération suivantes.
- ❖ **Crossover** : un point de croisement est choisi au hasard dans les gènes pour chaque paire de parents. Ensuite, les gènes entourés par les points de croisement sont échangés entre les parents pour former deux nouveaux individus.
- ❖ **Mutation** : certains gènes de nouveaux individus peuvent être soumis à une mutation avec une faible probabilité arbitraire. Cela implique que certains des bits de la chaîne peuvent être inversés.

#### **Régression symbolique multigénique (Multigene Symbolic Regression) :**

Un individu multigénique est constitué d'un ou plusieurs gènes. Les gènes sont acquis de manière incrémentielle par les individus afin de réduire la somme des erreurs quadratiques d'un modèle sur un ensemble de données.

Le modèle global est une combinaison linéaire pondérée de chaque gène. Les poids optimaux des gènes sont obtenus automatiquement en utilisant les moindres carrés pour régresser les gènes en fonction des données de sortie.

### Fonctionnement de la régression symbolique multigénique :

Dans la régression symbolique multigénique, chaque prédiction de la variable de sortie  $y$  est formée par la sortie pondérée de chacun des arbres/gènes de l'individu multigénique, plus un terme de biais. Chaque arbre est une fonction de zéro ou plus des  $N$  variables d'entrée  $x_1, \dots, x_N$

Mathématiquement, un modèle de régression multigénique peut être écrit comme suit :

$$\hat{y} = d_0 + d_1 \times \text{arbre1} + \dots + d_M \times \text{arbreM} \quad (3.25)$$

Où  $d_0$  est un terme de biais.  $d_1, \dots, d_M$  sont les poids des gènes et  $M$  est le nombre de gènes (arbres) qui composent l'individu actuel.

Le nombre et la structure des arbres évoluent au cours d'une exécution à l'aide de données d'apprentissage, c'est-à-dire un ensemble de valeurs d'entrée existantes et de valeurs de sortie correspondantes. Des données de test peuvent être utilisées, après l'exécution, pour évaluer les modèles et donner une indication de la manière dont les modèles se généralisent aux nouvelles données.

La figure ci-dessous montre un exemple de modèle multigénique tel que :

$$y = d_0 + d_1(0.41 x_1 + \tanh(x_2 x_3)) + d_2(0.45 x_3 + \text{sqrt}(x_2)) \quad (3.26)$$

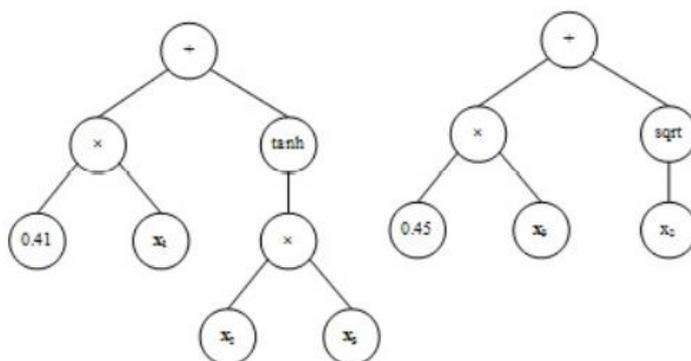


Figure 3.13 : modèle multigénique de la sortie  $y$  avec des entrées  $x_1, x_2$  et  $x_3$

### 3.9 Métriques d'évaluation d'un modèle

Pour comprendre comment le modèle se généralise sur un ensemble de données non vues, et pour évaluer ses performances, il est important d'utiliser divers moyens statistiques et métriques quantitatives sur lesquelles les performances du modèle peuvent être évaluées.

Les paramètres statistiques utilisés dans le cadre de ce mémoire de fin d'études sont : le coefficient de détermination ( $R^2$ ), la Racine de l'erreur quadratique moyenne (RMSE), l'erreur absolue moyenne (MAE) et l'erreur relative absolue (RAE).

❖ **Le coefficient de détermination  $R^2$ :**

Il représente la proportion de la variance de la variable dépendante expliquée à partir des variables indépendantes. Il permet ainsi d'évaluer la performance d'un modèle par rapport au niveau de variation présent dans les données. Il se calcule selon la formule suivante :

$$R^2 = 1 - \frac{\sum(y_i - \hat{y})^2}{\sum(y_i - \bar{y})^2} \quad (3.27)$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad (3.28)$$

Où :  $y_i$ :chaque point de données  
 $\hat{y}$ : la valeur prédite par le modèle  
 $\bar{y}$ : La valeur moyenne des données  
 $n$ : nombre de points de données).

Le meilleur score possible pour  $R^2$  est 1 ce qui montre que le modèle arrive à prédire exactement les valeurs réelles. Un modèle constant qui prédit toujours la même valeur de sortie de  $y$ , sans tenir compte des caractéristiques d'entrée, obtiendrait un score  $R^2$  de 0, comme il peut être négatif (car le modèle peut être arbitrairement plus mauvais).

❖ **La racine de l'erreur quadratique moyenne RMSE :**

C'est un autre critère permettant de tester la proximité entre la prédiction du modèle et les valeurs réelles. Plus cette valeur est faible, plus le modèle est bon. Cette erreur est calculée à partir de :

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y} - y_i)^2} \quad (3.29)$$

❖ **L'erreur absolue moyenne MAE :**

Elle se calcul en prenant la moyenne de la somme des valeurs absolues des erreurs selon la formule suivante :

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y} - y_i| \quad (3.30)$$

❖ **Erreur relative absolue RAE :**

L'erreur absolue relative (RAE) est un moyen de mesurer les performances d'un modèle prédictif. Elle est définie comme suit :

$$RAE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y})^2}{\sum_{i=1}^n y_i^2}} \quad (3.31)$$

## Chapitre 4 :

### Résultats et discussions

Avant d'élaborer de nouvelles approches d'études qui ont été rarement utilisées, il est strictement préconisé de valider la démarche de travail en utilisant des travaux antérieurement publiés.

Ce chapitre comprend donc essentiellement deux parties. A travers la première partie, on va valider la démarche et les méthodes adoptées dans la première section relative à l'étude du champ de Hassi Messaoud (la méthodologie est expliquée en détail dans la section 4.4.1), et ce en utilisant les données du champ de Reshadat issues la publication [30]. Dans un second temps, la même méthodologie avec des points supplémentaires seront appliqués au champ de Hassi Messaoud.

#### 4.1 Validation de la méthodologie en utilisant les données du champ de Reshadat

##### 4.1.1 Description du dataset du champ de Reshadat

Les données utilisées dans cette partie sont retirées à partir d'un article intitulé « Predicting liquid flow-rate performance through wellhead chokes with genetic and solver optimizers: an oil field case study » qui a été publié en 2018 [30]. Ces données ont été collectées à partir de Reshadat oil field. L'ensemble de données de 182 enregistrements couvre des débits de liquide de 100 à 30 000 STB/day. Chaque enregistrement comprend les mesures de quatre variables à savoir : le taux de production de liquide (Q), la pression à la tête de puits ( $P_u$ ), le diamètre de la duse(d), et le rapport gaz-liquide (GLR). 80 % de l'ensemble de données (146 points) ont été utilisées dans la phase de training (la création de modèle) et 20 % de l'ensemble de données (36points) ont été utilisées pour tester la précision des modèles.

Le tableau ci-dessous présente la distribution des valeurs des paramètres de la base de données du champ pétrolifère Reshadat

Tableau 4.1 : différente plage de paramètres des données de champ Reshadat

Paramètre	<b>d</b> (1/64 in)	<b>p<sub>u</sub></b> (psi)	<b>GLR</b> (scf/STB)	<b>Q</b> (STB/day)
Minimum	25.6	133	36	205
Maximum	64	881	885	34450
Moyenne	56.095	512.275	258.17	9445.14

### 4.1.2 Evaluation des modèles théoriques

L'équation de Gilbert (Equation 2.22) avec ses coefficients originaux ainsi que les corrélations ayant la même forme que celle de Gilbert (Achong, Baxendell, Pilehvari et Ros) ont été testées dans un premier temps, pour étudier leur applicabilité sur le champ pétrolifère de Reshadat.

Les résultats obtenus en calculant les différentes métriques d'évaluation des modèles sont présentés dans le tableau ci-dessous :

Tableau 4.2 : statistiques de test des modèles de régime critique

Modèle	R <sup>2</sup>	RMSE	MAE	RAE
Gilbert	0.71	3992.7	2772	0.33
Achong	0.983	972.30	853.1	0.082
Baxendell	0.907	2281.1	1501	0.19
Pilehvari	0.795	3309.4	2780	0.28
Ros	0.82	3159.3	2033	0.27

D'après les résultats obtenus dans le tableau (4.2), les modèles de Gilbert, Achong, Ros, Baxendell et Pilhevari sont capables de donner une bonne estimation du débit d'huile produit. Celui de Achong est le meilleur d'entre eux où il présente un coefficient de détermination de l'ordre 0.98, suivi du modèle de Baxendell ensuite Ros, Pelhvari et enfin Gilbert.

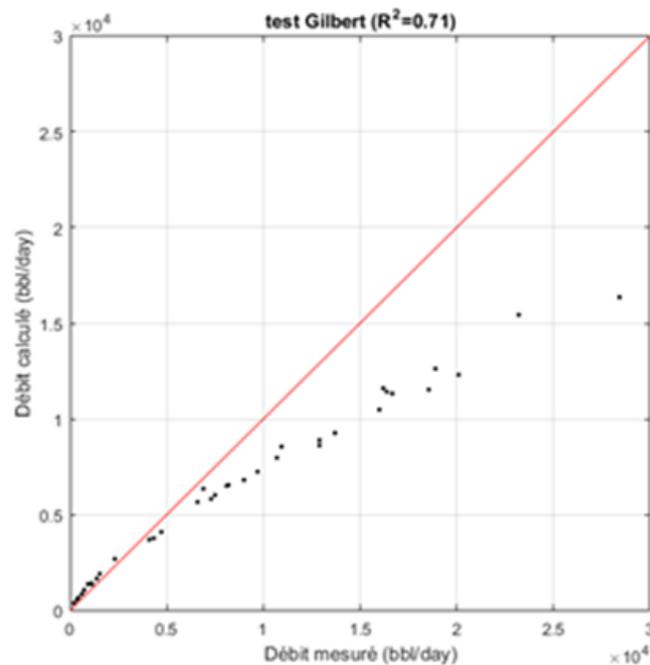


Figure 4.1 : cross plot de test de la corrélation de Gilbert  
Au champ de Reshadat

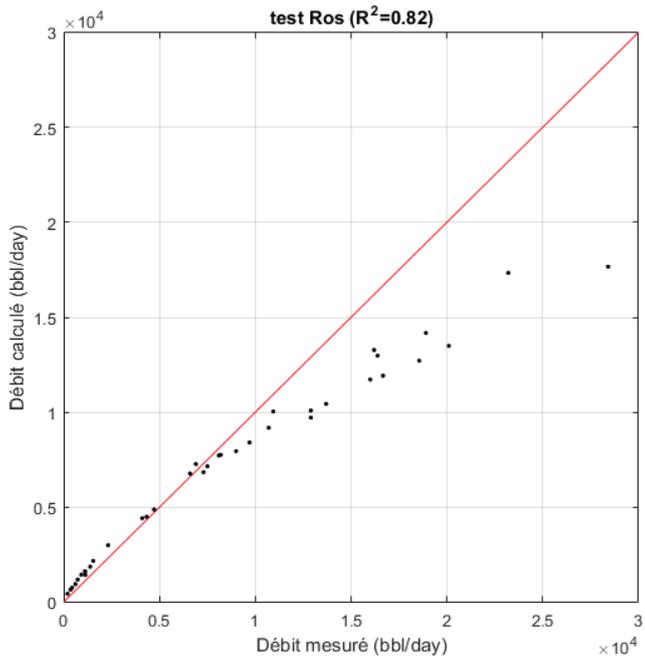


Figure 4.2 : cross plot de test du modèle de Ros au champ de Reshadat

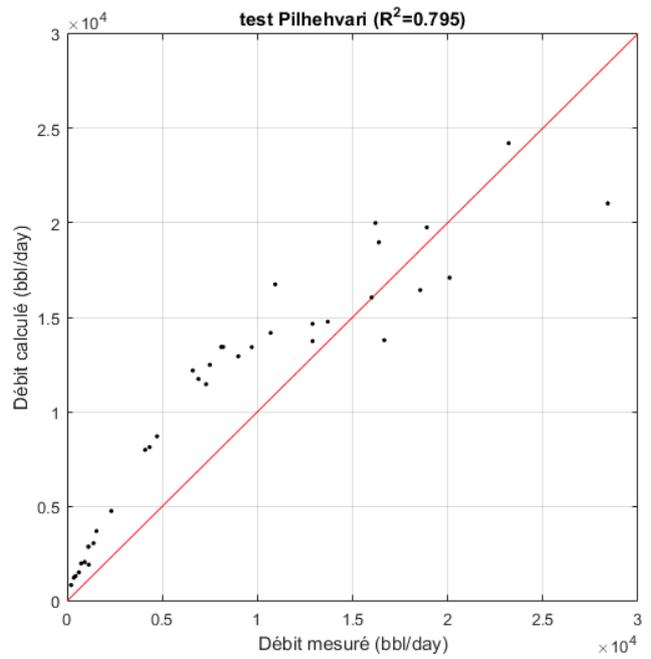


Figure 4.3 : cross plot de test du modèle de Pilehvari au champ de Reshadat

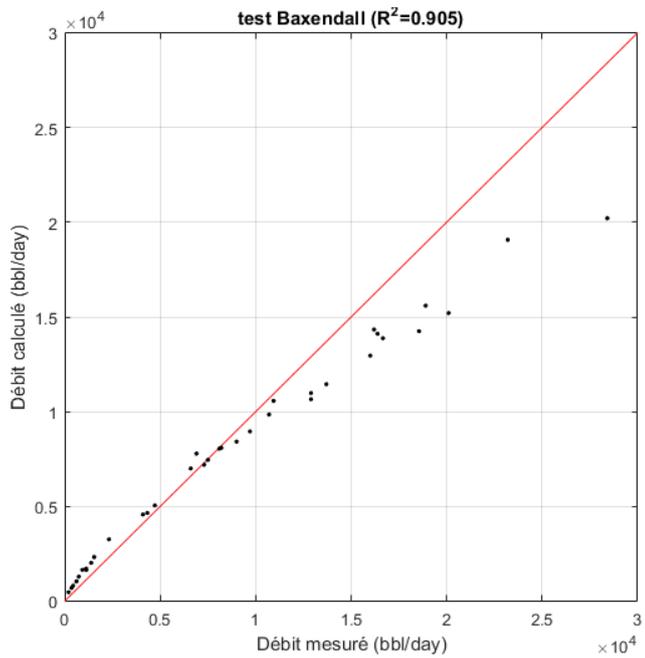


Figure 4.4 : cross plot de test du modèle de Baxendell au champ de Reshasat

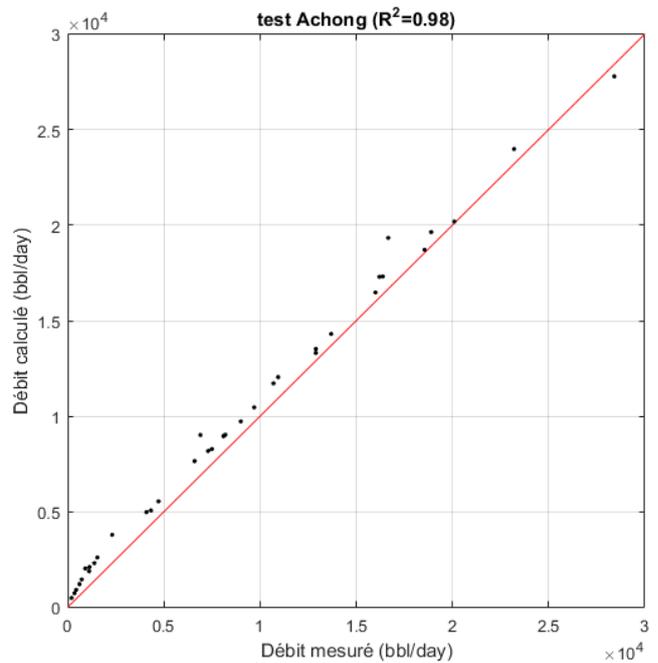


Figure 4.5 : cross plot de test du modèle de Achong au champ de Resadat

### 4.1.3 Adaptation du modèle de Gilbert et création de modèles intelligents

Dans cette partie ; et afin d'améliorer encore plus la capacité de prédiction du débit, la corrélation de Gilbert est adaptée suivant les données de champ Reshadat, l'algorithme de Levenberg-Marquardt est utilisé pour cette fin. Les modèles intelligents suivants sont également élaborés : Support vector regression SVR, Gaussian process regression GPR et Artificial neural network ANN. Ces derniers ont été créés en introduisant comme entrées les mêmes paramètres de la formule de Gilbert.

Les tableaux (4.3) et (4.4), représentent les nouveaux coefficients et les différentes métriques d'évaluation de la corrélation de Gilbert adaptée pour le champ de Reshadat ainsi que les modèles intelligents.

Tableau 4.3 : les nouveaux coefficients de l'équation de Gilbert adaptée au champ de Reshadat

a	b	c	e
0.0128	2.6849	0.7159	0.9943

Tableau 4.4 : statistiques de training et de test du modèle théorique de Gilbert adapté au champ Reshadat et des modèles intelligents

Modèle	Métriques training				Métriques test			
	R <sup>2</sup>	RMSE	MAE	RAE	R <sup>2</sup>	RMSE	MAE	RAE
Gilbert new	0.999	99.1	55.11	0.008	0.999	81.82	54.9	0.006
SVR	0.96	1323.2	667.7	0.11	0.82	3164	2049	0.265
GPR	0.999	756.96	306.1	0.002	0.997	366.96	210.2	0.02
ANN	0.999	194.74	100.9	0.02	0.999	172.07	133.1	0.01

- La corrélation de Gilbert adaptée suivant les données du champs de Reshadat, notée Gilbert new montre d'excellentes performances comparées aux modèles théoriques précédemment testés , le coefficient de détermination est de l'ordre 0.999 lors du training et du test, l'erreur relative absolue est de l'ordre de 0.006 sur l'ensemble des données de test , ce qui montre que ce modèle adapté est excellent pour la prédiction de débit pour ce champ.
- Presque tous les modèles intelligents ont montré d'excellentes performances lors du training et du test, plus particulièrement le modèle ANN et GPR ont enregistrés des coefficients de détermination de l'ordre de 0.99 sur l'ensemble des données de test.
- Les cross plot de training et de test de ces modèles sont illustrés dans les figures ci-dessous ou on remarque que les points sont presque tous alignés sur la ligne parfaite de 45°.

### 4.1.4 Conclusion I

Les modèles développés pour le champ de Reshadat ont montré d'excellentes performances, ce qui montre que la méthodologie appliquée est largement satisfaisante et nous permet de prédire les débits des puits d'une manière acceptable et juste pour un objectif de minitoring des puits.

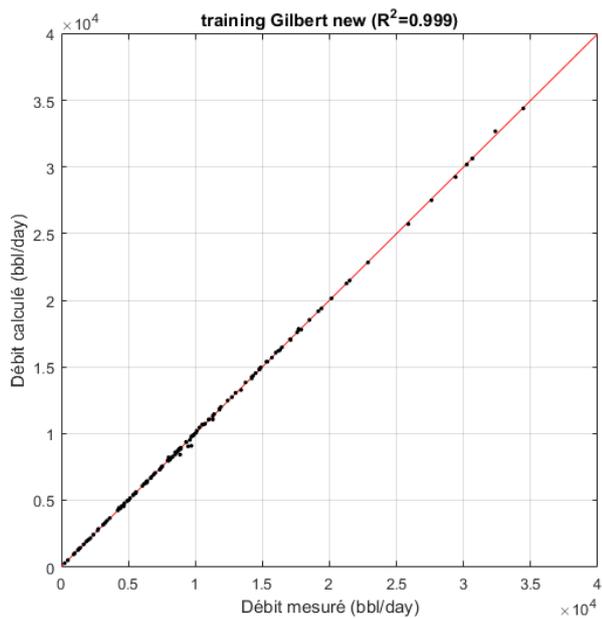


Figure 4.6 : cross plot de training du modèle de Gilbert adapté au champ de Reshadat

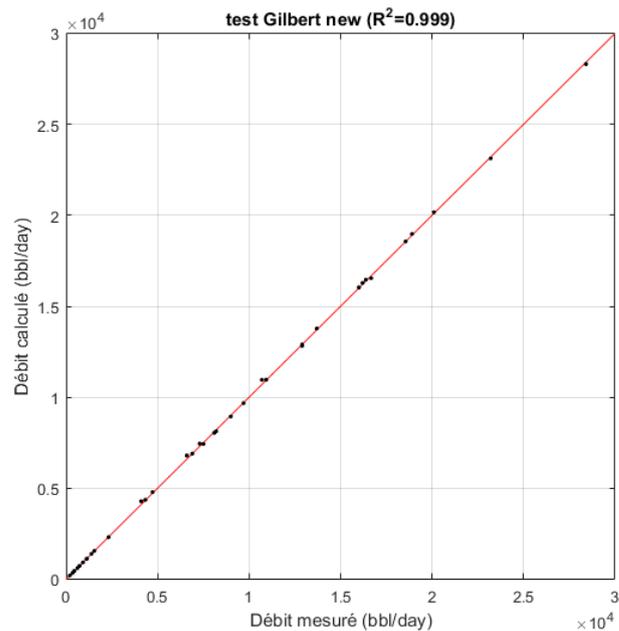


Figure 4.7 : cross plot de test du modèle de Gilbert adapté au champ de Reshadat

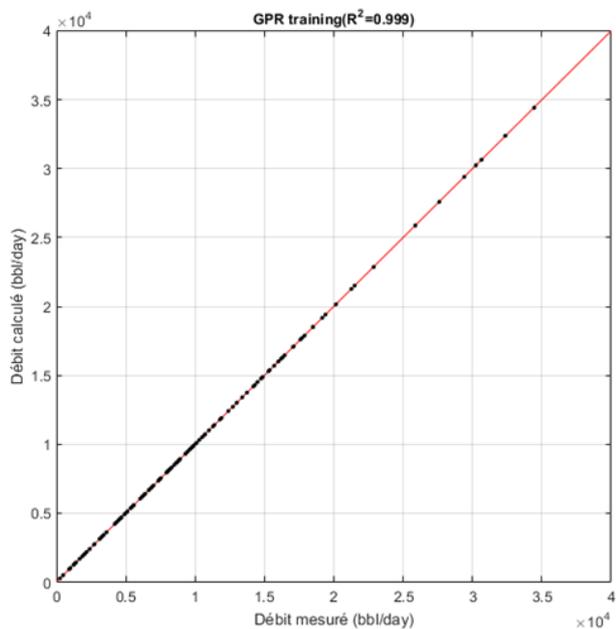


Figure 4.8 : cross plot du training du modèle GPR au champ de Reshadat

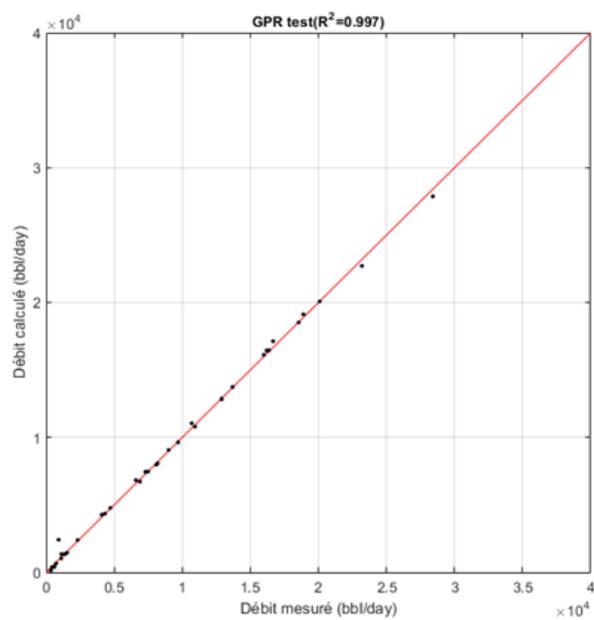


Figure 4.9 : cross plot du test du modèle GPR au champ de Reshadat

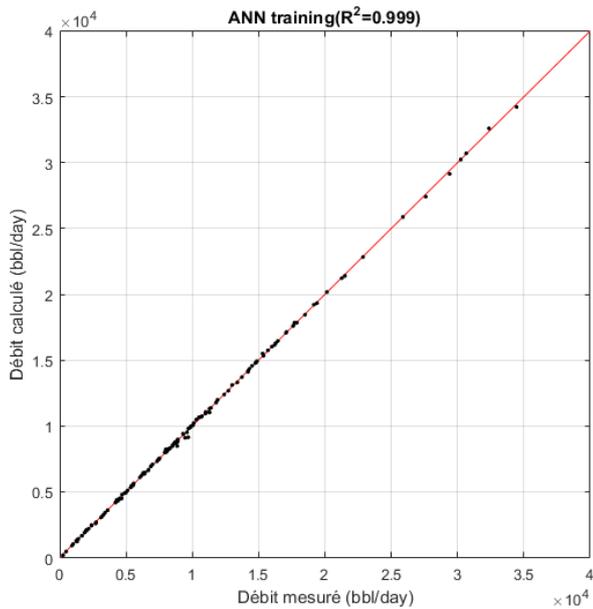


Figure 4. 10 ; cross plot de training de modèle ANN au champ de Reshadat

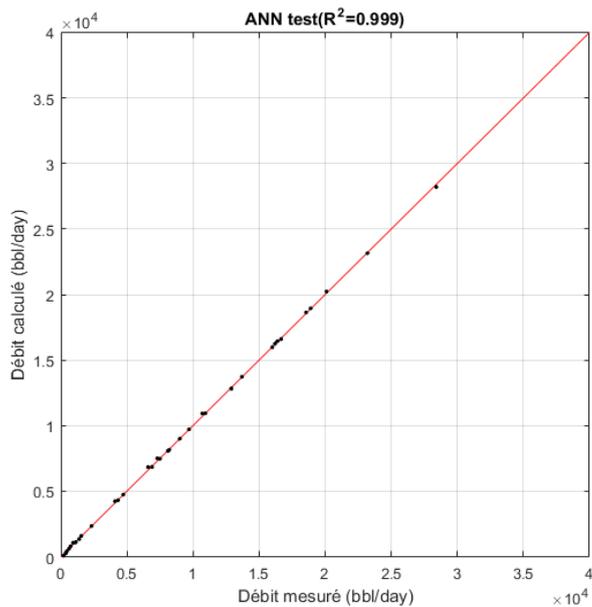


Figure 4. 11 : cross plot de test d modèle ANN au champ de Reshadat

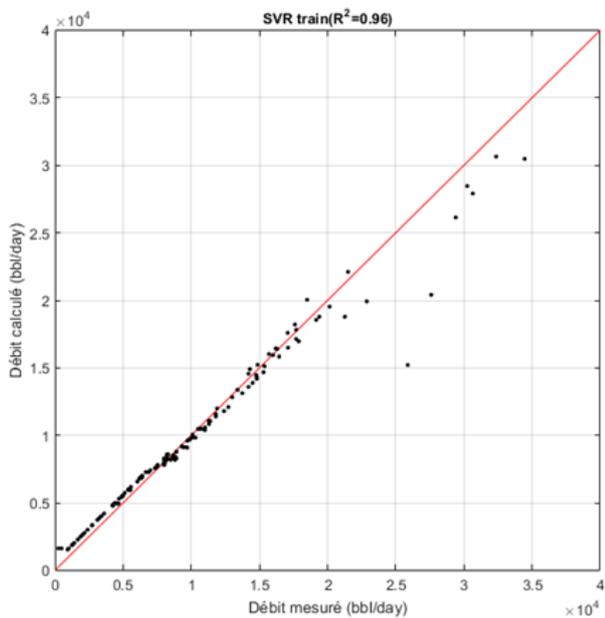


Figure 4. 12 : cross plot de training de modèle SVR au champ de Reshadat

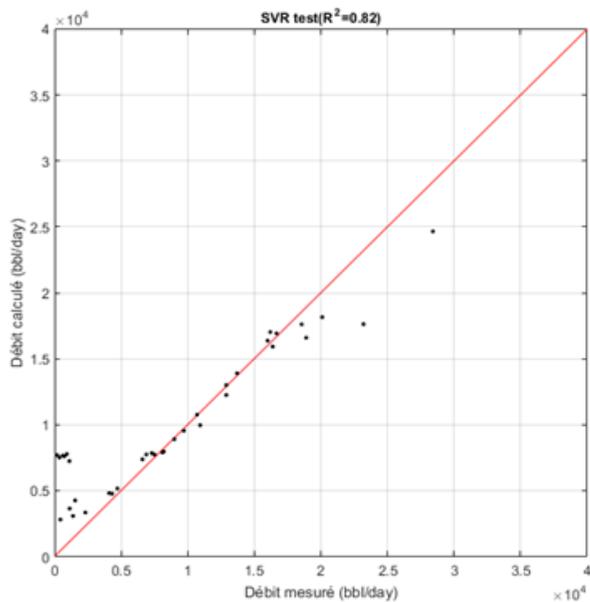


Figure 4.13 : cross plot de test de modèle SVR au champ de Reshadat

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

### 4.2.1 Description du dataset brut

Le dataset se présente sous forme d'un tableau Excel. Il comporte les données brutes de production de 1729 puits du champ de Hassi Messaoud, et ce depuis 1958 jusqu'à 2019. Ces données sont obtenues via des opérations de jaugeage qui s'effectuent en moyenne 4 fois par an pour chaque puits. Ceci peut différer d'un puits à un autre, en fonction de plusieurs facteurs parmi lesquels ; les problèmes qui peuvent apparaître au niveau de quelques puits durant leurs production ; ce qui implique parfois des opérations exceptionnelles, la disponibilité des installations et du personnel et bien sûr des conditions économiques du moment.

Le Dataset brut comporte initialement 67660 lignes de données. Il contient les variables (features) suivantes : le code du puits, la date du jaugeage, la taille de la duse, la pression en tête de puits (pression en amont de la duse), la pression du flow line (pression en aval de la duse), la pression au séparateur, le GOR, le débit de production d'huile, de gaz et d'eau, la densité du mélange produit. Ajouté à cela, quelques variables ne sont reportées que parfois dans le tableau, parmi lesquelles : la température de fluide, la pression au séparateur et la masse volumique d'huile.

### 4.2.2 Préparation des données

Ces données ne peuvent pas être utilisées tel qu'elles sont, car elles contiennent des valeurs manquantes, des erreurs de frappe, des outliers, des valeurs erronées... La préparation de ces données est donc une étape primordiale qui va permettre d'en tirer le maximum de valeur ainsi que de concevoir des modèles avec les meilleures performances.

En résumé, les étapes suivantes sont suivies pour la préparation des données :

- Elimination des lignes contenant des valeurs manquantes.
- Suppression des outliers, qui peuvent indiquer soit une erreur de mesure ou un résultat de variabilité présente dans les données. A titre d'exemple, des valeurs de débit de l'ordre de millions de barils par jour, des valeurs de diamètres ou de pressions très grandes...
- Correction des erreurs de frappe. Par exemple un diamètre  $\emptyset=2000$  (1/64 in), cette valeur est très grande par rapport aux dimensions d'une duse, il est clair qu'il s'agit d'une erreur de frappe, donc elle est corrigée en  $\emptyset=20$  (1/64 inch).
- Eliminations des lignes contenant des valeurs illogiques qui peuvent biaiser la création des modèles tel que :
  - Une pression de flow line supérieure à la pression de tête.
  - Un diamètre de duse noté nul dans le dataset.
- Création de nouvelles variables (ce qui est connu sous le nom de feature engineering) tel que le rapport de pression  $P_d/P_{up}$ , chute de pression à travers la duse  $\Delta P$ .
- Division de l'ensemble de données en données du régime critique et sous critique selon les valeurs du rapport de pression critique  $P_d/P_{up}$ . Le régime critique pour le champ de

HMD 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.

Le tableau ci-dessous montre les différentes plages de valeurs des paramètres utilisées. Il contient les valeurs maximales, minimales ainsi que la moyenne de chaque paramètre après le traitement de données.

Tableau 4.5 : Différentes plages des paramètres de données du champ de HMD

Paramètre	<b>d</b> (1/64 in)	<b>p<sub>u</sub></b> (psi)	<b>p<sub>d</sub></b> (psi)	GOR (Mscf/STB)	Q (bbl/day)	<b>Y</b>
Minimum	3	42.66	5.6880	0.0729	1,50955	0.681
Maximum	130	4629.2	1166	1620.7	7754.4	0.8575
Moyenne	17.2	532.9	277.3	6.692	618.1	0.8053

## Résultats et discussions

Quelques modèles visant à prédire le débit de production à travers les duses, développés dans le chapitre 2, ont été sélectionnés pour être appliqués dans la présente étude spécifique au champ de Hassi Messaoud.

Les modèles ont été sélectionnés en se basant principalement sur les critères suivants :

- La disponibilité des données nécessaires à leurs constructions.
- Sélection de modèles relativement simple et pratique à utiliser.
- Intégration d'un modèle contenant la densité du mélange produit pour voir l'influence d'une autre propriété de fluides sur le modèle.

En fonction du régime d'écoulement (critique ou sous-critique), les corrélations suivantes seront utilisées :

**Régime critique :**

- Corrélation de Gilbert (Equations 2.21 et 2.22)
- Corrélation de Towailib et Marhoun (Equation 2.29)

**Régime sous critique :**

- Corrélation de Gilbert (Equations 2.21 et 2.22)
- Corrélation de Elgibaly et Nashawi (Equation 2.33)
- Extension de la forme de Gilbert (Equation 2.41)
- Extensions utilisées à HMD (Equation 2.42 et 2.43)

### 4.3 Evaluation des modèles théoriques :

Les modèles sélectionnés précédemment sont dans un premier temps testés en utilisant les données du champ de Hassi Messaoud pour étudier leur applicabilité dans ce champ. Les

statistiques et les différentes erreurs de test de ces modèles sont reportées dans les tableaux (4.6) et (4.7). Les différents cross plot sont également illustrés ci-dessous.

**Tableau 4.6** : statistiques de test des modèles théoriques au régime critique (HMD)

Forme	modèle	R <sup>2</sup>	RMSE	MAE	RAE
1 <sup>ère</sup> forme	Gilbert	- 0.79	872	671.5	0.83
	Achong	-0.67	841	649	0.81
	Baxendell	-0.68	845.4	648.5	0.81
	Pilehvari	-1.57	1043.5	815	0.99
	Ros	-0.74	859	657	0.82
2 <sup>ème</sup> forme	Towailib & Marhoun	- 0.73	688.1	494.7	0.81

**Tableau 4.7** : statistiques de test des modèles théoriques au régime non critique (HMD)

Forme	modèle	R <sup>2</sup>	RMSE	MAE	RAE
1 <sup>ère</sup> forme	Gilbert	-1.06	831.5	596.67	1
2 <sup>ème</sup> forme	Elgibaly et Nashawi	-0.24	694.3	456.7	0.79
3 <sup>ème</sup> forme	Extension Gilbert	-0.54	774.1	539.86	0.88

D'après les résultats obtenus en testant les modèles sélectionnés pour les 2 régimes (critique et sous-critique), il paraît qu'aucun de ces modèles ne peut décrire le débit produit au champ de HMD. Les erreurs calculées sont très importantes, et les valeurs du coefficient de détermination sont toutes négatives, ceci apparaît clairement dans les cross plot, où on constate que le nuage de point s'écarte considérablement de la ligne parfaite (45°).

Ces résultats semblent cependant logiques, parce que tous ces modèles ; et avec leurs paramètres originaux, ont été développés sur la base de données d'autres champs, avec des propriétés d'écoulements et de fluides bien spécifiques, ce qui nous emmène à dire que le champ de HMD nécessite une étude spéciale.

Pour tenter de résoudre ce problème, on va essayer d'adapter ces différents modèles sur la base de données du champ de HMD, puis analyser les résultats et voir quel est le modèle qui s'adapte mieux à notre cas d'étude.

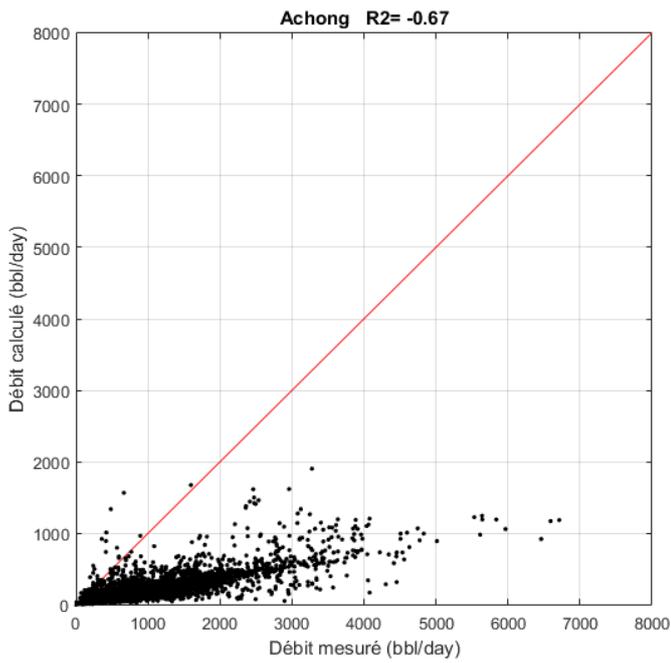


Figure 4.14: cross plot du test de la corrélation de Achong

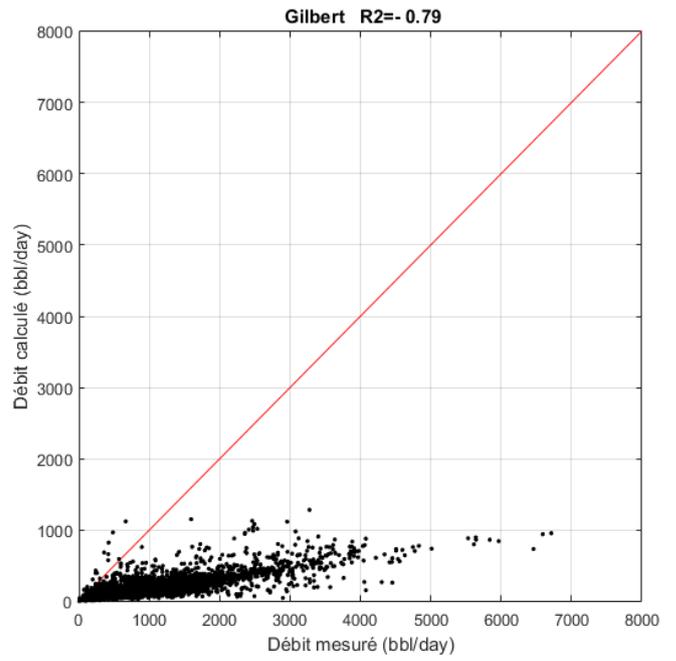


Figure 4.15 : cross plot du test de la corrélation de Gilbert (régime critique)

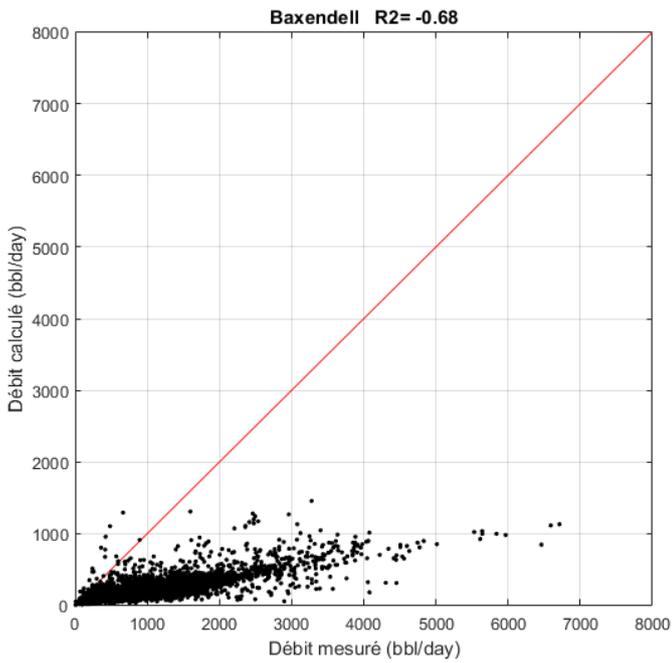


Figure 4.16 : cross plot du test de la corrélation de Baxendell

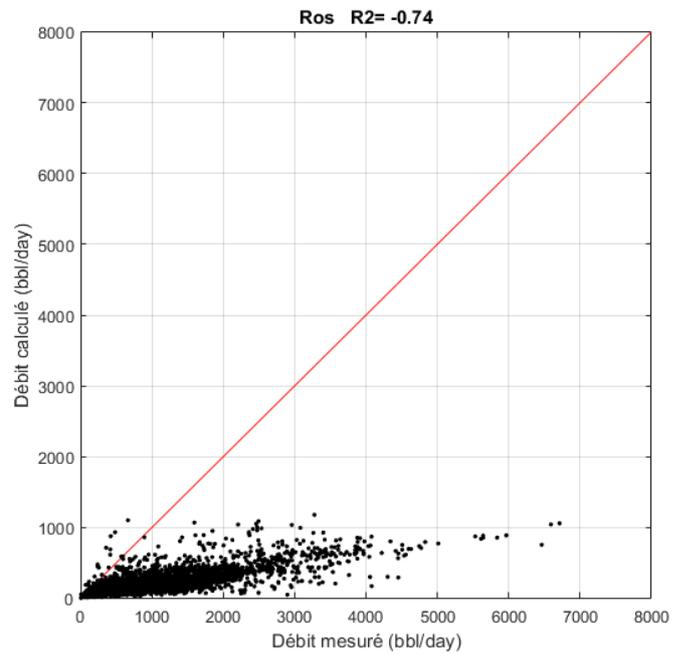


Figure 4. 17 : cross plot du test de la corrélation de Ros

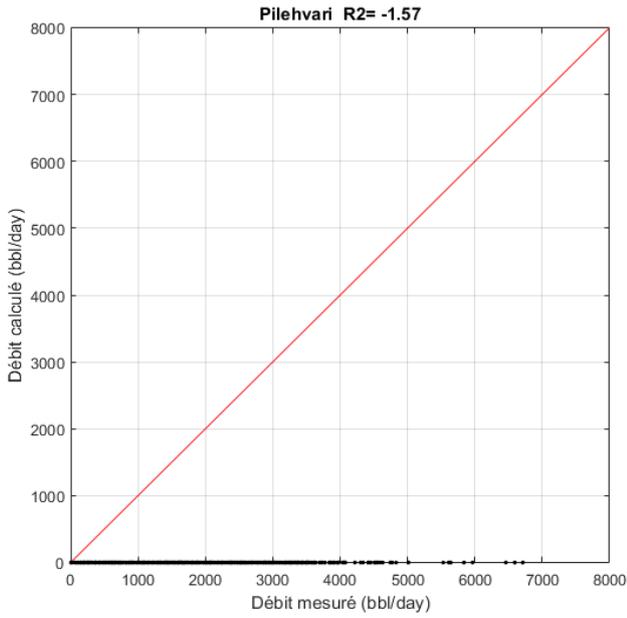


Figure 4.18 : cross plot du test de la corrélation de Pilehvari

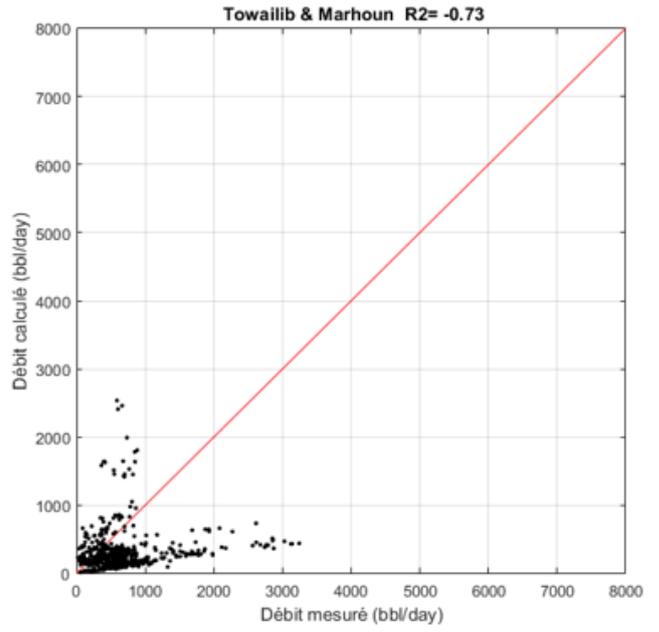


Figure 4.19 : cross plot du test de la corrélation de Towailib et Marhoun

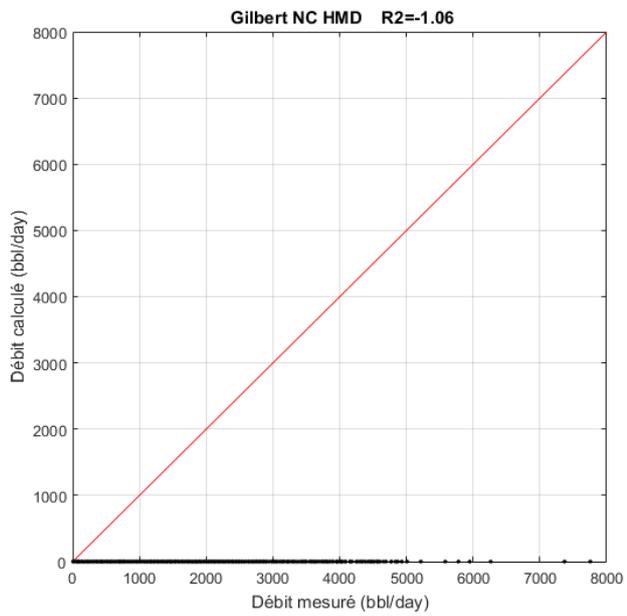


Figure 4. 20: cross plot du test de la corrélation de Gilbert au régime non critique

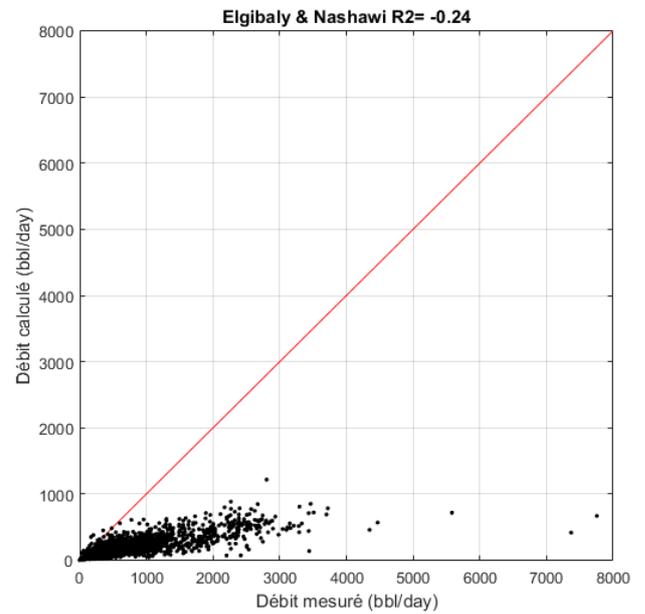


Figure 4.21: cross plot du test de la corrélation de Elgibaly et Nashawi

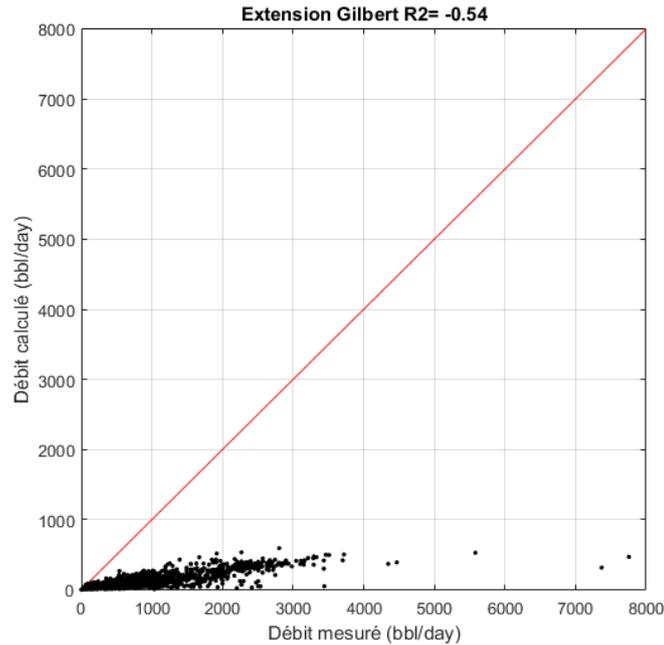


Figure 4. 22 : cross plot du test de l'extension de la corrélation de Gilbert

## 4.4 Adaptation des modèles théoriques et création de modèles intelligents

### 4.4.1 Méthodologie

Dans cette partie, on va essayer de développer quelques modèles pour l'estimation du débit de production des puits au champ de Hassi Messaoud en partant sur deux volets, l'adaptation des modèles théoriques et la création de quelques modèles intelligents.

#### ✓ Modèles théoriques

L'algorithme de régression non linéaire de Levenberg-Marquardt a été utilisé pour déterminer les valeurs de constantes et exposants des variables des modèles théoriques qui correspondent le mieux aux données mesurées à HMD. Les paramètres sont estimés en minimisant la somme des carrés des différences entre les points de données et la fonction du modèle.

#### ✓ Modèles intelligents

Chaque modèle intelligent est créé en utilisant comme entrées ; les mêmes variables du modèle théorique étudié. Par exemple, pour le modèle Gilbert, les variables utilisées comme entrée pour l'entraînement du modèle intelligent sont : la pression en amont de la duse, le diamètre et le GOR, et comme sortie ; le débit de production. Les modèles intelligents développés dans cette section sont :

- Support Vector Regression (SVR)
- Gaussian Process Regression (GPR)
- Artificial Neural Network (ANN)

Pour chaque modèle (théorique ou intelligent), les données initiales sont divisées en deux parties :

- 80% sont utilisées pour la création et l'apprentissage du modèle, on les appelle également données du training et c'est la notation qui sera utilisée dans la suite du document.
- 20 % sont utilisées pour tester la performance du modèle créé, et sa capacité de prédiction sur de nouvelles données, la notation utilisée est donc données de test.

#### 4.4.2 Nouveaux modèles : résultats et discussion

Dans cette partie, les différents résultats obtenus après l'adaptation des modèles théoriques et la création des modèles intelligents seront présentés.

##### 4.4.2.1 Régime critique

###### ❖ Modèle de Gilbert

Pour cette forme 46018 points de données sont utilisées dont 80% pour le training des modèles et 20% pour le test. Le débit à travers la corrélation de Gilbert s'écrit en termes du diamètre et pression en amont de la duse et du GOR. La corrélation de Gilbert adaptée en utilisant l'équation (2.20) est notée Gilbert\_HMD 1 tandis que celle adaptée en introduisant la puissance  $e$  ; c.-à-d. l'équation (2.21) est notée Gilbert\_HMD 2. Le GLR dans cette équation est remplacé par le GOR.

Les résultats concernant cette forme sont reportés dans les tableaux (4.8) et (4.9) qui contiennent respectivement les nouvelles constantes de la corrélation de Gilbert et les différentes statistiques relatives au training et au test du modèle théorique et des modèles intelligents. Les différents cross plot sont également illustrés.

**Tableau 4.8 :** les nouveaux coefficients de l'équation de Gilbert adaptée au champ de HMD (Régime critique)

Nouveaux coefficients pour l'équation de Gilbert				
	a	b	c	e
Gilbert_HMD 1	0.902	1.559	0.492	1
Gilbert_HMD 2	10.033	1.764	0.598	0.768

**Tableau 4.9 :** statistiques de training et de test du modèle théorique de Gilbert adapté et des modèles intelligents (Régime critique)

Modèle	Métriques training				Métriques test			
	R <sup>2</sup>	RMSE	MAE	RAE	R <sup>2</sup>	RMSE	MAE	RAE
Gilbert_HMD 1	0.645	380	229.44	0.36	0.637	391.7	233.9	0.375
Gilbert_HMD 2	0.69	354.84	220.74	0.34	0.68	367.9	232	0.35
SVR	0.79	310,21	173.48	0.29	0.68	366.8	211.4	0.34
GPR	0.81	275.4	164.4	0.26	0.69	359.6	215.3	0.34
ANN	0.80	304.5	179.7	0.28	0.80	289.9	176.4	0.27

- Le modèle de Gilbert adapté avec ses deux formes (Gilbert\_HMD 1 et Gilbert\_HMD 2) est plus représentatif des données du champ de Hassi Messaoud que le modèle original. Plus particulièrement, le modèle Gilbert\_HMD 2, avec l'introduction de la puissance  $e$  dans la formule est meilleur que le modèle Gilbert\_HMD 1 développé sans puissance.
- Le modèle Gilbert\_HMD 2 représente un coefficient de détermination de l'ordre de 0.68, et une erreur relative absolue relativement faible de l'ordre de 0.35, ce qui veut dire que ce modèle peut donner une bonne estimation du débit d'huile. Ceci apparaît clairement dans les cross plot des figures (4.25) et (4.26).
- Les modèles intelligents développés pour cette forme ont cependant montré de meilleures performances que les modèles théoriques adaptés lors du training (les modèles SVR, GPR, ANN ont montré des valeurs de coefficients de détermination de l'ordre de 0.8). Lors du test, les performances de SVR et GPR ont diminué un peu comparées au training.
- Lors du test, le modèle ANN a gardé presque les mêmes performances que celles enregistrées lors du training, il a également montré les plus petites erreurs parmi tous les autres modèles intelligents.

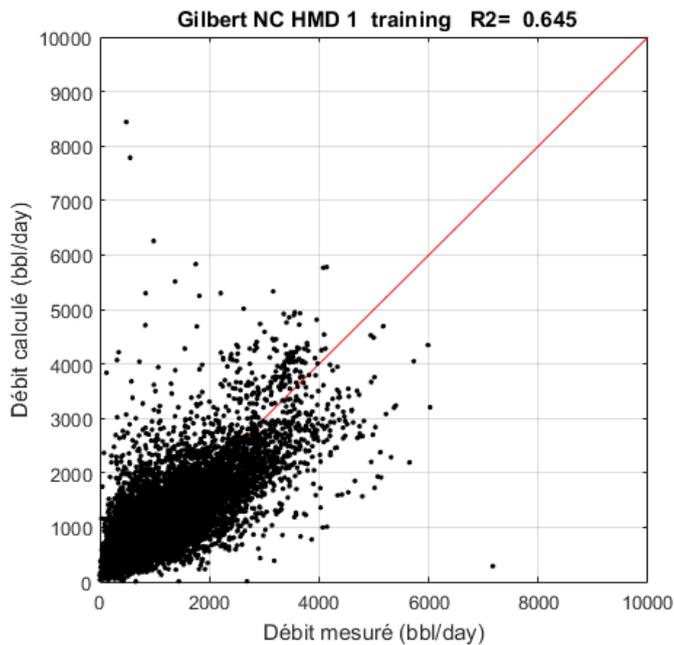


Figure 4.23 : cross plot de training du modèle Gilbert\_HMD 1

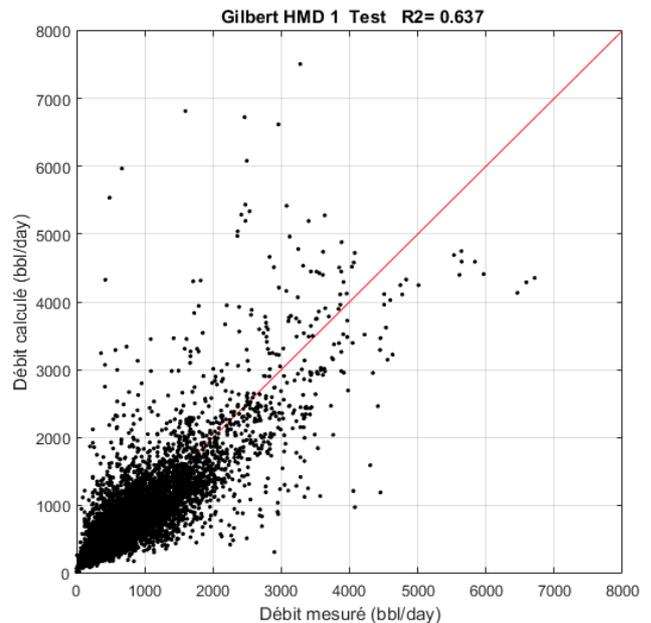


Figure 4.24 : cross plot de test du modèle Gilbert\_HMD 1

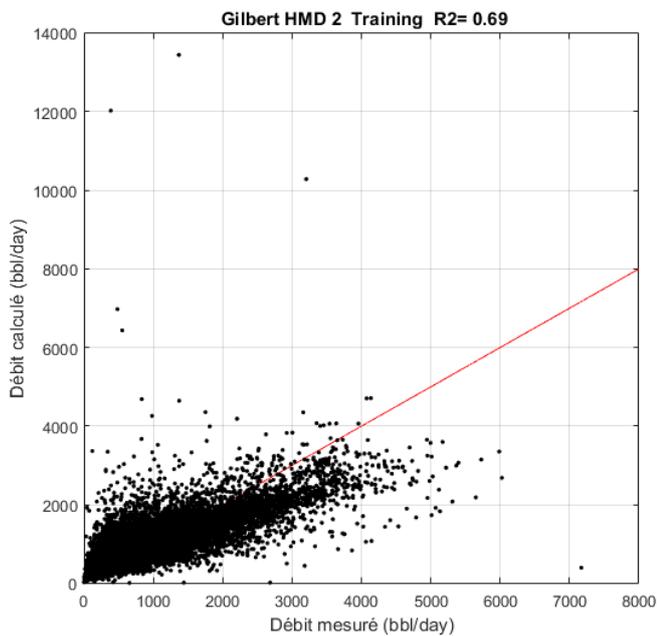


Figure 4.25 : cross plot de training du modèle Gilbert\_HMD 2

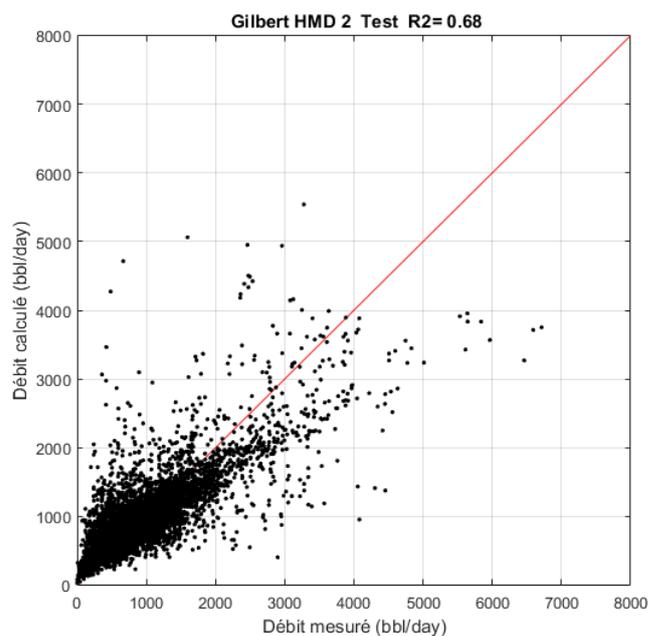


Figure 4.26 : cross plot de test du modèle Gilbert\_HMD 2

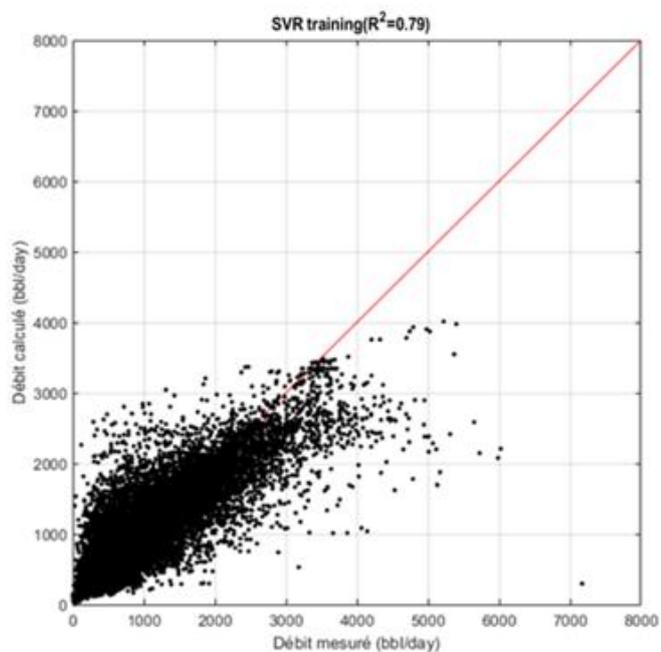


Figure 4.27 : cross plot de training du modèle SVR

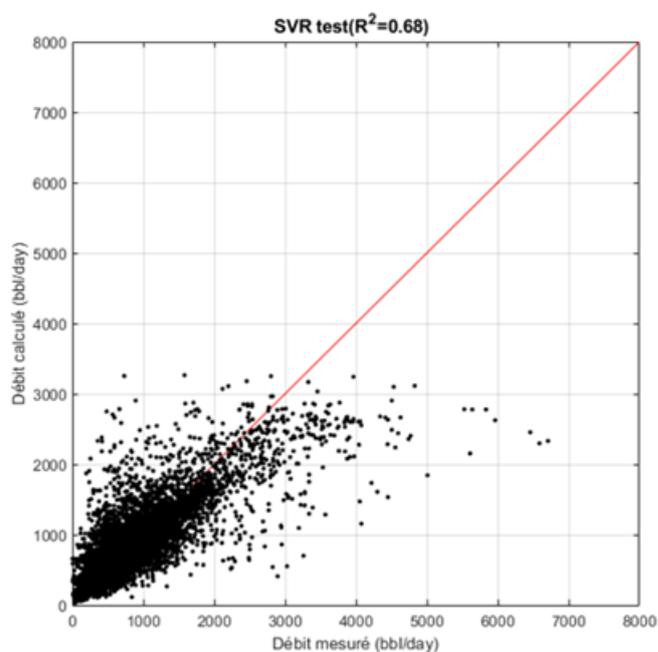


Figure 4.28 : cross plot de test du modèle SVR

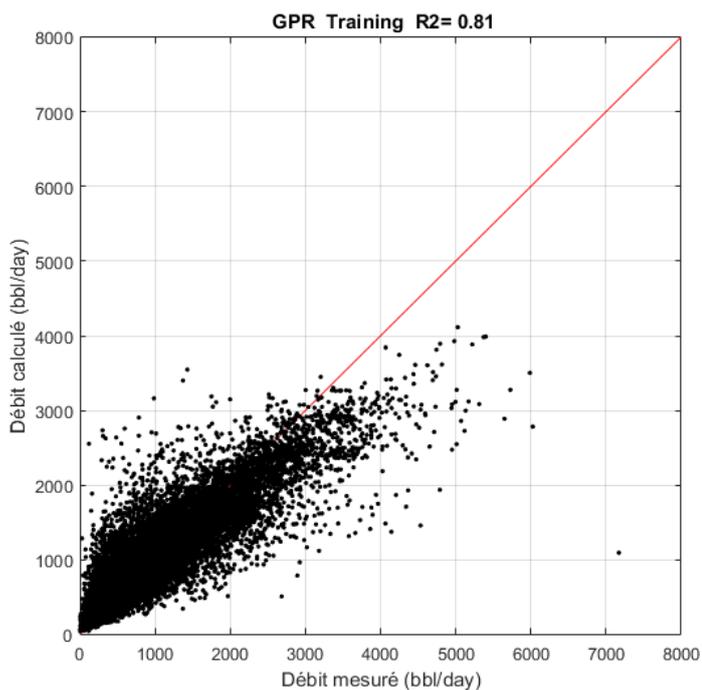


Figure 4.29 : cross plot de training du modèle GPR

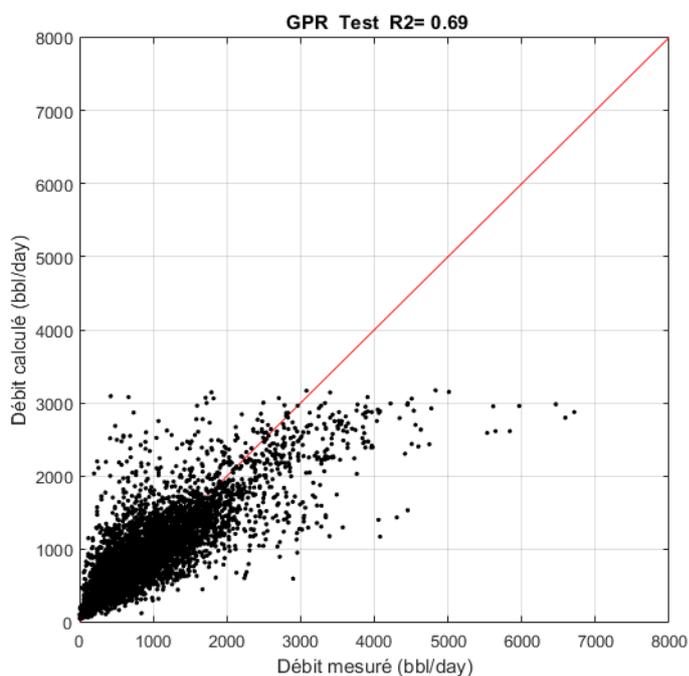


Figure 4.30 : cross plot de test du modèle GPR

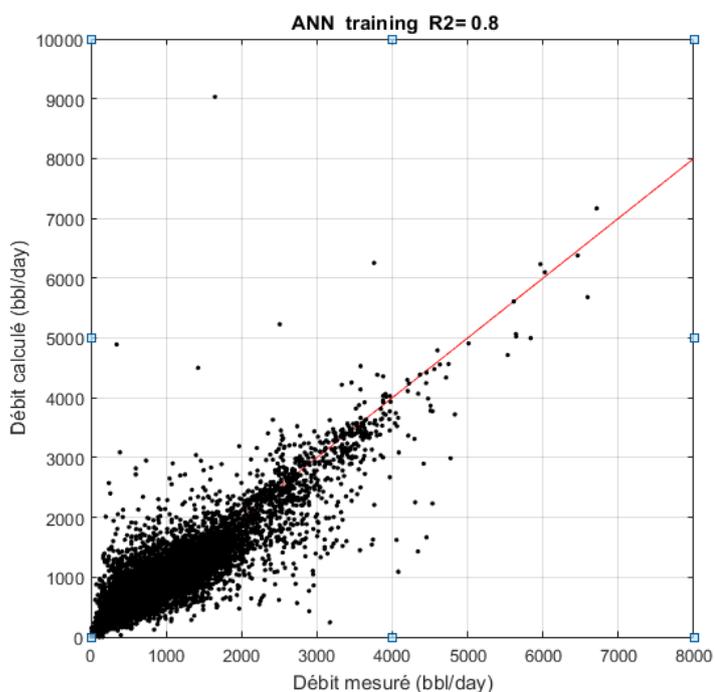


Figure 4.31 : cross plot de training du modèle ANN

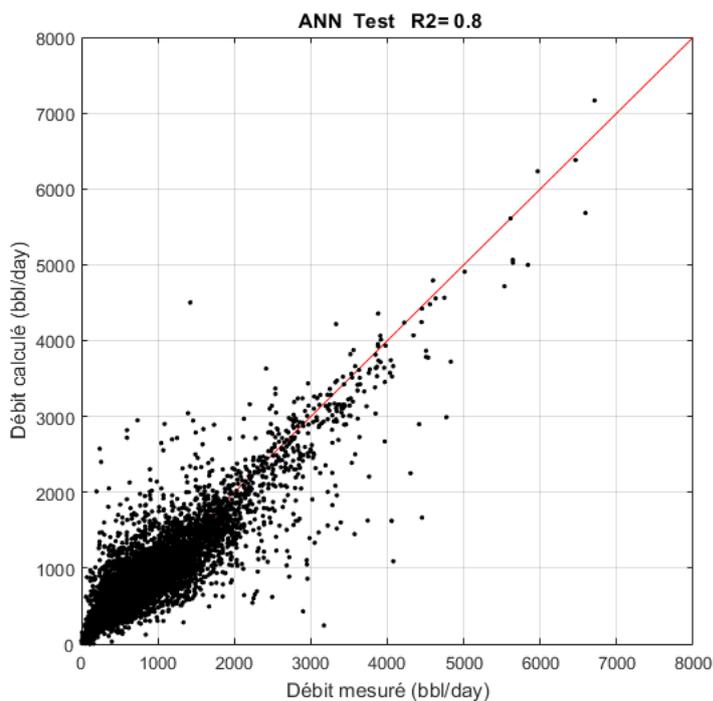


Figure 4.32 : cross plot de test du modèle ANN

### ❖ Modèle de Towailib et Marhoun

Dans cette forme (Equation 2.29), le débit est une fonction du diamètre de la duse, de la pression en amont de la duse ainsi que la densité du fluide produit.

Le modèle de Towailib et Marhoun adapté suivant les données de HMD est noté Towailib & M. \_HMD. 3105 points de données ont servi pour la création de ce modèle (80 % sont utilisées pour le training et le reste pour le test).

Les tableaux (4.10) et (4.11) représentent respectivement les nouveaux coefficients et les différentes métriques d'évaluation du modèle théorique et des modèles intelligents. Les différents cross plot sont également illustrés dans les figures ci-dessous.

**Tableau 4.10** : les nouveaux coefficients de l'équation de Towailib et Marhoun adaptée au champ de HMD

Nouveaux coefficients pour l'équation de Towailib et Marhoun				
	a	b	c	d
Towailib & M. _HMD	0.20	0.767	0.001	0.005

**Tableau 4. 11** : statistiques de training et de test du modèle théorique de Towailib et Marhoun adapté et des modèles intelligents

Modèle	Métriques training				Métriques test			
	R <sup>2</sup>	RMSE	MAE	RAE	R <sup>2</sup>	RMSE	MAE	RAE
Towailib & M. _HMD	0.35	480.41	325.1	0.54	0.35	420	284.5	0.49
SVR	0.55	400.60	222.29	0.46	0.47	380.11	253.77	0.45
GPR	0.91	174.5	122	0.20	0.39	407	280.5	0.48
ANN	0.84	236.5	137.3	0.27	0.8	234.1	134	0.27

- Le modèle de Towailib & M. \_HMD représente une erreur relative absolue de l'ordre de 0.49, le coefficient de détermination, quant à lui, est de l'ordre de 0.35. Cela veut dire que le modèle s'éloigne des résultats réels.
- La performance du modèle support vector machine regression SVR est relativement meilleure que celui de Towailib & M. \_HMD. Les erreurs ont relativement diminuées.
- Le gaussian process a montré les meilleures performances et les plus petites erreurs parmi tous les autres modèles lors du training. Cependant, sa performance a diminué et les erreurs ont relativement augmenté lorsqu'il a été testé sur de nouvelles données.
- Le modèle ANN a enregistré de très bonnes performances lors du training du modèle. Il a également montré les meilleures performances de prédiction comparé à tous les autres modèles.

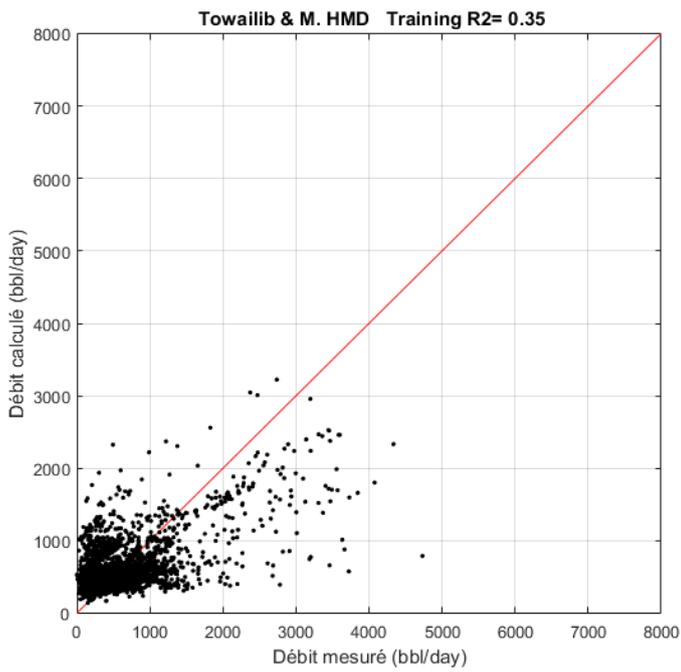


Figure 4.33 : cross plot de training du modèle de Towailib & M. \_HMD

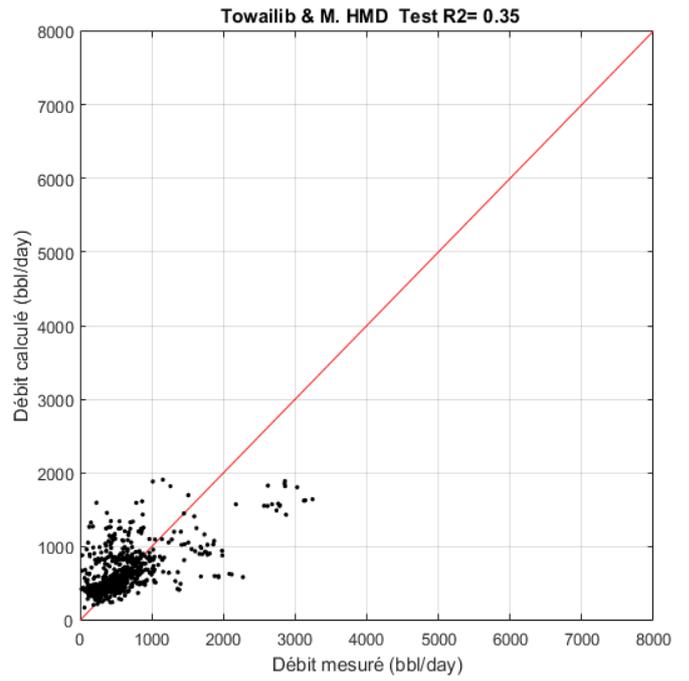


Figure 4. 34 : cross plot de test du modèle de Towailib &M. \_HMD

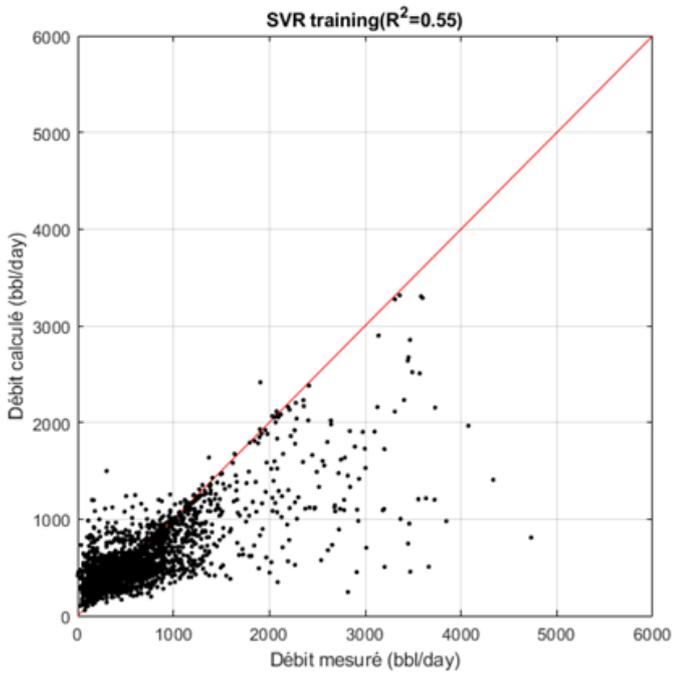


Figure 4.35 : cross plot de training du modèle SVR

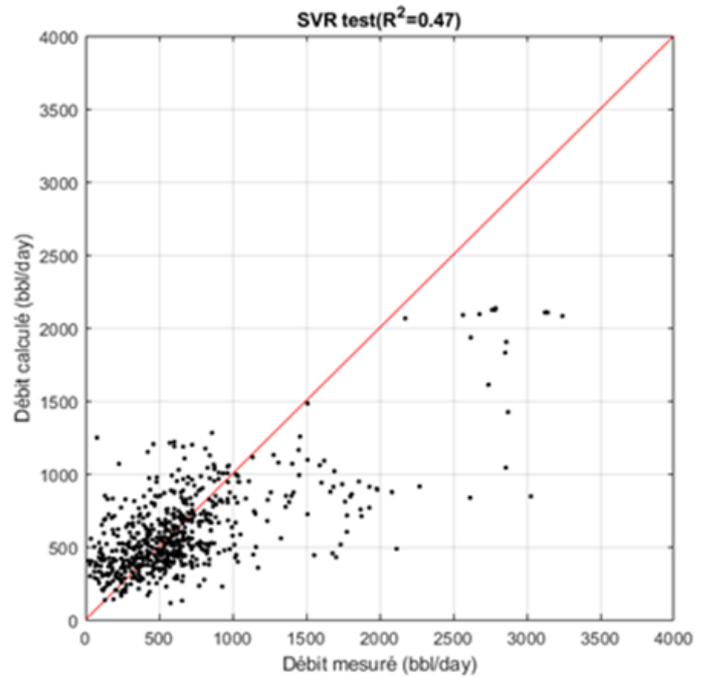


Figure 4. 36 : cross plot de test du modèle SVR

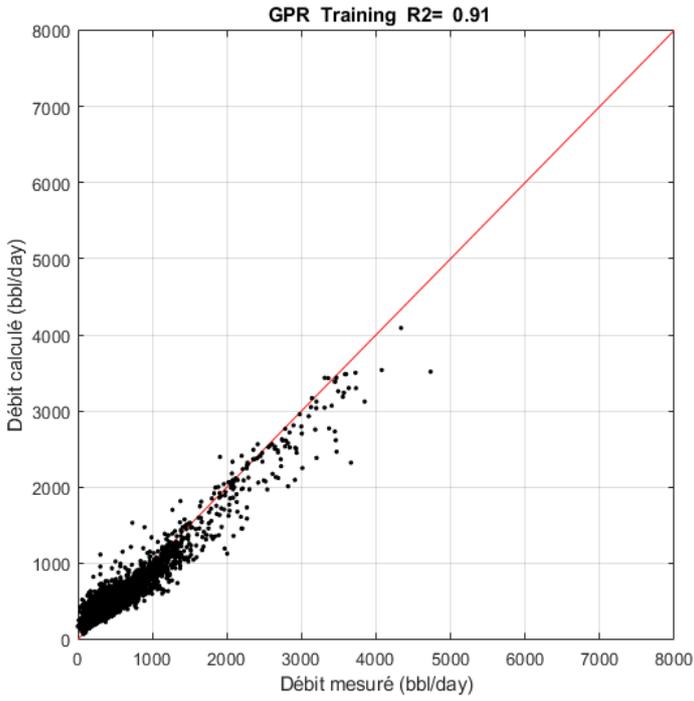


Figure 4.37 : cross plot de training du modèle GPR

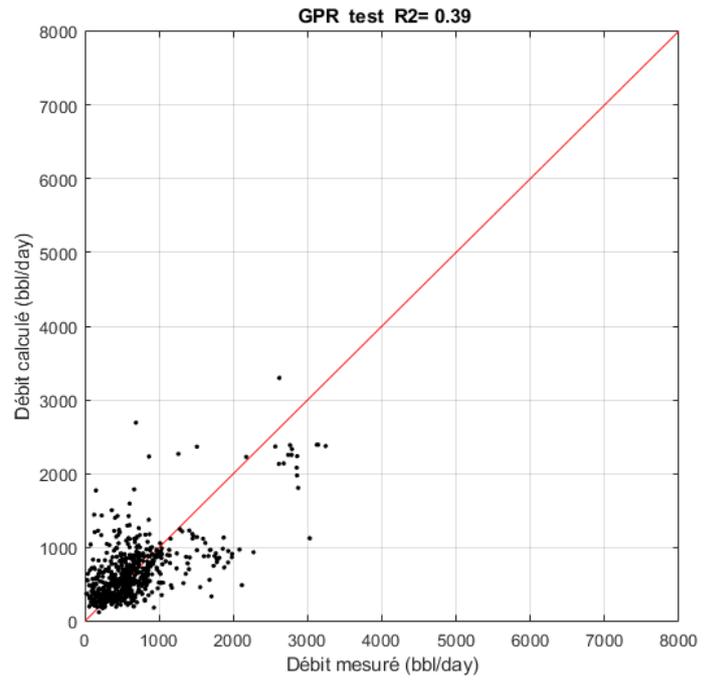


Figure 4.38 : cross plot de test du modèle GPR

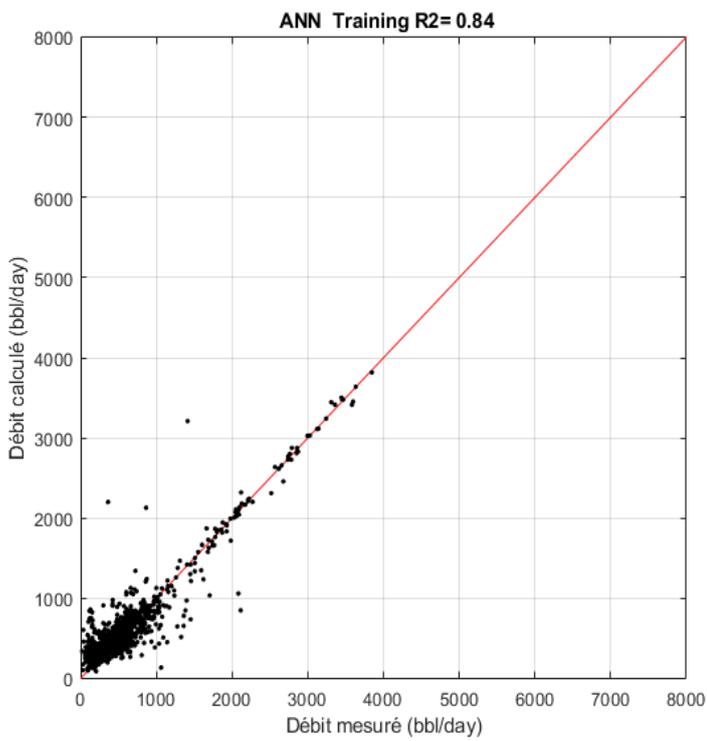


Figure 4.39 : cross plot de training du modèle ANN

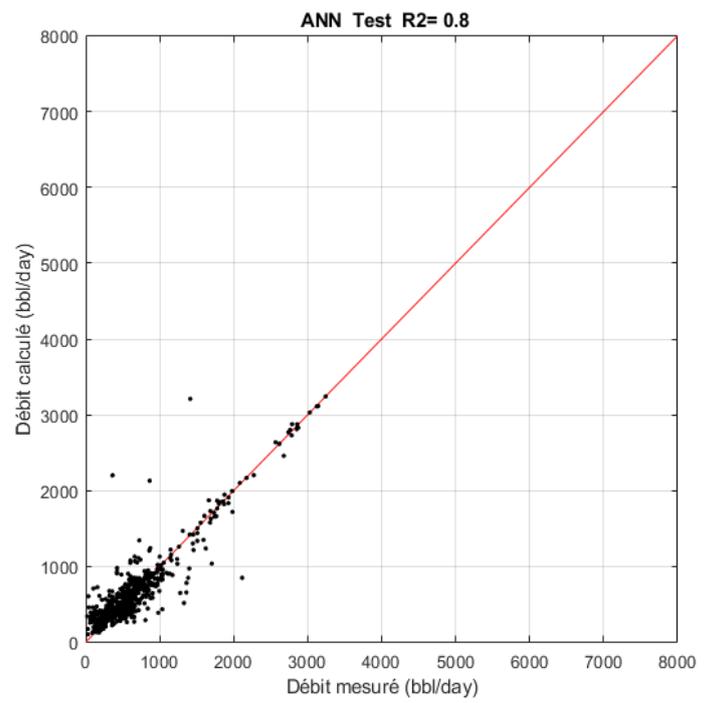


Figure 4.40 : cross plot de test du modèle ANN

#### 4.4.2.2 Régime non critique

Pour tous les modèles de ce régime ; 15603 Points de données ont été utilisés (80% ont servi pour la création des modèles et le reste pour le test).

##### ❖ Modèle de Gilbert

Comme il a été expliqué auparavant, l'équation de Gilbert (Equation 2.20) est développée pour le régime d'écoulement critique. Le débit est une fonction du diamètre de la duse, de la pression en amont de la duse et du GOR.

La même équation avec ses deux formes a été adaptée pour le régime non critique et les résultats sont montrés dans les tableaux (4.12) et (4.13).

Gilbert NC\_HMD 1 correspond à la première forme de l'équation de Gilbert (équation 2.20) adaptée au régime non critique et Gilbert NC\_HMD 2 représente la deuxième forme (équation 2.21) .

**Tableau 4.12 :** les nouveaux coefficients de l'équation de Gilbert adaptée au champ de HMD (Régime non critique)

Nouveaux coefficients pour l'équation de Gilbert au régime non critique				
	a	b	c	e
Gilbert NC_HMD 1	0.313	1.684	0.418	1
Gilbert NC_HMD 2	2.375	2.532	0.571	0.662

**Tableau 4.13 :** statistiques de training et de test du modèle théorique de Gilbert adapté au champ de HMD et des modèles intelligents (Régime non critique)

Modèle	Métriques du training				Métriques du test			
	R <sup>2</sup>	RMSE	MAE	RAE	R <sup>2</sup>	RMSE	MAE	RAE
Gilbert NC_HMD 1	0.65	332.5	197.4	0.40	0.71	333.5	204.6	0.38
Gilbert NC_HMD 2	0.69	315.2	188.9	0.38	0.72	327.6	198.8	0.37
SVR	0.76	276.89	149.94	0.34	0.68	351.06	191.46	0.4
GPR	0.95	131.8	71	0.16	0.7	337.5	193.8	0.38
ANN	0.8	259	159.32	0.31	0.83	253	156.32	0.28

- Le modèle Gilbert NC\_HMD 1 et Gilbert NC\_HMD 2 ont donné des résultats intéressants si on prend en compte le fait que l'équation de Gilbert est initialement développée pour le régime d'écoulement critique.
- Le modèle Gilbert NC\_HMD 2 est meilleur que Gilbert NC\_HMD 1, l'erreur relative absolue est de l'ordre de 0.37.
- Les performances de Gilbert NC\_HMD 2 sont relativement meilleures que le SVR dans la phase de test.
- Le gaussian process GPR a montré les meilleures performances parmi tous les autres modèles lors de la phase de training, avec une erreur relative absolue de 0.16 et un coefficient de détermination de 0.95. Cependant, sa performance a relativement diminué dans la phase de test.
- Le modèle ANN a montré de très bons résultats lors du training, et les plus petites erreurs comparées à tous les autres modèles lors de la phase de test. Les résultats donnés par ce modèle sont plus proches des valeurs réels.

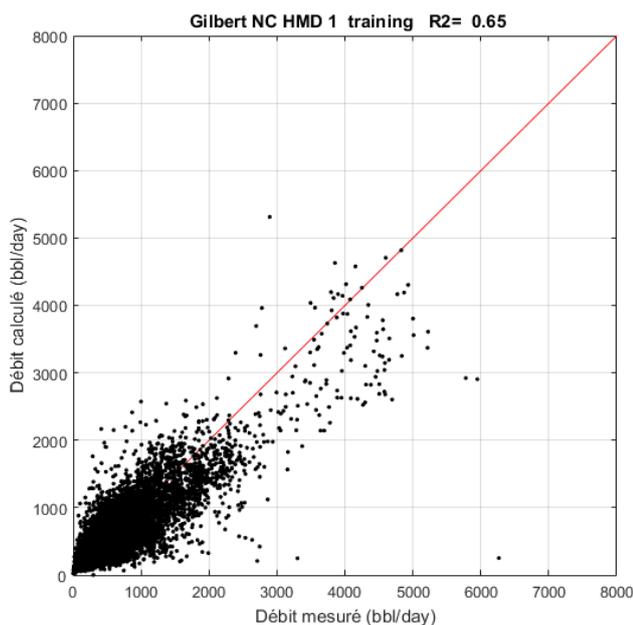


Figure 4.41 : cross plot de training du modèle Gilbert NC\_HMD 1

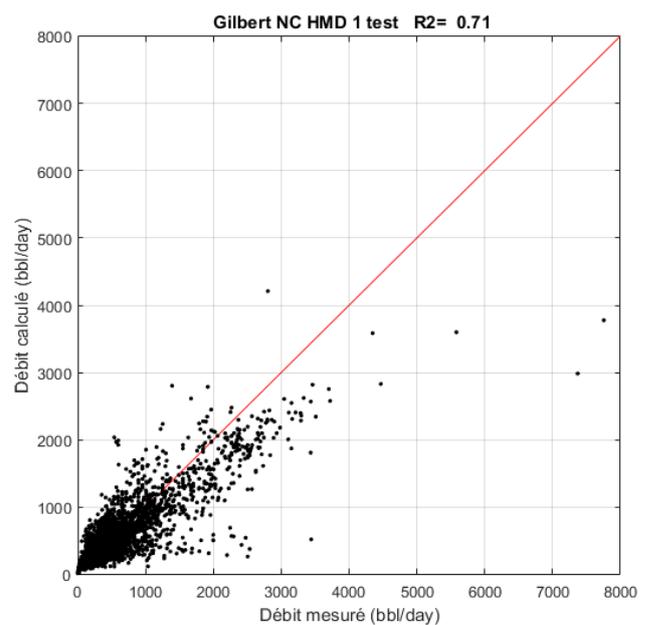


Figure 4.42 : cross plot de test du modèle Gilbert NC\_HMD 1

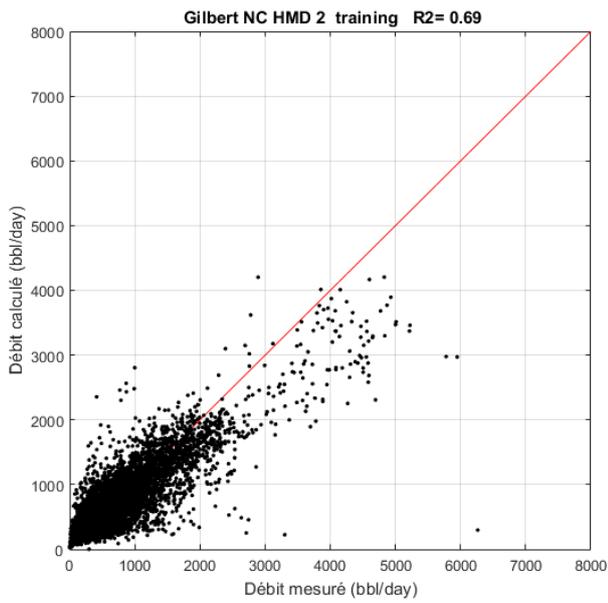


Figure 4.43 : cross plot de training du modèle Gilbert NC\_HMD 2

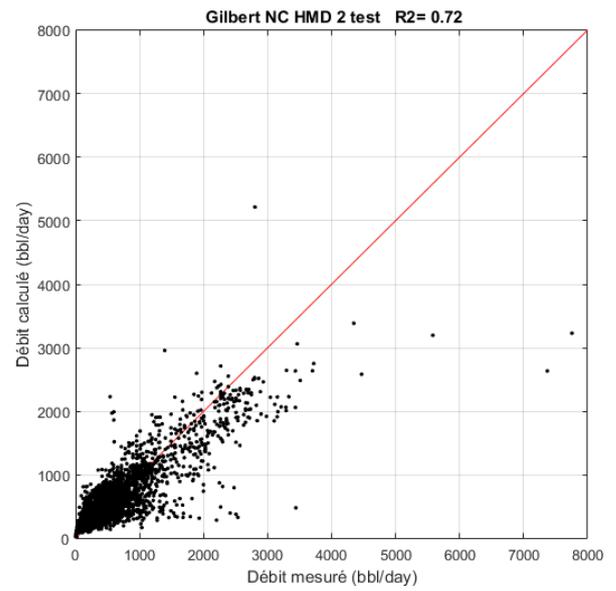


Figure 4.44 : cross plot de test du modèle Gilbert NC\_HMD 2

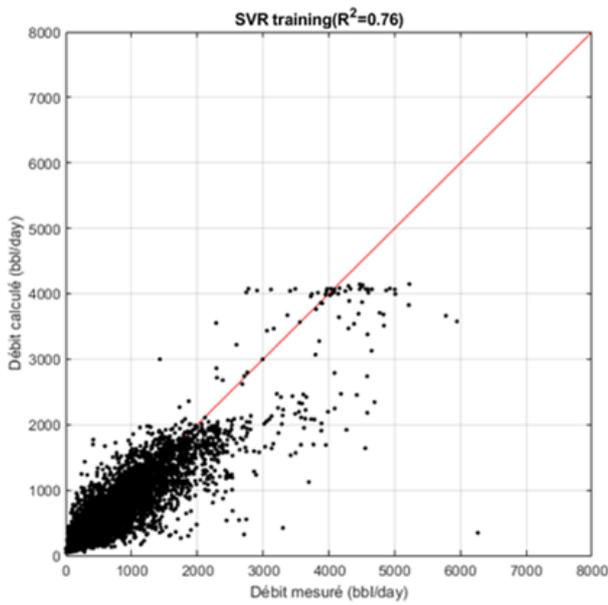


Figure 4.45 : cross plot de training du modèle SVR

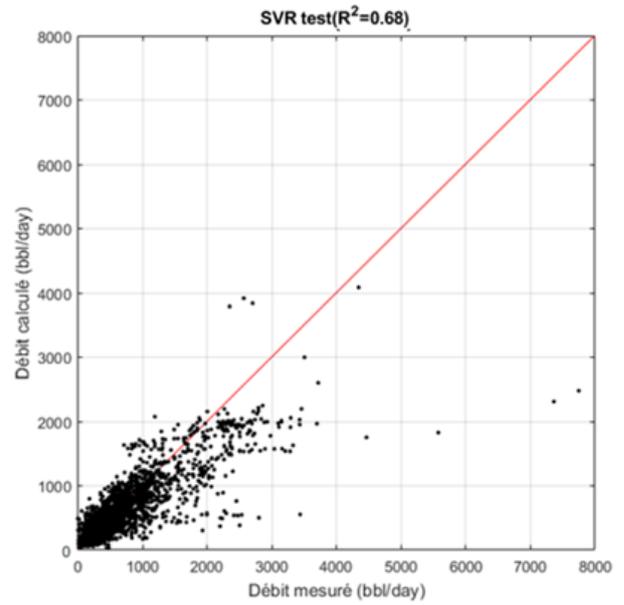


Figure 4.46 : cross plot de test du modèle SVR

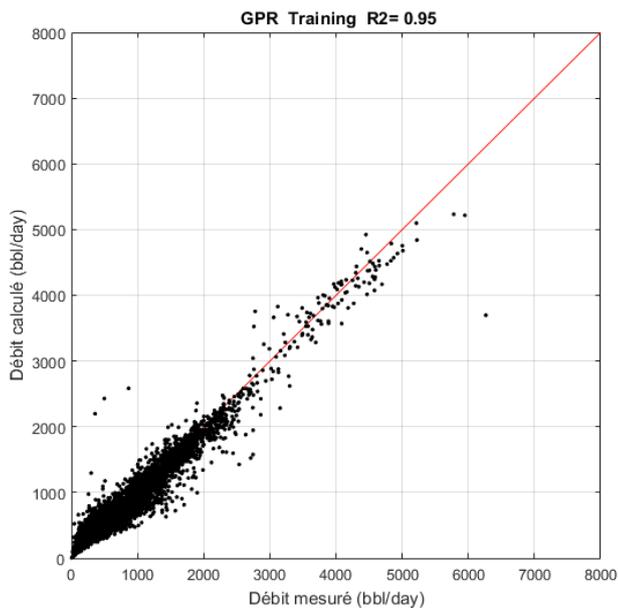


Figure 4.47 : cross plot de training du modèle GPR

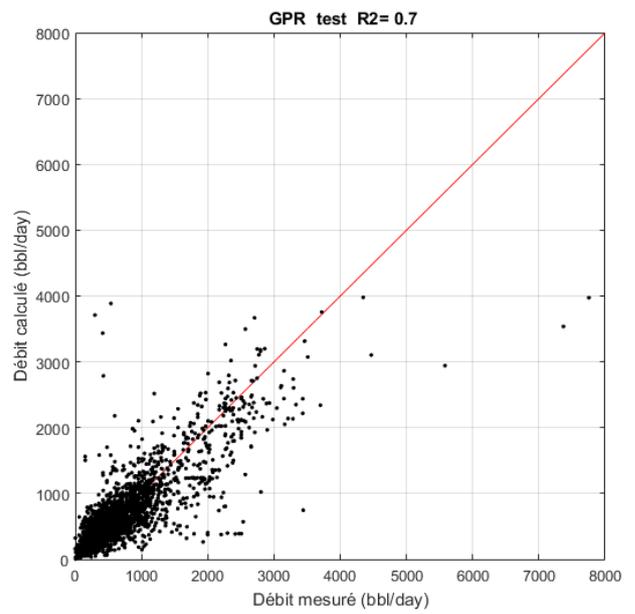


Figure 4.48 : cross plot de test du modèle GPR

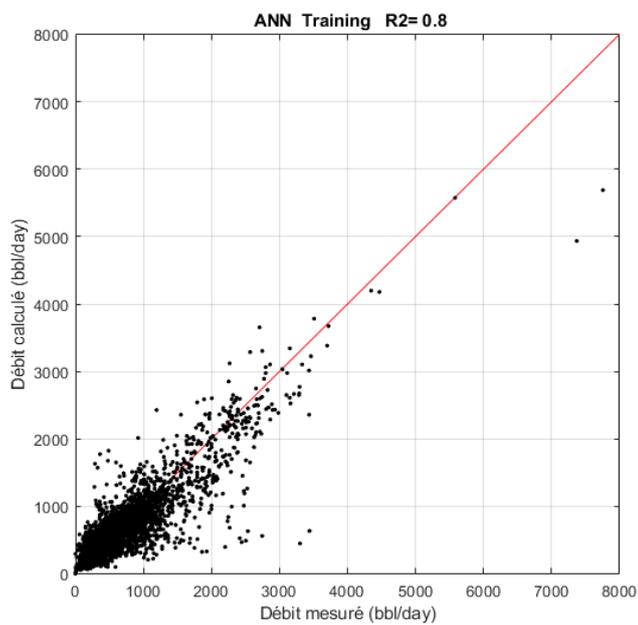


Figure 4. 49 : cross plot de training du ANN

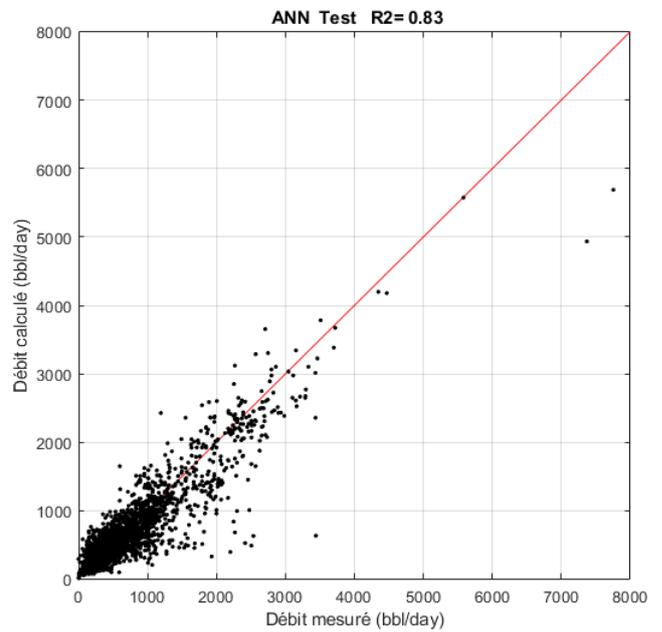


Figure 4. 50 : cross plot de test du modèle ANN

### ❖ Modèle de Elgibaly et Nashawi

L'expression du débit pour ce modèle (Equation 2.33) est une fonction du diamètre de la duse, de la chute de pression à travers la duse  $\Delta P$  et du GOR.

Le modèle de Elgibaly et Nashawi Adapté au champ de Hassi Messaoud est noté Elgibaly & N. \_HMD. Les nouveaux coefficients de ce modèle sont portés dans le tableau (4.14) et les résultats de calcul des différentes métriques d'évaluation des modèles sont représentés dans le tableau (4.15)

**Tableau 4.14:** les nouveaux coefficients de l'équation de Elgibaly et Nashawi adaptée au champ de HMD

Nouveaux coefficients pour l'équation Elgibaly et Nashawi			
	a	b	c
Elgibaly & N. _HMD	119.97	3.34	0.73

**Tableau 4.15 :** statistiques de training et de test du modèle de Elgibaly et Nashawi adapté au champ de HMD et des modèles intelligents

Modèle	Métriques training				Métriques test			
	R <sup>2</sup>	RMSE	MAE	RAE	R <sup>2</sup>	RMSE	MAE	RAE
Elgibaly & N. _HMD	0.64	336	196	0.41	0.69	344	206	0.39
SVR	0.69	311	162	0.38	0.64	371.8	201.3	0.42
GPR	0.8	253	144	0.31	0.68	353.6	196	0.39
ANN	0.75	285.8	162.5	0.34	0.75	312	174.3	0.35

- Le modèle Elgibaly & N. \_HMD a donné des résultats acceptables, l'erreur relative absolue est de l'ordre de 0.39 lors du test.
- le modèle gaussian process GPR a montré de bonnes performances lors du training, mais les erreurs ont relativement augmenté lors du test.
- Le modèle ANN a montré les meilleures performances par rapport aux autres modèles dans la phase de test.

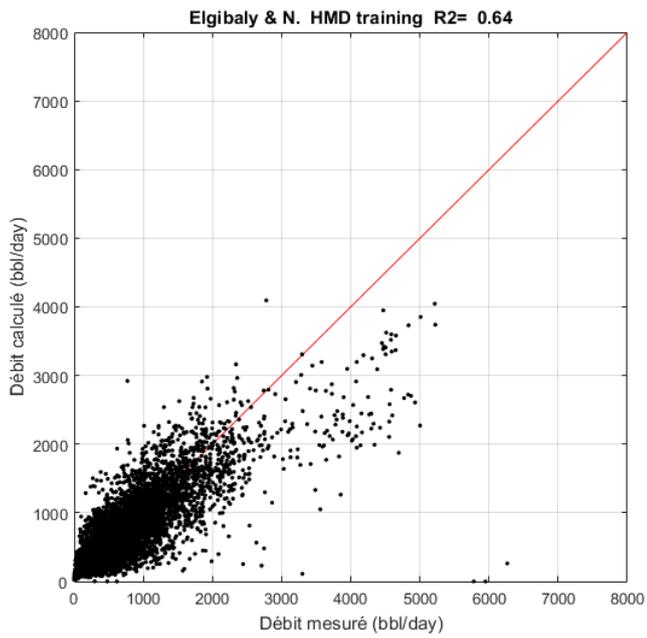


Figure 4.51 : cross plot de training du modèle de Elgibaly & Nashawi\_HMD

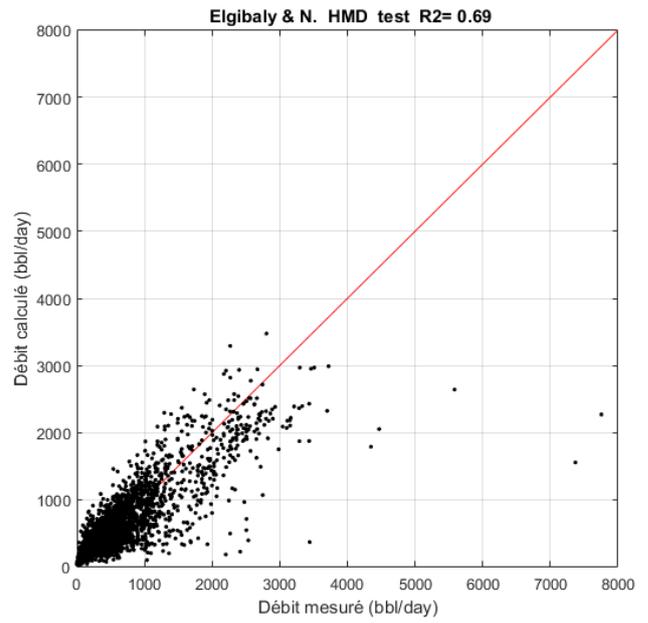


Figure 4.52: cross plot de test du modèle de Elgibaly & Nashawi .\_HMD

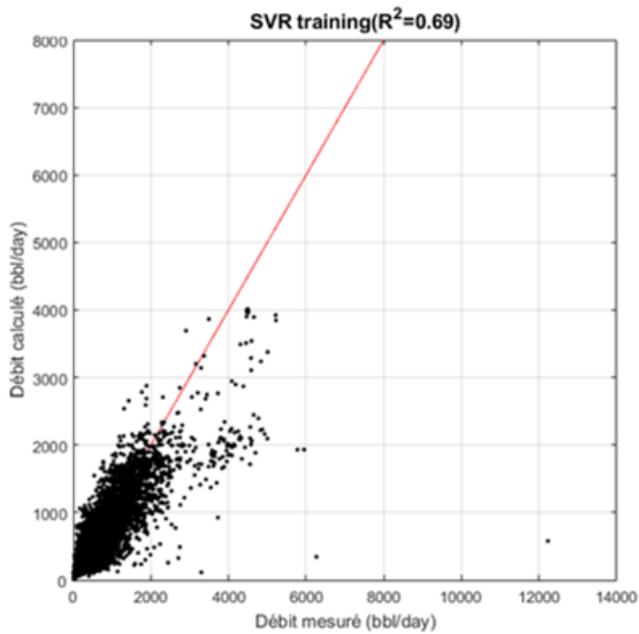


Figure 4.53 : cross plot de training du modèle SVR

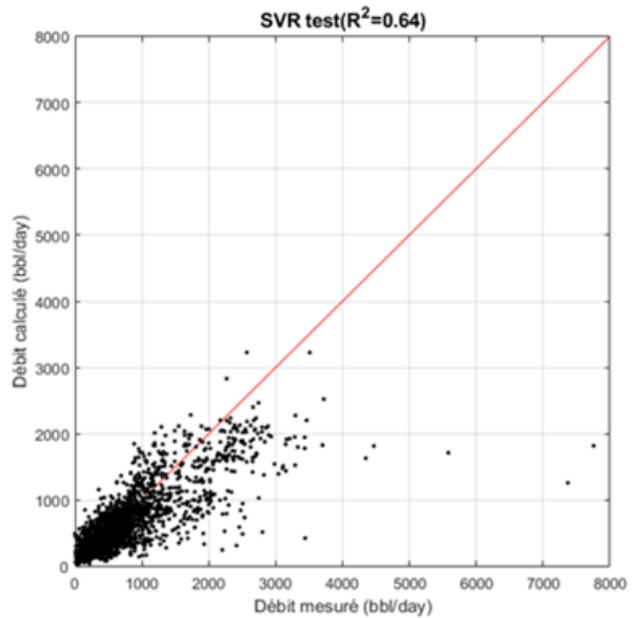


Figure 4.54 : cross plot de test du modèle SVR



Figure 4.55 : cross plot du training du modèle GPR

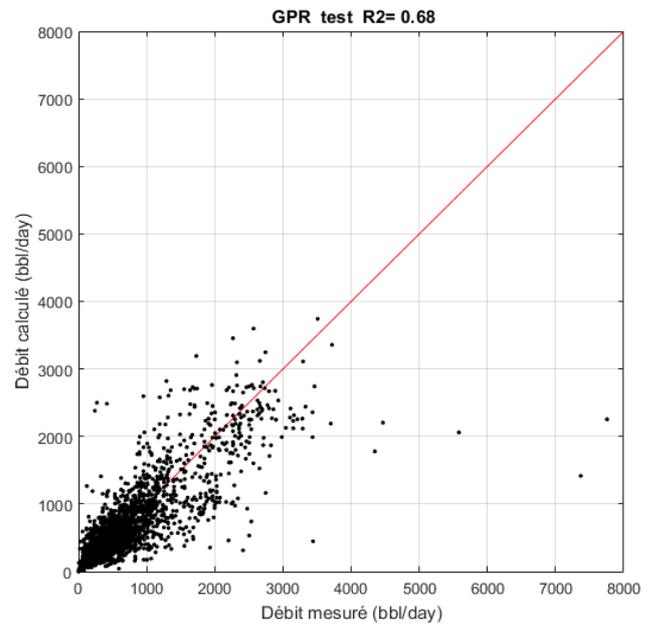


Figure 4.56: cross plot de test du modèle GPR



Figure 4. 57 : cross plot de training du modèle ANN

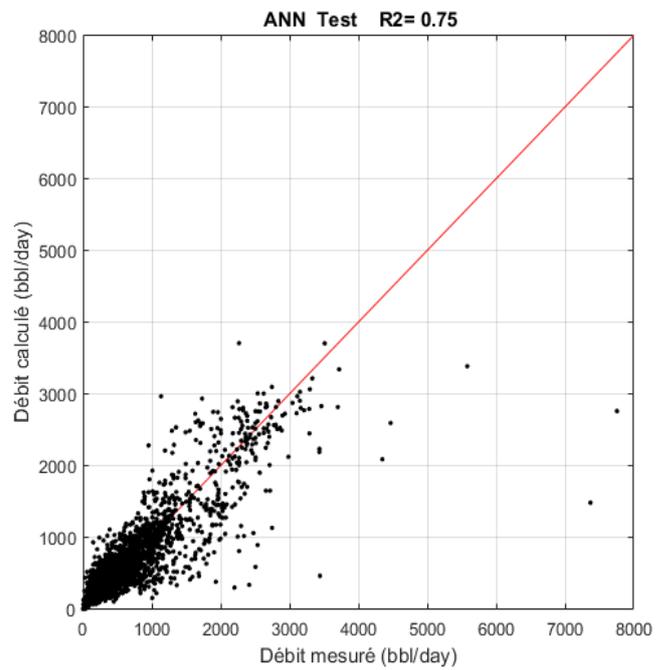


Figure 4.58 : cross plot de test du modèle ANN

### ❖ Modèle de l'extension de Gilbert

Son expression (équation 2.41) est une fonction du diamètre de la duse, pression amont, chute de pression à travers la duse et du GOR. Le modèle adapté au champ de Hassi Messaoud est noté Extension Gilbert \_HMD. Les nouveaux coefficients ainsi que le calcul des différentes métriques d'évaluation des modèles sont reportées dans les tableaux (4.16) et (4.17) respectivement.

**Tableau 4.16** : les nouveaux coefficients de l'équation de l'extension de Gilbert adaptée au champ de HMD

	a	b	c	e	f
Extension Gilbert _HMD	2.128	0.502	0.202	1.688	0.386

**Tableau 4.17** : statistiques de training et de test du modèle extension de Gilbert adapté et des modèles intelligents

Modèle	Métriques training				Métriques test			
	R <sup>2</sup>	RMSE	MAE	RAE	R <sup>2</sup>	RMSE	MAE	RAE
Extension Gilbert _HMD	0.69	313.89	181.43	0.38	0.74	320	192.64	0.36
SVR	0.70	281.10	142.36	0.34	0.65	368	198.63	0.42
GPR	0.90	174	102	0.21	0.73	325	189	0.36
ANN	0.88	199.3	124.7	0.23	0.84	247.3	133.8	0.27

- Le modèle Extension Gilbert \_HMD a montré de bonnes performances, l'erreur relative absolue est de l'ordre de 0.36 lors du test, le coefficient de détermination quant à lui est de l'ordre de 0.74, ce qui veut dire que ce modèle arrive à bien décrire le débit de production d'huile à Hassi Messaoud pour le régime d'écoulement sous critique.
- Les performances se sont un peu améliorées lors du training avec l'utilisation de support vector regression SVR. Pour la phase de test, il paraît que le modèle extension Gilbert \_HMD est relativement meilleur que le SVR.
- Le modèle gaussian process GPR a montré les meilleures performances pendant la phase de training lorsqu'il est comparé aux autres modèles ; à savoir extension Gilbert \_HMD, SVR et ANN. Ses performances ont un peu diminuées lors du test, les résultats obtenu pour cette phase sont presque similaires à Extension Gilbert \_HMD.
- Le modèle ANN présente de très bonnes performances lors du training ainsi que les meilleures performances lors du test comparé à tous les autres modèles développés pour cette forme.

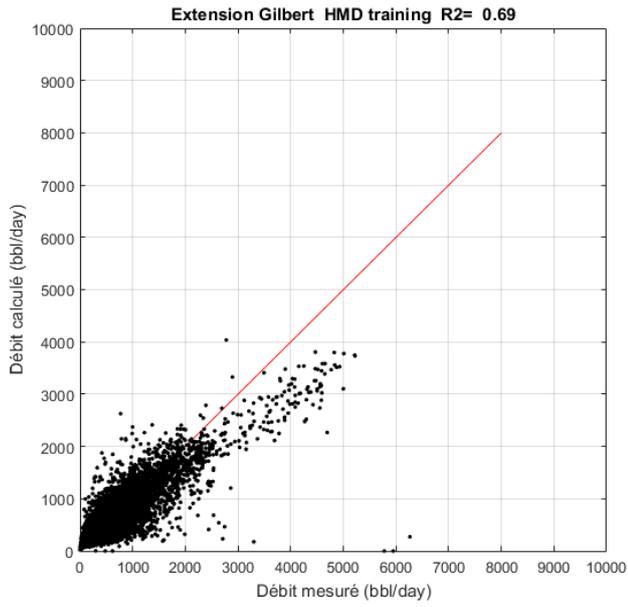


Figure 4. 59 : cross plot de training du modèle extension\_HMD

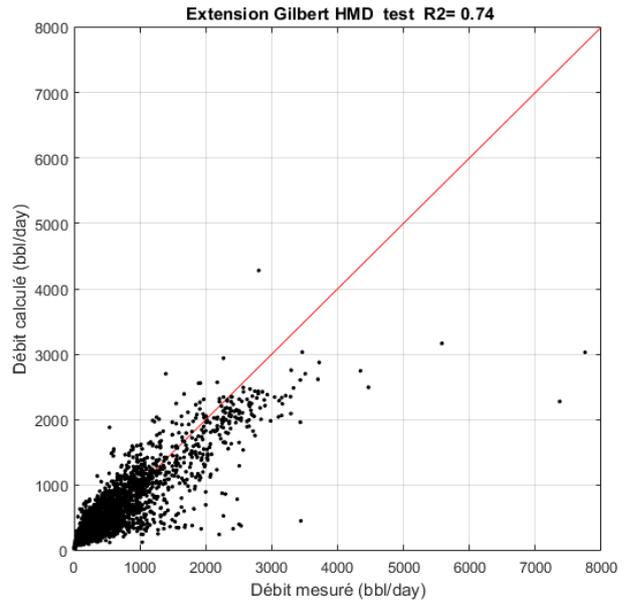


Figure 4.60 : cross plot de test du modèle extension Gilbert\_HMD

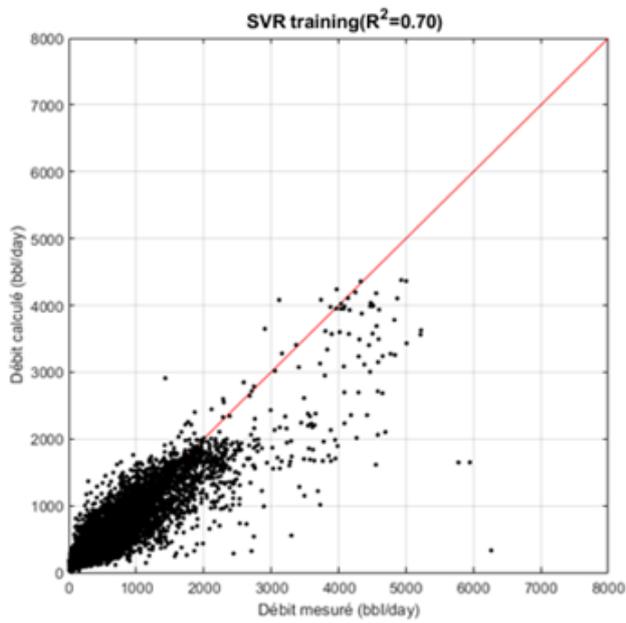


Figure 4.61 : cross plot de training du modèle SVR

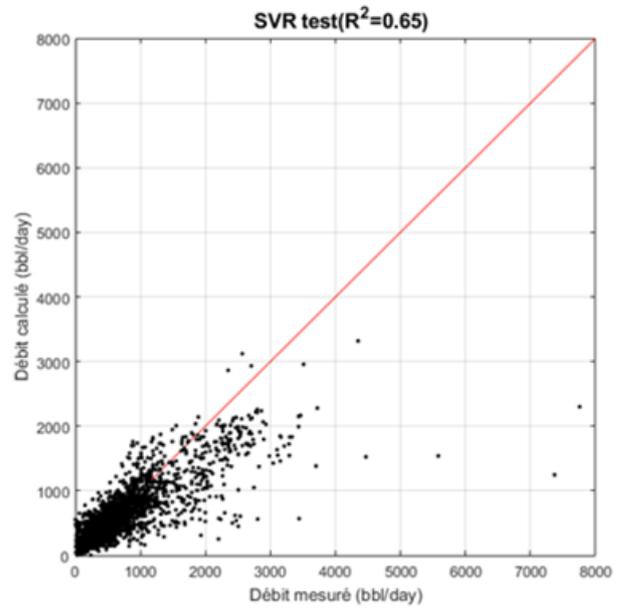


Figure 4. 62 : cross plot de test du modèle SVR

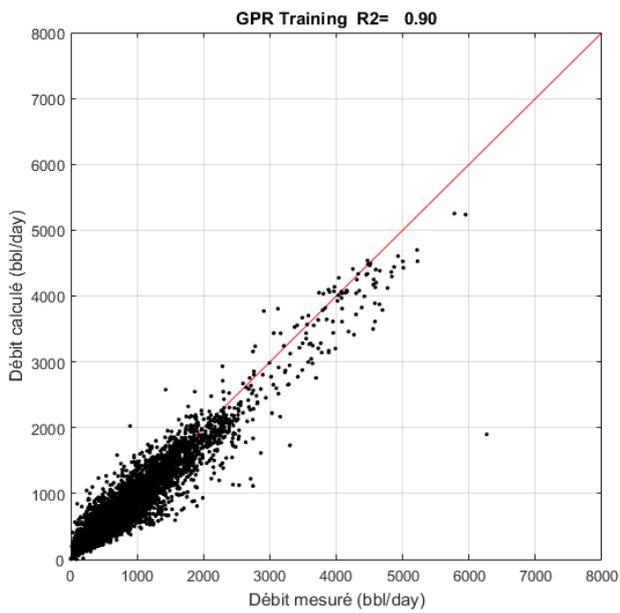


Figure 4.63 : cross plot de training du modèle GPR

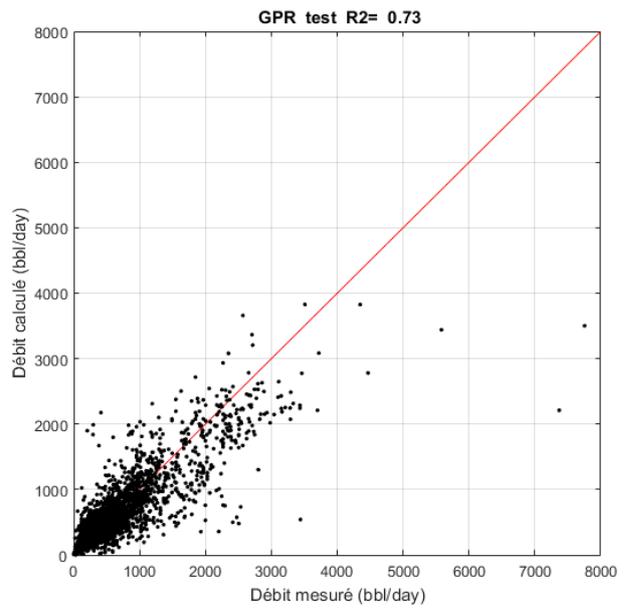


Figure 4.64 : cross plot de test du modèle GPR

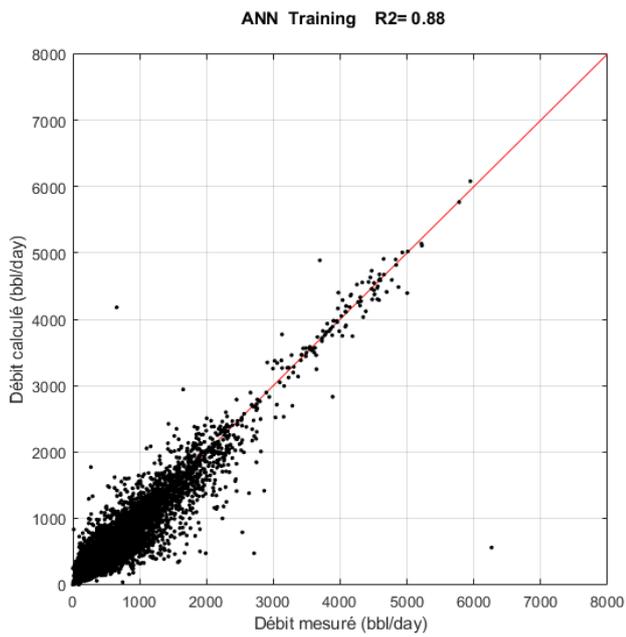


Figure 4.65 : cross plot de training du modèle ANN

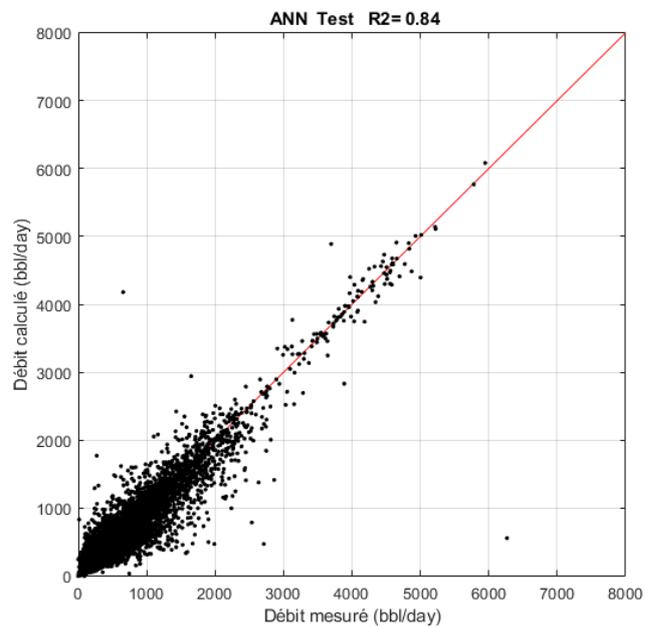


Figure 4.66 : cross plot de test du modèle ANN

#### ❖ Modèles des extensions utilisées à HMD

Les extensions sont exprimées en fonction du débit décrit par l'équation de Gilbert, multiplié fois une fonction en terme du rapport entre la pression en amont et en aval de la duse et l'inverse de ce rapport (Equation 2.42 et 2.43)

Les tableaux (4.18) et (4.19), représentent les nouveaux coefficients et les différentes métriques d'évaluation des deux extensions ajustées au cas de Hassi Messaoud.

**Tableau 4.18:** les nouveaux coefficients des extensions utilisées à HMD

	a	b	c	e
Extension 1_HMD	0.8298	1.7318	0.4288	1
Extension 2_HMD	0.5419	1.7118	0.4233	0.5589

**Tableau 4.19 :** statistiques de training et de test des extensions utilisées à HMD adaptées et des modèles intelligents

Modèle	Métriques training				Métriques test			
	R <sup>2</sup>	RMSE	MAE	RAE	R <sup>2</sup>	RMSE	MAE	RAE
Extension 1_HMD	0.66	332.75	187.68	0.40	0.71	336.97	198.62	0.38
Extension 2_HMD	0.67	327.61	187.39	0.4	0.72	328.26	197.13	0.37
SVR	0.73	291	146	0.35	0.63	378.9	206	0.42
GPR	0.90	171.5	102.6	0.2	0.72	330	192.8	0.37
ANN	0.83	239	124	0.28	0.88	215.9	130.1	0.24

- Les modèles Extension 1\_HMD et Extension 2\_HMD montrent de bonnes performances. Le modèle Extension 2\_HMD est relativement meilleur que le modèle Extension 1\_HMD.
- Le modèle gaussian process GPR présente le minimum d'erreurs pendant le training, puis les erreurs ont relativement augmenté pour la phase de test.
- Le modèle ANN présente de très bonnes performances, surtout pour la phase de test, où il a gardé sa capacité de généralisation sur de nouvelles données en enregistrant le minimum d'erreurs parmi tous les autres modèles.

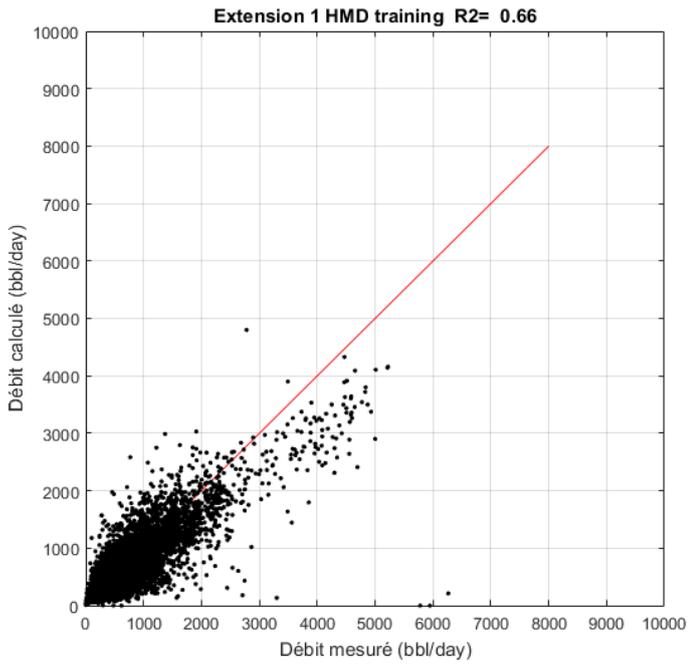


Figure 4.67 : cross plot de training du modèle Extension 1\_HMD

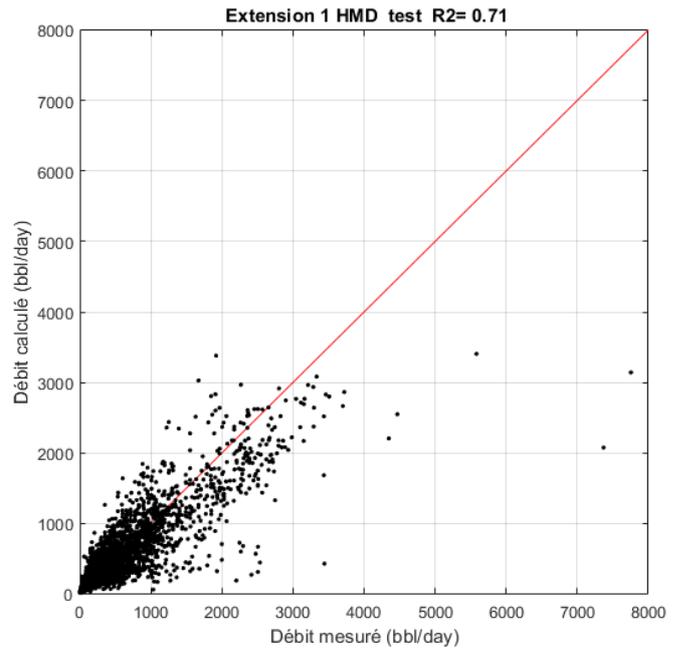


Figure 4.68 : cross plot de test du modèle Extension 1\_HMD

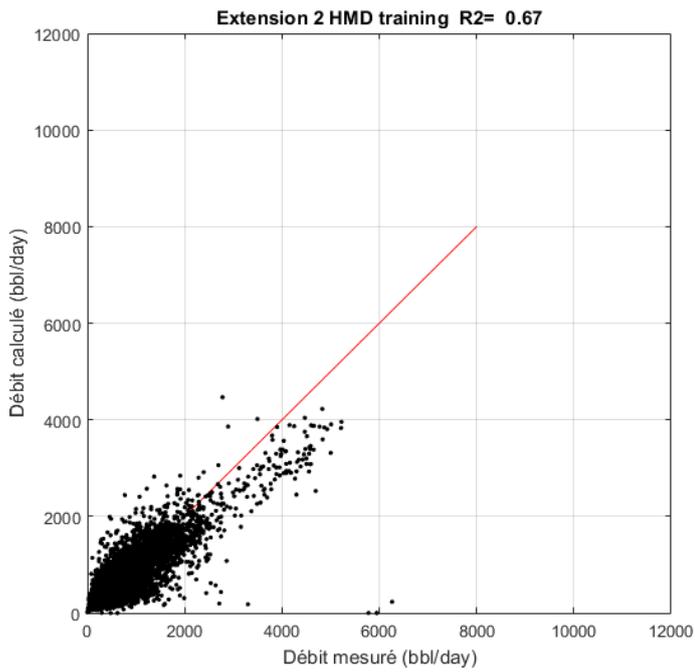


Figure 4.69 : cross plot du training de modèle Extension 2\_HMD

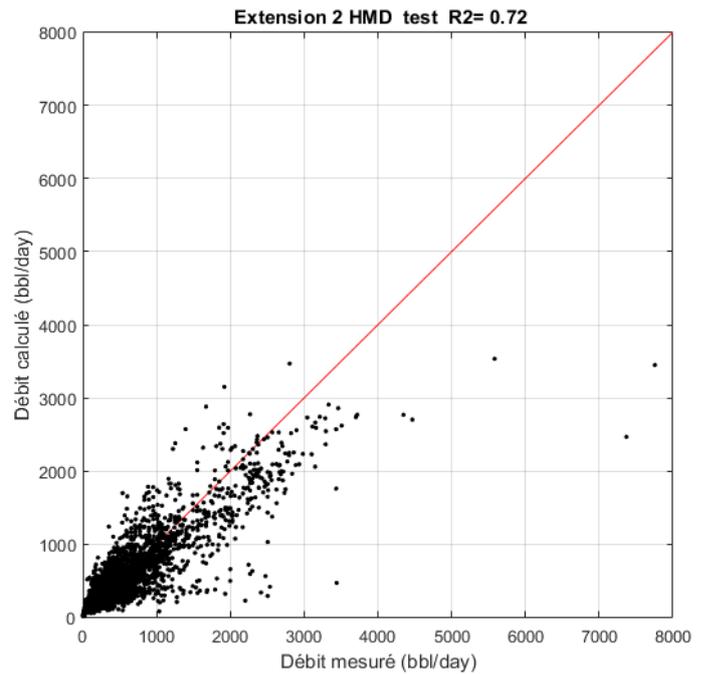


Figure 4.70 : cross plot de test du modèle Extension 2\_HMD

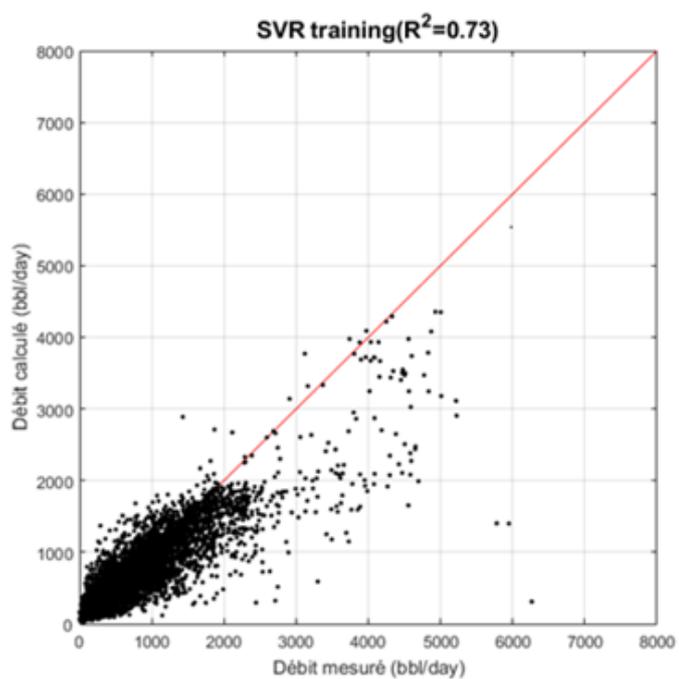


Figure 4.71 : cross plot de training du modèle SVR

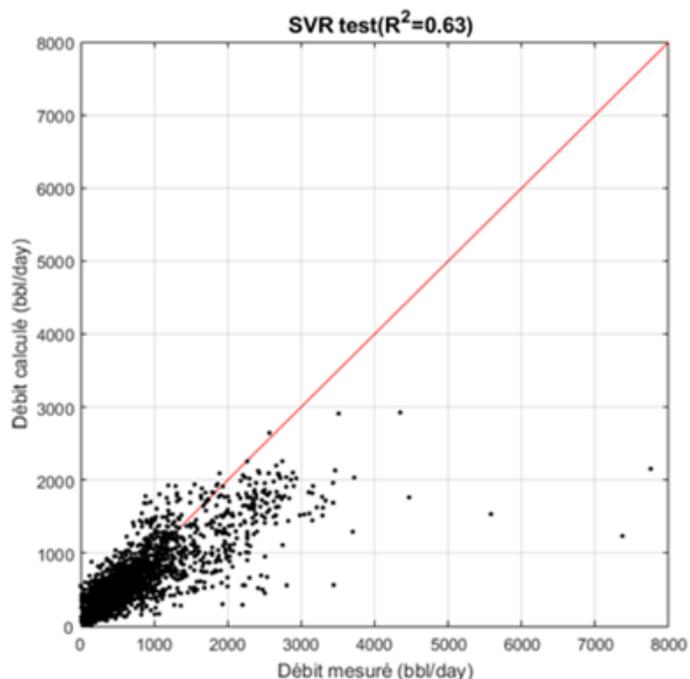


Figure 4.72 : cross plot de test du modèle SVR

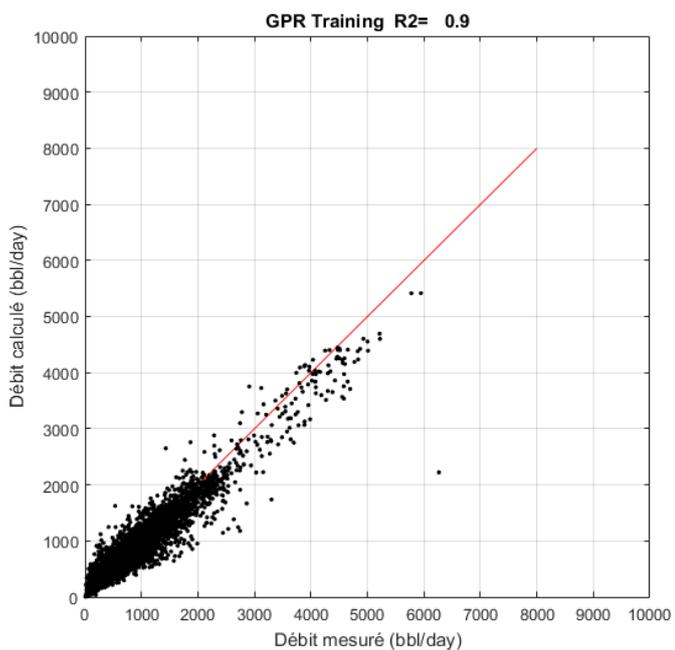


Figure 4.73 : cross plot de training du modèle GPR

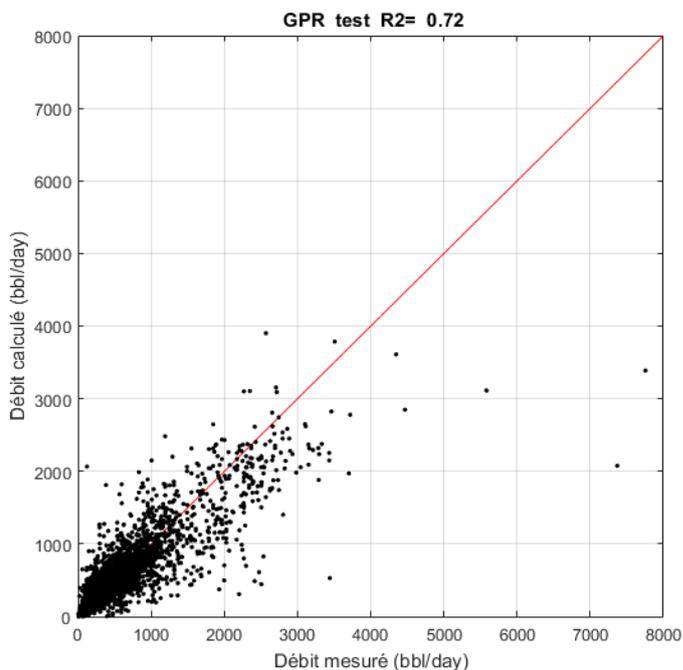


Figure 4.74 : cross plot de test du modèle GPR



Figure 4.75 : cross plot de training du modèle ANN

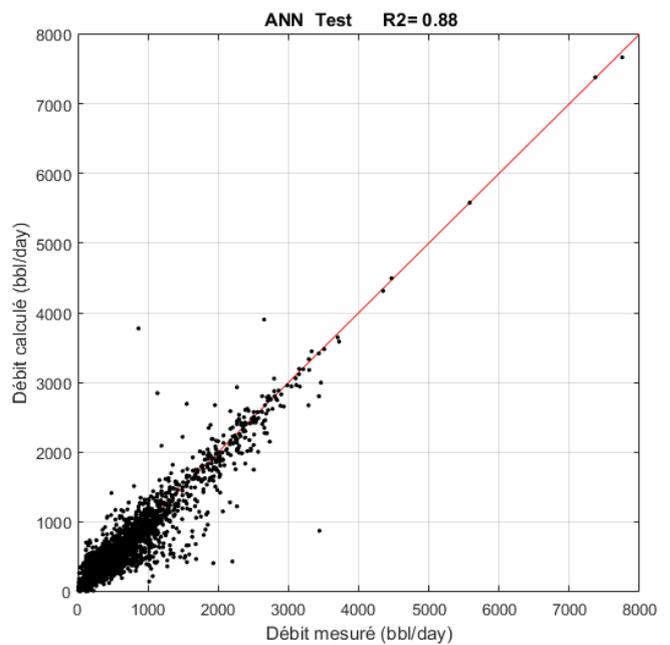


Figure 4.76 : cross plot de test du modèle ANN

### 4.3.3 Conclusion 2

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

- Pour le régime critique ; et en comparant les deux modèles théoriques adaptés, il paraît que le modèle de Gilbert ; et plus spécifiquement le modèle noté Gilbert\_HMD 2 arrive à donner une bonne estimation du débit de production d'huile au champ de Hassi Messaoud avec une erreur relative absolue de l'ordre de 0.35.
- Il a été retrouvé que la corrélation de Gilbert adaptée au régime non critique ; notée Gilbert\_NC\_HMD dans la présente étude, donne une bonne estimation du débit de production ; bien que la corrélation de Gilbert soit initialement développée pour l'étude du régime critique. L'erreur relative absolue pour ce modèle est de l'ordre de 0.37 et le coefficient de détermination quant à lui est de l'ordre de 0.72.
- De façon générale, tous les modèles théoriques adaptés pour le régime non critique ont montré de bonnes performances, plus particulièrement le modèle Extension Gilbert\_HMD a enregistré relativement le minimum d'erreurs avec une erreur relative absolue de l'ordre de 0.36 et un coefficient de détermination de l'ordre 0.74.
- Tous les modèles ANN développés dans cette partie ont montré de très bonnes performances de prédiction avec le minimum d'erreurs lorsqu'ils sont comparés aux autres types de modèles élaborés, que ce soit les modèles théorique ou bien les autres modèles intelligents (SVR et GPR).

## 4.5 Adaptation suivant les intervalles de GOR

Afin de tester l'efficacité des modèles théoriques adaptés précédemment et aussi améliorer leurs performances, une subdivision en intervalles de GOR a été envisagée.

Pour chaque régime d'écoulement (critique et sous-critique), le meilleur modèle théorique adapté a été sélectionné, ainsi le modèle de Gilbert adapté ; noté « Gilbert\_HMD 2 » a été utilisé pour l'étude du régime critique, et le modèle de l'extension de Gilbert noté « Extension Gilbert\_HMD » a été choisi pour l'étude du régime non critique.

Pour chaque intervalle de GOR, le modèle adapté correspondant est dans un premier temps testé, puis réadapté sur chaque intervalle afin d'améliorer ses performances.

Les nouveaux modèles seront ensuite comparés aux modèles ANN initialement développés pour chaque régime.

Les différents cross plot obtenus dans cette partie sont illustrés dans l'annexe B.

### 4.5.1 Régime critique

#### 4.5.1.1 Test sur différents intervalles de GOR

Les résultats relatifs au test du modèle Gilbert\_HMD 2 sont représentés dans le tableau (4.20).

Tableau 4. 20 : résultats de test du modèle Gilbert\_HMD 2 sur différents intervalles de GOR

Métriques de test du modèle Gilbert HMD 2					
Intervalle de GOR ( $\text{sm}^3/\text{sm}^3$ )	Nombre de points	$R^2$	RMSE	MAE	RAE
<b>GOR <math>\leq</math> 100</b>	1703	0.55	527.32	362.99	0.44
<b>100 &lt; GOR <math>\leq</math> 200</b>	9475	0.70	379.46	249.5	0.29
<b>200 &lt; GOR <math>\leq</math> 300</b>	4790	0.62	443.84	284.31	0.33
<b>300 &lt; GOR <math>\leq</math> 400</b>	3442	0.646	489.63	293.13	0.34
<b>400 &lt; GOR <math>\leq</math> 800</b>	9764	0.71	278.29	173.29	0.31
<b>800 &lt; GOR <math>\leq</math> 1300</b>	5963	0.729	228.47	166.19	0.31
<b>1300 &lt; GOR <math>\leq</math> 4000</b>	8899	0.69	242.22	186.67	0.36
<b>4000 &lt; GOR</b>	1981	0.11	246.82	207.18	0.59

- D'après les résultats de test du modèle Gilbert\_HMD 2 sur différents intervalles de GOR, on a remarqué que ce dernier s'applique bien lorsque le GOR est compris entre 100 et 4000  $\text{sm}^3/\text{sm}^3$  avec une erreur relative absolue qui varie entre 0.29 et 0.36.
- Lorsque les valeurs du GOR sont inférieurs à 100  $\text{sm}^3/\text{sm}^3$  ou strictement supérieurs à 4000  $\text{sm}^3/\text{sm}^3$ , les erreurs commencent à devenir relativement importantes, comme c'est montré dans le tableau (4.20) où on voit que le coefficient de détermination et l'erreur relative absolue sont de l'ordre de 0.55 et 0.44 respectivement pour un GOR inférieur à 100  $\text{sm}^3/\text{sm}^3$  et de l'ordre de 0.11 et 0.59 lorsque le GOR est supérieur à 4000  $\text{sm}^3/\text{sm}^3$ .

#### 4.5.1.2 Adaptation sur différents intervalles de GOR

Puisque le modèle de Gilbert\_HMD 2 s'adapte mal pour des valeurs de GOR inférieurs à 100  $\text{sm}^3/\text{sm}^3$  et supérieur à 4000  $\text{sm}^3/\text{sm}^3$ , la corrélation de Gilbert (équation 2.22) est réadaptée sur ces deux intervalles afin de trouver des modèle décrivant mieux le débit de production. La corrélation de Gilbert est également réadaptée sur les intervalles de GOR restants afin d'améliorer la capacité de prédiction des modèles dérivant ces intervalles. Les tableaux (4.21) et (4.22) représentent respectivement les nouveaux coefficients obtenus ainsi que les résultats de training et de test des nouveaux modèles.

Tableau 4.21: nouveaux coefficients du modèle de Gilbert adapté sur chaque intervalle de GOR (régime critique)

Nouveaux coefficients pour l'équation de Gilbert adapté pour chaque GOR					
N°	Intervalles de GOR ( $\text{sm}^3/\text{sm}^3$ )	a	b	c	e
1	<b>GOR <math>\leq</math> 100</b>	1.3689	0.8227	0.1474	0.9088
2	<b>100 &lt; GOR <math>\leq</math> 200</b>	0.3814	1.580411	0.05209	0.7760
3	<b>200 &lt; GOR <math>\leq</math> 300</b>	32.1445	1.714537	0.7647	0.7966
4	<b>300 &lt; GOR <math>\leq</math> 400</b>	27.1446	1.88746	0.7705	0.7732
5	<b>400 &lt; GOR <math>\leq</math> 800</b>	5.5648	1.8834	0.6566	0.8376
6	<b>800 &lt; GOR <math>\leq</math> 1300</b>	8.1924	1.8402	0.7590	0.8984
7	<b>1300 &lt; GOR <math>\leq</math> 4000</b>	19.2229	1.7876	0.8519	0.9152
8	<b>4000 &lt; GOR</b>	71.3204	1.8971	0.9990	0.9015

Tableau 4. 22 : résultats de training et de test des modèles adaptés pour chaque intervalle de GOR (régime critique)

Intervalle de GOR ( $\text{sm}^3/\text{sm}^3$ )	Métriques de training du modèle				Métriques de test du modèle			
	R <sup>2</sup>	RMSE	MAE	RAE	R <sup>2</sup>	RMSE	MAE	RAE
<b>GOR <math>\leq</math> 100</b>	0.70	443.59	259.58	0.36	0.71	422.98	300.2	0.35
<b>100 &lt; GOR <math>\leq</math> 200</b>	0.69	398.77	245.78	0.30	0.70	377.78	243.7	0.29
<b>200 &lt; GOR <math>\leq</math> 300</b>	0.63	408.36	266.19	0.31	0.65	422.05	270.17	0.32
<b>300 &lt; GOR <math>\leq</math> 400</b>	0.63	353.49	227.57	0.30	0.707	445.47	272.91	0.31
<b>400 &lt; GOR <math>\leq</math> 800</b>	0.70	315.9	193.3	0.305	0.728	272.7	161.8	0.305
<b>800 &lt; GOR <math>\leq</math> 1300</b>	0.78	233.49	136.98	0.28	0.79	202.33	124.5	0.28
<b>1300 &lt; GOR <math>\leq</math> 4000</b>	0.81	174.46	101.31	0.27	0.76	213.35	121.08	0.32
<b>4000 &lt; GOR</b>	0.75	109.37	51.68	0.348	0.72	138.32	77.87	0.33

- Pour les intervalles  $\text{GOR} \leq 100$  et  $4000 < \text{GOR}$ , et en comparant les résultats de test du modèle Gilbert\_HMD 2 initialement développé et le test des nouveaux modèles décrivant chaque intervalle (tableaux 4.20 et 4.22), on remarque que les nouveaux modèles ont nettement amélioré la capacité de prédiction, le coefficient de détermination a augmenté de 0.55 à 0.71 et de 0.11 à 0.72 respectivement et les autres erreurs ont clairement diminuées.

- Pour les autres intervalles, bien que le modèle Gilbert\_HMD 2 donne une bonne estimation du débit de production, les nouveaux modèles adaptés sur chaque intervalle ont permis d'améliorer encore plus la capacité de prédiction, les erreurs relatives absolues sont comprises entre 0.28 et 0.32.
- En analysant les valeurs des différents coefficients a, b, c et e, on remarque que le coefficient a n'a pas une tendance claire, il prend différentes valeurs selon l'intervalle de GOR étudié. Le coefficient b est compris principalement entre 1.58 et 1.89 sur tous les intervalles sauf pour le premier intervalle où il prend une valeur de 0.82. le coefficient c prend des valeurs comprises entre 0.65 et 0.9 pour tous les intervalles sauf pour les deux premiers. La tendance de e est relativement stable, ce coefficient prend des valeurs entre 0.77 et 0.91. la différence de tendance entre les coefficients b et c peut être expliqué par les nombres de point dans chaque intervalle.

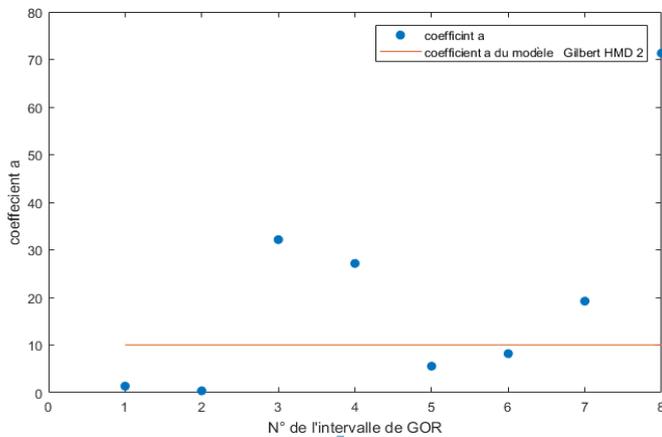


Figure 4.77 : Evolution du coefficient a en fonction de l'intervalle de GOR (RC)

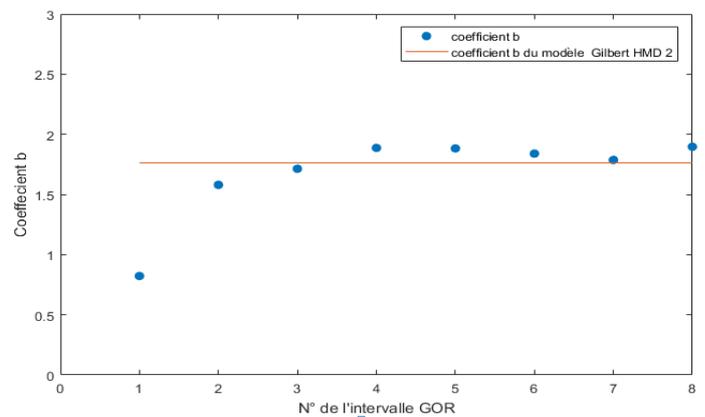


Figure 4.78 : Evolution du coefficient b en fonction de l'intervalle de GOR (RC)

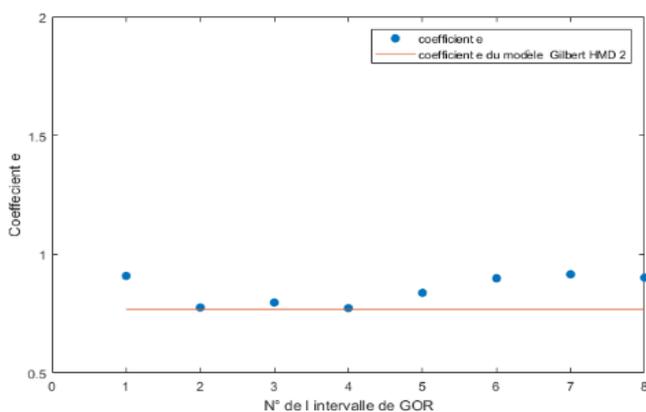


Figure 4.79 : Evolution du coefficient e en fonction de l'intervalle de GOR (RC)

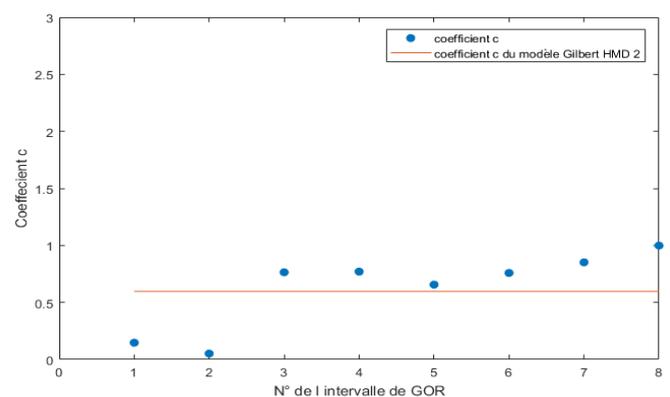


Figure 4.80 : Evolution du coefficient c en fonction de l'intervalle de GOR (RC)

### 4.5.1.3 Comparaison entre le modèle ANN et les modèles adaptés sur différents intervalles de GOR :

Parmi les modèles intelligents développées dans la première section, c'est le modèle ANN qui a montré les meilleures performances pendant le test, c'est pour cela, il a été sélectionné pour être comparé aux nouveaux modèles adaptés sur chaque intervalle de GOR.

Tableau 4.23 : résultats de test du modèle de Gilbert adapté sur chaque intervalle GOR et le test du modèle ANN

Intervalle de GOR	Métriques de test du modèle adapté pour chaque GOR				Métriques de test du modèle ANN initialement développé			
	R <sup>2</sup>	RMSE	MAE	RAE	R <sup>2</sup>	RMSE	MAE	RAE
<b>GOR ≤ 100</b>	0.71	422.98	300.2	0.35	0.889	263.16	179.86	0.22
<b>100 &lt; GOR ≤ 200</b>	0.70	377.78	243.7	0.29	0.79	320.82	216.14	0.25
<b>200 &lt; GOR ≤ 300</b>	0.65	422.05	270.17	0.32	0.72	379.82	245.21	0.285
<b>300 &lt; GOR ≤ 400</b>	0.707	445.47	272.91	0.31	0.84	330.66	211.97	0.23
<b>400 &lt; GOR ≤ 800</b>	0.728	272.7	161.8	0.305	0.78	243.26	149.84	0.27
<b>800 &lt; GOR ≤ 1300</b>	0.79	202.33	124.5	0.28	0.80	194.41	120.24	0.267
<b>1300 &lt; GOR ≤ 4000</b>	0.76	213.35	121.08	0.32	0.82	184.85	107.07	0.278
<b>4000 &lt; GOR</b>	0.72	138.32	77.87	0.33	0.77	124.6	73.96	0.29

- D'après la comparaison entre les modèles adaptés sur chaque intervalle de GOR et le modèle ANN initialement développé, il est clair que les résultats du modèle ANN sont plus proches des valeurs réelles, le coefficient de détermination est compris entre 0.72 et 0.889, et les erreurs relatives absolue ont relativement diminuées (entre 0.22 et 0.29). Ces résultats semblent logiques, puisque les modèles ANN construisent des relations très complexes entre les variables ce qui conduit à des performances relativement meilleures.
- Bien que les performances du modèle ANN soient meilleures que les modèles adaptés pour chaque intervalle de GOR, ces derniers restent très pratiques et donnent une bonne estimation du débit si l'on souhaite travailler plutôt avec des modèles théoriques (l'erreur relative absolue est comprise entre 0.28 et 0.35).

## 4.5.2 Régime non critique

### 4.5.2.1 Test sur différents intervalles de GOR

Les résultats relatifs au test du modèle « Extension Gilbert \_HMD » sont représentés dans le tableau (4.24).

Tableau 4.24 : résultats de test du modèle Extension Gilbert \_HMD sur différents intervalles de GOR régime non critique

Métriques de test du modèle l'extension Gilbert HMD					
Intervalle de GOR ( $\text{sm}^3/\text{sm}^3$ )	Nombre de points	R <sup>2</sup>	RMSE	MAE	RAE
<b>GOR ≤ 200</b>	4142	0.75	303.08	198.42	0.29
<b>200 &lt; GOR ≤ 400</b>	2178	0.706	461.88	266.23	0.3
<b>400 &lt; GOR ≤ 800</b>	2879	0.78	253.08	168.61	0.29
<b>800 &lt; GOR ≤ 1300</b>	2182	0.823	179.18	134.95	0.27
<b>1300 &lt; GOR</b>	4221	0.41	207.55	159.71	0.53

Les résultats représentés dans le tableau (4.24) montrent que le modèle « Extension Gilbert \_HMD » donne une bonne estimation du débit sur tous les intervalles (l'erreur relative absolue est comprise entre 0.27 et 0.3) sauf pour le dernier où  $1300 < \text{GOR}$  (l'erreur relative absolue est de l'ordre de 0.53).

### 4.5.2.2 Adaptation de l'extension de Gilbert sur différents intervalles de GOR

Afin de trouver un bon modèle pour décrire le débit de production lorsque le  $1300 < \text{GOR}$  et améliorer la capacité de prédiction sur les autres intervalles, le modèle de l'extension de Gilbert (équation 2.41) sera cette fois adapté pour chaque intervalle de GOR au régime non critique. Les résultats de training et de test des différents modèles sont reportés dans le tableau (4.26) et les nouveaux coefficients dans le tableau (4.25).

Tableau 4.25 : nouveaux coefficients de l'extension de Gilbert adaptée sur chaque intervalle de GOR (régime non critique)

Nouveaux coefficients de l'extension de Gilbert adapté pour chaque GOR						
N°	Intervalles de GOR ( $\text{sm}^3/\text{sm}^3$ )	a	b	c	e	f
1	<b>GOR ≤ 200</b>	0.0626	0.7220	0.2352	1.6292	0.0200
2	<b>200 &lt; GOR ≤ 400</b>	5.1080	0.7386	0.2010	1.5713	0.6180
3	<b>400 &lt; GOR ≤ 800</b>	3.5603	0.6512	0.2882	1.6024	0.5684
4	<b>800 &lt; GOR ≤ 1300</b>	3.6636	0.6201	0.2713	1.6038	0.5483
5	<b>1300 &lt; GOR</b>	20.0185	0.6750	0.2639	1.6087	0.7804

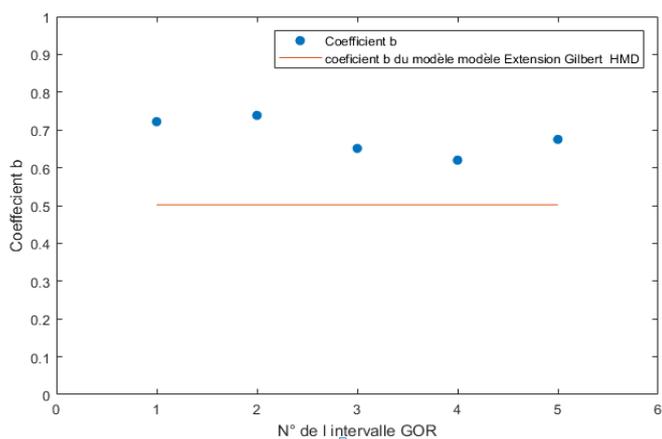


Figure 4.81 : Evolution du coefficient b en fonction de l'intervalle de GOR (régime non critique)

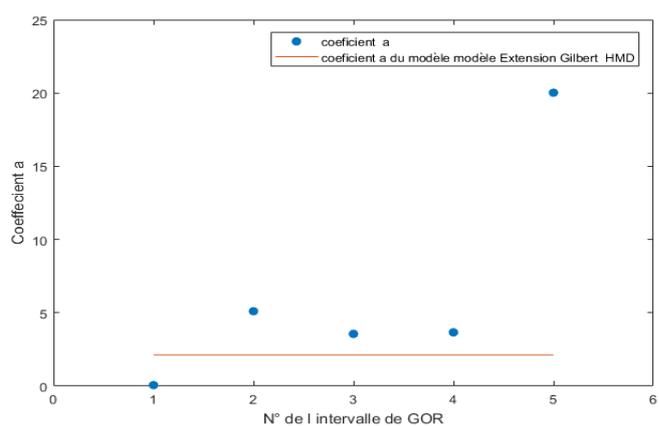


Figure 4.82 : Evolution du coefficient a en fonction de l'intervalle de GOR (régime non critique)

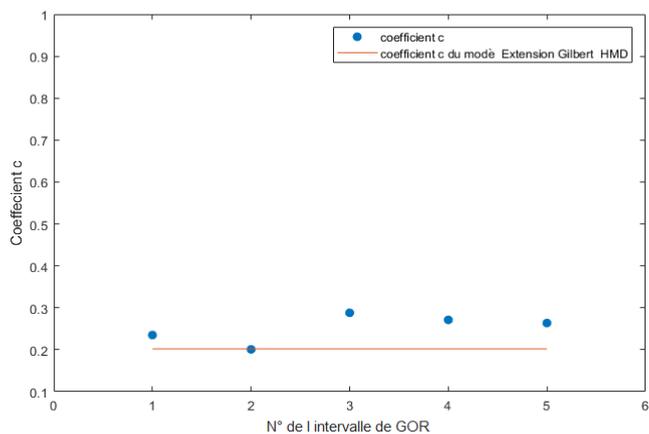


Figure 4.83 : Evolution du coefficient c en fonction de l'intervalle de GOR (régime non critique)

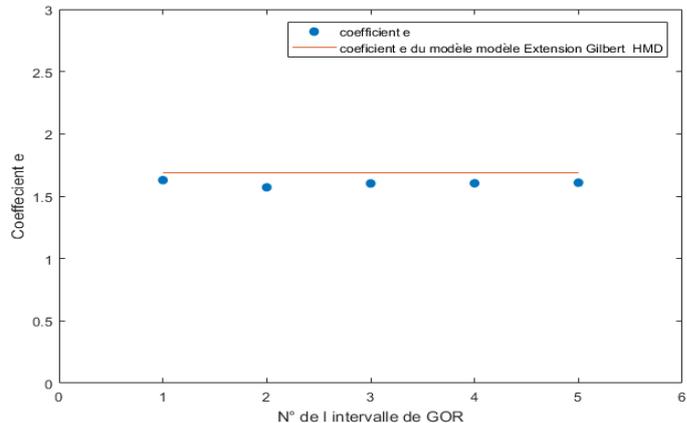


Figure 4.84 : Evolution du coefficient e en fonction de l'intervalle de GOR (régime non critique)

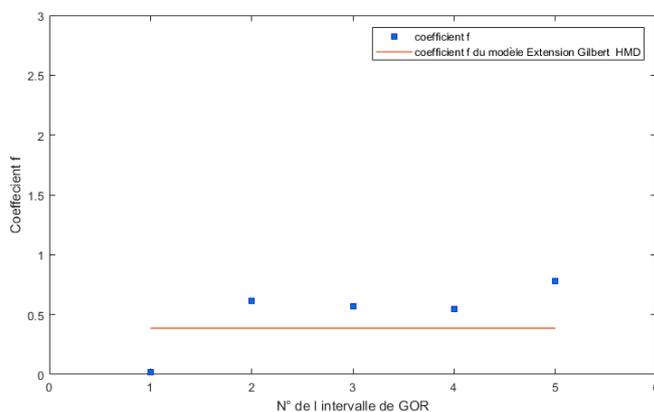


Figure 4.85 : Evolution du coefficient f en fonction de l'intervalle de GOR (régime non critique)

Tableau 4.26 : résultats de training et de test des modèles adaptés pour chaque intervalle de GOR (régime non critique)

Intervalle de GOR (sm <sup>3</sup> /sm <sup>3</sup> )	Métriques de training du modèle				Métriques de test du modèle			
	R <sup>2</sup>	RMSE	MAE	RAE	R <sup>2</sup>	RMSE	MAE	RAE
<b>GOR ≤ 200</b>	0.70	387.23	184.55	0.37	0.816	262.42	172.16	0.28
<b>200 &lt; GOR ≤ 400</b>	0.77	327.41	214.60	0.29	0.80	377.85	228.07	0.295
<b>400 &lt; GOR ≤ 800</b>	0.78	239.05	152.16	0.29	0.805	239.47	156.97	0.28
<b>800 &lt; GOR ≤ 1300</b>	0.826	179.53	121.82	0.27	0.85	164.49	112.66	0.25
<b>1300 &lt; GOR</b>	0.817	124.92	77.21	0.29	0.80	120.71	75.6	0.30

- Les nouveaux coefficients obtenus b, c et e ont une tendance relativement claire. Le coefficient b est compris entre 0.62 et 0.73 sur tous les intervalles, c est de l'ordre de 0.2, e est de l'ordre de 1.6. Le coefficient a n'a pas une tendance claire, et f prend une valeur de 0.5 en moyenne pour les intervalles 2,3 et 4.
- L'adaptation de l'extension de Gilbert a permis d'améliorer les performances de prédiction sur tous les intervalles, y compris pour le dernier intervalle (1300 < GOR) qui a montré des performances médiocres lors du premier test (tableau 4.24). Les coefficients de détermination lors de la phase de test des différents modèles adaptés sont de l'ordre de 0.8 ce qui montre que ces modèles sont plus proches de la réalité et permettent de donner une très bonne estimation du débit produit pour les puits fonctionnant en régime non critique.

#### 4.5.2.3 Comparaison entre le modèle ANN et les modèles adaptés sur différents intervalles de GOR

Le modèle ANN développé précédemment lors de l'étude du régime non critique est le meilleur modèle intelligent parmi les autres, pour ceci il a été comparé aux nouveaux modèles adaptés sur chaque intervalle de GOR au régime non critique. Les résultats de comparaison sont représentés dans le tableau (4.27)

D'après ce tableau, le test du modèle ANN développé pour le régime non critique a donné des résultats relativement meilleurs que ceux calculés par les modèles adaptés, le coefficient de détermination est compris entre 0.826 et 0.89, ce qui montre que le débit calculé par ce modèle est plus proche des valeurs mesurées.

Tableau 4. 27 : résultats de test du modèle de Gilbert adapté sur chaque GOR et le test de test du modèle ANN

Intervalle de GOR	Métriques de test du modèle adapté pour chaque GOR				Métriques de test du modèle ANN initialement développé			
	R <sup>2</sup>	RMSE	MAE	RE	R <sup>2</sup>	RMSE	MAE	RE
<b>GOR ≤ 200</b>	0.816	262.42	172.16	0.28	0.87	220.39	152.17	0.23
<b>200 &lt; GOR ≤ 400</b>	0.80	377.85	228.07	0.295	0.89	283.37	191.16	0.22
<b>400 &lt; GOR ≤ 800</b>	0.805	239.47	156.97	0.28	0.878	189.12	136.94	0.22
<b>800 &lt; GOR ≤ 1300</b>	0.85	164.49	112.66	0.25	0.89	138.31	96.32	0.21
<b>1300 &lt; GOR</b>	0.80	120.71	75.6	0.30	0.826	112.11	69.45	0.29

### 4.5.3 Conclusion 3

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

- L'adaptation de la corrélation de Gilbert pour le régime critique et de l'extension de Gilbert pour le régime non critique suivant différents intervalles de GOR a permis de développer des modèles plus représentatifs que les modèles théoriques initialement ajustés sur toutes les données ; où il a été remarqué que les erreurs ont considérablement diminuées.
- Les deux modèles ANN développés initialement pour l'étude du régime critique et sous-critique ont montré des performances relativement meilleures que les modèles théoriques adaptés sur chaque GOR, les résultats obtenus en utilisant les réseaux de neurones pour le calcul du débit semblent donc plus proches des valeurs réelles ce qui montre que cet outil est prometteur pour la prédiction du débit de production.
- Si l'on souhaite travailler avec des modèles théoriques, il est recommandé d'utiliser les modèles adaptés suivant chaque intervalle de GOR, puisque ces derniers ont été améliorés par rapport aux modèles initiaux et présentent donc un minimum d'erreurs.

## 4.6 Modélisation par puits via la programmation génétique

L'élaboration des modèles servant à décrire le débit de production dans les sections précédentes a pris en considération à chaque fois le régime d'écoulement du puits (critique et sous-critique). Or, si on ne connaît pas ce régime, il est relativement difficile de choisir quel modèle utiliser. Pour ceci, on va essayer d'appliquer la programmation génétique PG sur quelques puits afin de développer différents modèles pour chaque puits, le but dans un premier temps étant de développer un modèle sans prendre en compte le régime d'écoulement et dans un second développer des modèles plus représentatifs pour décrire la production de chaque puits individuellement.

La programmation génétique a été appliquée sur 4 puits différents. A chaque fois différents paramètres sont introduits dans les modèles, les résultats de training et de test de ces modèles sont représentés dans le tableau (4.28). Les cross plot ainsi que les équations obtenues sont montrés dans l'annexe C.

$P_u, d, GOR, P_d, \Delta P$  et  $r$ , représentent respectivement : la pression de tête, le diamètre de la duse, le rapport gaz/huile produit, la pression en aval de la duse, la chute de pression à travers la duse et le rapport de pression critique  $p_d/p_u$

Tableau 4. 28 : statistiques de training et test de GP pour chaque puits avec différentes variables

Puits	variables	Métriques de test du modèle PG				Métriques de test du modèle PG			
		R <sup>2</sup>	RMSE	MAE	RAE	R <sup>2</sup>	RMSE	MAE	RAE
1	$P_u, d, GOR$	0.89	112.79	82.5	0.15	0.783	162.45	118.3	0.20
	$P_u, d, GOR, P_d$	0.91	99.1	71.7	0.13	0.76	170.	119.99	0.21
	$P_u, d, GOR, P_d, \Delta P$	0.91	100.3	75.5	0.135	0.64	209.84	149.1	0.26
	$P_u, d, GOR, P_d, \Delta P, r$	0.90	103.3	73.79	0.13	0.74	176.6	119.4	0.22
2	$P_u, d, GOR$	0.89	450.02	334.45	0.18	0.80	432.58	361.37	0.19
	$P_u, d, GOR, P_d$	0.9	421.03	345.95	0.18	0.77	463.52	359.13	0.21
	$P_u, d, GOR, P_d, \Delta P$	0.85	506.19	409.66	0.21	0.76	475.68	368.74	0.21
	$P_u, d, GOR, P_d, \Delta P, r$	0.89	441.48	354.03	0.18	0.76	474.99	406.14	0.21
3	$P_u, d, GOR$	0.91	244.88	184.8	0.18	0.85	135.97	113.41	0.10
	$P_u, d, GOR, P_d$	0.91	242.74	182.55	0.18	0.58	227.33	200.76	0.17
	$P_u, d, GOR, P_d, \Delta P$	0.91	248.4	179.62	0.18	0.47	255.76	233.32	0.2
	$P_u, d, GOR, P_d, \Delta P, r$	0.92	227.98	174.19	0.17	0.72	184.21	152.13	0.14
4	$P_u, d, GOR$	0.72	114.42	87.22	0.22	0.7	64.56	55.18	0.13
	$P_u, d, GOR, P_d$	0.76	106.99	83.23	0.21	0.75	59.38	50.39	0.12
	$P_u, d, GOR, P_d, \Delta P$	0.76	106.77	80.34	0.21	0.75	59.29	49.05	0.12
	$P_u, d, GOR, P_d, \Delta P, r$	0.72	114.46	87.83	0.22	0.71	63.92	51.61	0.13

- Les modèles obtenus par la programmation génétique présentent des erreurs relatives absolue comprises entre 0.1 et 0.26 lors de la phase de test ce qui montre que l'étude de chaque puits individuellement peut donner des résultats très intéressants, même si cette approche n'est pas pratique puisque généralement on cherche des modèles plutôt généraux qu'on peut facilement intégrer dans d'autres programmes de calcul (simulateurs) à titre d'exemple.
- D'après l'analyse des erreurs calculées lors de la phase de test pour ces 4 puits, il paraît que les variables qui influent le plus sur les modèles développés sont en premier lieu  $P_u, d, GOR$  avec une erreur relative absolue moyenne de 0.15 et un coefficient de détermination moyen de 0.78.

## 4.7 Utilisation des modèles LSTM pour la prédiction du débit :

Dans cette section, une autre méthode data driven a été utilisée pour évaluer sa capacité de prédiction du débit d'huile à travers chaque puits. Ceci est réalisé grâce à une architecture spéciale de réseaux de neurones récurrents qui sont les modèles LSTM.

Cette approche est utilisée pour la modélisation des problèmes qui se présentent sous forme de séries temporelles ; puisque le débit est un paramètre qui évolue en fonction du temps, on va le considérer comme une série chronologique, un modèle time series sera ainsi implémenté en utilisant un algorithme LSTM (long short term memory).

Le puits **Well1289** a été choisi pour une étude de cas au champ de Hassi Messaoud. On dispose de l'historique du puits depuis 1970 jusqu'à 2019 avec 129 points de jaugeage.

L'implémentation réalisée pour le training du modèle contient 16 unités LSTM empilées, suivies par une couche dense, et une couche de sortie avec un seul nœud (une variable de sortie). Le débit est considéré comme la seule entrée pour ce modèle. 3 points jaugeage sont à chaque fois utilisés pour la prédiction du point suivant.

Les résultats de training et de test du modèle sont représentés dans le tableau (4.29). Les figures 4.87 et 4.88 montrent les cross plot du training et du test du modèle LSTM et la figure 4.89 montre le profil de production réel (en bleu) de ce puits ainsi que le débit prédit (en rouge) par le modèle LSTM.

Tableau 4. 29 : statistiques de training et de test du modèle LSTM

Modèle	Métriques training				Métriques test			
	R <sup>2</sup>	RMSE	MAE	RAE	R <sup>2</sup>	RMSE	MAE	RAE
LSTM	0.81	287.4	158.2	0.19	0.857	144.8	115	0.18

- Les résultats obtenus montrent de bonnes performances de prédiction du modèle LSTM implémenté. Le modèle arrive à prédire le débit réel avec une erreur relative absolue de 0.18 et un coefficient de détermination de l'ordre de 0.85, ce qui indique que ce modèle est très proche de la réalité. Les performances obtenues sont très encourageantes si on prend en compte le fait que le modèle a été entraîné sur la base d'une seule variable (le débit),
- Ces résultats indiquent que l'implémentation de ce type de modèles intelligents sur différents puits peut donner de très bonnes estimations du débit de production avec une capacité de prédiction future. Ces modèles peuvent devenir plus performants si l'on dispose d'un historique de production plus riche.

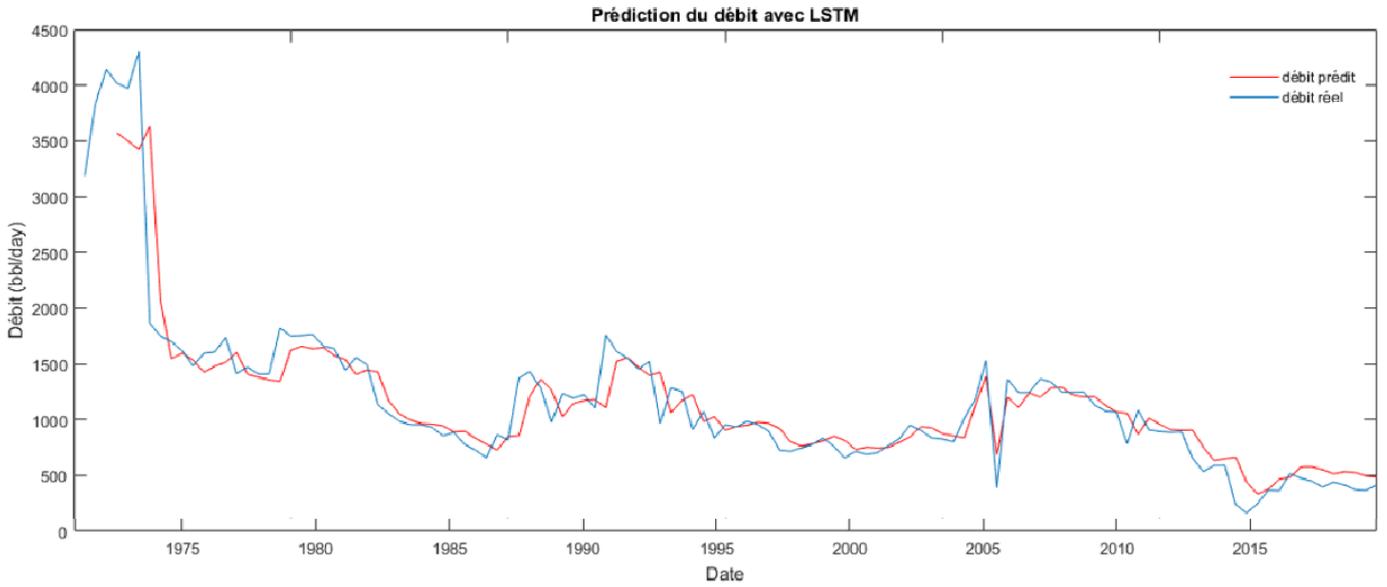


Figure 4.86 : Tracé du profil de production réel et prédit du puits Well1289

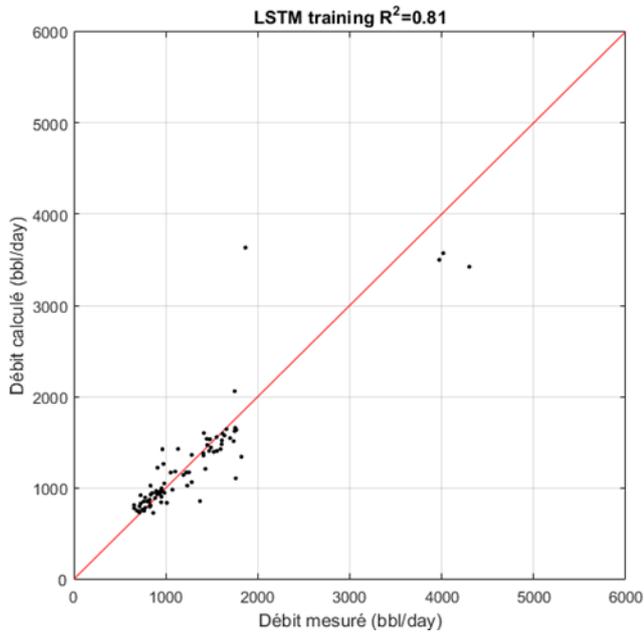


Figure 4. 87 : cross plot du training de LSTM

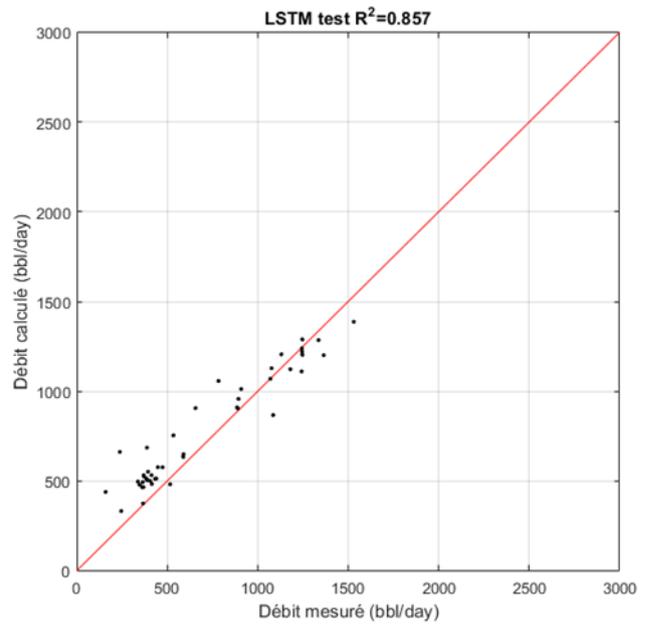


Figure 4. 88 : cross plot du test de LSTM

## 4.8 Influence de l'équation de Gilbert adaptée sur le point de fonctionnement d'un puits :

Un puits modèle a été choisi parmi ceux existants dans PIPESIM afin d'étudier l'influence du choix de l'équation de duse sur la détermination du point de fonctionnement. Ce puits est productif incliné et avec un système d'activation gaz lift. La pression de réservoir est prise à 4200 psia et le GOR est de l'ordre de 460,47 scf/STB. Trois équipements de surface ont été rajoutés à ce modèle à savoir :

- Une duse qui est décrite par Engine keyword tool.
- Un flowline.
- Un Sink (comme un séparateur), sa pression est d'ordre 200 psia.

On a réalisé une analyse nodale sur ce puits, le nœud est fixé au niveau de la tête de puits. L'étude est effectuée pour différents diamètres de la duse et de flowlines. Les formules de Gilbert et Gilbert\_HMD 2 sont utilisées pour modéliser l'écoulement à travers la duse.

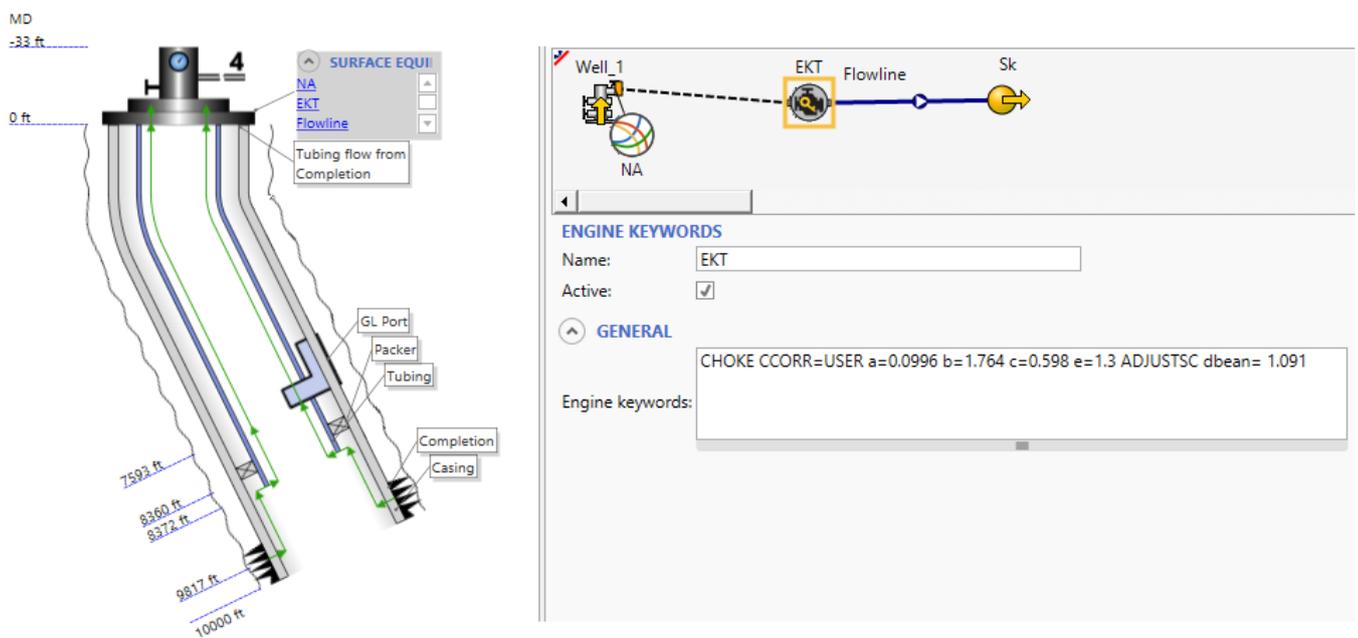


Figure 4.89 : modèle puits en PIPESIM

Les résultats obtenus après l'analyse nodale sont montrés dans le tableau et les figures ci-dessous.

Tableau 4. 30: le point de fonctionnement en utilisant l'équation de Gilbert originale et celle adaptée à HMD

Diamètre duse. 1/64in	Diamètres Flowline, in	Gilbert_HMD 2		Gilbert		Différence De débit
		Débit. STB/d	Pression. psia	Débit. STB/d	Pression. psia	
0.789	1.9	1940.053	1719.41	1890.232	1721.798	49.821
	2.375	3226.096	1644.53	3006.836	1658.749	219.26
	3.26	5947.199	1434.35	4799.731	1529.996	1147.468
	4.525	8978.19	1135.402	5562.194	1467.522	3415.996
	7.9	10147.79	999.376	5562.194	1467.522	4585.596
	10	10147.79	999.376	5562.194	1467.522	4585.596
1.091	1.9	1946.727	1719.085	1932.883	1719.749	13.844
	2.375	3257.336	1642.467	3191.208	1646.823	66.128
	3.26	6155.719	1415.948	5704.878	1455.359	450.841
	4.525	9797.056	1043.122	8380.114	1198.998	1416.942
	7.9	12152.41	711.9588	8380.114	1198.998	3772.296
	10	12152.41	711.9588	8055.219	1233.207	4097.191
2	1.9	1950.604	1718.896	1949.633	1718.944	0.971
	2.375	3275.901	1641.237	3270.932	1641.566	4.969
	3.26	6286.329	1404.272	6246.576	1407.839	39.753
	4.525	10374.47	971.9912	10133.32	1001.216	241.15
	7.9	13408.63	441.667	12739.93	601.2416	668.7
	10	13673.78	362.0659	12826.72	582.7945	847.06
3	1.9	1953.57	1718.09	1953.47	1718.101	0.1
	2.375	3188.025	1646.43	3187.542	1646.461	0.483
	3.26	6308.885	1401.705	6303.899	1402.155	4.986
	4.525	10247.88	987.7513	10216.29	991.81	31.59
	7.9	13373.66	449.947	13262.08	480.313	111.58
	10	13855.57	292.135	13700.8	350.69	154.77
7	1.9	1953.84	1718.08	1953.856	1718.082	-0.016
	2.375	3189.199	1718.083	3189.254	1646.35	-0.055
	3.26	6318.099	1400.87	6318.48	1400.839	-0.381
	4.525	10287.4	982.64	10288.57	982.4931	-1.17
	7.9	13440.52	430.6938	13440.7	430.62	-0.18
	10	13923.92	263.5789	13920.25	246.8163	3.67
64	1.9	1953.872	1718.082	1953.872	1718.082	0
	2.375	3189.223	1646.345	3189.324	1646.345	-0.101
	3.26	6319.048	1400.788	6319.048	1400.788	0
	4.525	10291.37	982.1315	10291.37	982.1314	0
	7.9	13447	428.77	13447.01	428.77	-0.01
	10	13929.74	260.9578	13929.74	260.9569	0

Le tableau ci-dessus contient les coordonnées (débit, pression) du point de fonctionnement du système obtenu pour chaque cas. On remarque que :

- la différence du débit entre le modèle de Gilbert original et Gilbert\_HMD 2 est plus marquée lorsque le diamètre de la duse est inférieur à 7 (1/64 in), avec un diamètre de flowline supérieur à 2.375

-Pour des diamètres de duses supérieurs ou égales à 7(1/64 in), les deux équations ont donné presque le même point de fonctionnement.

-Pour les cas où les diamètres de flowline sont d'ordre 1.9 ou 2.375 avec n'importe quel diamètre de la duse, la différence entre le débit estimé dans le cas de Gilbert et Gilbert\_HMD 2 ne présente pas de grandes valeurs.

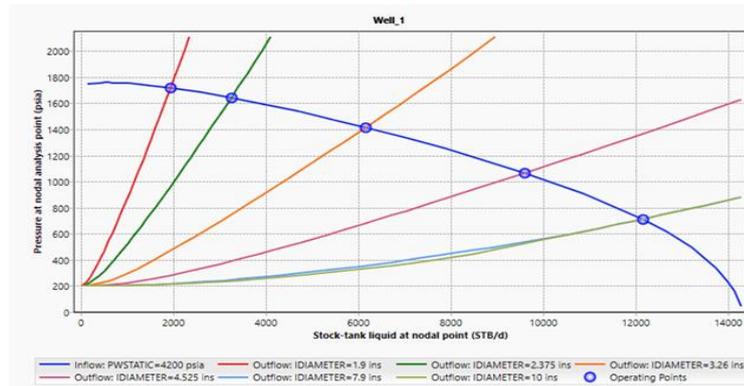


Figure 4.90 : points de fonctionnement où le diamètre de la duse est 1.019 dans le cas de Gilbert\_HMD 2

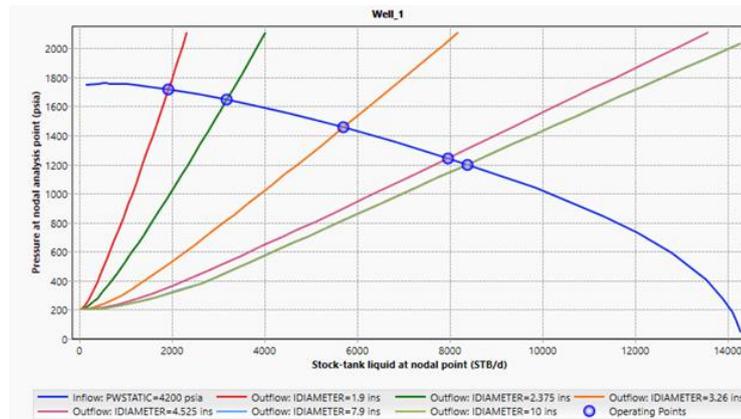


Figure 4.91 : points de fonctionnement où le diamètre de la duse est 1.019 dans le cas de Gilbert

#### 4.8.1 Conclusion 4:

D'après les résultats obtenus précédemment, et pour des diamètres de duses inférieurs à 7 (1/64 in), il existe une différence remarquable entre le débit calculé en utilisant la formule originale de Gilbert et celle adaptée au champ de Hassi Messaoud. Ceci nous emmène à dire que le choix de la formule de duse dans ces conditions est crucial car cette différence peut conduire à une mauvaise estimation du point de fonctionnement pouvant probablement conduire au dysfonctionnement du puits.

## Conclusion et recommandations

Dans ce mémoire de fin d'études, différents modèles servant à prédire le débit d'huile à travers les duses ont été développés pour le champ de Hassi Messaoud.

D'après les résultats obtenus, il paraît que le modèle Gilbert HMD 2 ; qui est une adaptation de la corrélation de Gilbert semble mieux approprié pour la prédiction du débit pour Les puits s'écoulant sous un régime critique.

Pour l'étude du régime non critique, l'adaptation du modèle de Gilbert a permis de donner une bonne estimation du débit, ceci montre que ce modèle peut être appliqué efficacement pour le cas de HMD, ces résultats étaient un peu surprenant puisque la corrélation de Gilbert est initialement développée pour l'étude du régime critique. Ajouté à cela, l'étude de chaque puits séparément en utilisant la programmation génétique a révélé que les variables utilisées dans l'équation de Gilbert sont celles qui ont plus d'influence sur le résultat final, ce point vient confirmer la possibilité d'utiliser l'équation de Gilbert adaptée pour le régime non critique pour le cas de ce champ. Ensuite, l'introduction de la chute de pression à travers le modèle extension Gilbert HMD a permis d'apporter une amélioration sur la capacité de prédiction du débit pour ce régime d'écoulement.

En outre, l'adaptation de ces modèles pour plusieurs intervalles de GOR selon les deux régimes d'écoulements a permis de réduire les erreurs et d'améliorer la capacité de prédiction de ces modèles théoriques sur l'ensemble des données de test.

Parmi les modèles intelligents développés pour chaque régime d'écoulement, l'outil de réseau neuronal ANN a montré les meilleures performances et a prouvé son efficacité pour l'estimation du débit, il serait donc très intéressant d'implémenter des modèles se basant sur cet outil.

Le degré de précision peut être augmenté encore plus si l'on considère l'étude de chaque puits individuellement, c'est ce qu'il a été montré d'après l'utilisation de la programmation génétique et des modèles LSTM.

Voici la liste de quelques recommandations qu'on peut tirer d'après les résultats de la présente étude :

- Si l'on souhaite travailler avec des modèles théoriques pour l'estimation du débit, il est recommandé d'utiliser les modèles adaptés selon chaque intervalle de GOR, surtout pour les intervalles de **GOR  $\leq$  100** et **GOR  $>$  4000** ( $\text{sm}^3/\text{sm}^3$ ) pour les puits s'écoulant sous un régime critique, et l'intervalle **GOR  $>$  1300** ( $\text{sm}^3/\text{sm}^3$ ) pour les puits produisant sous un régime non critique.
- Intégration des modèles intelligents, spécialement les réseaux de neurones pour l'estimation du débit afin d'obtenir des résultats plus proches de la réalité.
- Des efforts restent encore à faire lors du processus de collecte de données au champ de HMD, où il faut veiller à ce que les mesures soient plus représentatives.
- Effectuer plus de tests pour avoir davantage de points de mesure, ce qui permettra de mieux exploiter les modèles intelligents et produire des résultats plus performants.
- lors des opérations de jaugeages, il serait intéressant d'effectuer des mesures des propriétés des fluides produits à différentes conditions de pression et de température, et ce afin de tenter l'utilisation d'autres modèles, que ce soit de modèles théoriques qui incluent ces

caractéristiques ou même de modèles intelligents, ceci pourrait produire des modèles plus riches et plus performants.

- Considérer l'étude de chaque puits individuellement puisque chaque puits a son propre comportement et historique pour plus de précision.
- Utilisation des modèles LSTM pour les puits qui comportent un bon historique de production puisque ces modèles permettent en plus de générer des prédictions futures.
- Considérer l'utilisation des coefficients de l'équation de Gilbert adaptée au champ de Hassi Messaoud lors des simulations puisque cette dernière reste plus représentative de ce champ que l'équation de Gilbert originale.

## Références

- [1] Dr. James P. Brill and Dr.H.Dale Beggs, « Two-Phase Flow In Pipes », sixth Edition, January,1991
- [2]Tangren, R.F.; Dodge, C.H.; Seifert, H.S.: Compressibility effects in two-phase flow. J. Appl. Phys. 20(7), 637–645 (1949)
- [3] Gilbert, W.E.: Flowing and gas-lift well performance. API Drill. Prod. Pract. 20, 126–157 (1954)
- [4] Ros, N.C.J.: An analysis of critical simultaneous gas/liquid flow through a restriction and its application to flowmetering. Appl. Sci. Res. 9(Section A), 374–388 (1960)
- [5] Poettmann, F.E.; Beck, R.L.: New charts developed to predict gas– liquid flow through chokes. Word Oil 184(3), 95–100 (1963)
- [6]Omana, R.A.: Multiphase flow through chokes. M.S. Thesis, University of Tulsa (1968)
- [7] Ashford, F.E.; Pierce, P.E.: Determining multiphase pressure drops and flow capacities in down-hole safety valves. J. Pet. Technol. 27(9), 1145–1152 (1975)
- [8 ]Al-Attar, H.H.; Abdul-Majeed, G.H.: Revised bean performance equation for east Baghdad oil wells. SPE Prod. Eng. 3(1), 127–131 (1988)
- [9]Abdul-Majeed, G.H.; Maha, R.A.A.: Correlations developed to predict two-phase flow through wellhead chokes. J. Can. Pet. Technol. 30(06), 47–55 (1991)
- [10] Elgibaly, A.A.M.; Nashawi, I.S.: New correlations for critical and subcritical two-phase flow through wellhead chokes. J. Can. Pet. Technol. 37(06), 36–43 (1998)
- [11 ]Al-Towailib A.I. and Al-Marhoun, M.A.: "New Correlation for Two-phase Flow Through Chokes," M.S. Thesis, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia (1992).
- [12] Ghorbani, H.; Wood, D.A.; Moghadasi, J.; Choubineh, A.; Abdizadeh, P.; Mohamadian, N.: Predicting liquid flow-rate performance through wellhead chokes with genetic and solver optimizers: an oil field case study. J. Pet. Exp. Prod. Technol. 9, 1355–1373 (2019)
- [13] Boyun Guo, PhD Xinghui Liu, PhD Xuehao Tan, PhD : Petroleum Production Engineering, Second Edition, Gulf Professional Publishing is an imprint of Elsevier, 2017
- [14] Dr.James P.Brill Dr. H. Dale Beggs : Two Phase Flow In Pipes,6<sup>th</sup> edition, 1978
- [15] B. R. Munson, D. F. Young, and T. H. Okiishi. Fundamentals of Fluid Mechanics. John Wiley & Sons, Inc, third edition, 1998.
- [16]Sakir Amiroudine, Professeur à l'université Bordeaux I, Jean-Luc Battaglia Professeur à l'université Bordeaux I, Mécanique Des Fluides, édition Dunod, Paris, 2011
- [17] H. Beggs, Production optimization : using NODAL analysis. Tulsa, Okla : OGCI Publications, 1991.

- [18] Mihael Golan, Curtis H. Wwhitson, Well Performance, NTNU
- [19] Nour ElHouda Tellache, Meriem Waffa Hassen, Mohamed Otmanine, Mohamed Khodjal, Improved Multiphase Flow Rate Models for Chokes in the Algerian HMD Oil Field
- [20] PIPESIM, online help version 2017.2
- [21]Yogendra Narayan, Pandey Ayush, Rastogi Sribharath, Kainkaryam Srimoyee, Bhattacharya Luigi Saputelli, Machine Learning in the Oil and Gas Industry, Apress
- [22] Mohammad Ali Kargarpour, Oil and gas well rate estimation by choke formula: semi-analytical approach, Journal of Petroleum Exploration and Production Technology (2019)
- [23]Juliana Yotham Mwalyepelo, Analysis of selected Multiphase metering concepts, Norwegian University of Science and Technology
- [24]J. Waltrich Craft and Hawkins ,Performance evaluation of multiphase flow models applied to virtual flow metering K. Mokhtari & P. Department of Petroleum Engineering, Louisiana State University, USA
- [25 ] J. Jansen, Nodal Analysis of Oil and Gas Production Systems. Richardson SPE, 2017
- [26]Mohammad Rasheed Khan, Zeeshan Tariq , Abdulazeez Abdulraheem, Application of Artificial Intelligence to Estimate Oil Flow Rate in Gas-Lift Wells
- [27] Seidi, S.; Sayahi, T.: A new correlation for prediction of sub-critical two-phase flow pressure drop through large-sized wellhead chokes. J. Nat. Gas Sci. Eng. 26, 264–278 (2015)
- [28] Nasriani, H.R.; Kalantariasl, A.: Choke performance in high-rate gas condensate wells under subcritical flow condition. Energy Sour., Part A: Recover., Util., Environ. Eff. 37(2), 192–199 (2015)
- [29] Pejman Shoeibi omrani, Iulian Dobrovolschi et Stephane Belfroid , Peter Kronberger et Esteban Munoz, Improving Accuracy Of Virtual Flow Metring And Back Allocation Through Machine Learning, SPE- 192819-MS 2018
- [30] Hamzeh Ghorbanil · David A. Wood · Jamshid Moghadasi · Abouzar Choubineh · Peyman Abdizadeh · Nima Mohamadian, Predicting liquid flow-rate performance through wellhead chokes with genetic and solver optimizers: an oil field case study, Journal of Petroleum Exploration and Production Technology,2018
- [31] Mohammad Khamis -Murtada Elhaj - Abdulazeez Abdulraheem, Optimization of choke size for two-phase fow using artificial intelligence,
- [32] Henri P. Gavin, The Levenberg-Marquardt algorithm for nonlinear least squares curve-fitting problems, Department of Civil and Environmental Engineering Duke University September 18, 2020
- [34] <https://www.educba.com/support-vector-regression/>
- [35]: <https://www.mathworks.com/help/stats/gaussian-process-regression-models.html>
- [36] MERZOUG Ahmed, Simulation de l'écoulement multiphasique. Et application des algorithmes métaheuristiques pour l'optimisation de la méthode de gaz lift , Mémoire de fin d'études (Master), 2020 .

## Annexe A : graphes de la partie théorique

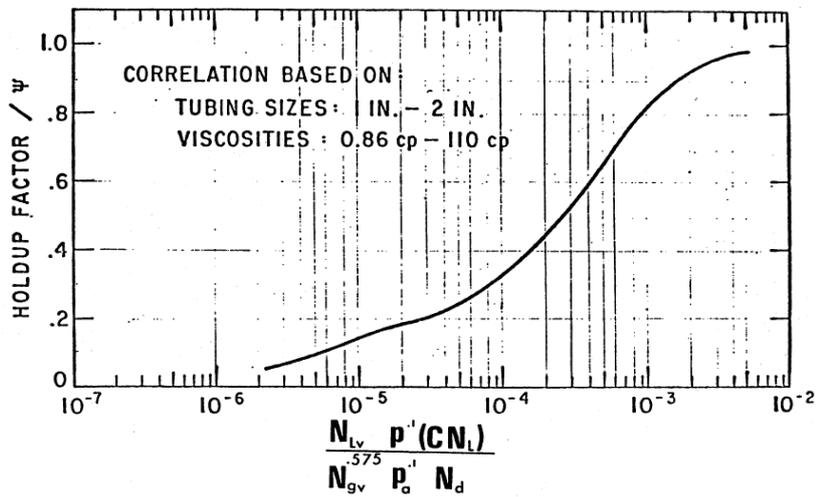
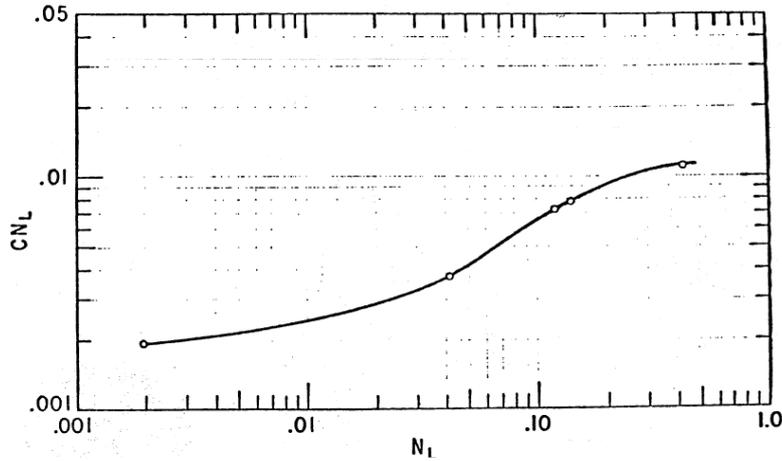


Figure An. A 2 : graphe  $H_L / \Psi$  en fonction de  $\left( \frac{N_{LV} p'(CN_L)}{N_{gV}^{.575} p_a' N_d} \right) (p/p_a)^{0.1} (CN_L/N_d)$

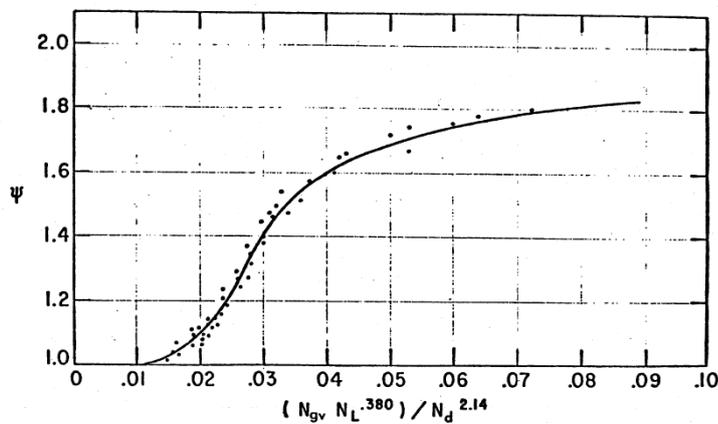


Figure An. A 3 : graphe  $\Psi$  en fonction de  $(N_{gV} N_L^{.380}) / N_d^{2.14}$

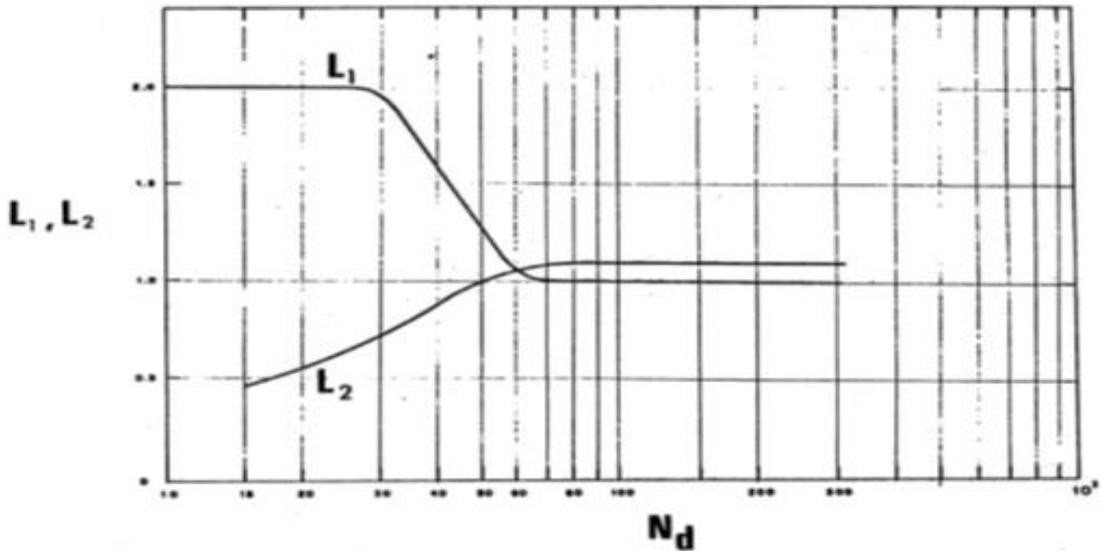


Figure An. A 4 : graphe du nombre de régime d'écoulement

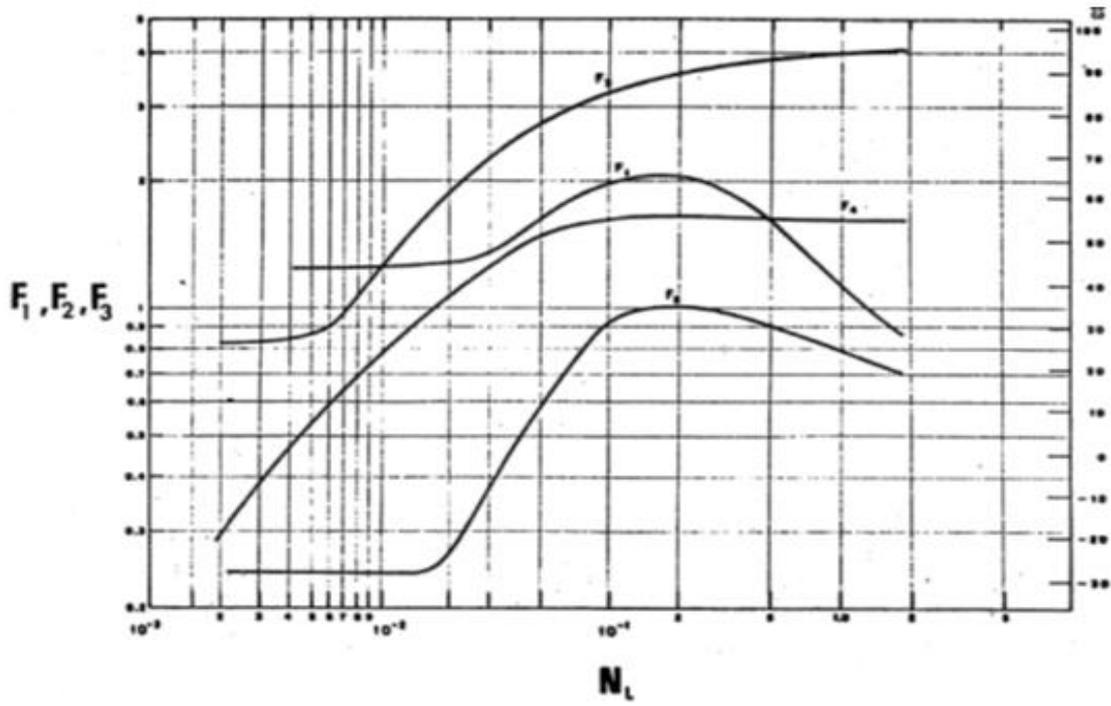


Figure An. A 5 : graphe du nombre de vitesse de glissement pour bubble flow

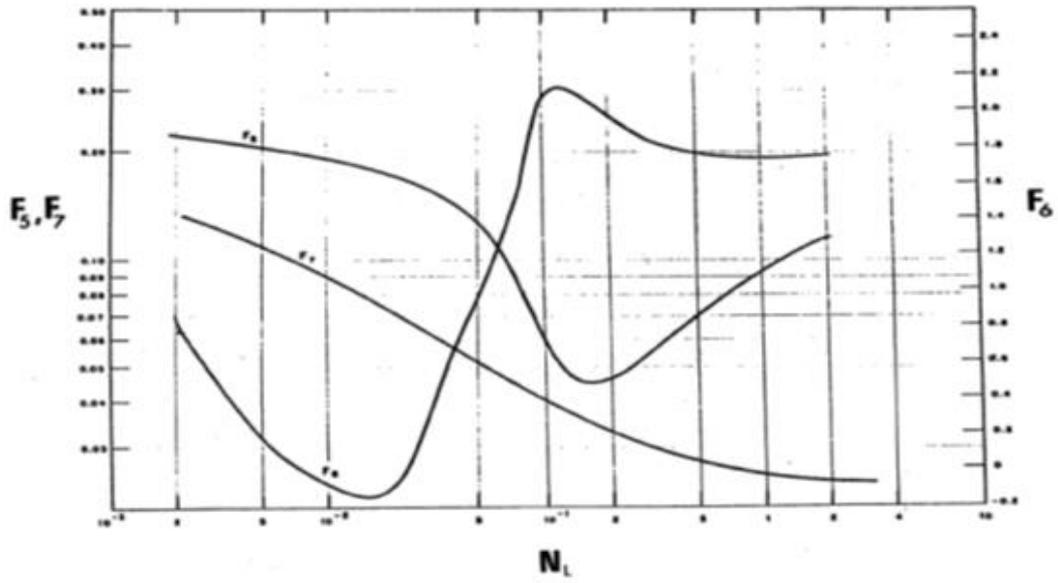


Figure An. A 6 : graphe du nombre de vitesse de glissement pour slug flow

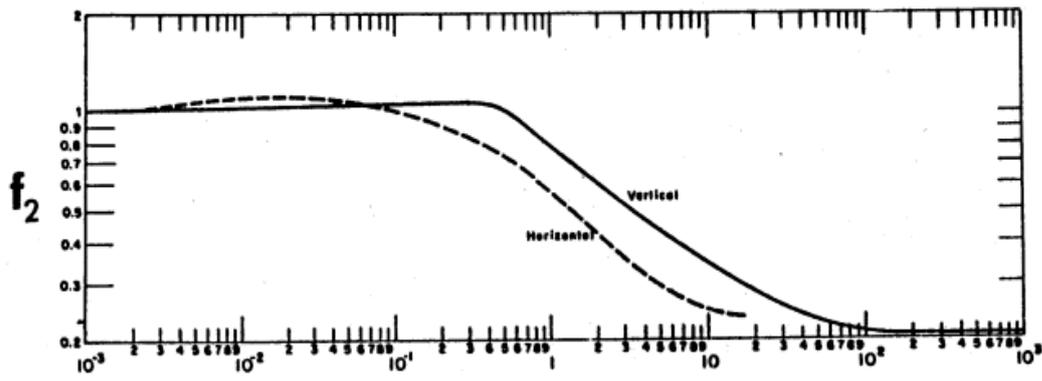


Figure An. A 7 : correction du facteur de friction bubble flow

## Annexe B : cross plot après subdivision en intervalles GOR

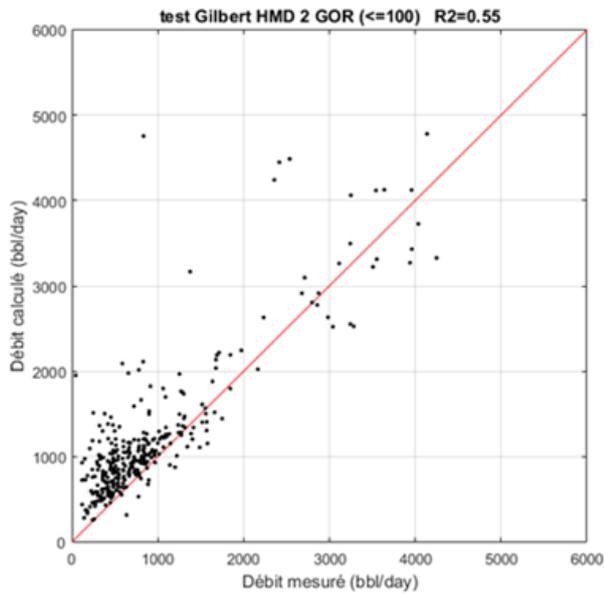


Figure An.B 1 : cross plot test Gilbert HMD 2 sur les GOR $\leq 100$ , régime critique.

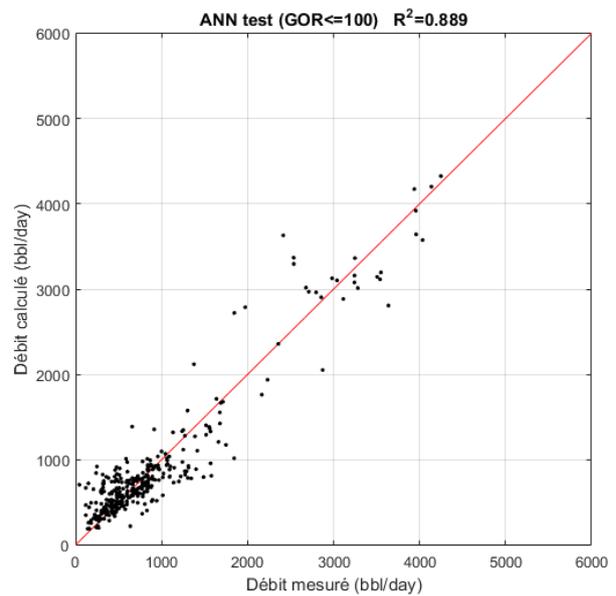


Figure An.B 2 ; cross plot test ANN sur les GOR $\leq 100$ , régime critique.

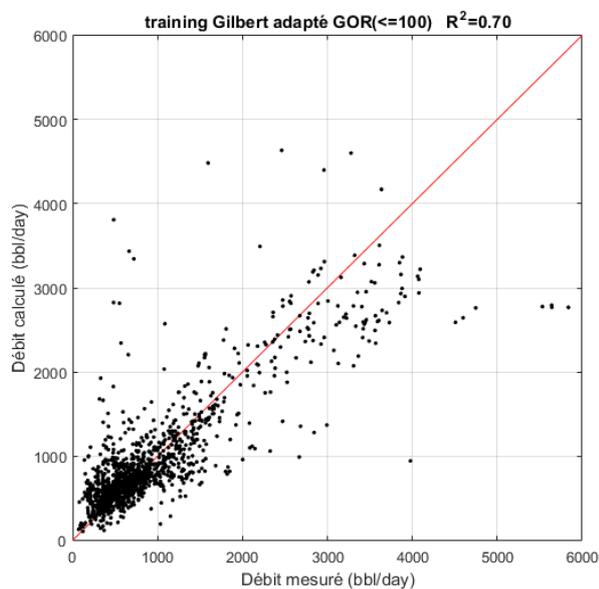


Figure An.B 3 : cross plot Gilbert adapté training sur Les GOR $\leq 100$ , régime critique

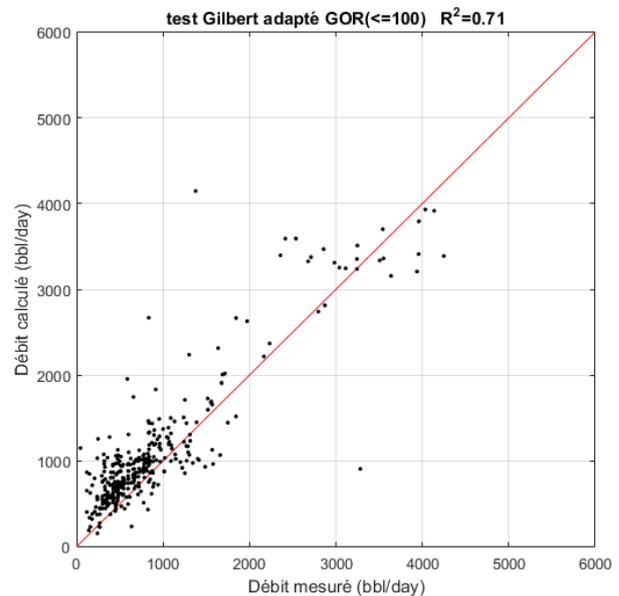


Figure An.B 4 : cross plot Gilbert adapté test sur les GOR $\leq 100$ , régime critique

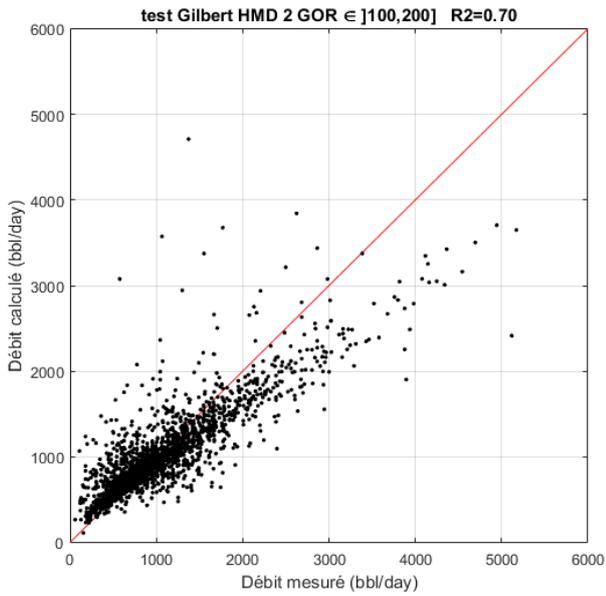


Figure An.B 5 : cross plot test Gilbert HMD 2 sur  $GOR \in ]100,200]$ , régime critique

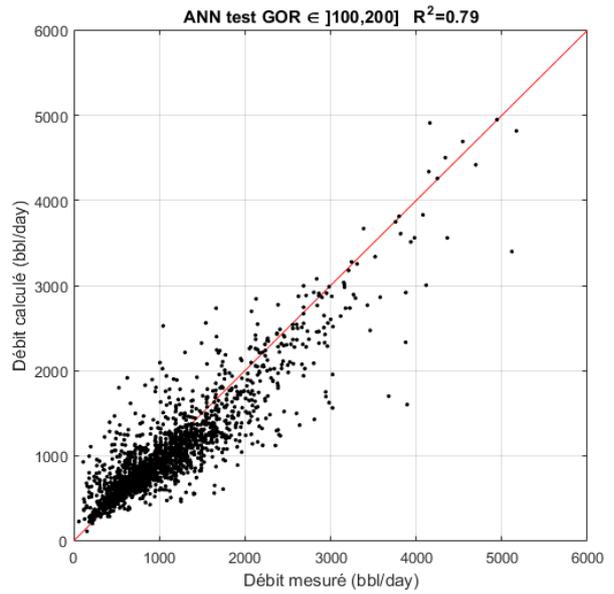


Figure An.B 6 : cross plot test ANN sur  $GOR \in ]100,200]$ , régime critique

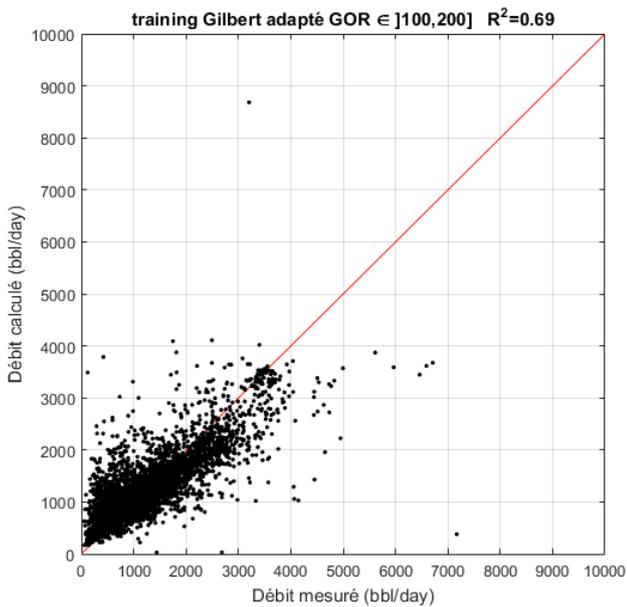


Figure An.B 7 : cross plot training Gilbert adapté sur  $GOR \in ]100,200]$ , régime critique

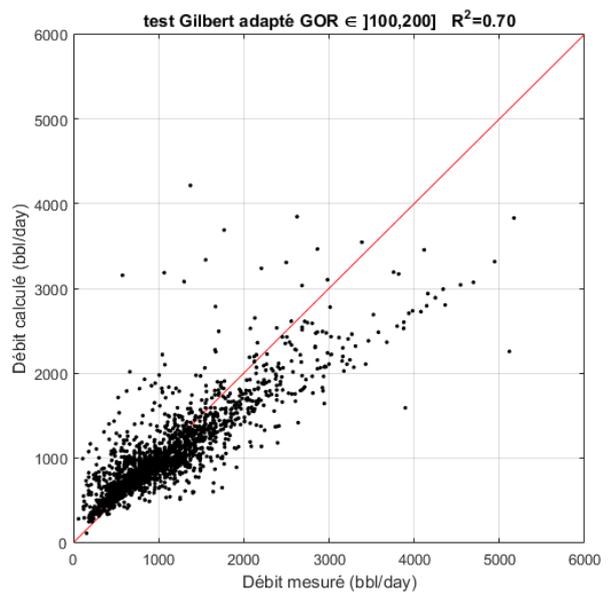


Figure An.B 8 : cross plot test Gilbert adapté sur  $GOR \in ]100,200]$ , régime critique

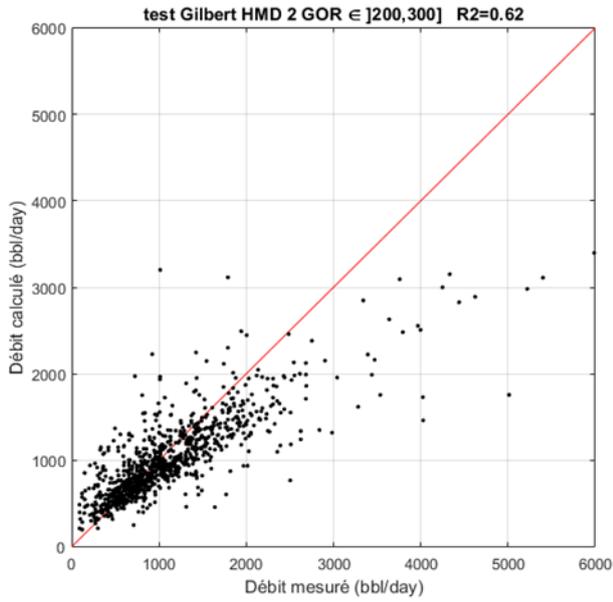


Figure An.B 9 : cross plot test Gilbert HMD 2 sur  $GOR \in ]200,300]$ , régime critique

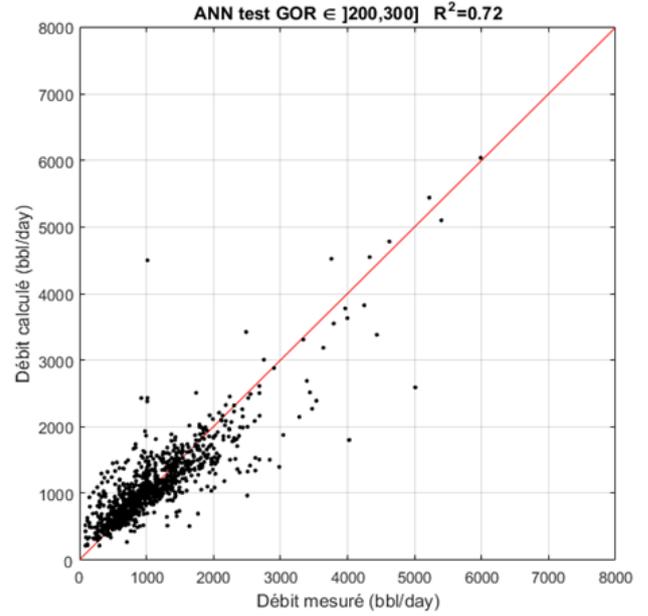


Figure An.B 10 : cross plot test ANN sur  $GOR \in ]200,300]$ , régime critique



Figure An.B 11 : cross plot training Gilbert adapté sur  $GOR \in ]200,300]$ , régime critique

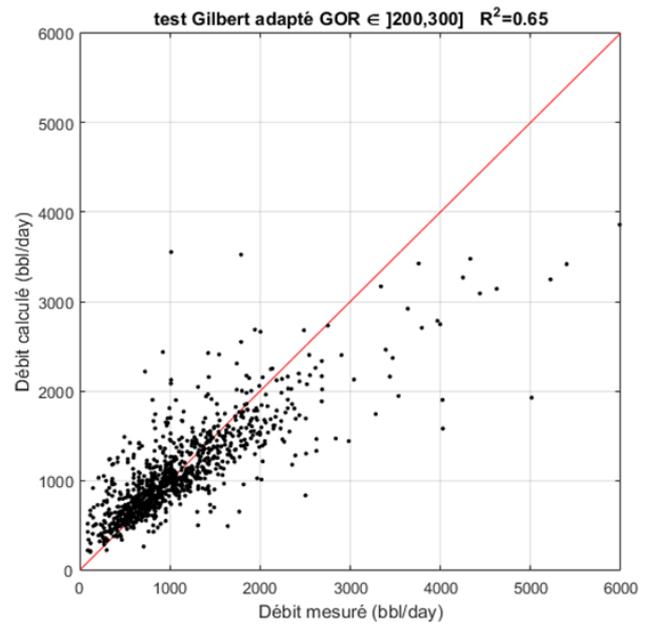


Figure An.B 12 : cross plot test Gilbert adapté sur  $GOR \in ]200,300]$ , régime critique

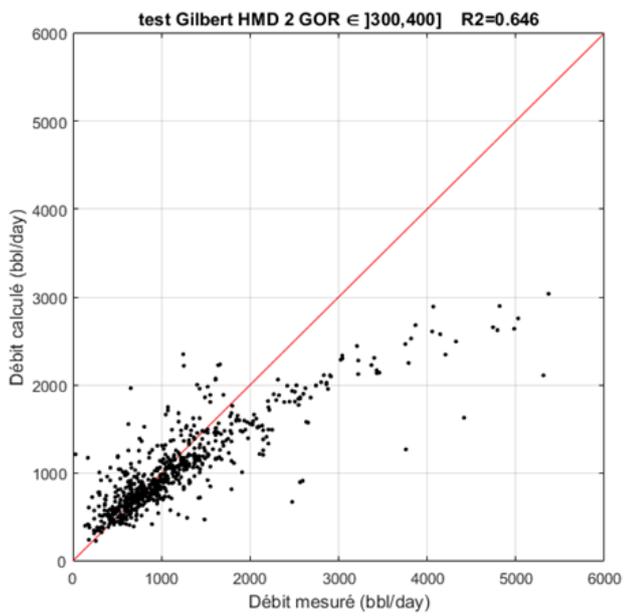


Figure An.B 13 : cross plot test Gilbert HMD 2 sur  $GOR \in ]300,400]$ , régime critique

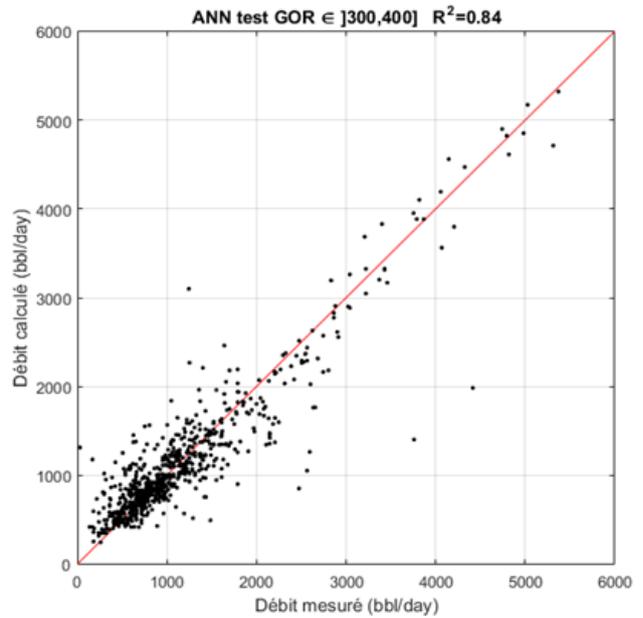


Figure An.B 14 : cross plot test ANN sur  $GOR \in ]300,400]$ , régime critique



Figure An.B 15 : cross plot training Gilbert adapté sur  $GOR \in ]300,400]$ , régime critique

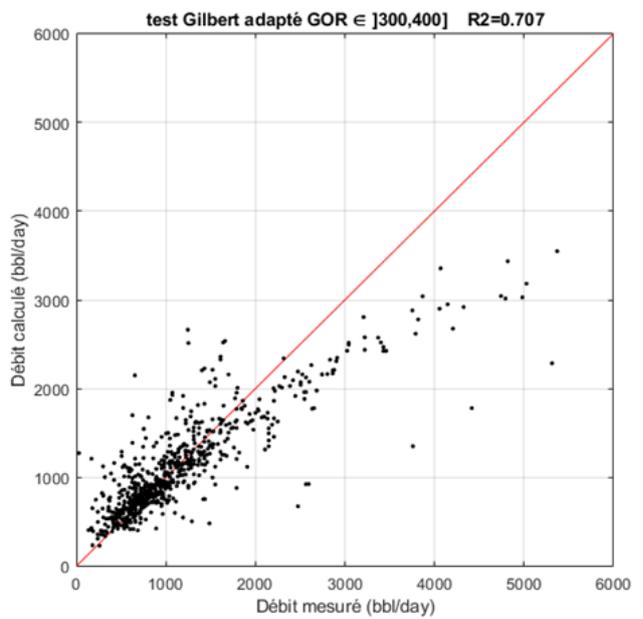


Figure An.B 16 : cross plot test Gilbert adapté sur  $GOR \in ]300,400]$ , régime critique

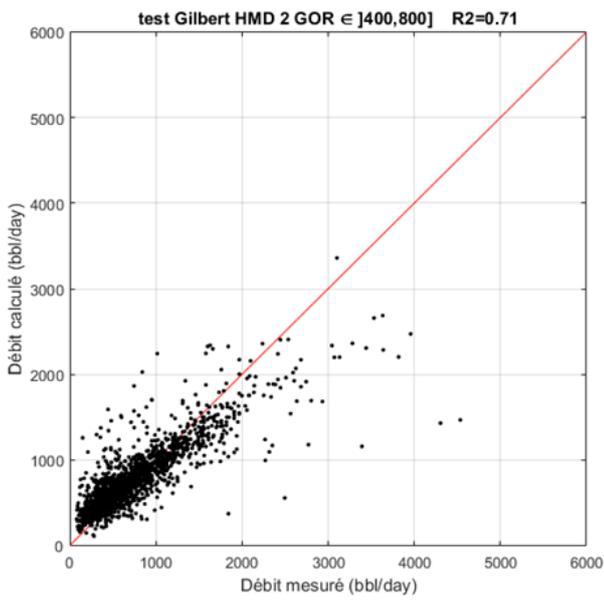


Figure An.B 17 : cross plot test Gilbert HMD 2 sur  $GOR \in ]400,800]$ , régime critique

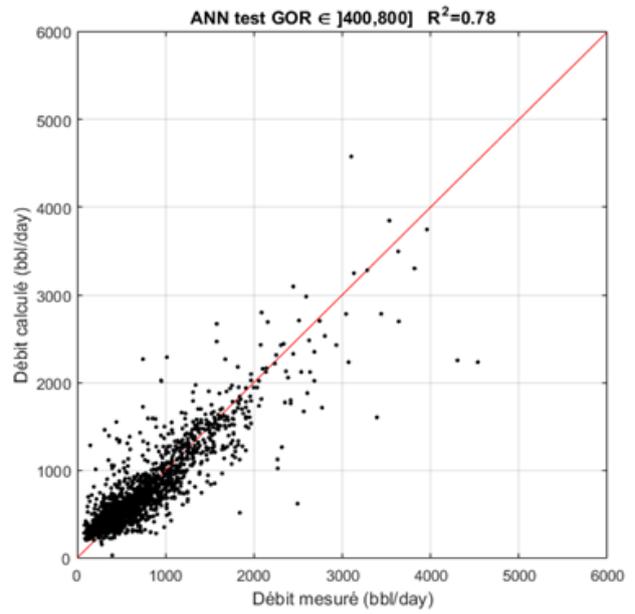


Figure An.B 18 : cross plot test ANN sur  $GOR \in ]400,800]$ , régime critique

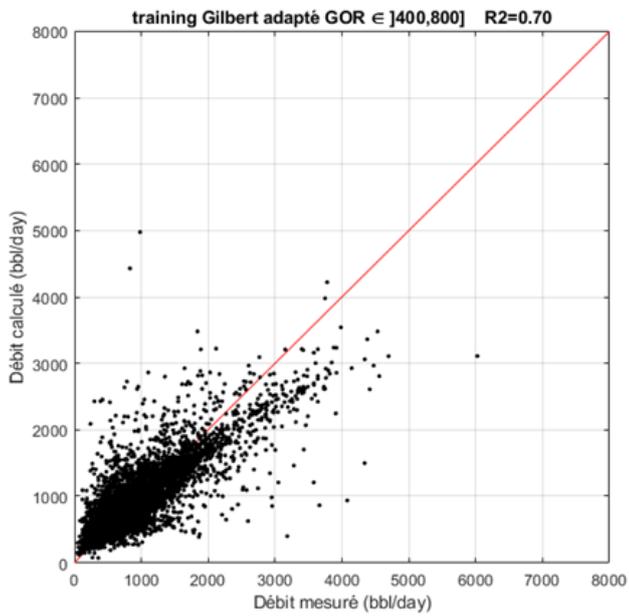


Figure An.B 19 : cross plot training Gilbert adapté sur  $GOR \in ]400,800]$ , régime critique

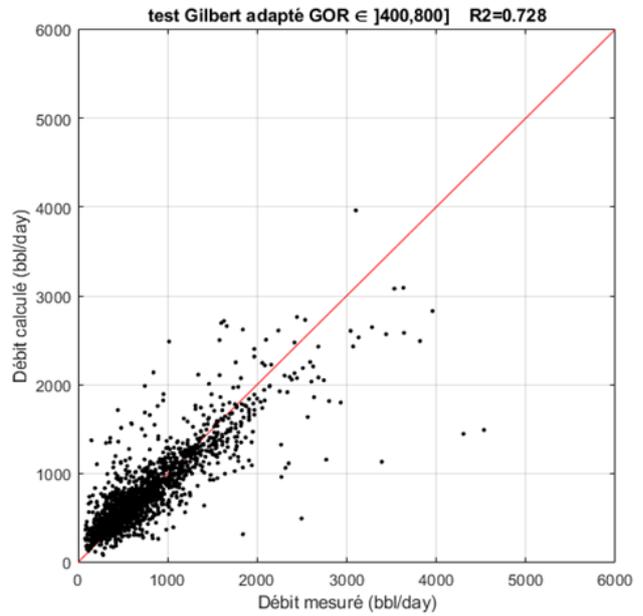


Figure An.B 20 : cross plot test Gilbert adapté sur  $GOR \in ]400,800]$ , régime critique

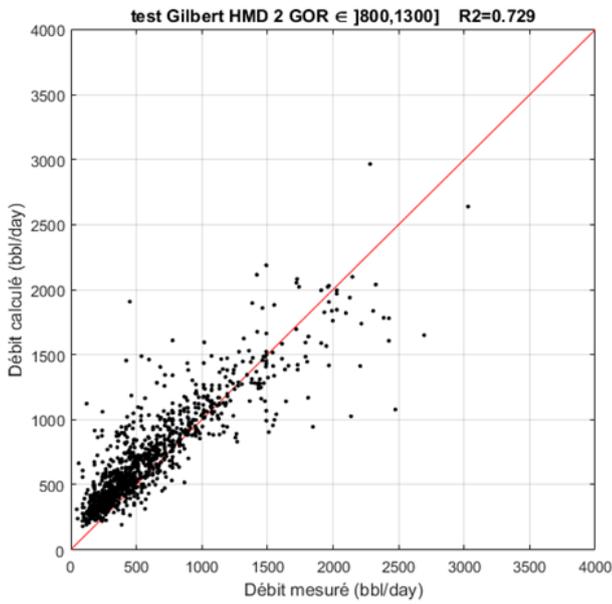


Figure An.B 21 : cross plot test Gilbert HMD 2 sur  $GOR \in ]800,1300]$ , régime critique

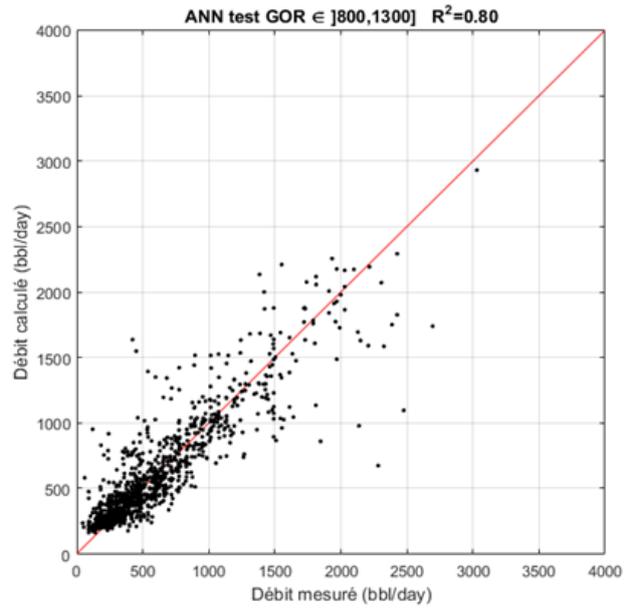


Figure An.B 22 : cross plot test ANN sur  $GOR \in ]800,1300]$ , régime critique

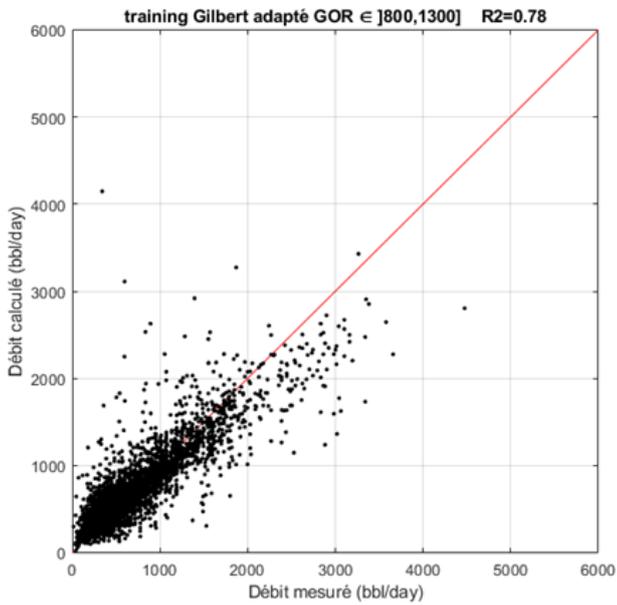


Figure An.B 23 : cross plot training Gilbert adapté sur  $GOR \in ]800,1300]$ , régime critique

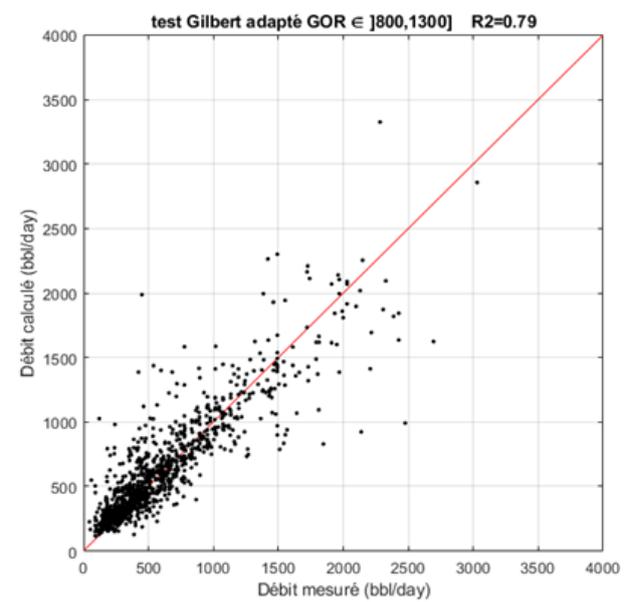


Figure An.B 24 : cross plot test Gilbert adapté sur  $GOR \in ]800,1300]$ , régime critique

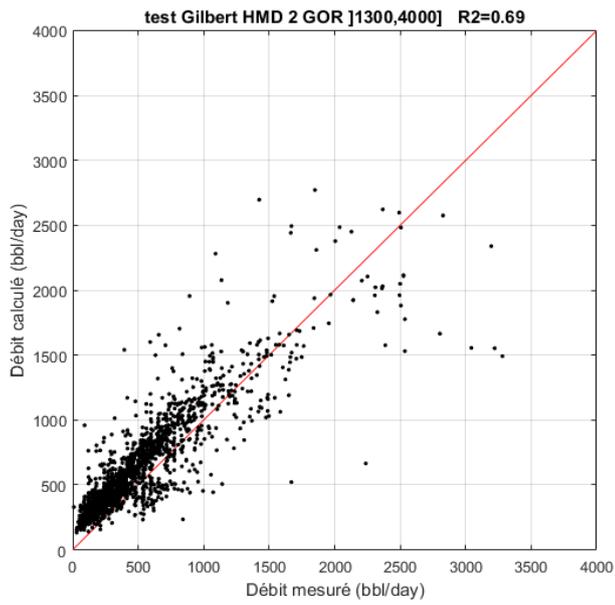


Figure An.B 25 : cross plot test Gilbert HMD 2 sur  $GOR \in ]1300,4000]$ , régime critique

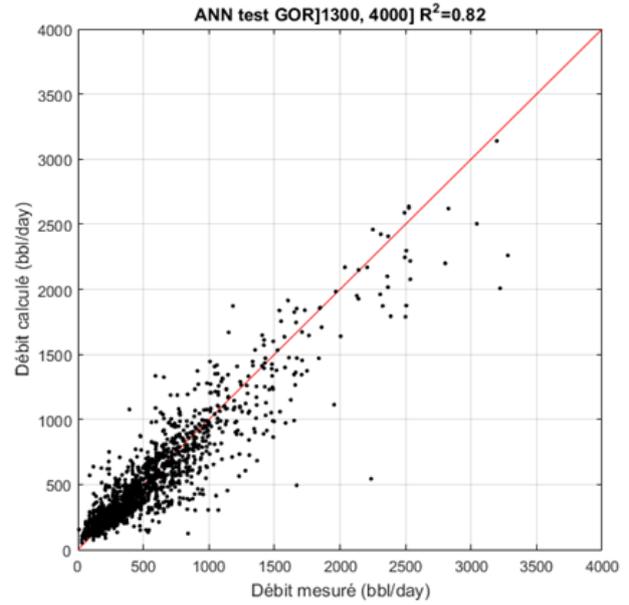


Figure An.B 26 : cross plot test ANN sur  $GOR \in ]1300,4000]$ , régime critique

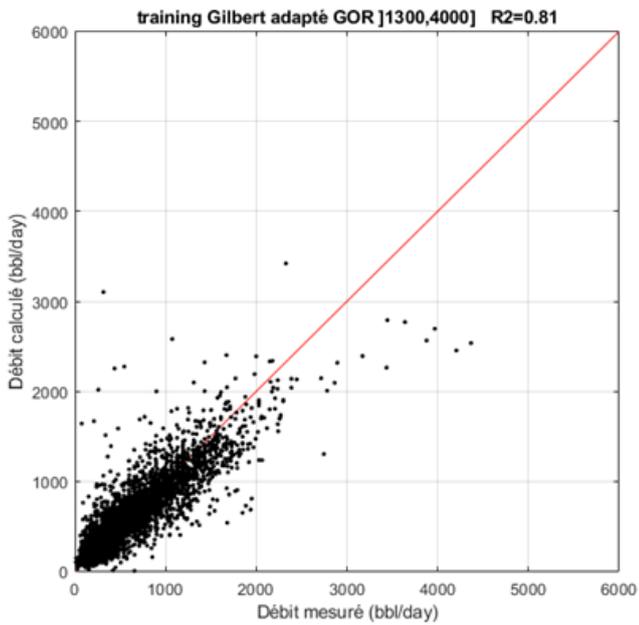


Figure An.B 27 : cross plot training Gilbert adapté sur  $GOR \in ]1300,4000]$ , régime critique

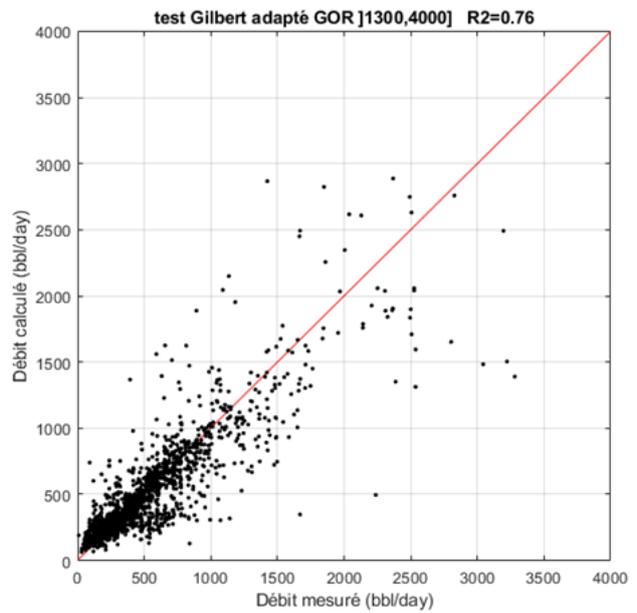


Figure An.B 28 : cross plot test Gilbert adapté sur  $GOR \in ]1300,4000]$ , régime critique

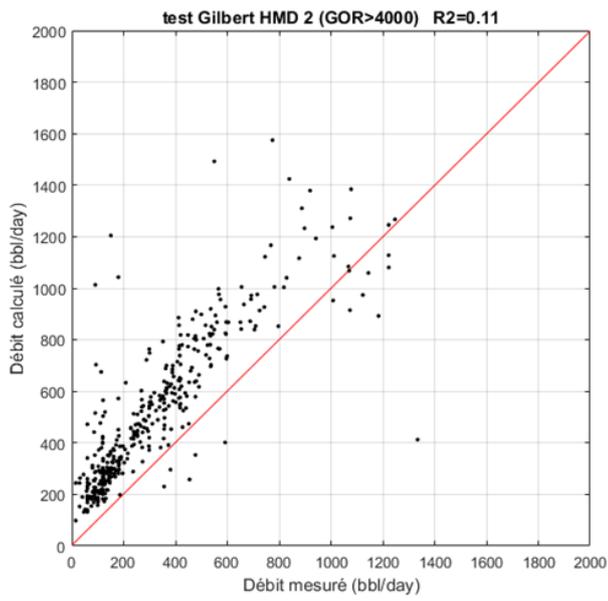


Figure An.B 29 : cross plot test Gilbert HMD 2 sur GOR>4000, régime critique

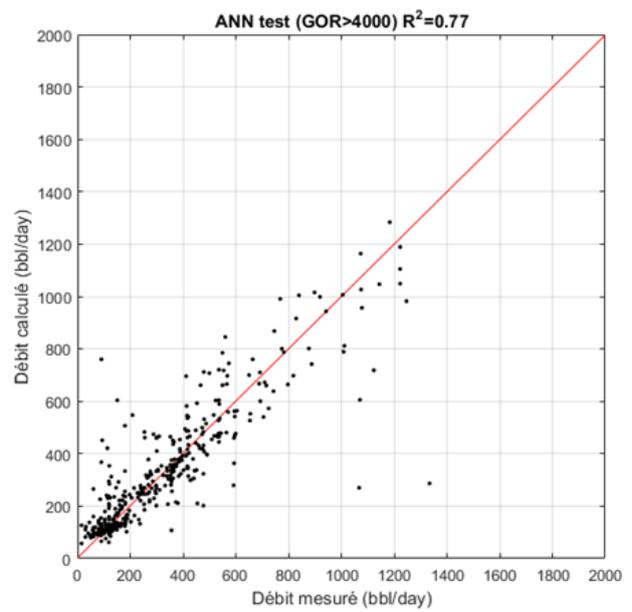


Figure An.B 30 : cross plot test ANN sur GOR>4000, régime critique

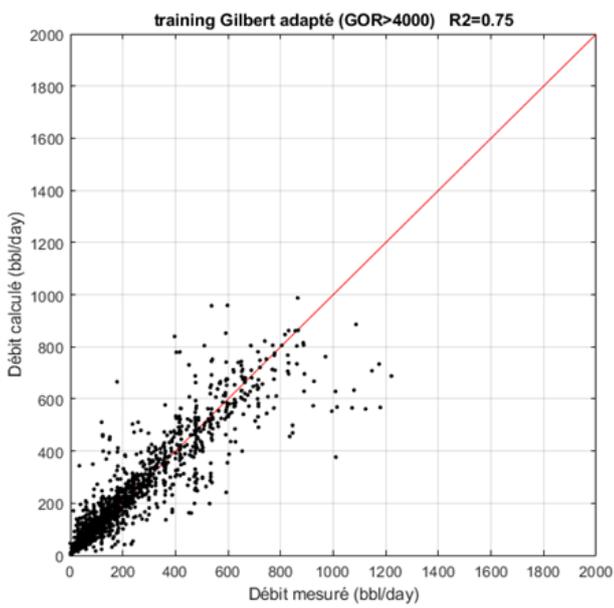


Figure An.B 31 : cross plot training Gilbert adapté sur GOR>4000, régime critique

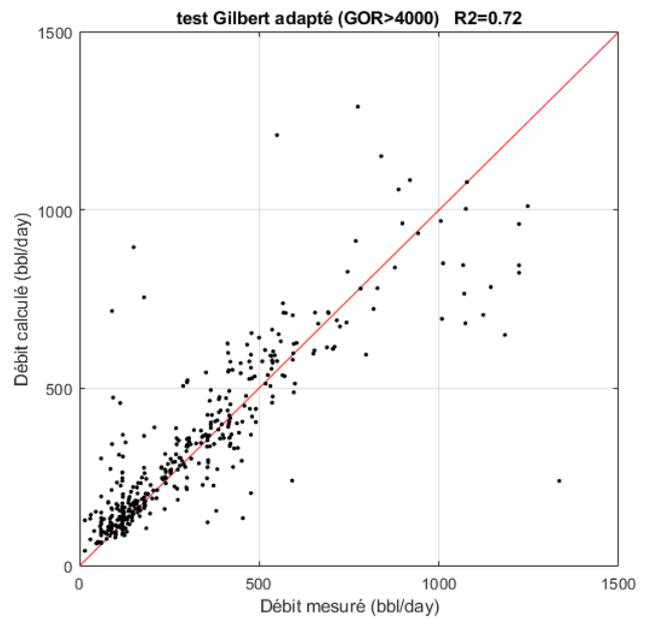


Figure An.B 32 : cross plot test Gilbert adapté sur GOR>4000, régime critique

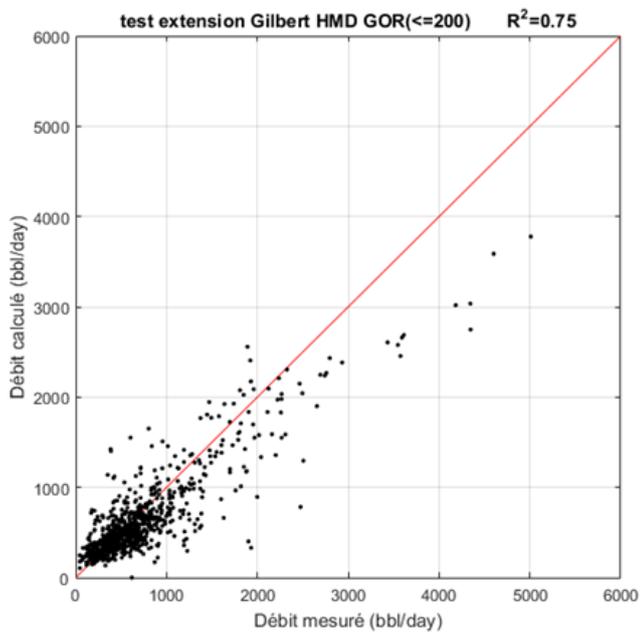


Figure An.B 33 : cross plot test extension Gilbert HMD sur  $GOR \leq 200$ , régime sous critique

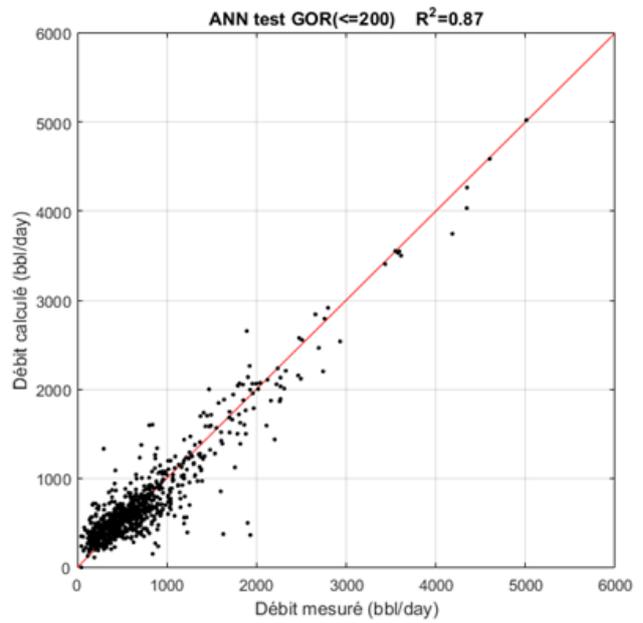


Figure An.B 34 : cross plot test ANN sur  $GOR \leq 200$ , régime sous critique

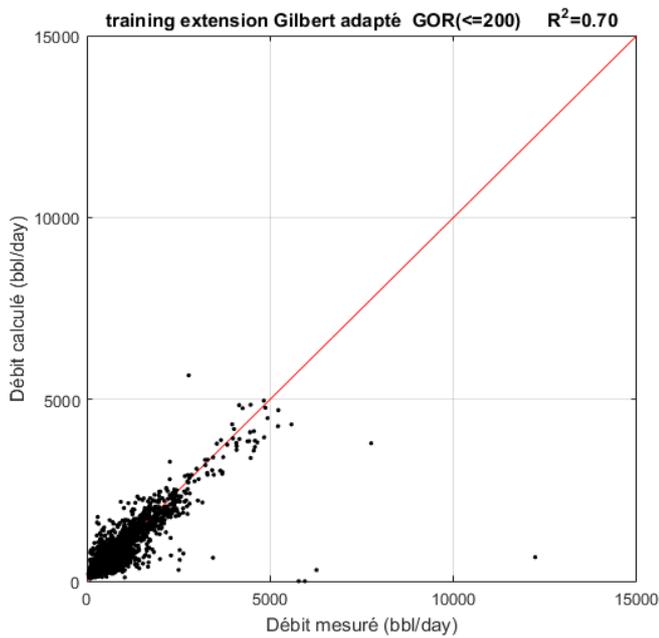


Figure An. B 35 : cross plot training extension Gilbert adapté sur  $GOR \leq 200$ , régime sous critique

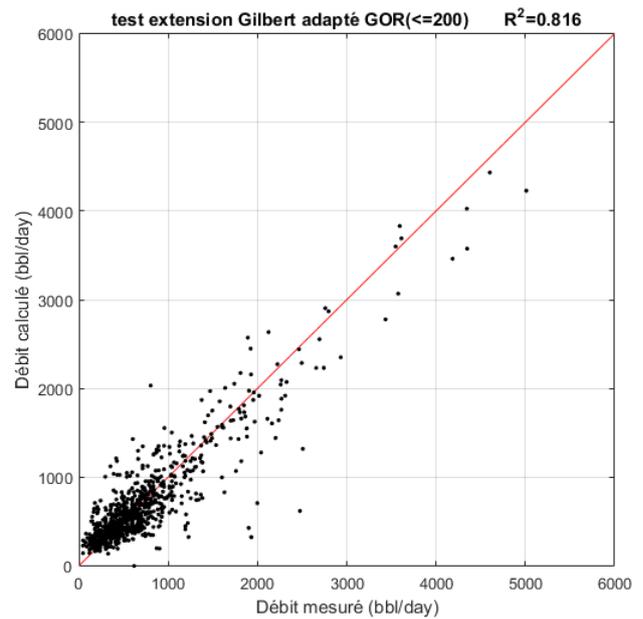


Figure An. B 36 : cross plot test extension Gilbert adapté sur  $GOR \leq 200$ , régime sous critique

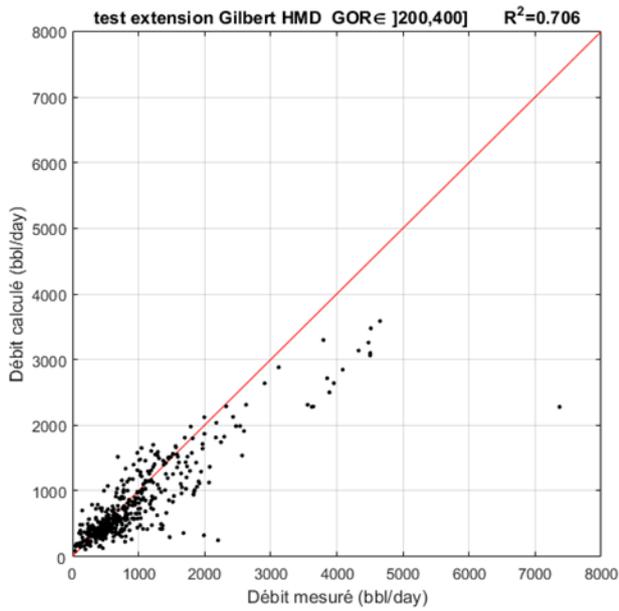


Figure An. B 37 : cross plot test extension Gilbert HMD sur  $GOR \in ]200, 400]$ , régime sous critique

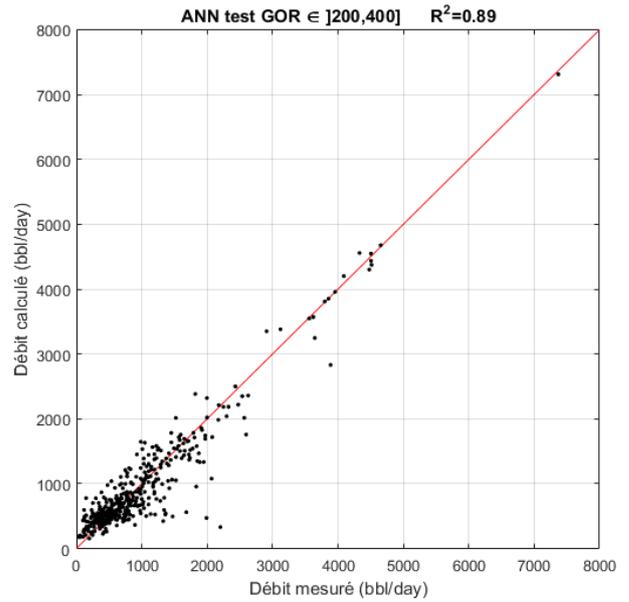


Figure An. B 38 : cross plot test ANN sur  $GOR \in ]200,400]$ , régime sous critique

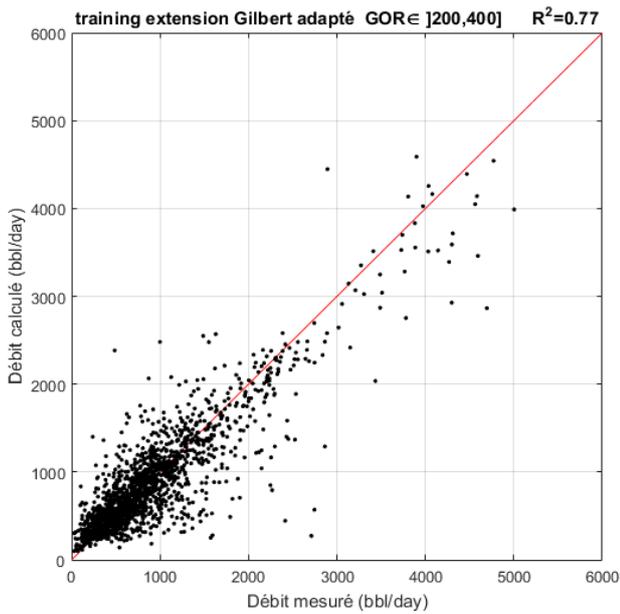


Figure An.B 39 : cross plot training extension Gilbert adapté sur  $GOR \in ]200,400]$ , régime sous critique

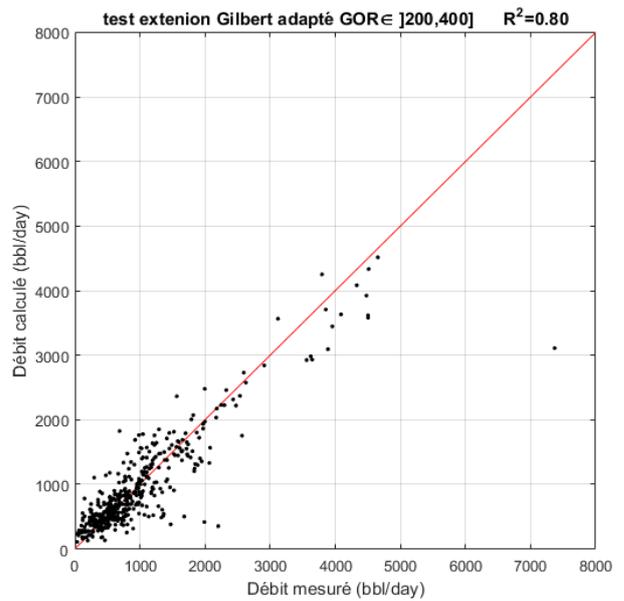


Figure An.B 40 : cross plot test extension Gilbert adapté sur  $GOR \in ]200,400]$ , régime sous critique

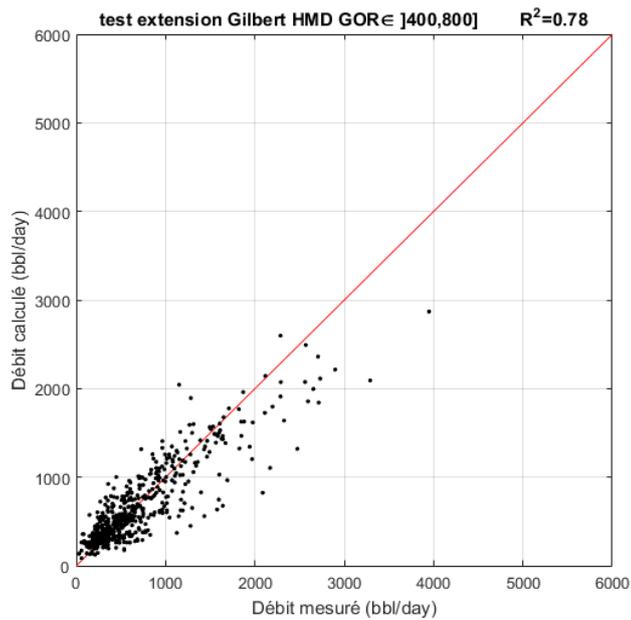


Figure An.B 41 : cross plot test extension Gilbert HMD sur  $GOR \in ]400,800]$ , régime sous critique

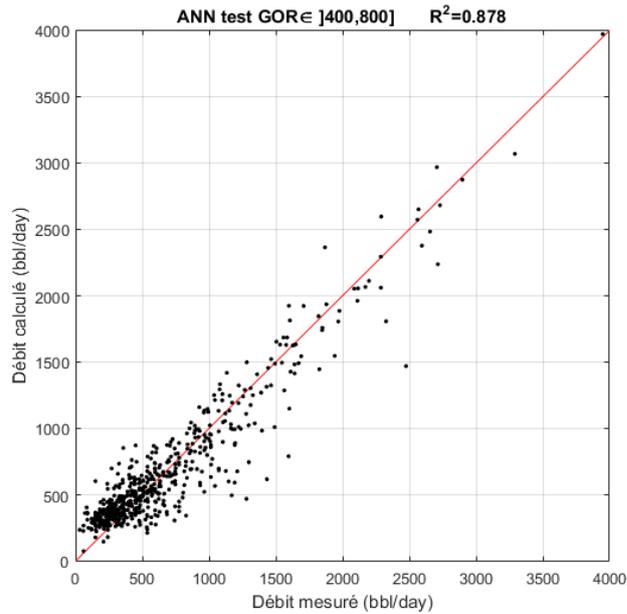


Figure An.B 42 : cross plot test ANN sur  $GOR \in ]400,800]$ , régime sous critique

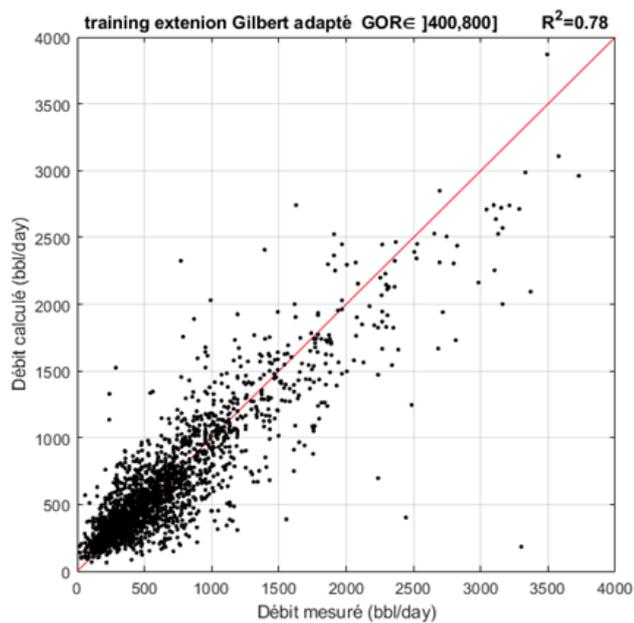


Figure An.B 43 : cross plot training extension Gilbert adapté sur  $GOR \in ]400,800]$ , régime sous critique

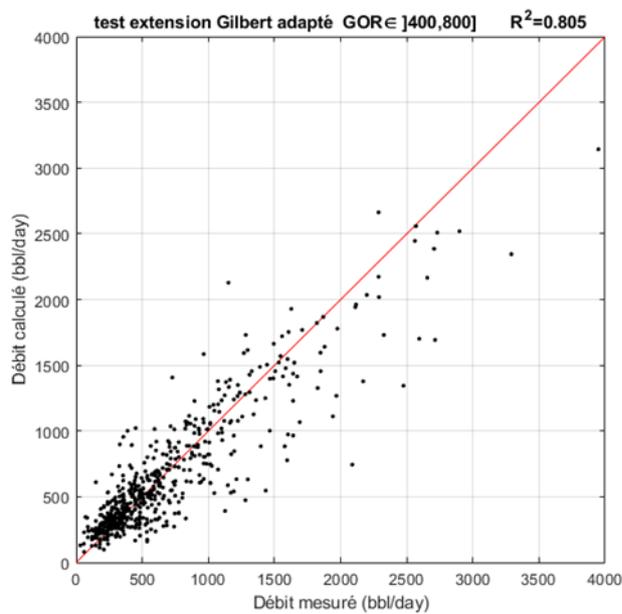


Figure An.B 44 : cross plot extension Gilbert adapté sur  $GOR \in ]400,800]$ , régime sous critique

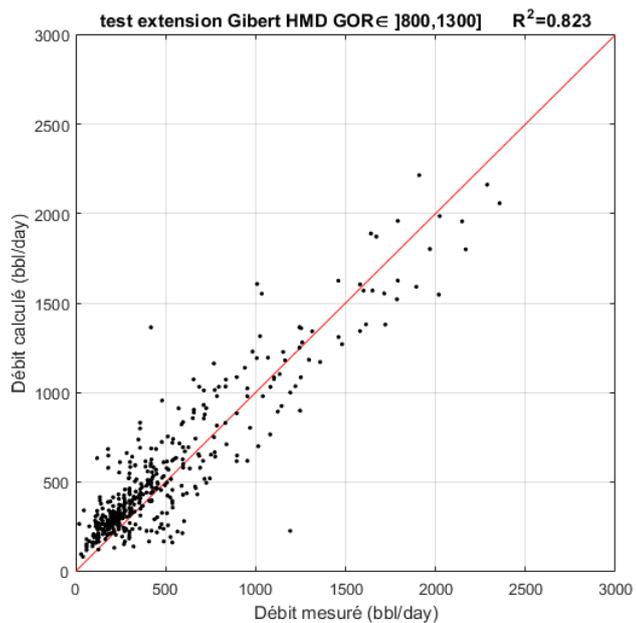


Figure An.B 45 : cross plot test extension Gilbert HMD sur  $GOR \in ]800,1300]$ , régime sous critique

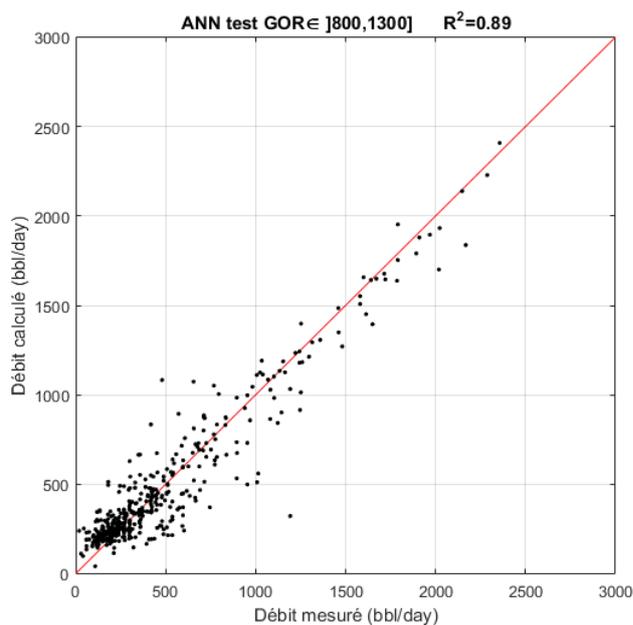


Figure An.B 46 : test ANN sur  $GOR \in ]800,1300]$ , régime sous critique

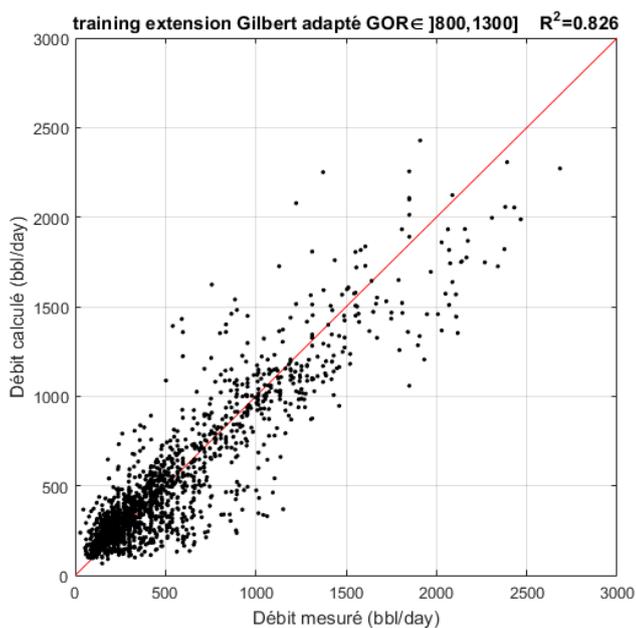


Figure An.B 47 : cross plot training extension Gilbert adapté sur  $GOR \in ]800,1300]$ , régime sous critique

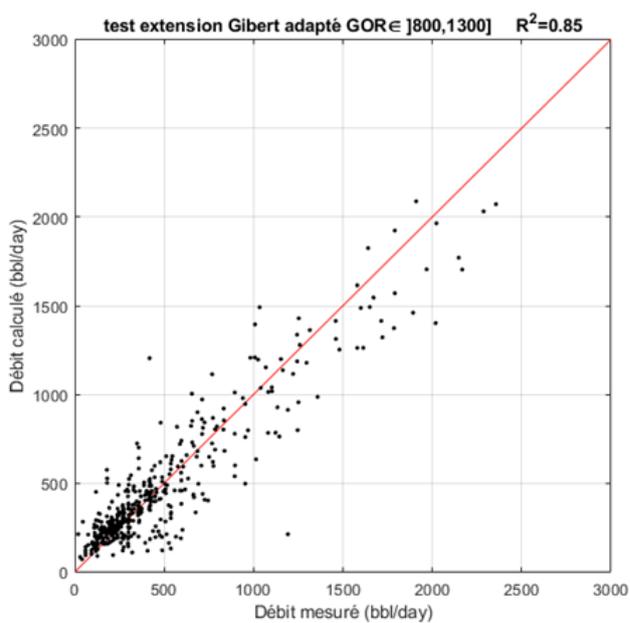


Figure An.B 48 : cross plot test extension Gilbert adapté sur  $GOR \in ]800,1300]$ , régime sous critique

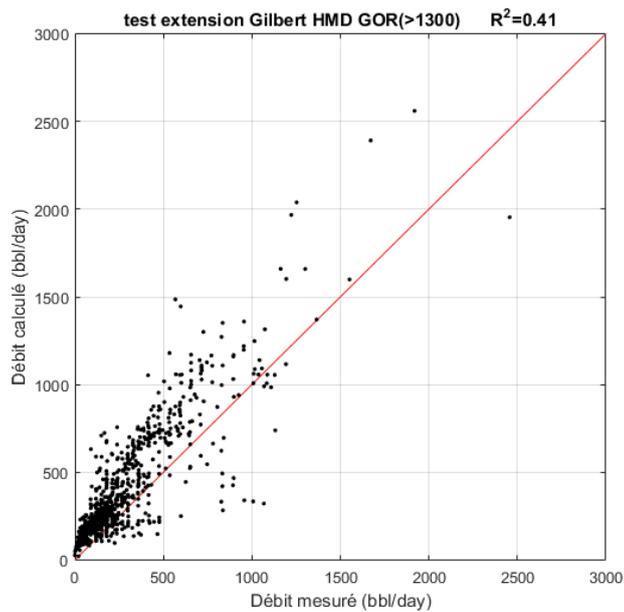


Figure An.B 49 : cross plot test extension Gilbert HMD sur GOR>1300, régime sous critique

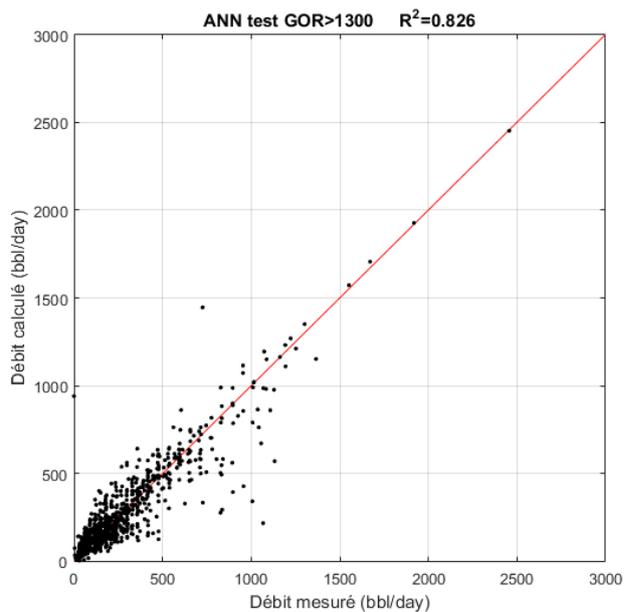


Figure An.B 50 :cross plot test ANN sur GOR>1300, régime sous critique

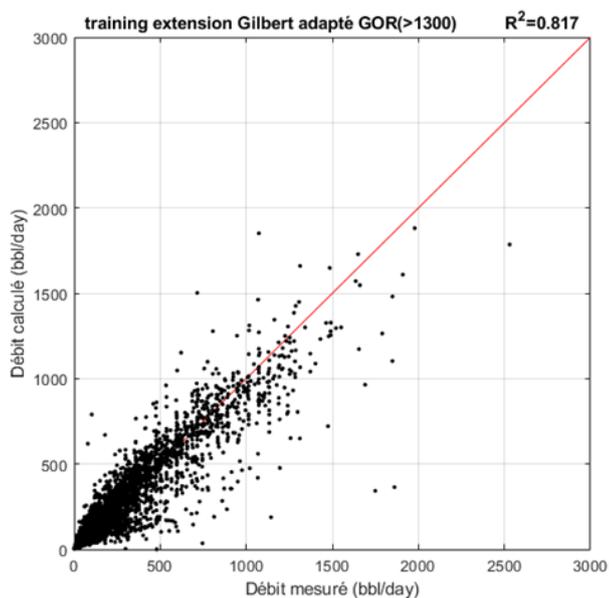


Figure An. B 51 : cross plot training extension Gilbert adapté sur GOR>1300, régime sous critique

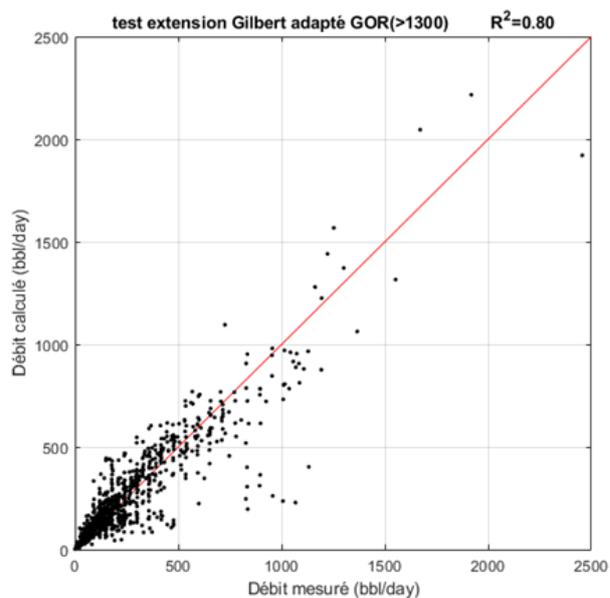


Figure An. B 52 : cross plot test extension Gilbert adapté sur GOR>1300, régime sous critique

## Annexe C : cross plot et équations de la programmation génétique

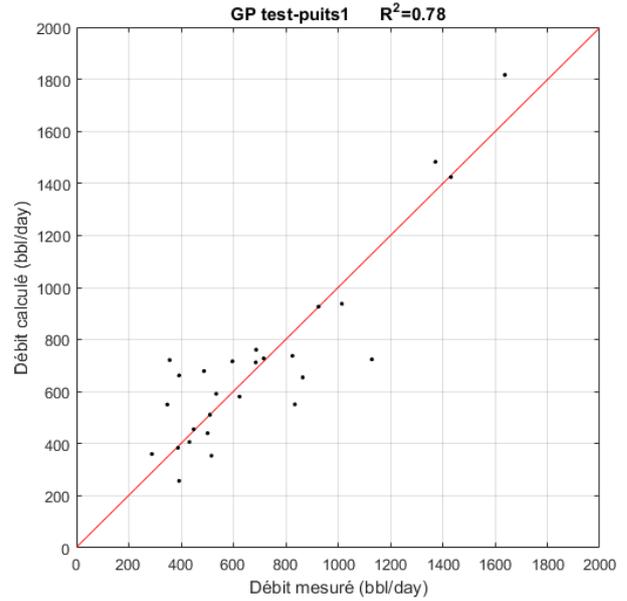
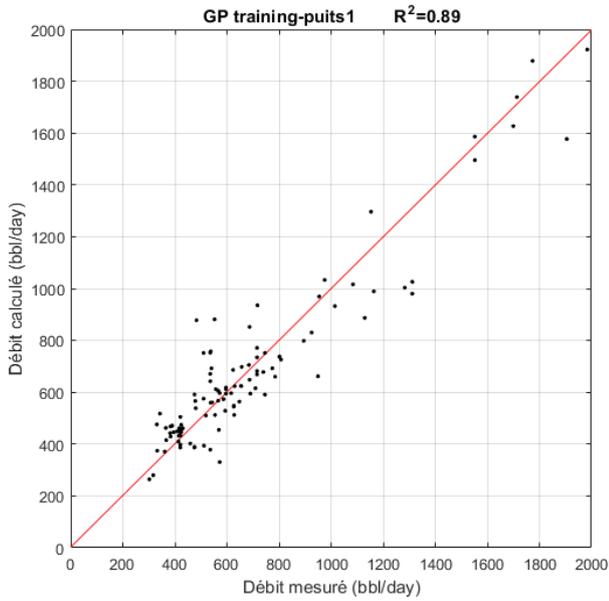


Figure An. C 1 : cross plot training GP de puits 1 avec les variables ( $P_u, d, GOR$ )

Figure An.C 2 : cross plot test GP de puits 1 avec les variables ( $P_u, d, GOR$ )

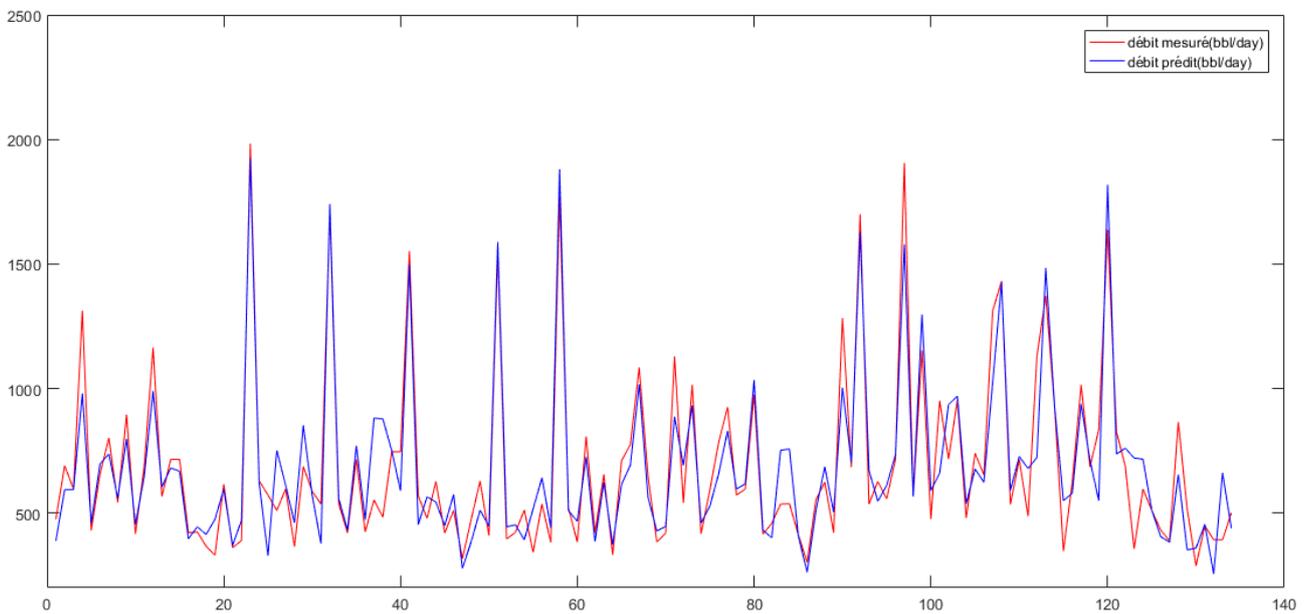


Figure An.C 3 : Débit mesuré vs débit prédit en (bbl/day) du puits1

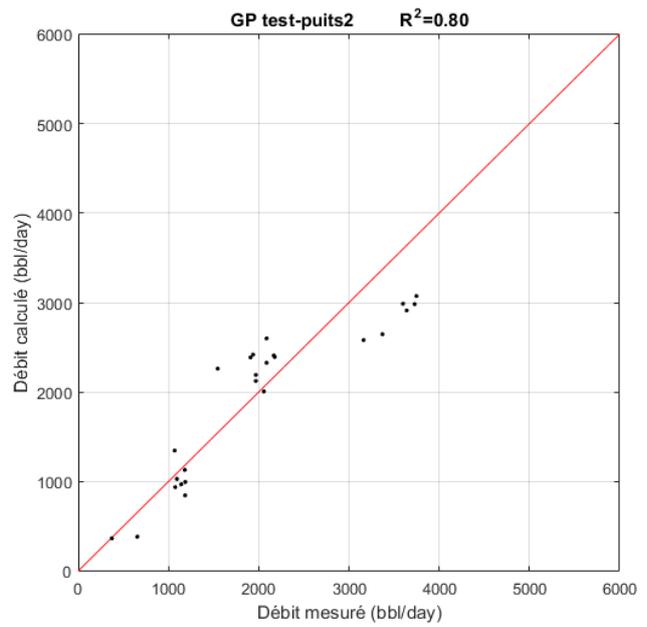
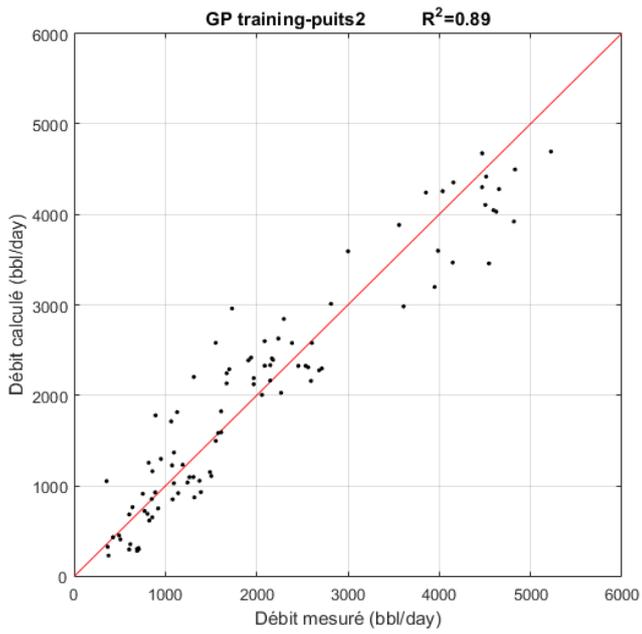


Figure An.C 4 : cross plot training GP de puits 2 avec les variables ( $P_w, d, GOR$ )

Figure An.C 5 : cross plot test GP de puits 2 avec les variables ( $P_w, d, GOR$ )

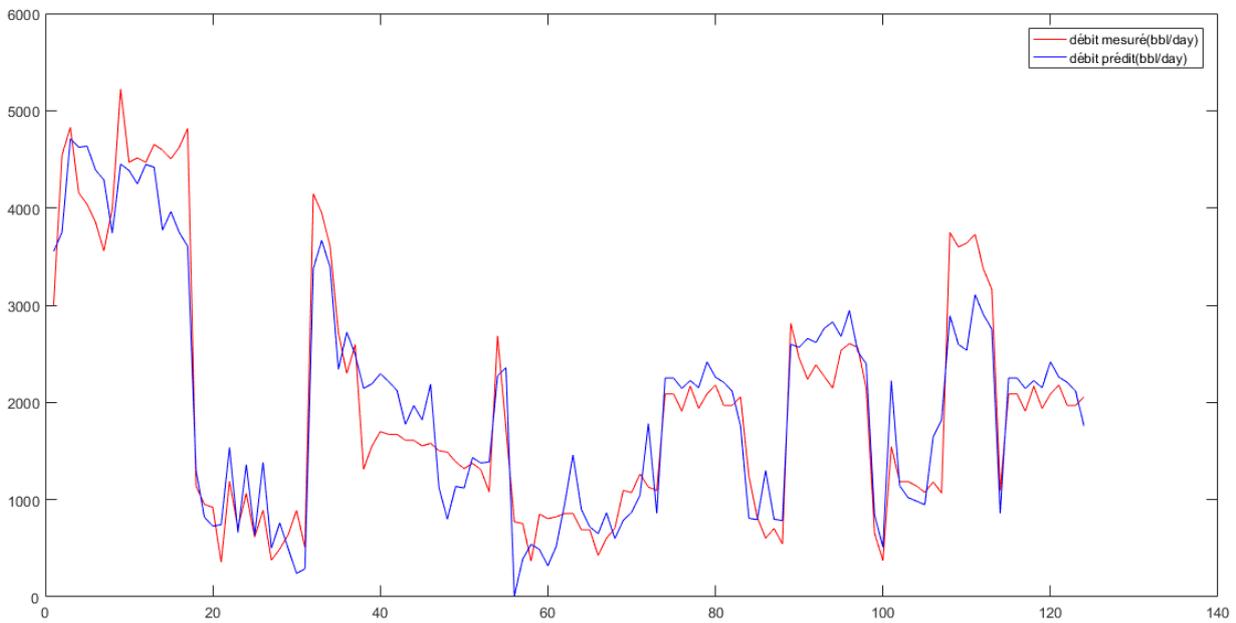


Figure An.C 6 : Débit mesuré vs débit prédit en (bbl/day) de puits 2

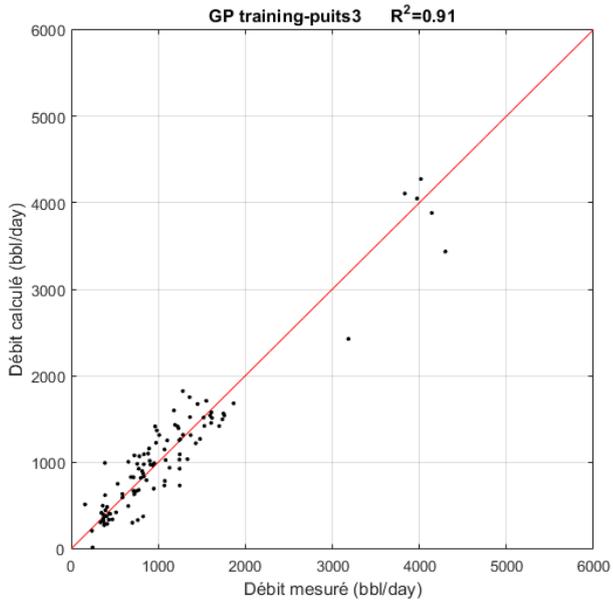


Figure An.C 7 : cross plot training GP de puits 3 avec les variables ( $P_w, d, GOR$ )

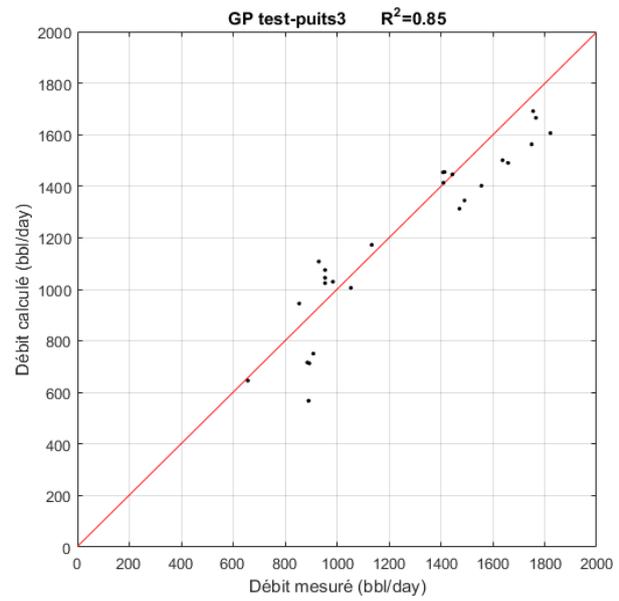


Figure An.C 8 : cross plot test GP de puits 3 avec les variables ( $P_w, d, GOR$ )

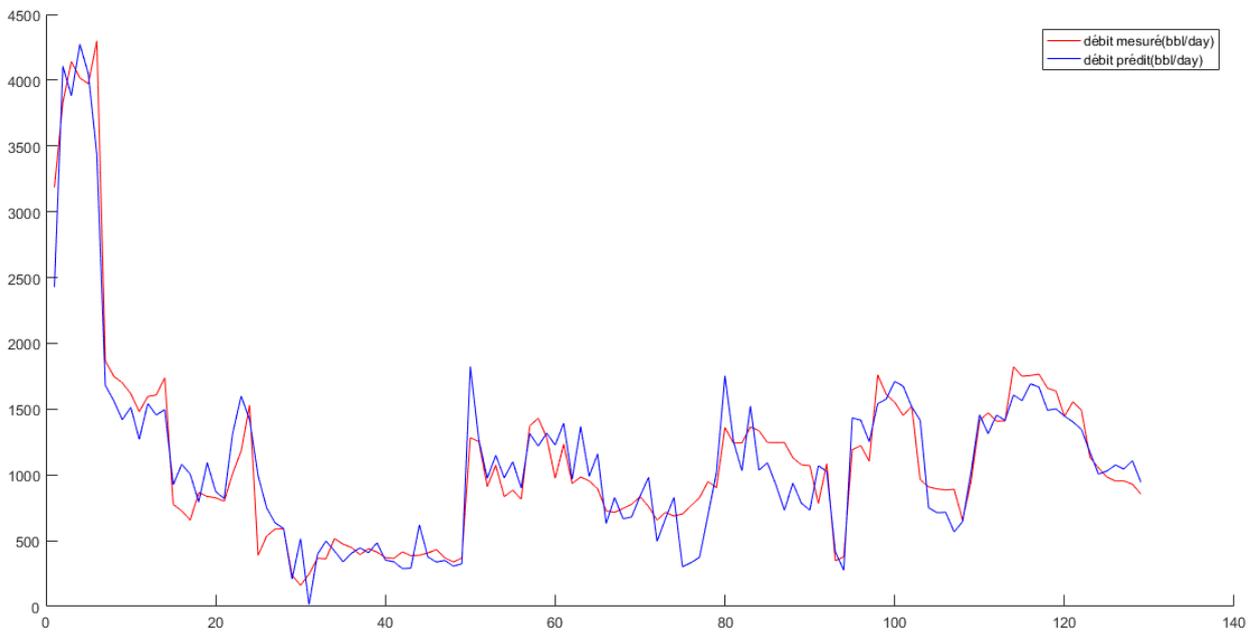


Figure An.C 9 : Débit mesuré vs débit prédit en (bbl/day) de puits 3

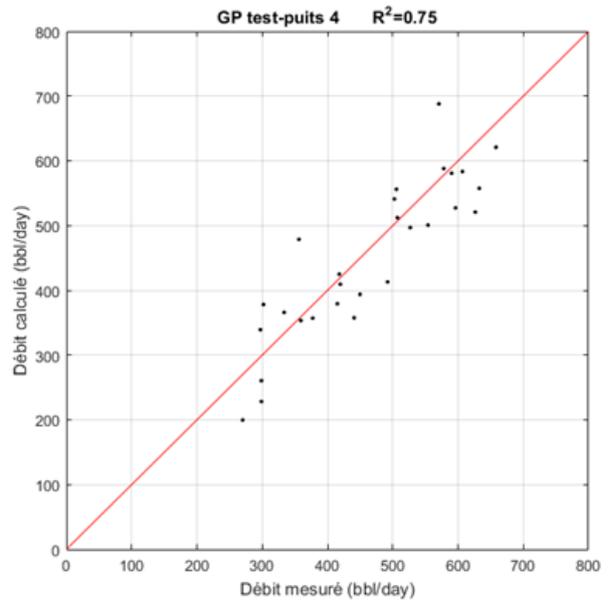
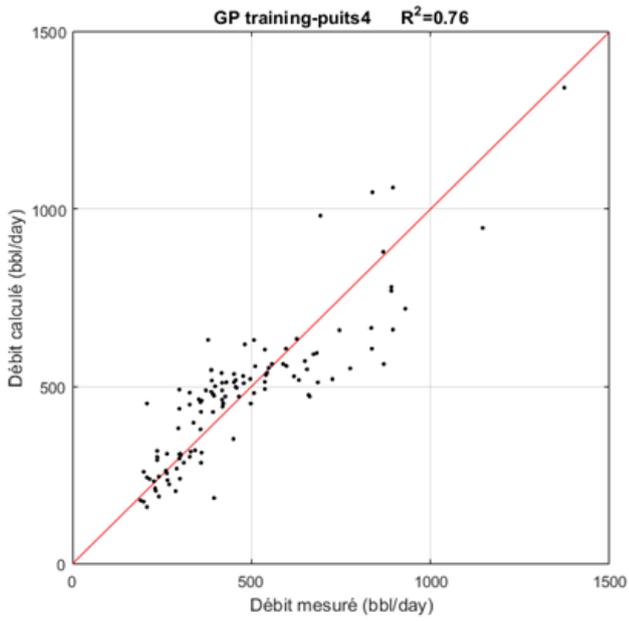


Figure An.C 10 : cross plot training GP de puits 4 avec les variables ( $P_w, d, GOR, P_d, \Delta P$ )

Figure An.C 11 : cross plot test GP de puits 4 avec les variables ( $P_w, d, GOR, P_d, \Delta P$ )

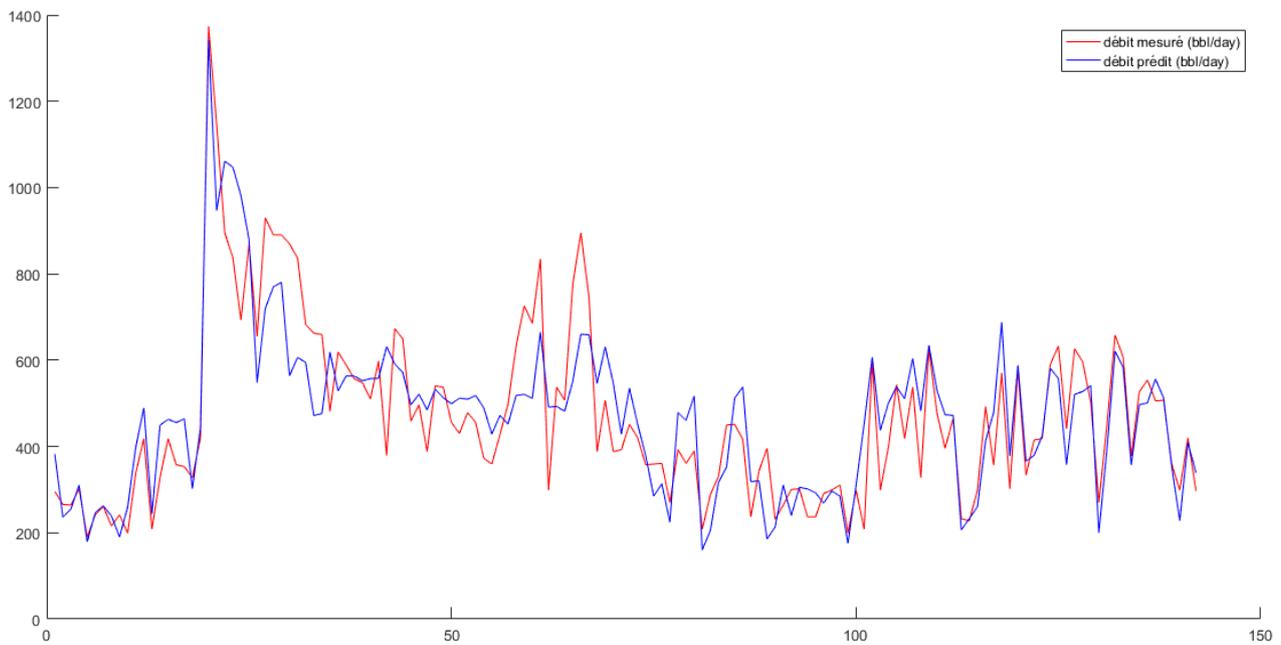


Figure An.C 12 : Débit mesuré vs débit prédit en(bbl/day) de puits 4

Tableau An. C12 : les équations générées par GP pour puits1

PUITS	Variables	Equation
PUITS1	$x1 = P_u$ $x2 = d$ $x3 = GOR$	$Q=0.38389376760659726928182067240414*x1*x2 - 831.11029756546440694364719092846*\cos(\cos(\sin(x2^2))) - 109.34703082070831214878126047552*\sin(\sin(\cos(x2))) - 79.134220633313944404108042363077*\cos(x1) - 109.34703082070831214878126047552*x2 - 79.134220633313944404108042363077*(\cos(x2) + x3^{(1/2)})^{(1/2)} - 1328.0743083513358486329062231143*(x1 - 6.3534)^{(1/4)} + 0.38389376760659726928182067240414*x1*\cos(x2) + 0.76778753521319453856364134480827*x2*\cos(x1) + 0.76778753521319453856364134480827*x(1)^2 + 6929.8366655034496943699195981026$
	$x1 = P_u$ $x2 = d$ $x3 = GOR$ $x4 = P_d$	$Q=67.560810537164954325817234348506*\cos((x1 + x2)^{(1/2)}) - 0.44657516362670168774684498202987*x4 - 67.560810537164954325817234348506*x2 - 78.830287028876753652184561360627*\cos(x2 - \cos(x2)) + 204.90621627322136077964387368411*\sin(\sin(\sin(2.4153*x2))) - 0.44657516362670168774684498202987*(3.0996*x4 + x3*x4)^{(1/2)} + 0.44657516362670168774684498202987*x1*x2 - 210.96187703350946662794740404934*(x1 + 2*x2 + \cos(x2 + x4))^{(1/2)} + 3822.8027282011853458243422210217$
	$x1 = P_u$ $x2 = d$ $x3 = GOR$ $x4 = P_d$ $x5 = \Delta P$	$Q=0.63708374338211548604249401250854*x2*x5 - 16.298834752279510951211705105379*x2 - 8.1494173761397554756058525526896*\cos(x2) - 8.1494173761397554756058525526896*x1 - 0.68283231940427757677980480366386*x1*\cos(x2) - 1.3656646388085551535596096073277*x5*\cos(x2) + 0.024159737495879195917636295121156*x4^2 + 0.0079172517761023630811223483760841*x1*x5*\cos(x2) + 0.69074957118037993986092715203995*x5*\cos(9.5957*x5)*\cos(x2) + 1325.0892316327126536634750664234$
	$x1 = P_u$ $x2 = d$ $x3 = GOR$ $x4 = P_d$ $x5 = \Delta P$ $x6 = r$	$Q=0.60365378445850443353826175751919*x1 - 0.11623479502031509869031111747972*\sin(x5^2*\cos(x6)) - 235.01215452160687391369719989598.*\cos(x2) + 0.082513883670745385989775400048529*x1^2*x2^{(1/8)} - 0.11623479502031509869031111747972*x1^2 - 5.4226749379781553628276924428064*x2*x3^{(1/4)}*x6 + 0.000047158745391964072217038522749632*x1^2*x3^{(1/2)}*\cos(x2) + 781.83525355075611687346216246985$

Tableau An.C13 : les équations générées par GP pour le puits2

PUITS	Variables	Equation
PUITS2	$x1 = d$ $x2 = P_u$ $x3 = GOR$	$Q=14.601750790162130400062778790016*x1 +$ $3.6504376975405326000156946975039*x2 +$ $0.011708307258575364739601291919069*x1*x2 +$ $0.005932391188588736885675591764544*x1*x3 -$ $0.0036483741132791337041796353446443*x1^2*x2 +$ $0.0000034954761189721471353495860223681*x1*x3^2 -$ $0.00088533634502691266159046801220711*x1^2*x3 +$ $3.6504376975405326000156946975039*x1^2 -$ $0.000078237559301054515874945805009588*x2^2 -$ $0.000078237559301054515874945805009588*x3^2 +$ $0.0000017477380594860735676747930111841*x1*x2*x3 +$ $0.0014789719309713854171473013110472$
	$x1 = d$ $x2 = P_u$ $x3 = GOR$ $x4 = P_d$	$Q=4.5505477407496799908237505860598*x1 -$ $0.30919464097367105166469514188066*x3 +$ $20.270274734449501922028957778821*x4 +$ $(7311543613544067*(x1-x4*(x1 + x4 + 9.5584)))/$ $288230376151711744 + (2057459203254213*x3*(x1 *x3 -$ $2*x3 + (x4 + 9.5584)^2))/9444732965739290427392 -$ $(6775857472448609*x1*(x3 - 9.5584*x4)*(x1 - x2 +$ $x1^2))/1180591620717411303424 +$ $20.271543228183963066803621733136$
	$x1 = P_u$ $x2 = d$ $x3 = GOR$ $x4 = P_d$ $x5 = \Delta P$	$Q= (40344859931235915*x3/72057594037927936 -$ $(6577759576987677*x1)/18014398509481984 +$ $(56735664620178433*x4)/18014398509481984 -$ $(12539476260797689*x5)/9007199254740992 -$ $(7348873539244181*x1*x3)/288230376151711744 +$ $(3056784180760387*x1*x4)/9007199254740992 -$ $(7348873539244181*x1*x5)/144115188075855872 -$ $(97895279692558651*x2*x4)/288230376151711744 +$ $(50114263864302957*x4*x5)/144115188075855872 +$ $(2904932503049625*x4^2)/9007199254740992 +$ $2540285618479433/2199023255552$
	$x1 = P_u$ $x2 = d$ $x3 = GOR$ $x4 = P_d$ $x5 = \Delta P$ $x6 = r$	$Q=8.4211929094764226277902707806788*x1-$ $0.017912459668608344404150400919207*x2 -$ $0.40565599606006906216459917402517*x3 -$ $0.017912459668608344404150400919207*x5 +$ $0.017912459668608344404150400919207*x6 -$ $0.0000188226924122028002389281353679*x1*x2 +$ $0.017912459668608344404150400919207*x1*x4+$ $0.000028234038618304200358392203051849*x2*x3+$ $0.000028234038618304200358392203051849*x2*x4+$ $16.630601591088506552296166773885*x1*x6+$ $0.0000188226924122028002389281353679*x3*x4+$ $16.630601591088506552296166773885*x2*x6 -$ $0.0000094113462061014001194640676839498*x1*x2^2 +$ $0.0000094113462061014001194640676839498*x3^2 -$ $0.017903048322402243004030936851523*x4^2 -$ $0.0000094113462061014001194640676839498*x1*x2*x4 -$ $27.487876526373259810087645860222$

Tableau An.C14 : les équations générées par GP pour puits3

PUITS	Variables	Equation
PUITS3	$x1 = d$ $x2 = P_u$ $x3 = GOR$	$Q = (7885386247472217 * x1) / 1152921504606846976 - (10100802517882265331 * x2) / 295147905179352825856 + (7885386247472217 * x3) / 288230376151711744 + (18175835535574787431 * x1 * x2) / 73786976294838206464 - (2016156172313611695 * x1 * x3) / 147573952589676412928 - (2502707039275857 * x2 * x3) / 73786976294838206464 - (2206109351694395 * x1 * x2^2) / 18446744073709551616 + (37800456666386177 * x1^2 * x2) / 147573952589676412928 + (145852781643477106001 * x1^2) / 295147905179352825856 + (28976019259608597 * x2^2) / 73786976294838206464 + (2502707039275857 * x3^2) / 295147905179352825856 + 49702387287401 / 2305843009213693952$
	$x1 = d$ $x2 = P_u$ $x3 = GOR$ $x4 = P_d$	$Q = 0.33616774721768560851842266856693 * x1 + 1.2099040117833614457587554369411 * x2 - 0.015938211829856348253987556747545 * x3 + 1.2020225368876881047034555300898 * x4 - 0.0079826476103254471322623331275281 * x1 * x3 + 0.33619483060848015452895977807444 * x1 * x4 - 0.0000072500778441337430498740516948697 * x2 * x3 - 0.00039202006866155645215671343706276 * x3 * x4 - 0.00039202006866155645215671343706276 * x1 * x4^2 + 0.0000072500778441337430498740516948697 * x3^2 - 0.00072987335573402088323920785910281 * x4^2 + 0.0000072500778441337430498740516948697 * x4^3 - 0.0000072500778441337430498740516948697 * x1 * x2 * x4 + 0.0000072500778441337430498740516948697 * x1 * x3 * x4 + 0.000038755132015785731416750159450402$
	$x1 = P_u$ $x2 = d$ $x3 = GOR$ $x4 = P_d$ $x5 = \Delta P$	$Q = 14.080425531651594339113860165913 * x1 - 0.90923493083829706635076370324633 * x2 - 0.13410228644916732942959924912429 * x3 + 3.486645456864340353320130816428 * x4 + 1.433660723970758085465959652538 * x5 - 0.000012424277177066898980952770747965 * x1 * x3 + 0.13410228644916732942959924912429 * x1 * x4 + 0.000012424277177066898980952770747965 * x1 * x5 + 0.0012529344580591591853295854730455 * x2 * x4 - 0.00018327958151242830096300663166725 * x3 * x4 - 0.0000062121385885334494904763853739826 * x3 * x5 + 0.000012424277177066898980952770747965 * x1^2 * x4 + 0.0000062121385885334494904763853739826 * x3^2 + 0.00012307898887866834864069807628506 * x4^2 - 0.0000062121385885334494904763853739826 * x1 * x3 * x4 + 33.791948739897220593081448962447;$
	$x1 = P_u$ $x2 = d$ $x3 = GOR$	$Q = (804936026889913227917 * x1) / 1180591620717411303424 + (12905376785416909 * x2) / 36028797018963968 + (6141107884732403 * x4) / 9007199254740992 + (6141107884732403 * x1 * x4) / 18014398509481984 - (8734222267701901 * x2 * x3) / 590295810358705651712 - (623161015952103 * x3 * x6) / 36028797018963968 + (12905376785416909 * x1^2) / 36028797018963968 +$

$x4 = P_d$ $x5 = \Delta P$ $x6 = r$	(3063945402260055*x1*x2*(x1 + x2+ x4)) /295147905179352825856 - (8734222267701901*x1^2*x2*x4/1180591620717411303424 + (3063945402260055*x3*x6*(x1 + x3 - x5)/295147905179352825856 - (623161015952103*x1*x3*x6)/36028797018963968 + 3988106736895773/73786976294838206464
---	---

Tableau An.C15 : les équations générées par GP pour puits 4

PUITS	Variables	Equation
PUITS4	$x1 = d$ $x2 = P_u$ $x3 = GOR$	Q=0.62681269427576099100146711506909*x2 - 2.8842512275918947631248023810713*x1 - 0.050859788125222162025185212996803*x3 - 0.00000022807312240518609769508099997803*x1*x2 - 0.0000019991293398181777021266934891075*x1*x3 - 0.000025734509472741215807353988434545*x2*x3 + 1.6293560200035699760893486151634*x1^2 + 0.0000017710562174129916044316124891294*x3^2 + 101.30595929055274049407839309644
	$x1 = d$ $x2 = P_u$ $x3 = GOR$ $x4 = P_d$	Q=47.401825993515352308804722269997*x1 + 0.61691547274690350377568393014371*x2 - 0.079713634375867759018774449941702*x3 - 0.80573775968110964328781925134979*x4 - 0.000016380828661954892038333901627123*x2*x3 + 0.0000027301381103258153397223169378538*x3*x4 + 0.0000027301381103258153397223169378538*x2^2 + 0.0000027301381103258153397223169378538*x3^2 + 7.3237264554104495317687906208448
	$x1 = P_u$ $x2 = d$ $x3 = GOR$ $x4 = P_d$ $x5 = \Delta P$	Q=(1612150465096827*x1/35184372088832 - (5801489702263749*x3)/72057594037927936 + (5801489702263749*x5)/36028797018963968 + 0.00014093671885040353338272058059122*x4*(x1 - x2 - x4 + x1*x5) - (5660417586140487*(x2 - x3)*(x3 - x5))/2361183241434822606848 + 3.0396501438577049395775575568734
	$x1 = P_u$ $x2 = d$ $x3 = GOR$ $x4 = P_d$ $x5 = \Delta P$ $x6 = r$	Q=(2267025282017101*x4/9007199254740992 - (2267025282017101*x3)/18014398509481984 - (2267025282017101*x1)/18014398509481984 - (8860853784682997*x1*(2*x1 + x3))/36028797018963968 - (7945128531450207*(x1 - x4)*(2*x1 + x3 - x4))/147573952589676412928 + (4210656503132709*x1*(x1 + 2*x5 + x6 - (x2 - x5)*(x4 - x5)))/36893488147419103232 - (8860853784682997*x1*(2*x1 - x4))/36028797018963968 + (2267025282017101*x1*(2*x3 - 2*x4)) /18014398509481984 + 213441617514503/274877906944

