

Prédiction des Paramètres Physiques des Couches Pétrolifères par Analyse des Réseaux de Neurones et Analyse Faciologique.

Rafik Baouche

▶ **To cite this version:**

Rafik Baouche. Prédiction des Paramètres Physiques des Couches Pétrolifères par Analyse des Réseaux de Neurones et Analyse Faciologique.. Sciences de la Terre. université M'hamed Bougara. Boumerdès, 2015. Français. NNT : . tel-01150323

HAL Id: tel-01150323

https://insu.hal.science/tel-01150323

Submitted on 12 May 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N° Ordre...../Faculté/UMBB/2014

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE M'HAMED BOUGARA-BOUMERDES



FACULTE DES SCIENCES PHYSIQUES
LABORATOIRE LIMOSE

THESE de DOCTORAT

Présentée par

BAOUCHE RAFIK

Filière **: Sciences Physiques**

Option : **Physique de la couche pétrolifère**

Prédiction des **P**aramètres **P**hysiques des **C**ouches **P**étrolifères par **A**nalyse des **R**éseaux de **N**eurones et **A**nalyse **F**aciologique.

Devant le jury:

A. MEZGHICHE	Université M'hamed Bougara, Boumerdès	Président
K. BADDARI	Université de Bouira	Directeur de thèse
T. AIFA	Université de Rennes 1	Co-Directeur de thèse
R. CHAOUCHI	Université de M'hamed Bougara, Boumerdès	Examinateur
H. SHOUT	Université Mentouri, Constantine	Examinateur
N. DJARFOUR	Université M'hamed Bougara, Boumerdès	Examinateur

Année Universitaire : 2014/2015

Je dédie cette thèse à mes parents, à mes enfants, à ma femme qui m'a soutenu et pour sa patience, à ma sœur Souhila et à tous ceux que j'aime

Il est bien plus beau de savoir quelque chose de tout que de savoir tout d'une chose. Pascal, Blaise (1623-1662)

Je crois que l'avenir de l'humanité est dans le progrès de la raison par la science. Zola, Émile (1840-1902)

Remerciements

Ce travail n'aurait jamais vu le jour sans l'aide de personnes auxquelles je veux adresser mes sincères remerciements.

Je voudrais en premier lieu exprimer tous mes remerciements aux Professeurs :

Kamel Baddari, pour m'avoir donné l'opportunité d'effectuer ce travail dans un domaine qui m'a toujours passionné et pour m'avoir accueilli dans son équipe et pour qui sans lui, ce travail n'aurait jamais pu être réalisé **ainsi qu'à mon Co-encadrant** :

Tahar Aïfa pour m'avoir encadré tout au long de ce travail, qu'il trouve ici l'expression de ma plus profonde gratitude pour m'avoir soutenu et encouragé dans les moments difficiles.

Je le remercie également pour les lectures attentives qu'il a faites de ce travail aux différentes étapes de sa réalisation et aux bons conseils qu'il m'a prodigué lorsqu'un problème ou un autre me bloquait.

Tous mes remerciements vont également à **Mrs Mezghiche**, **Mr Chaouchi**, **Mr Shout** et **Mr Djarfour** membres du jury pour m'avoir fait l'honneur d'examiner ce travail avec un regard empli d'expérience.

Mohand Kessal pour m'avoir aidé à me réinscrire dans cette branche, qu'il trouve ici tous mes remerciements, sans oublier **Mme Messad** qui m'a apporté son soutien moral.

Mes remerciements vont également à tous ceux qui m'ont aidé à concrétiser ce travail en m'accueillant dans le laboratoire **« LIMOSE »** auquel je dois tout et principalement le Directeur **Mr Mezghiche**, avec qui j'ai eu beaucoup d'entretiens constructifs sur le plan recherche en soft computing et m'a encouragé tout au long de ce travail.

Sans oublier Mr Le vise doyen de la Faculté des Sciences **Mr Djillali** qui m'a beaucoup aidé dans les formalités de la soutenance, ainsi que tout le service de la PG.

Enfin, je remercie mes amis ***Karim Allek, Khaled Loumi****, et beaucoup d'autres pour m'avoir assisté à chaque fois que j'en avais besoin, leur présence à mes côtés est sans doute le plus beau témoignage d'amitié. Je me souviendrai toujours.*

LISTE DES PUBLICATIONS

A. Articles dans des revues internationales

1. Baddari K., M. Djeddi, R. Baouche, 1994. Analyse statistique des paramètres pétrophysiques des réservoirs carbonatés du Sud-Est Constantinois. *Bulletin du service Géologique de l'Algérie*, 5(2), *167-187.*
2. Baouche R., A. Nedjari, 2004. The use of t*he Modular* Dynamic Tool in petrophysical parameters evaluation: application to the Bir-Berkine reservoirs – Algeria. *Notes et Mémoires serv. Géol. Maroc*, 527, 23-30.
3. Baouche R., A. Nedjari, 2006. Petrophysical Analysis in Reservoir Characterization. Application in the Triassic Hamra Gas Field Algeria, 25, 18-24.
4. Baouche R., A. Nedjari, R. Chaouchi, S. El Adj, 2009. A sedimentological approach to refining reservoir architecture using the well log data and core analysis in the Saharan Platform of Algeria. *Journal Wseas Transactions on Environment and Development*, 5(8), *178-188.*
5. Baouche R., Nedjari A., El Aadj S., Chaouchi R., (2009) « Facies Analysis of Triassic Formations of the Hassi R'Mel in Southern Algeria Using Well Logs: Recognition of Paleosols Using Log Analysis ». *The Open Geology Journal*, 3, 39-57.
6. Baouche R., K. Baddari, 2011 (*soumis*). Using the petrophysical logs and Circumferential Borehole Image Log to Naturally fractured reservoirs. A case study. El Agreb Field, Algeria. *Journal of American Association of Petroleum Geologist.*
7. Baouche R., A. Nedjari, R. Chaouchi, *2012.* Analysis and Interpretation of Environment Sequence Models in the Triassic Province of Algeria. *Journal, The Geology of Southern Libya, 2,.301-318.*
8. Aïfa T., R. Baouche, K. Baddari, 2014. Neuro-Fuzzy system to predict Permeability and Porosity from well log data: a case study of Hassi R'Mel Gas Field, Algeria. Journal of Petroleum Science and Engineering, 123, 217-229.
9. Baouche R., K. Baddari, 2014 (*soumis*). Intelligent systems for prediction of Nuclear Magnetic Resonance porosity and permeability from conventional well log data: A case study. Sahara Fields. Algeria. Arabian Journal of Geosciences.
10. Baouche R., K. Baddari, 2014 (*soumis*). Facies Analysis and Permeability/ porosity Prediction from well log data using the Non Parametric Regression with Multivariate Analysis and Neural Network in the_Reservoirs of the Hassi R'Mel Southern Field (Algeria). Journal of African Earth sciences.

B. Articles de conférences (2011 à 2014)

1. Baouche R, A. Nedjari, R. Ait Ouali, 2011. *Geological* Modeling of the Saharan Platform in Meso-Cenozoic. 3D reconstruction, cross sections and volume computation on surfaces Algeria. 5 th Symposium Applied Geophysics Maghreb. Algiers, April 12-14, 219-221.
2. Baouche R., K. Baddari, M. Djeddi, R. Chaouchi, 2012. Using the Oil-Base Micro Imager to characterize naturally Fracture Reservoirs in Hassi Messaoud Wells: OMKZ353, OMKZ341 and OMK441. Algeria. ISHC6, 13-15 octobre 2012, Algiers, 2, 8
3. Baouche R., K. Baddari, 2014. Facies Analysis and Permeability/Porosity Prediction from well log data using the Non Parametric Regression with Multivariate Analysis and Neural Network in the_Reservoirs of the Hassi R'Mel Southern Field (Algeria). The 7th International Symposium of hydrocarbons and Chemistry. Boumerdès, 05-07 Mai 2014, Algiers, 212-214.

Résumé

La caractérisation des *réservoirs* argilo-gréseux par les données de diagraphies est un moyen pratique de la description des réservoirs dans les champs pétroliers**.** Au cours des dernières années, plusieurs études ont été menées dans le domaine de l'ingénierie pétrolière en appliquant l'intelligence artificielle. Ce travail représente une méthode basée sur la pétrophysique qui utilise des diagraphies de puits et des données de modules de base pour prédire et enregistrer les données en profondeur dans les réservoirs argilo-gréseux de la formation du Trias dans le champ de Hassi R'Mel (Sahara algérien).

Dans l'étude des gisements de pétrole, la prédiction de la perméabilité absolue et de la porosité est un élément fondamental dans les descriptions de réservoirs ayant un impact direct sur les autres paramètres pétrophysiques, les programmes d'injection d'eau et la bonne gestion de réservoir d'une manière plus efficace. Les formations du Trias du champ de Hassi R'Mel sont composées de grès et de sable schisteux avec de la dolomie. Les enregistrements diagraphiques de 10 puits de ce champ sont le point de départ pour la caractérisation de son réservoir. Ce travail présente un modèle hybride "neuro-fuzzy" basé sur l'utilisation des données de diagraphies pour l'estimation de la porosité et de la perméabilité. Une approche de la logique floue (fuzzy logic) est utilisée pour comparer la perméabilité carotte et la perméabilité calculée à partir des réseaux de neurones ainsi que celles de la porosité, développées dans ce modèle sur la base des données disponibles au niveau des puits. La logique floue est utilisée pour le choix des meilleurs rapports de forage associés à la porosité et la base de données de perméabilité. Le réseau neuronal est utilisé comme méthode de régression non linéaire pour développer une transformation entre diagraphies de puits sélectionnés et mesures de porosité et de perméabilité. Cette technique de méthode intelligente est utilisée comme un outil puissant pour l'estimation des propriétés des réservoirs d'après les paramètres diagraphiques et dans les projets de développement pétrolier et de gaz naturel.

***Mots clés**:* Diagraphie; modélisation; perméabilité; porosité; logique floue; Hassi R'Mel.

Abstract

Characterization of the shaly sand reservoirs by well log data is a practical way of reservoir descriptions in the oil fields. During the last few years several studies were conducted in the field of petroleum engineering by applying artificial intelligence. This work represents a petrophysical-based method that uses well loggings and core plug data to predict well log data recorded at depth in shaly sand reservoir of the Triassic Formation in Hassi R'Mel field (Algerian Sahara). In the study of oil reservoirs, the prediction of absolute permeability is a fundamental key in reservoir descriptions which has a direct impact on, amongst others, effective completion designs, successful water injection programs and more efficient reservoir management. The Triassic Formations of the Hassi R'Mel field are composed of sandstones and shaly sand with dolomite. Logs from the 10 wells are the starting point for the reservoir characterization. This work presents a hybrid neuro-fuzzy model based on the use of well log data in porosity and permeability estimation. A fuzzy logic approach is used to calibrate the calculated permeability and core permeability and neural network was developed in this model based on data available in the field. Fuzzy analysis is based on fuzzy logic and is used to get the best related well logs with core porosity and permeability data. Neural network is used as a nonlinear regression method to develop transformation between the selected well logs and core measurements. Porosity and permeability are predicted in these wells using the linear regression and multilayer perceptron models are constructed. Their reliabilities are compared using regression coefficients for predictions in uncored sections. This method of intelligent technique is used as a powerful tool for reservoir properties estimation from well logs in oil and natural gas development projects.

***Keywords**:* Well log; modelling; permeability; porosity; neuro-fuzzy; Hassi R'Mel.

ملخص

خلال السنوات القليلة الماضية وقد أجريت العديد من الدراسات في مجال هندسة البترول من خلال تطبيق الذكاء الاصطناعي. ويمثل هذا العمل وسيلة القائم على البتروفيزيائية يستخدم جيدا والبيانات المكونات الأساسية للتنبؤ بشكل جيد البيانات المسجلة في العمق في خزان تســـجي رمل [شلي] من العصر الترياسي تشكيل في حقل حاسي الرمل تسجيل، فـــي دراسة المكامن النفطية، والتنبؤ النفاذية المطلقة هي المفتاح الأساسي في وصف الخزان، رالجـزائ ولها تأثير مباشر على، من بين أمور أخرى، تصاميم الانتهاء فعالة والبرامج حقن المياه ناجحة وإدارة المكامن أكثر كفاءة. وتتكون الترياسي تشكيلات حقول حاسي الرمل من الحجر الرملي و الرملية مع الدولوميت. السجلات من 10 بئرا هي نقطة البداية لتوصيف المكامن. تقدم هذه الورقة نموذجا العصبية غامض الهجين على أساس استخدام البيانات سجل جيد في المسامية والنفاذية التقدير. ويستخدم نهج المنطق الضبابي لمعايرة محسوبة النفاذية والأساسية النفاذية وتم تطوير الشبكة العصبية في هذا النموذج على أساس البيانات المتاحة في هذا المجال. ويستند التحليل غامض على المنطق الضبابي، ويستخدم لاختيار أفضل السجلات ذات الصلة جيدا مع المسامية الأساسية والبيانات النفاذية. يستخدم الشبكة العصبية كوسيلة من وسائل الانحدار غير الخطية لتطوير التحول بين سجلات جيدا مختارة وتالقياســـــــا. المســـاميةالأساسية ونفاذية من المتوقع في هذه الآبار باستخدام الانحدار الخطي ونماذج متعددة الطبقات المستقبلات هي التي شيدت وتتم مقارنة المصداقية من خلال معاملات الانحدار ل ويستخدم التنبؤات في محفور الامم المتحدة طريقة من تقنية ذكية كأداة قوية لخصائص المكمن تقدير من سجلات الآبار في مشروعات النفط وتطوير الغاز الطبيعي.

الكلمات الرئيسية **:** تصميم النفاذية;العصبية غامض; حفرالسجلات**;** حاسي الرمل

Table des matières

REMERCIEMENTS……..*iii*
RESUME……..*iv*
ABSTRACT……*vi*
TABLE DES MATIERES……*viii*
LISTE DES FIGURES……*xi*
LISTE DES TABLEAUX……*xiii*
ANNEXES……..xvi

Introduction générale ……..1
1. Problématique……..1
2. Objectifs de l'étude……2
3. Organisation du manuscrit……2

Chapitre I

Caractérisation des réservoirs triasiques du champ de Hassi R'Mel

I.1- Rappels sur le champ de Hassi R'Mel……5
I.2- Contexte géologique du gisement de Hassi R'Mel……6
I.3- Caractéristiques diagraphiques des réservoirs triasiques de Hassi R'Mel……7
I.4- Résultats des électrofaciès issus de l'analyse manuelle ……8
I.5- Le milieu de dépôt……9
I.5-1- Le milieu de dépôt……9
I.5-2- Environnements de dépôt……10
I.5-3- Synthèse sur la région d'étude……10

Chapitre II

Les réseaux de neurones

II.1. Rappel sur les réseaux de neurones ……12
II.2. Historique sur les réseaux de neurones……12
II.3. Les réseaux de neurones……13

II.3.1. Principe du neurone artificiel…….....13
II.3.2. Définition…….…….....16
II.4. Architecture des réseaux de neurones……17
II.4.1. Les réseaux de neurones non bouclés……17
II.4.2. Les réseaux de neurones bouclés……18
II.5. Apprentissage des réseaux de neurones…….20
II.5.1. Type d'apprentissage……..21
II.5.2. Algorithme d'apprentissage……23
II.6. Modélisation à l'aide des réseaux de neurones……...25
II.6.1. Modèle "boîte noire"…….....…..25
II.6.2. Modèle "boîte grise" ou hybride…….26
II.7. Conception d'un réseau de neurones…….....….....27
II.7.1. Détermination des entrées/sorties du réseau de neurones……..…….27
II.7.2. Choix et préparation des échantillons…….28
II.7.3. Elaboration de la structure du réseau……..28
II.7.4. Apprentissage…….....29
II.7.5. Validation et Tests……..29
II.8. Analogie de MLR avec le réseau de neurones……...30

Chapitre III

Les systèmes flous et neuro flous

IIIa. La logique Floue "Fuzzy logic**"**

IIIa. 1. Rappel sur la logique floue …….……..……33
IIIa. 2. Les principes de base de la logique floue ……..……..……..34
IIIa. 3. Système de logique floue …….……..…….....34
IIIa. 4. Logique floue: opérations de base …….……..…….…….....36
IIIa. 5. Méthode de fuzzification ……...……37
IIIa. 6. Méthode de Mamdani …….....….....39
IIIa. 7. Méthode de Sugeno ……...……40
IIIa. 8. Mécanismes de défuzzification …….....40
IIIa. 9. Conclusion sur la logique floue …….....……...42

IIIb. La technique "Neuro-Fuzzy"

IIIb.1. Rappel sur les systèmes neuro-flous....…….....…….....42

IIIb.2. Les systèmes flous……..…….....43
IIIb.3. Réseaux de neurones (Neural Networks)…….....…….....45
IIIb.4. Les systèmes neuro-fuzzy…….…….…….……45
IIIb.5. Types de Systèmes neuro-fuzzy …….....46
IIIb.6. Systèmes Cooperative "Neuro-Fuzzy"……...……..……..…47
IIIb.7. Systèmes Concurrent "Neuro-Fuzzy"……...…….…….……...47
IIIb.8. Systèmes Hybrid "Neuro-Fuzzy"……..……..….48
IIIb.9. Architecture "FALCON"…….....…….....49
IIIb.10. Architecture "ANFIS"……...…….....50
IIIb.11. Architecture "GARIC"…….……51
IIIb.12. Architecture "NEFCON"……...…53
IIIb.13. Architecture "EFuNN"……...…...53
IIIb.14. Discussion et Applications……..…..55
IIIb.15. Synthèse des résultats ……..…55

Chapitre IV

Application de la technique Neuro-Fuzzy à la prédiction de la Porosité et Perméabilité à partir des données de diagraphies de puits: cas du champ de Hassi R'Mel, Algérie.

IV.1. Rappel sur les réservoirs triasiques de *Hassi R'Mel Est*…….……....57
IV.2. Zones d'intérêt à huile…….....58
IV.3. Back propagation neural networks (BPNN)…….....59
IV.4. modèle Fuzzy logic (FL) …….. 61
IV.5. modèle Neuro-flou (NF) …….....66
IV.6. Données utilisées et méthodologie.…….…….....68
IV.6.1. Organisation des données……..…..70
IV.6.2. Normalisation des données …….....70
IV.6.3. Fuzzy clustering……..70
IV.6.4. Construction du Système Fuzzy Inférence (FIS)……72
IV.7. Résultats et Discussion…….…...85
IV.8. Analyse et Interprétation des résultats obtenus…….…….....94

Chapitre V

Utilisation des Méthodes Multivariées et les réseaux de neurones pour la prédiction des Faciès

V.1. Application de l'analyse Multivariée dans l'estimation des Faciès.....96
V.2. Cas de la Formation Hassi R'Mel Sud……..97
V.3. Méthodologie utilisée……..……..98
V.4. Données disponibles……..…….…….....100
V.5. Analyse des résultats obtenus…….....101
V.5.1. Analyse des électrofaciès.....….....101
V.5.1.1. Analyse en composantes principales…….....103
V.5.1.2. Analyse des clusters…….....….....107
V.5.1.3. Analyse discriminante…….……..108
V.5.2. Corrélation de la perméabilité…….....….....110
V.5.2.1. General Additive Model (GAM)…….....…….……..110
V.5.2.2. Réseaux de neurones feed forward (Nnet).....…….112
V.5.2.3. Perméabilité dans les puits et HFU (Hydraulic Flux Unit)…….……..114
V.5.3 Classification basée sur les lithofaciès**.**.....…….….118
V.6. Comparaison à d'autres techniques de perméabilité prévisionnelle.....….....123
V.7. Analyse des méthodes utilisées……..…….....…...126

Conclusion générale…….....126
Perspectives de recherches et travaux envisagés…….....….....127
Références bibliographiques…….....130
Appendix A & Nomenclature……...……137
Annexes……138

LISTE DES FIGURES

Figure 1. Carte de situation du champ de Hassi R'Mel…….……..7
Figure 2. Expressions des séquences d'ordre 2....…….……11
Figure 3. Neurone artificiel……..…….....14
Figure 4. Différents types de fonction de transfert pour le neurone artificiel……15
Figure 5. Réseau de neurones à n entrées, une couche de Nc : neurones cachés et N_0 : neurones de sortie…….17
Figure 6. Réseau de neurone bouclé……..19
Figure 7. Erreur moyenne sur la base d'apprentissage en fonction du nombre d'itérations……..21

Figure 8. Diagramme schématique d'un modèle neuronal " boîte noire "…….26
Figure.9 Organigramme de conception d'un réseau de neurones…….....….....30
Figure.10 Model de logs synthétiques pour données de réseaux de neurones……...32
Figure.11 La méthode de contrôle-Analyse logique floue…….....35
Figure 12. Contrôleur en logique floue …….....36
Figure 13. Un exemple de fonction floue d'adhésion de la logique…….……..36
Figure 14. Interprétation graphique des opérateurs flous…….....38
Figure 15. Démonstration graphique de méthodes de "deffuzification"…..…….....41
Figure 16. Systems Coopérative……47
Figure 17. Systems Concurrent…….…….48
Figure 18. Architecture "FALCON"……..50
Figure 19. "ANFIS" architecture……...51
Figure 20. "GARIC" architecture……...52
Figure 21. "NEFCON" architecture…….…....53
Figure 22. "EFuNN. " architecture……54
Figure 23. Hassi R'Mel-Est (HRE) Lieu de champ montrant les principaux puits.....58
Figure 24. Model de règle "If-Then" (Matlab User's Guide, 2012)…….....61
Figure 25. (a) Architecture neuronale du classificateur NF ……..…….…..….63
Figure 26. Diagramme de procédure de formation dans un réseau de neurones supervisé……..……65
Figure 27. Représentation schématique du flux d'informations dans un système NF……...……67
Figure 28. "Substractive clustering" du model de fuzzy perméabilité ……..71
Figure 29. " Substractive clustering" du model de fuzzy porosité…….……..72
Figure 30. FIS Fonctions " Membership Gaussian " Générées par Fuzzy Inférence Systems pour le model données d'entrée pour la perméabilité……75, 76,77
Figure 31. FIS Fonctions " Membership Gaussian " Générées par Fuzzy Inférence systems pour le model données d'entrée pour la porosité……..78,79, 80
Figure 32. Porosité Carotte pour tous les puits de Hassi R'Mel …….....82
Figure 33. Porosité carotte de tous les puits de Hassi R'Mel……82
Figure 34. Membership fonctions pour la perméabilité…….83
Figure 35. Membership fonctions pour la porosité……83
Figure 36. Structure du modèle NF pour la perméabilité…….….84
Figure 37. Structure du modèle NF pour la porosité…….84
Figure 38. Density vs. Neutron Cross plot dans le réservoir Triasique de Hassi R'Mel.…….....85
Figure 39. Prévisions des épaisseurs des couches pour le puits: HR-167. U1, U2, M1, M2, L1, L2 et L3……87

Figure 40. Prédictions des Logs pour le puits HR-205. ……...88
Figure 41. Prédictions des Logs pour le puits HR-209…….……88
Figure 42. Prédictions des Logs pour le puits HR-199. ……89
Figure 43.Distributions Bin pour la porosité et perméabilité dans le réservoir HR'Mel……...90
Figure 44. Comparaison de la porosité (a) et la perméabilité (b) et des prédictions pour NN/NF concernant la porosité et la perméabilité noyau, respectivement…….…....92
Figure 45. Vérification d'erreur (erreur quadratique moyenne) pour la porosité (a) et la perméabilité (b) en utilisant la méthode NF……...93
Figure 46. Analyse des faciès dans le champ de Hassi R'Mel Sud…….…….99
Figure 47 Electrofaciès des Puits du Trias de Hassi R'Mel…….……102
Figure 48 Scatter plot de: GR, ΔT, RLLD, Rhob, RHOCMA et Nphi, Sw…….....…….104
Figure 49a Projection des variables sur le facteur plan (1x2) avec perméabilité κ_H….....…...…..105
Figure 49b Projection des variables sur le facteur plan (1x2) avec porosité CPor…….106
Figure 50a Analyse typologique des faciès d'environments (Root 1 vs Root 2)……107
Figure 50b Analyse typologique des faciès d'environments (Root 1 vs Root 3)……..108
Figure 51 Feed Forward Neural Net in multiple layers …….….113
Figure 52 Porosité mesurée et Porosité Prédite /Perméabilité vs. Profondeur. Puits HRS-7, Trias Supérieur. Neural Network Algorithme……..…….…..115
Figure 53 Plot Scatter : log RQI vs log PhiZ où HFU données. Puits: HRS-7……...……..…116
Figure 54 Prédictions de la perméabilité et de la porosité en fonction de la classification zonale pour le HRS-7: κ_H (perméabilité carotte)…….....….....…...117
Figure 55 Prédictions de la perméabilité et de la porosité en fonction de la classification zonale pour le HRS-8: κ_H (perméabilité carotte)…….....….....118
Figure 56a Plot Moyennes avec erreurs de porosité en fonction de la Lithologie……..……120
Figure 56b Plot Moyennes avec erreurs de perméabilité en fonction de la Lithologie...……...….121

LISTE DES TABLEAUX

Tab.1 - Légendes des figures employées pour les faciès de Hassi R'Mel (Trias)……...8
Tab.2. Réponses des diagraphies en face des faciès……9
Tab.3 - Descriptives Statistiques (Puits de Hassi R'Mel)……70
Tab.4 - Tableau montrant l'entrée (a) et la sortie (b) des paramètres de la fonction de "membership" dérivés par TS-FIS pour la porosité…….....73
Tab.5 - Tableau montrant l'entrée (a) et la sortie (b) des paramètres de la fonction de "membership" dérivés par TS-FIS pour la perméabilité…….....74

Tab.6 – Comparaison des coefficients de corrélation et RMSE pour les trois méthodes utilisées dans ce travail……..…….....94

Tab.7a – Erreurs quadratiques moyennes (MSE) et erreurs absolues moyennes (MAE) pour la porosité…….....100

Tab.7b – Erreurs quadratiques moyennes (MSE) et erreurs absolues moyennes (MAE) pour la perméabilité……101

Tab.8 – Corrélations des diagraphies et principaux composants pour la porosité……...102

Tab.9 – Corrélations des diagraphies et principaux composants pour la perméabilité…….……...103

Tab.10a – Composants des Vecteurs propres vs Variables diagraphiques pour la variable porosité (CPOR)…….105

Tab.10b – Composants des Vecteurs propres vs Variables diagraphiques pour la variable perméabilité (CPerm)…….....…….....106

Tab.11 – Classification préalable des données en sous-groupes relativement homogènes……..…109

Tab.12a – Les corrélations de vecteurs dans une matrice représentant les valeurs prédictives et la valeur prédite (Perméabilité)....…….....111

Tab.12b – Les corrélations de vecteurs dans une matrice représentant les valeurs prédictives et la valeur prédite (Porosité)……..……112

Tab.13 – Description des différents types de lithofaciès…….……....119

Tab.14 – Erreurs de classification pour lda (analyse linéaire discriminante) et qda (analyse quadratique discriminante) …….……...122

Introduction Générale

1 Problématique

Cette dernière décennie a connu des avancées significatives dans la transformation des données géoscientifiques et ainsi, dans les perspectives favorables, générer des modèles structurels précis et créer des modèles de réservoir avec les propriétés associées. Cela a été rendu possible par (i) l'amélioration de l'intégration des données, (ii) la quantification des incertitudes, (iii) l'utilisation efficace de la modélisation géophysique pour mieux décrire la relation entre les données d'entrée et les propriétés des réservoirs, et (iv) l'utilisation des méthodes statistiques non conventionnelles. Les techniques de "Soft Computing" telles que les réseaux de neurones et la logique floue et leur utilisation appropriée dans de nombreux problèmes géophysiques et géologiques ont joué un rôle clé dans les progrès réalisés ces dernières années. Cependant, il y a un consensus de l'opinion que nous avons seulement commencé à gratter la surface en réalisant pleinement des avantages de la technologie "Soft Computing". De nombreux défis restent à relever lorsque nous sommes confrontés à la caractérisation des réservoirs où l'hétérogénéité et la fracturation importe en explorant dans des réservoirs avec de mauvaise qualité. La qualité des données utilisées ou le contrôle limité de la couverture sismique executés dans les champs d'exploration permettent de quantifier l'incertitude et l'intervalle de confiance des estimations des résultats de mesure. Parmi les problèmes inhérents que nous devons surmonter on peut noter : (i) l'échantillonnage des données bien insuffisant et réparti inégalement, (ii) la non-unicité de cause à effet dans les propriétés du sous-sol, (iii) les différentes échelles des données sismiques, (iv) les variations des paramètres de diagraphie et leur gestion dans le réservoir en cours de caractérisation.

Ce document passe en revue les dernières applications en géosciences de "Soft Computing" (**SC**) avec un accent particulier sur l'exploration. Le rôle de l'informatique souple comme une méthode efficace de fusion des données sera mise en évidence. C'est un consortium de méthodologies de calcul [logique floue (**FL**), neuro-informatique (**NC**), calcul génétique (**GC**) et raisonnement probabiliste (**PR**) y compris algorithmes génétiques (**GA**), systèmes chaotiques (**CS**), réseaux de l'évidence (**NE**), théorie de l'apprentissage (**LT**)] qui fournissent collectivement une base pour la conception, l'élaboration et le déploiement de systèmes intelligents. Le modèle de rôle pour le "Soft Computing" est l'esprit humain. Parmi les

principales composantes de "Soft Computing", les réseaux de neurones artificiels, la logique floue et les algorithmes génétiques dans le "domaine de l'exploration", plus précisément, les applications de l'exploration de la Terre de SC dans divers aspects.

2. Objectifs de la thèse

L'objectif de ce travail est de décrire et de caractériser la formation triasique de Hassi R'Mel qui est productive au niveau des unités du Trias en utilisant les données de base classiques disponibles, et les données de diagraphie enregistrées au niveau des différents puits explorés. La distribution des paramètres des réservoirs distincts concernant les propriétés pétrophysiques sont pris en considération pour une porosité et perméabilité efficaces. Les diagraphies de puits sont analysées pour chaque puits, et les résultats sont corrélés avec des informations sur les données de base (carotte) afin de prédire des estimations fiables entre les paramètres étudiés. L'ensemble des données enregistrées est prédit dans ces puits à l'aide de l'analyse par la logique floue (fuzzy logic) et les modèles de perméabilité sont construits et leurs fiabilités sont comparées par les coefficients de régression pour les prévisions dans les intervalles de formations non carottées.

3. Organisation du manuscrit

Cette thèse est composée de quatre chapitres et organisée de la façon suivante:

Chapitre I : Caractérisation des réservoirs triasiques du champ de Hassi R'Mel

Le champ de Hassi R'Mel est considéré, de par ses dimensions et ses réserves, comme étant l'un des plus grands au monde. Il s'étend sur une superficie d'environ 3700 km², produisant ainsi du gaz à condensât avec présence d'un anneau d'huile assez important sur sa périphérie est et sud.

Il se présente sous forme d'une structure anticlinale orientée NNE-SSW, située dans la partie NW du bassin triasique, à environ 550 km au sud d'Alger, à 100 km au NW de la ville de Ghardaïa et à 80 km au SE de la ville de Laghouat. Ce champ est constitué de quatre (4) réservoirs gréseux d'âge triasique, nommés A, B, C et la série inférieure, séparés entre eux par des bancs argileux.

Toutes les études faites auparavant sur ce champ s'intéressaient uniquement aux réservoirs supérieurs (A, B et C). Ces réservoirs présentent d'excellentes caractéristiques pétrophysiques

associées à des épaisseurs utiles importantes, alors que la série inférieure était plus ou moins prometteuse vu la présence de faciès argileux et volcanique.

Chapitre II : Les réseaux de neurones

Les calculs à base de réseaux de neurones artificiels (**ANN**) ont suscité un intérêt considérable au cours des dernières décennies, et sont appliquées avec succès dans un large éventail de problématiques, à des domaines aussi divers que la médecine, les finances, l'ingénierie, la géologie et la physique, à des problèmes de dynamique complexe et la prédiction du comportement complexe, classification ou de contrôle. Plusieurs architectures, des stratégies et des algorithmes d'apprentissage ont été introduites dans ce domaine très dynamique. Ces nouveaux outils pour l'étude des réservoirs sont évalués et testés au cours des processus de forage et d'exploitation pétrolière par des analyses. La prédiction de paramètres pétrophysiques à travers divers outils et technologies basés sur les réponses de diagraphies et d'analyse (théorie et applications) sont de nos jours les plus utilisés.

Chapitre IIIa : La Technique "Fuzzy Logic"

La technique " Fuzzy Logic (**FL**) " qui est capable d'exprimer les caractéristiques d'un système de règles compréhensibles humaines a également été utilisée dans la prédiction des paramètres pétrophysiques tels que la porosité et la perméabilité. Un ensemble flou " fuzzy " décrit le degré d'appartenance d'un élément dans un ensemble à être un nombre réel entre 0 et 1, ce que permet l'observation humaine, expression et expertise pour modéliser de façon plus précise. Une fois que les ensembles flous ont été définis, il est alors possible de les utiliser dans la construction de règles pour les systèmes experts flous, en effectuant une inférence floue. Cette approche semble être ainsi adaptée à l'analyse des diagraphies car elle permet l'intégration des connaissances intelligentes et humaines à traiter chaque cas individuel.

Chapitre IIIb : Les techniques "Neuro-Fuzzy"

Les systèmes **NF** hybrides combinent les avantages des systèmes flous (qui traitent de la connaissance explicite) avec ceux des systèmes de réseaux neuronaux (**NN**) (qui traitent de la connaissance implicite). D'autre part, la logique floue (**FL**) augmente la capacité de généralisation d'un système de réseau neuronal (**NN**) en fournissant une sortie plus fiable et nécessaire lorsque l'extrapolation est au-delà de la limite des données d'apprentissage.

Ces deux techniques sont utilisées de façon combinée et sont appelés systèmes "neuro-fuzzy " ou systèmes hybrides. Ce chapitre résume d'une façon générale les étapes décrivant les techniques les plus connues hybrides neuro-flous, ses avantages et ses inconvénients.

Chapitre IV : Technique neuro-floue appliquée aux réservoirs de Hassi R'Mel Sud pour la détermination de la porosité et de la perméabilité

La technique neuro-floue "**NF**" est utilisée, dans ce travail, pour déterminer la porosité et la perméabilité de la formation réservoir triasique de Hassi R'Mel à l'aide des données de puits disponibles, ainsi que les données de base de perméabilité et de porosité carotte.

Les résultats finaux obtenus montrent l'importance de l'utilisation de Fuzzy Inférence Systems (**FIS**) dans la résolution des problèmes qui sont liés à la prédiction des paramètres physiques. D'autres figures, affichent les données d'essais et de production du système d'inférence floue (**FIS**). La performance du modèle est évaluée par l'erreur quadratique moyenne des ensembles de données. Les coefficients de corrélation entre les valeurs mesurées et celles prédites montrent que cette technique " **Neuro-Fuzzy** " a permis de faire une approche de la prédiction de la porosité et de la perméabilité obtenue avec une très bonne performance : pour la porosité (R^2 = 0.9879) et la perméabilité (R^2 = 0.9689).

Chapitre V : Prédiction des faciès par les méthodes Multivariées

L'application des techniques de régression non paramétriques pour la prédiction de la perméabilité à l'aide des diagraphies de puits dans chaque électrofaciès semble encore la mieux adaptée. En effet, trois approches non paramétriques sont examinées par " Conditionnelles alternatifs attentes " (**ACE**) dans lesquelles le modèle généralisé d'additif (**GAM**), les réseaux de neurones (**NN**) et leurs avantages et inconvénients sont explorés. Les résultats sont comparés à trois autres approches de prédiction de la perméabilité qui utilisent le partitionnement des données en fonction de couches réservoirs, informations lithofaciès et des unités de flow hydraulique (**HFU**). L'examen des taux d'erreur associés à l'analyse discriminante pour les puits non carottés indique que la classification des données de caractérisation sur la base d'électrofaciès est plus robuste par rapport à d'autres approches. Ce chapitre montre la prédiction de la porosité et de la perméabilité où le modèle de l'**ACE** semble être le mieux adapté parmi les trois approches non paramétriques.

I.1- Rappels sur le champ de Hassi R'Mel

Cette étude a fait l'objet d'une recherche et développement de la série inférieure dans la région sud de Hassi R'Mel par la Division Petroleum Engineering et Développement (PED) dont les principaux points s'articulent autour des objectifs suivants:

Une étude de caractérisation de la série inférieure basée sur :

- La sédimentologie de la région dans le but de déterminer les différents types de faciès, leur extension et leur milieu de dépôt.
- Les corrélations diagraphiques pour suivre l'extension latérale de la série inférieure dans tout le champ de Hassi R'Mel Sud.
- L'évolution verticale et latérale de chaque propriété pétrophysique et pétrographique permet de délimiter les meilleurs drains à hydrocarbure.

Sur la base de ce travail, la modélisation des différentes propriétés pétrophysiques du réservoir ont permis de conclure sur un éventuel réservoir prometteur.

L'estimation des réserves en place :

La caractérisation et la modélisation du réservoir de la série inférieure de la région de Hassi R'Mel Sud permettent l'approche d'un **modèle pétrophysique** type de la région étudiée, montrant ainsi l'intérêt pétrolier important de cette série localisée dans la zone structuralement haute de la partie sud du champ de Hassi R'Mel ; cela permettra d'aboutir au développement de ce réservoir à travers les futurs puits à forer.

I.2 - Contexte géologique du gisement de Hassi R'Mel

Le champ de Hassi R'Mel (2°55'-3°00'E; 33°15'-33°45'N) est situé à environ 50 km au sud de Laghouat (Fig.1). Il est situé à une altitude d'environ 760 m et s'étend sur 80 km dans une direction NS et à 60 km WE (Courel et al., 2000). La découverte de gaz à condensats dans le puits de H-1 (Hamel et al., 1988) ainsi que l'excellente qualité du réservoir et son apparente continuité ont contribué au développement de l'exploration dans la région. D'autres forages ont donc été établis dans le nord du champ (Bordj Nili; NL1 avec NL5). Les réservoirs de grès préservent leurs bonnes caractéristiques mais sont structurellement plus faibles, donc à sec ou envahi d'eau. Vers le NE, les forages Lg1, Ph1 et Pg1 révèlent une détérioration de la qualité du réservoir.

L'examen des profils sismiques qui a été entrepris avait pour objectif une étude sismique stratigraphique de telle manière que les grands marqueurs soient localisables. L'épaisse couche de sel couvrant les formations du Trias atténuent énormément les signaux sismiques.

Le domaine étudié contient en moyenne onze puits. La principale industrie de la région est essentiellement du condensat de gaz seuls les puits forés dans l'est de la structure de Hassi R'Mel ont trouvé du pétrole (Boudjema, 1987). Cette huile a été découverte en 1956 par S.N. Repal Co., mais l'exploitation du pétrole n'a commencé que dans les années 70.

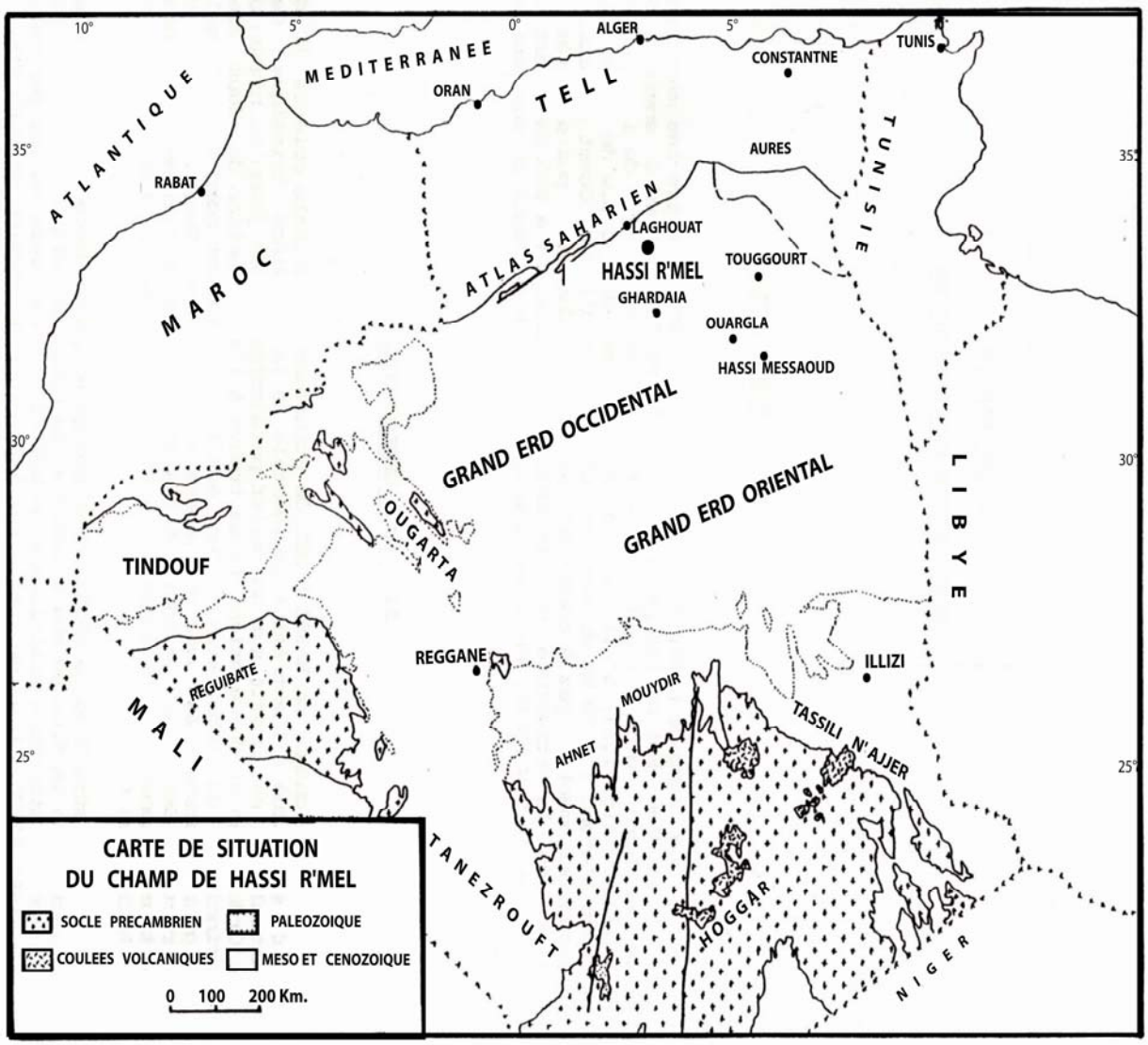


Fig.1. Carte de situation du champ de Hassi R'Mel (d'après Hamel, 1988).

I.3- Caractéristiques diagraphiques des réservoirs triasiques de Hassi R'Mel

L'analyse faciologique présentée avec le modèle semi-automatique en utilisant le logiciel Petrolog (développé par la compagnie Weatherford), nous permettra dans une deuxième étape d'étudier d'une manière générale les faciès rencontrés. Ce qui nous permettra par ailleurs de faire une analyse des différentes formations de Hassi R'Mel.

Dans une première étape, l'utilisation de l'analyse faciologique, à partir des diagraphies, pour les formations triasiques de Hassi R'Mel, a été faite manuellement. Après examen des électrobancs représentatifs, une itération entre les diagrammes lithologiques (Baouche et al., 2010) a permis de définir les lithologies principales et ainsi recenser une dizaine de lithologies. Les lithologies fluctuent entre les points lithologiques: grès, argile et dolomie avec

dans certains cas des andésites. L'utilisation automatique des colonnes lithologiques, à partir du Logiciel Petrolog (2010), a permis de réaliser un traitement semi-automatique des données disponibles et donc de généraliser le modèle au niveau de tous les puits étudiés. Les résultats obtenus sont confrontés à ceux des carottes et des analyses de déblais.

Le traitement manuel des données de diagraphies différées enregistrées au niveau des puits de Hassi R'Mel Sud, NE et NW, a permis donc d'établir des crossplots neutron-densité et sonique-densité des formations triasiques. La définition des zones d'électrofaciès à partir des réponses diagraphiques a permis de distinguer 6 zones de faciès différents, caractérisées chacune par des valeurs diagraphiques propres.

N°	**Couleur**	**Lithologies**
1	Indigo	Halite
2	bleu	Halite argileuse
3	Jaune	Grès
4	Bleu ciel	Grès faiblement argileux
5	Vert clair	Grès argileux
6	Orange	Dolomies argileuses
7	Violette	Dolomies argilo-gréseuses
8	Grenat	Dolomies faiblement argilo-gréseuses
9	Vert foncé	Argiles gréseuses
10	Rouge	Andésites

Tab.1. Légendes des figures employées pour les faciès de Hassi R'Mel (Trias).

I.4- Résultats des électro faciès issus de l'analyse manuelle

De la base au sommet des sondages, il est possible, à partir du puits HRS-7, de distinguer au maximum trois zones d'électrofaciès, ayant les caractéristiques diagraphiques suivantes (Tab.2):

	Radioactivité GR (API)	porosité neutron (%)	Densité (g/cc)
Argiles gréseuses	110 à 167	20 à 34	2.42 à 2.66
Grès faiblement argileux	30 à 64	8 à 23	2.33 à 2.58
Grès argileux	44 à 133	8 à 28	2.38 à 2.68
Andésites	45 à 65	20 à 36	2.44 à 2.67
Dolomie argileuse	42 à 134	22 à 36	2.43 à 2.67
dolomie argilo-gréseuse	45 à 114	11 à 28	2.56 à 2.67

Tab.2. Réponses des diagraphies en face des faciès (Baouche et al., 2010)

La formation supérieure correspondant aux évaporites est constituée essentiellement de la halite, en intercalation avec les argiles, avec une radioactivité naturelle qui varie de 0 à 13 API, une porosité neutron qui varie de -1 à 0 % et une densité qui varie de 2.04 à 2.14 g/cc. Les courbes de neutron et de densité ne reflètent pas l'argilosité à cause du pouvoir de résolution de l'outil Gamma Ray au niveau des couches minces d'argile.

I.5-1 Le milieu de dépôt

L'identification du milieu de dépôt à partir des diagraphies peut être effectuée suivant les modèles de séquences obtenus pour chaque type d'environnement en fonction des réponses diagraphiques enregistrées pour chaque modèle. La restitution des environnements de dépôts à partir des expressions des différentes séquences (d'ordre 2) rencontrées sont-elles mêmes fonction des différents environnements sédimentaires.

A partir de l'analyse des résultats obtenus au niveau des sondages ayant faits l'objet de cette étude, du point de vue sédimentlogique, la sédimentation du Trias est de type continental, silicoclastique et évaporitique avec la présence d'un matériel éruptif, correspondant à des coulées magmatiques (andésites) dont l'épaisseur peut varier d'un puits à un autre.

La formation I (correspondant à la "série inférieure") est caractéristique d'un système fluviatile de type méandriforme, distal où domine un matériel fin et argileux.

L'ensemble sédimentaire de la **formation II** s'est formée dans un système fluviatile en tresse, à dominante gréseuse. Tous les réservoirs de Hassi R'Mel, représentant un intérêt pétrolier, s'y trouvent localisés. Ces réservoirs qui sont situés dans des niveaux gréseux (généralement

3), agencés de manière complexe, présentent une géométrie tabulaire en couches de grande extension, limités au mur par des niveaux argileux imperméables.

La formation III, de la série argilo-salifère (**S4**), s'est déposée dans un environnement de lagune salée, évaporitique et constitue le terme ultime de l'ensemble du Trias : limité au toit par le repère "**D2**" qui est daté du Lias et à la base du Trias par la **discordance Hercynienne.**

Il est à noter enfin que des changements de faciès et des épaisseurs affectent ces niveaux réservoirs caractérisés par des variabilités, en terme pétrophysique, plus importantes. Les nouveaux découpages sédimento-pédogénétiques étant établis à la base des accidents sédimentaires où les coupures utilisées sont fondées sur les discontinuités et non sur les faciès. Des corrélations établies à partir des enregistrements diagraphiques et des accidents pédogénétiques expliquent bien l'extension des réservoirs dans cette région de Hassi R'Mel.
Grâce à cette analyse faciologique, les résultats obtenus permettent une bonne approche de l'agencement vertical des faciès et une caractérisation et identification de divers ordres séquentiels ($2^{ème}$, $3^{ème}$ et $4^{ème}$) correspondant à des séquences d'environnements, des membres ou des formations**,** mais basé sur la lithologie, comme celui des pétroliers.

I.5-2 Les environnements de dépôt

L'analyse du remplissage a permis l'identification d'environnements continentaux : des dépôts fluviatiles (tresses et méandriformes), des dépôts de plaine d'inondation et des dépôts de Sebkha (Fig.2).

I.5-3 Synthèse sur la région d'étude

Les études réalisées, à partir des études de lames minces, des carottes et des diagraphies, au niveau des sondages des puits du champ de **Hassi R'Mel** ont donné les résultats suivants :

- **La sédimentologie du Trias**

La sédimentation rencontrée au niveau du Trias de Hassi R'Mel Sud est continentale et composée de matériel silicoclastique et évaporitique. Cette sédimentation est entrecoupée parfois par d'importantes coulées magmatiques guidées par les accidents. De même que ces dépôts triasiques ont subi des modifications pédogénétiques avec une intensité variable, ils constituent ainsi des discontinuités qui ont contribué à la délimitation des séquences de $2^{ème}$,

de 3ème et de 4ème ordre, et qui sont la base du découpage sédimento-pédogénétiques (Ait Ouali et al., 1994).

Ce découpage montre 3 formations coiffées par des encroûtements complexes (dalles), continus à l'échelle de la région. Les formations renferment des séquences de 2ème ordre à pédogenèse simple. La pédogenèse affecte aussi bien les dépôts sédimentaires que le matériel éruptif.

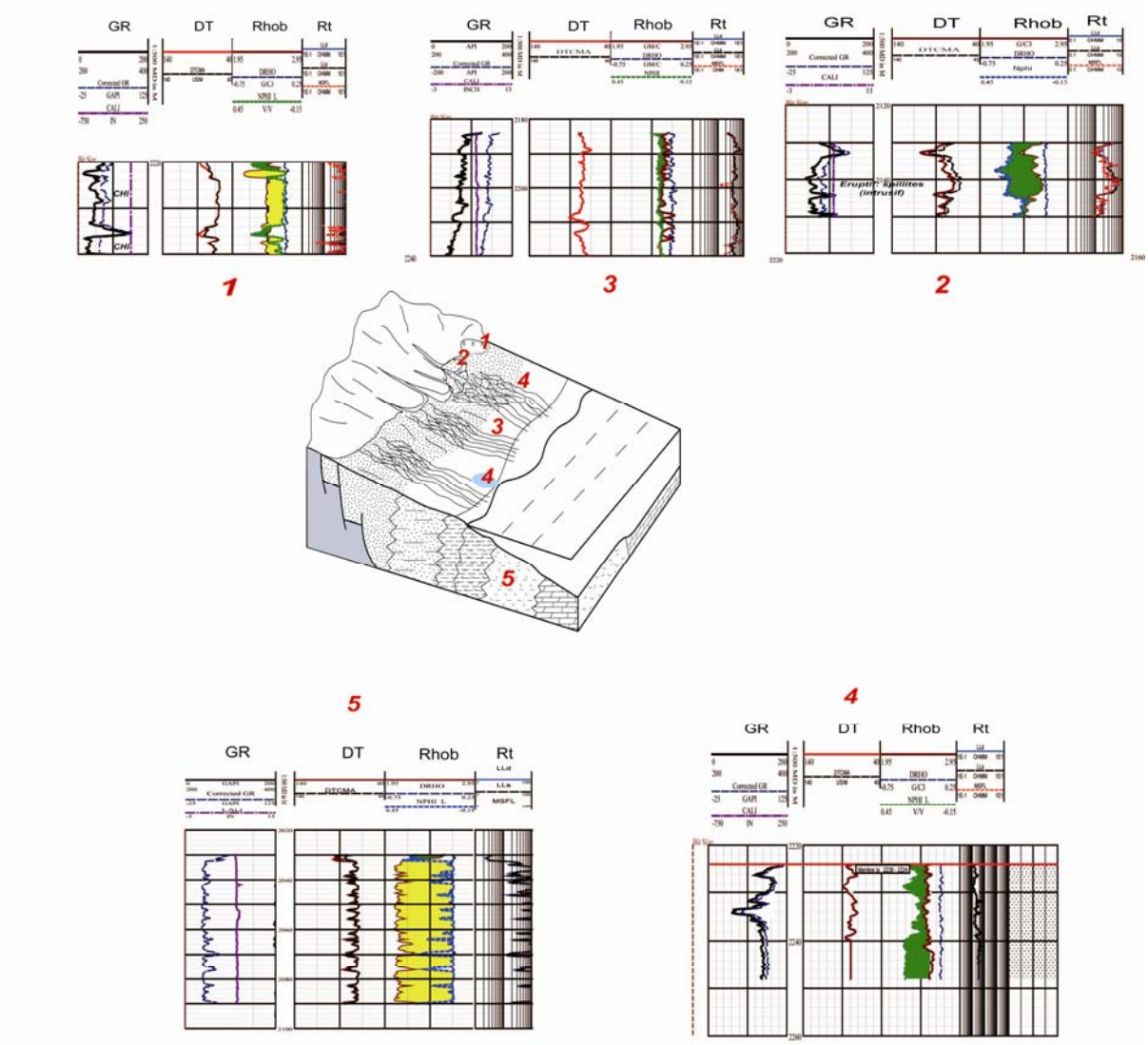


Fig.2. Expression des différentes séquences d'ordre 2. (1) Chenaux fluviatiles ; (2) Dolérites (éruptives) ; (3) Paléosols ; (4) ; Plaine d'inondation ; (5) Sebkha (évaporites), (Baouche et al., 2010)

II.1 Rappel sur les réseaux de neurones

Un neurone est une fonction algébrique non linéaire, paramétrée, à valeurs bornées.

Les réseaux de neurones sont fabriqués de structures cellulaires artificielles et constituent une approche permettant d'aborder sous des angles nouveaux les problèmes de perception, de mémoire, d'apprentissage et de raisonnement. Ils s'avèrent aussi des alternatives très prometteuses pour contourner certaines des limitations des ordinateurs classiques. Grâce à leur traitement parallèle de l'information et à leurs mécanismes inspirés des cellules nerveuses (neurones), ils infèrent des propriétés émergentes permettant de solutionner des problèmes jadis qualifiés de complexes.

Dans ce chapitre, nous nous intéresserons à l'application des réseaux de neurones aux géosciences.

Dans un premier temps, nous rappellerons les définitions et notations de base relatives aux réseaux de neurones. Nous poursuivrons en exposant les types et la méthodologie d'apprentissage.

Nous présenterons aussi, d'une manière générale, les étapes de conception d'un réseau de neurones : le choix des entrées et sorties, l'élaboration de la base de données, de la structure du réseau, etc.

Dans ce travail, l'application est destinée à montrer que dans le domaine des géosciences, les réseaux de neurones sont susceptibles d'apporter des solutions efficaces et élégantes notamment dans la modélisation et la prédiction des paramètres pétrophysiques des roches réservoirs.

II.2 Historique sur les réseaux de neurones

Les recherches effectuées dans le domaine du connexionnisme ont pris leur départ lors de la présentation en 1943 par McCulloch et Pitts d'un modèle simplifié de neurone biologique communément appelé neurone formel. Ils ont montré que théoriquement les réseaux de neurones formels simples peuvent réaliser des fonctions logiques, arithmétiques et symboliques complexes.

En 1949, Hebb initie, dans son ouvrage "The Organisation of Behavior", la notion d'apprentissage. Deux neurones entrant en activité simultanément vont être associés (c'est-à dire que leurs contacts synaptiques vont être renforcés). On parle de loi de Hebb et d'associationnisme.

Un modèle du Perceptron a été développé en 1958 par Rosenblatt. C'est un réseau de neurones inspiré du système visuel qui possède deux couches de neurones : une couche de perception (servant à recueillir les entrées) et une couche de décision représentant ainsi le premier modèle pour lequel on peut définir un processus d'apprentissage.

Widrow et Hoff (1962), en s'inspirant du perceptron, développent, dans la même période, le modèle de l'Adaline (Adaptive Linear Element) qui sera le dernier du modèle de base des réseaux de neurones multicouches.

Les recherches sur les réseaux de neurones en 1969 ont été pratiquement abandonnées lorsque Minsky et Papert ont publié leur livre (1988) sur les " Perceptrons " et ont démontré que les limites théoriques du perceptron, en particulier, sont incapables de traiter les problèmes non linéaires par ce modèle.

Hopfield développe un modèle qui utilise des réseaux totalement connectés basés sur la règle de Hebb (1982), pour définir les notions d'attracteurs et de mémoire associative. En 1984 c'est la découverte des cartes de Kohonen avec un algorithme non supervisé basé sur l'auto-organisation et suivi une année plus tard par la machine de Boltzman (1985).

Une révolution survient alors dans le domaine des réseaux de neurones artificiels: une nouvelle génération de réseaux de neurones, capables de traiter avec succès des phénomènes non-linéaires : le perceptron multicouche ne possède pas les défauts mis en évidence par Minsky et Papert. Proposé pour la première fois par Werbos, le Perceptron multicouche apparaît en 1986 introduit par Rumelhart, et, simultanément, sous une appellation voisine, chez Le Cun (1985). Ces systèmes reposent sur la rétropropagation du gradient de l'erreur dans des systèmes à plusieurs couches, chacune de type Adaline de Bernard Widrow, proche du Perceptron de Rumelhart.

L'utilisation des réseaux de neurones dans divers domaines ne cesse de croître et les applications sont multiples et variées.

II.3 Les réseaux de neurones

II.3.1. Principe du neurone artificiel

Chaque neurone artificiel est un processeur élémentaire. Il reçoit un nombre variable d'entrées en provenance de neurones en amont ou des capteurs composant la machine dont il fait partie. A chacune de ses entrées est associé un poids représentatif de la force de la connexion.

Chaque processeur élémentaire est doté d'une sortie unique qui se ramifie ensuite pour alimenter un nombre variable de neurones en aval. A chaque connexion est associé un poids.
Il est commode de représenter graphiquement un neurone comme indiqué sur la figure 3. Cette représentation est à l'origine de la première vague d'intérêt pour les neurones formels, dans les années 1940 à 1970 (McCulloch et al., 1943 ; Minsky et al., 1988).

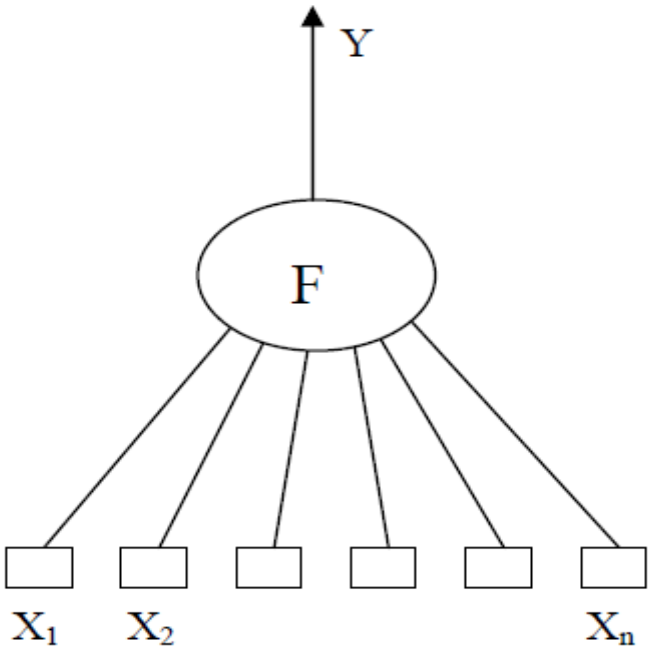
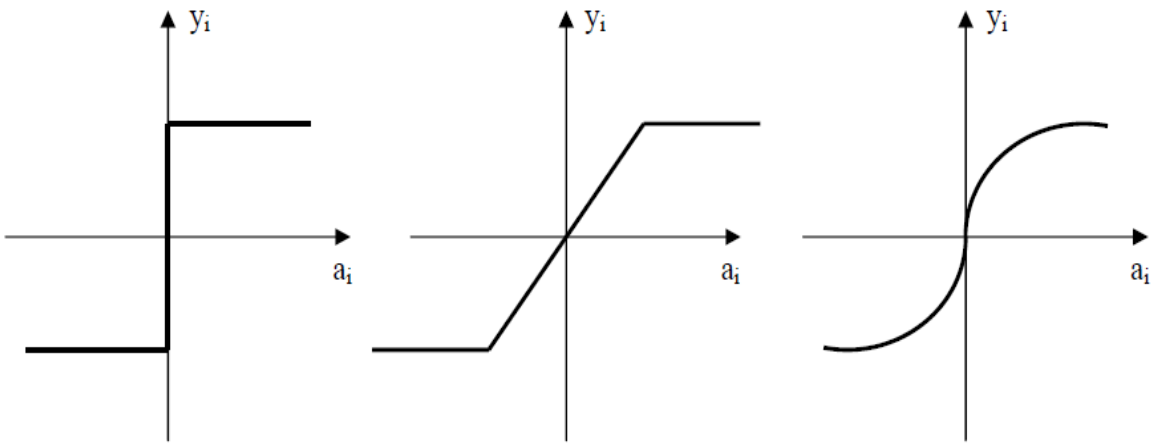


Fig.3. Neurone artificiel (Arbib, 1995).

Le neurone réalise alors trois opérations sur ses entrées :

- **Pondération** : multiplication de chaque entrée par un paramètre appelé poids de connexion,
- **Sommation** : une sommation des entrées pondérées est effectuée
- **Activation** : passage de cette somme dans une fonction, appelée fonction d'activation.

La valeur calculée est la sortie du neurone qui est transmise aux neurones suivants.



Fonction à seuil *Fonction linéaire par morceaux* *Fonction de type sigmoïde*

Fig.4. Différents types de fonction de transfert pour le neurone artificiel (Arbib, 1995).

La fonction f est appelée fonction d'activation (Fig.4). Elle peut être une fonction à seuil, une fonction linéaire ou non-linéaire. La fonction sigmoïde se présente comme une approximation continûment dérivable de la fonction d'activation linéaire par morceaux ou de la fonction seuil. Elle présente l'avantage d'être régulière, monotone, continûment dérivable, et bornée entre 0 et 1 :

$$f(x)=\frac{1}{1+\exp(-x)} \qquad (3.1)$$

La fonction f peut être paramétrée de manière quelconque. Deux types de paramétrages sont fréquemment utilisés:

Les paramètres sont attachés aux entrées du neurone : la sortie du neurone est une fonction non linéaire d'une combinaison des entrées $\{xi\}$ pondérées par les paramètres $\{wi\}$, qui sont alors souvent désignés sous le nom de poids.

$$y=th\left[W_0+\sum_{i=1}^{n-1}W_i X_i\right] \qquad (3.2)$$

Les paramètres sont attachés à la non-linéarité du neurone : ils interviennent directement dans la fonction f.

Les modèles linéaires et sigmoïdaux sont bien adaptés aux algorithmes d'apprentissage impliquant une rétropropagation du gradient car leur fonction d'activation est différentiable ; ce sont les plus utilisés. Le modèle à seuil est sans doute plus conforme à la "réalité"

biologique mais pose des problèmes d'apprentissage. Enfin le modèle stochastique est utilisé pour des problèmes d'optimisation globale de fonctions perturbées ou encore pour les analogies avec les systèmes de particules.

II.3.2. Définition

Un réseau de neurones peut être considéré comme un modèle mathématique de traitement réparti, composé de plusieurs éléments de calcul non-linéaire (neurones), opérant en parallèle et connectés entre eux par des poids.
Les réseaux de neurones artificiels sont des réseaux fortement connectés de processeurs élémentaires fonctionnant en parallèle. Chaque processeur élémentaire calcule une sortie unique sur la base des informations qu'il reçoit.

Les neurones artificiels sont souvent utilisés sous forme de réseaux qui diffèrent selon le type de connections entre les neurones, une cinquantaine de types peut être dénombrée. En guise d'exemples nous citons : le perceptron de Rosemblat, les réseaux de Hopfield, etc.

Ces derniers sont les plus utilisés dans le domaine de la modélisation et de la commande des procédés. Ils sont constitués d'un nombre fini de neurones qui sont arrangés sous forme de couches. Les neurones de deux couches adjacentes sont interconnectés par des poids. L'information dans le réseau se propage d'une couche à l'autre, on dit qu'ils sont de type " feed-forward ". Nous distinguons trois types de couches :

Couche d'entrée : les neurones de cette couche reçoivent les valeurs d'entrée du réseau et les transmettent aux neurones cachés. Chaque neurone reçoit une valeur, il ne fait pas donc de sommation.

Couches cachées : chaque neurone de cette couche reçoit l'information de plusieurs couches précédentes, effectue la sommation pondérée par les poids, puis la transforme selon sa fonction d'activation qui est en général une fonction sigmoïde. Par la suite, il envoie cette réponse aux neurones de la couche suivante.

Couche de sortie : elle joue le même rôle que les couches cachées, la seule différence entre ces deux types de couches est que la sortie des neurones de la couche de sortie n'est liée à aucun autre neurone.

II.4. Architecture des réseaux de neurones

Suivant la fonction du graphe de leurs connexions, il existe deux structures de réseau, c'est-à-dire du graphe dont les nœuds sont les neurones et les arêtes les "connexions" entre eux :

- Les réseaux de neurones statiques (acycliques, ou non bouclés).
- Les réseaux de neurones dynamiques (récurrents, ou bouclés).

II.4.1. Les réseaux de neurones non bouclés

Un réseau de neurones non bouclé réalise une (ou plusieurs) fonction algébrique de ses entrées par composition des fonctions réalisées par chacun de ses neurones. Dans un tel réseau (Fig. 5), le flux d'information circule des entrées vers les sorties sans retour en arrière. Si l'on représente le réseau comme un graphe dont les nœuds sont les neurones et les arêtes les "connexions" entre eux, le graphe d'un réseau non bouclé est acyclique.

Tout neurone dont la sortie est une sortie du réseau est appelé "neurone de sortie".

Les autres effectuent des calculs intermédiaires et sont des "neurones cachés".

Il existe deux types de réseaux de neurones : les réseaux complètement connectés et les réseaux à couches. Le réseau de neurones à une couche cachée et une sortie linéaire est un cas particulier de ce dernier type.

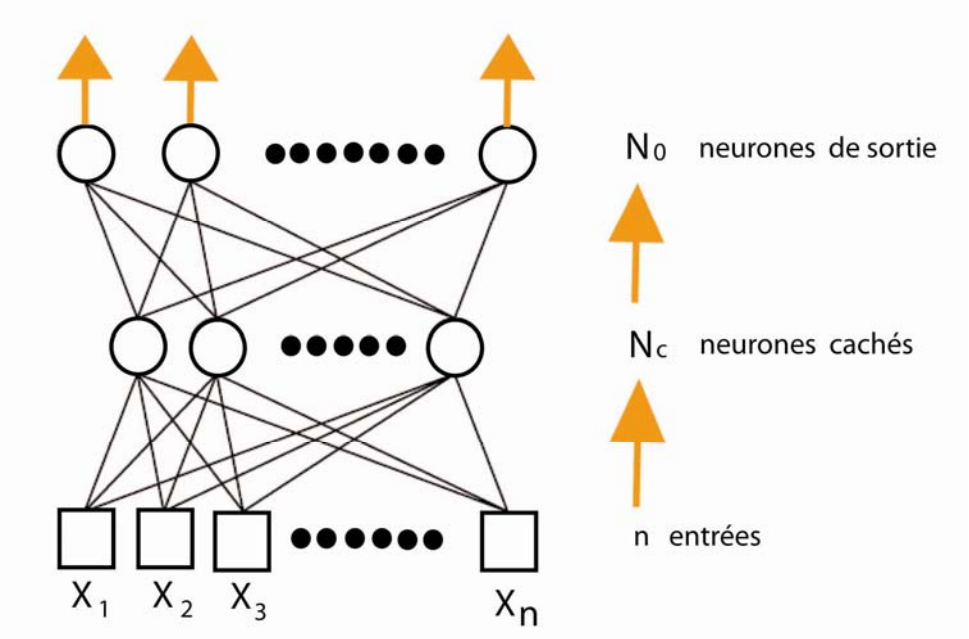


Fig.5. Réseau de neurones à n entrées, une couche de Nc : neurones cachés et N_0 : neurones de sortie (Arbib, 1995).

Les réseaux de neurones complètement connectés

Dans un réseau complètement connecté, les entrées puis les neurones (cachés et de sortie) sont numérotés, et, pour chaque neurone :

- Ses entrées sont toutes les entrées du réseau ainsi que les sorties des neurones de numéro inférieur.
- Sa sortie est connectée aux entrées de tous les neurones de numéro supérieur.
- Les réseaux de neurones à couches

Dans une architecture de réseaux à couches, les neurones cachés sont organisés en couches, les neurones d'une même couche n'étant pas connectés entre eux. De plus les connexions entre deux couches de neurones non consécutives sont éliminées.

Une telle architecture est historiquement très utilisée, surtout en raison de sa pertinence en classification.

- ***Remarque :***

Dans un réseau de neurones non bouclé, le temps ne joue aucun rôle fonctionnel : si les entrées sont constantes, les sorties le sont également. Le temps nécessaire pour le calcul de la fonction réalisée par chaque neurone est négligeable et on peut considérer ce calcul comme instantané.

Pour cette raison, les réseaux non bouclés sont souvent appelés "réseaux statiques", par opposition aux réseaux bouclés ou "dynamiques". Ils sont utilisés en classification, reconnaissance des formes (caractères, parole, ...) ainsi qu'en prédiction.

II.4.2. Les réseaux de neurones bouclés

L'architecture la plus générale pour un réseau de neurones est le "réseau bouclé", dont le graphe des connexions est cyclique : lorsqu'on se déplace dans le réseau en suivant le sens des connexions, il est possible de trouver au moins un chemin qui revient à son point de départ (un tel chemin est désigné sous le terme de "cycle"). La sortie d'un neurone du réseau peut donc être fonction d'elle-même; cela n'est évidemment concevable que si la notion de temps est explicitement prise en considération.

Ainsi, à chaque connexion d'un réseau de neurones bouclé (ou à chaque arête de son graphe) est attaché, outre un poids comme pour les réseaux non bouclés, un retard, multiple entier

(éventuellement nul) de l'unité de temps choisie. Une grandeur, à un instant donné, ne pouvant pas être fonction de sa propre valeur au même instant, tout cycle du graphe du réseau doit avoir un retard non nul.

Les connexions récurrentes ramènent l'information en arrière par rapport au sens de propagation défini dans un réseau multicouche. Ces connexions sont le plus souvent locales.

Pour éliminer le problème de la détermination de l'état du réseau par bouclage, on introduit sur chaque connexion "en retour" un retard qui permet de conserver le mode de fonctionnement séquentiel du réseau (Fig. 6).

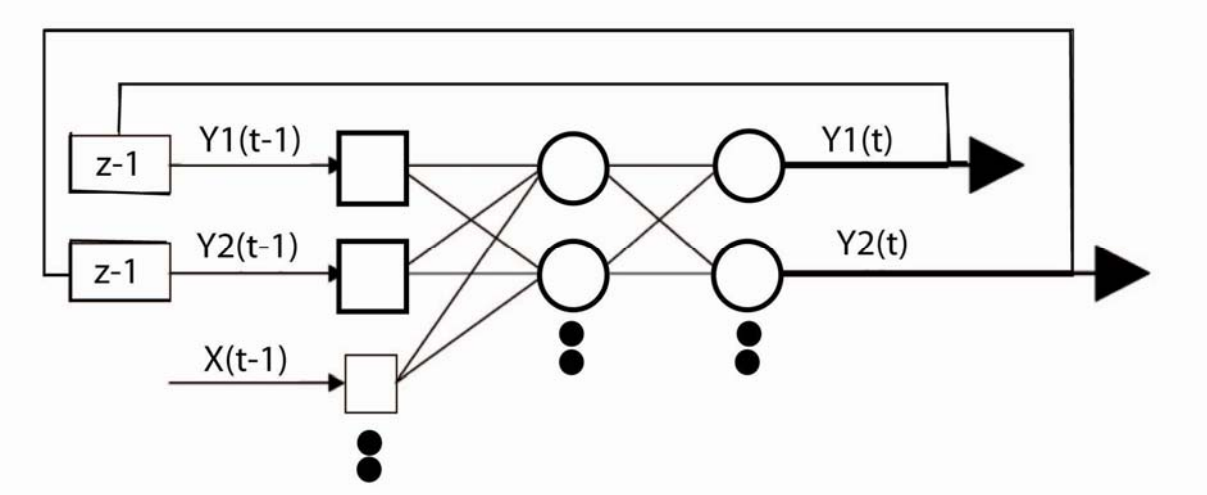


Fig.6. Réseau de neurone bouclé (Arbib, 1995).

Le graphe des connexions de réseaux récurrents est cyclique. Ces réseaux sont décrits par un système d'équations avec des différences.

- ***Forme canonique des réseaux récurrents***

Il a été montré (Nerrand et al., 1993) que tout réseau bouclé peut être mis sous une forme particulière, appelée forme canonique, qui est la représentation d'état minimal de la fonction réalisée par ce réseau. Cette forme canonique est constituée d'un graphe acyclique, et de connexions à retard unité reliant certaines sorties de ce graphe à ses entrées. La fonction réalisée par un réseau de neurones ayant cette structure particulière est décrite par les équations aux différences suivantes :

$$x(k+1)=\Phi(x(k),u(k+1)) \qquad (3.3)$$

$$y(k+1)=\Psi(x(k+1),u(k+1)) \qquad (3.4)$$

où x(*k*) est le vecteur d'état à l'instant *k*, u(*k*) est le vecteur des variables de commande exogènes, y(*k*) le vecteur des sorties, Ψ et Φ sont deux fonctions qui dépendent de la structure de la partie acyclique du réseau.

II.5. Apprentissage des réseaux de neurones

Le point crucial du développement d'un réseau de neurones est son apprentissage. Il s'agit d'une procédure adaptative par laquelle les connexions des neurones sont ajustées face à une source d'information (Hebb, 1949; Grossberg, 1982; Rumelhart et al., 1986).

Dans le cas des réseaux de neurones artificiels, on ajoute souvent à la description du modèle l'algorithme d'apprentissage. Le modèle sans apprentissage présente en effet peu d'intérêt.

Dans la majorité des algorithmes actuels, les variables modifiées pendant l'apprentissage sont les poids des connexions. L'apprentissage est la modification des poids du réseau dans l'optique d'accorder la réponse du réseau aux exemples et à l'expérience. Les poids sont initialisés avec des valeurs aléatoires. Puis des exemples expérimentaux représentatifs du fonctionnement du procédé dans un domaine donné, sont présentés au réseau de neurones. Ces exemples sont constitués de couples expérimentaux de vecteurs d'entrée et de sortie. Une méthode d'optimisation modifie les poids au fur et à mesure des itérations pendant lesquelles on présente la totalité des exemples, afin de minimiser l'écart entre les sorties calculées et les sorties expérimentales. Afin d'éviter les problèmes d'apprentissage, la base d'exemples est divisée en deux parties : la base d'apprentissage et la base de test. L'optimisation des poids se fait sur la base d'apprentissage, mais les poids retenus sont ceux pour lesquels l'erreur obtenue sur la base de test est la plus faible. En effet, si les poids sont optimisés sur tous les exemples de l'apprentissage, on obtient une précision très satisfaisante sur ces exemples mais on risque de ne pas pouvoir généraliser le modèle à des données nouvelles. A partir d'un certain nombre d'itérations, le réseau ne cherche plus l'allure générale de la relation entre les entrées et les sorties du système, mais s'approche trop près des points et " apprend " le bruit (Pollard et al., 1992). En figure 7, on peut observer qu'au début de l'apprentissage, pour les premières itérations, l'erreur sur la base d'apprentissage est grande et peut légèrement augmenter étant donné que les poids initiaux sont choisis aléatoirement. Ensuite, cette erreur

diminue avec le nombre d'itérations. L'erreur sur la base de test diminue puis augmente à partir d'un certain nombre d'itérations. Les poids retenus sont ceux qui minimisent l'erreur sur la base de test.

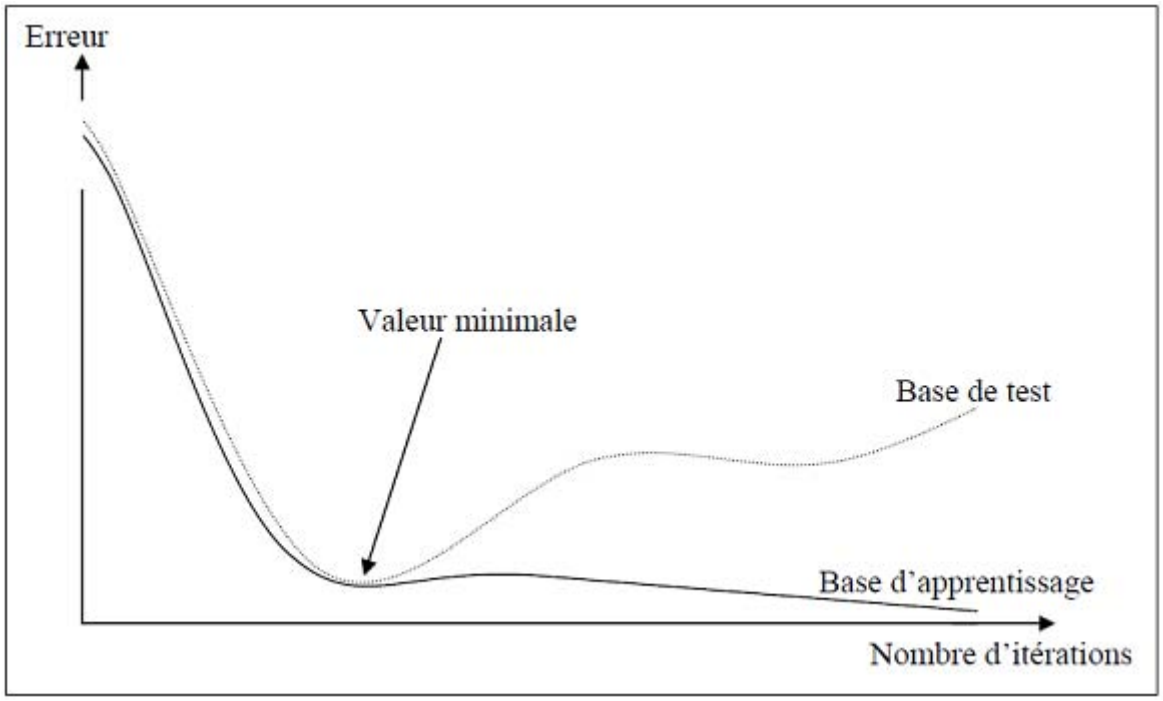


Fig.7. Erreur moyenne sur la base d'apprentissage en fonction du nombre d'itérations (Arbib, 1995).

- ***Surapprentissage***

Il arrive qu'à faire apprendre un réseau de neurones toujours sur le même échantillon, celui-ci devient inapte à reconnaître autre chose que les éléments présents dans l'échantillon.

Le réseau ne cherche plus l'allure générale de la relation entre les entrées et les sorties du système, mais cherche à reproduire les allures de l'échantillon. On parle alors de surapprentissage : le réseau est devenu trop spécialisé et ne généralise plus correctement.

Ce phénomène apparaît aussi lorsqu'on utilise trop d'unités cachées (de connexions), la phase d'apprentissage devient alors trop longue (trop de paramètres réglables dans le système) et les performances du réseau en généralisation deviennent médiocres.

II.5.1. Type d'apprentissage

Il existe de nombreux types de règles d'apprentissage qui peuvent être regroupées en trois catégories (Hassoum, 1995) : les règles d'apprentissage supervisé, non supervisé, et renforcé. Mais l'objectif fondamental de l'apprentissage reste le même : soit la classification,

l'approximation de fonction ou encore la prévision (Weiss et Kulikowski, 1991 ; Ammar, 2007). Dans l'optique de la prévision, l'apprentissage consiste à extraire des régularités (à partir des exemples) qui peuvent être transférées à de nouveaux exemples.

- ***Apprentissage supervisé***

Un apprentissage est dit supervisé lorsque l'on force le réseau à converger vers un état final précis, en même temps qu'on lui présente un motif. Ce genre d'apprentissage est réalisé à l'aide d'une base d'apprentissage, constituée de plusieurs exemples de type entrées-sorties (les entrées du réseau et les sorties désirées ou encore les solutions souhaitées pour l'ensemble des sorties du réseau).

La procédure usuelle dans le cadre de la prévision est l'apprentissage supervisé (ou à partir d'exemples) qui consiste à associer une réponse spécifique désirée à chaque signal d'entrée. La modification des poids s'effectue progressivement jusqu'à ce que l'erreur (ou l'écart) entre les sorties du réseau (ou résultats calculés) et les résultats désirés soient minimisés.

Cet apprentissage n'est possible que si un large jeu de données est disponible et si les solutions sont connues pour les exemples de la base d'apprentissage.

- ***Apprentissage renforcé***

L'apprentissage renforcé est une technique similaire à l'apprentissage supervisé à la différence qu'au lieu de fournir des résultats désirés au réseau, on lui accorde plutôt un grade (ou score) qui est une mesure du degré de performance du réseau après quelques itérations.

Les algorithmes utilisant la procédure d'apprentissage renforcé sont surtout utilisés dans le domaine des systèmes de contrôle (White et Sofge, 1992; Sutton, 1992).

- ***Apprentissage non supervisé***

L'apprentissage non supervisé consiste à ajuster les poids à partir d'un seul ensemble d'apprentissage formé uniquement de données. Aucun résultat désiré n'est fourni au réseau.

Qu'est-ce que le réseau apprend exactement dans ce cas ? L'apprentissage consiste à détecter les similarités et les différences dans l'ensemble d'apprentissage. Les poids et les sorties du réseau convergent, en théorie, vers les représentations qui capturent les régularités statistiques des données (Hinton, 1992 ; Ammar, 2007). Ce type d'apprentissage est également dit compétitif et (ou) coopératif (Grossberg, 1988). L'avantage de ce type d'apprentissage réside dans sa grande capacité d'adaptation reconnue comme une auto-organisation, "self-organizing" (Kohonen, 1987 ; Andréassian, 2004). L'apprentissage non supervisé est surtout utilisé pour le traitement du signal et l'analyse factorielle.

II.5.2. Algorithme d'apprentissage

L'algorithme d'apprentissage est la méthode mathématique qui va modifier les poids de connexions afin de converger vers une solution qui permettra au réseau d'accomplir la tâche désirée. L'apprentissage est une méthode d'identification paramétrique qui permet d'optimiser les valeurs des poids du réseau.

Plusieurs algorithmes itératifs peuvent être mis en œuvre, parmi lesquels on note :

L'algorithme de rétropropagation, la méthode Quasi-Newton, Algorithme de BFGS, etc.

- ***Algorithme de rétropropagation***

L'algorithme de rétropropagation (ARP) ou de propagation arrière " back propagation " est l'exemple d'apprentissage supervisé le plus utilisé à cause de l'écho médiatique de certaines applications spectaculaires telles que la démonstration de Sejnowski et Rosenberg (1986) dans laquelle l'ARP est utilisé dans un système qui apprend à lire un texte. Un autre succès fut la prédiction des cours du marché boursier (Refenes et al., 1994; Lee et al., 1996) et plus récemment la détection de la fraude dans les opérations par cartes de crédit (Dorronsoro et al., 1997).

La technique de rétropropagation du gradient est une méthode qui permet de calculer le gradient de l'erreur pour chaque neurone du réseau, de la dernière couche vers la première. L'historique des publications montre que l'ARP a été découvert indépendamment par différents auteurs mais sous différentes appellations (Grossberg, 1998). Le principe de la rétropropagation peut être décrit en trois étapes fondamentales : acheminement de l'information à travers le réseau; rétropropagation des sensibilités et calcul du gradient; ajustement des paramètres par la règle du gradient approximé. Il est important de noter que l'ARP souffre des limitations inhérentes à la technique du gradient à cause du risque d'être piégé dans un minimum local. Il suffit que les gradients ou leurs dérivées soient nuls pour que le réseau se retrouve bloqué dans un minimum local. Ajoutons à cela la lenteur de la convergence surtout lorsqu'on traite des réseaux de grande taille (c'est-à-dire pour lesquels le nombre de poids de connexion à déterminer est important).

Pour rendre l'optimisation plus performante, on peut utiliser des méthodes de second ordre telles que les méthodes dites de Quasi-Newton ou de Newton modifiée.

- ***Méthodes Quasi-Newton***

Les méthodes de quasi-Newton consistent à chaque itération k à remplacer le problème par :

$$(P_k)\ :\ \begin{cases} résoudre & F''(x_k)xd_k = -F'(x_k) \\ trouver\ \alpha_k\ qui\ \min imise & \phi(\alpha) = F(x_k + \alpha d_k) \\ calculer & x_{k+1} = x_k + \alpha_k d_k \end{cases} \qquad (5.3)$$

par le problème :

$$(P'_k)\ :\ \begin{cases} calculer une approximation D_k\ de\ \left[F''(x_k)\right]^{-1} \\ calculer la direction de descente\ \ d_k = -D_k.F'(x_k) \\ trouver \alpha_k\ qui \min imise & \phi(\alpha) = F(x_k + \alpha d_k) \\ calculer & x_{k+1} = x_k + \alpha_k d_k \end{cases} \qquad (5.4)$$

Le principe des méthodes de résolution de type Quasi-Newton est de générer une séquence de matrices symétriques définies positives qui soient des approximations, toujours améliorées, de la matrice Hessienne réelle ou de son inverse. On recherche une méthode telle que, dans le cas d'un problème quadratique, la matrice B_k converge vers la valeur exacte des dérivées secondes (constantes dans ce cas), de sorte qu'en fin de convergence, on retrouve une convergence de type Newton. Si l'on applique la méthode à une fonction quelconque, B_k peut être considéré, à chaque instant, comme une approximation (définie positive) du Hessien.

- ***Algorithme de BFGS***

L'algorithme de BFGS (du nom de ses inventeurs : Broyden, Fletcher, Goldfarb et Shanno) (Press et al., 1988) prend implicitement en compte les dérivées secondes et s'avère donc nettement plus performante que la méthode de rétropropagation. Le nombre d'itérations est nettement plus faible et les temps de calcul réduits d'autant.

L'algorithme de BFGS est une règle d'ajustement des paramètres qui a l'expression suivante :

$$\theta^k = \theta^{k-1} \pm \mu_k M_k \nabla J\left(\theta^{k-1}\right) \qquad (5.5)$$

où M_k est une approximation, calculée itérativement, de l'inverse de la matrice Hessienne. L'approximation de l'inverse du Hessien est modifiée à chaque itération suivant la règle suivante :

$$M_k = M_{k+1} + \left[1 + \left(\frac{\gamma_{k-1}^T M_{k-1} \gamma_{k-1}}{\delta_{k-1}^T \gamma_{k-1}}\right)\right] \frac{\delta_{k-1}^T \delta_{k-1}}{\delta_{k-1}^T \gamma_{k-1}} \pm \frac{\delta_{k-1}\gamma_{k-1}^T M_{k-1} + M_{k-1}\gamma_{k-1}\delta_{k-1}^T}{\delta_{k-1}^T \gamma_{k-1}} \qquad (5.7)$$

avec : $\gamma_{k-1} = \nabla J\left(\theta^k\right) \pm \nabla J\left(\theta^{k-1}\right)$ *et* $\delta_{k-1} = \theta^k \pm \theta^{k-1}$

Nous prenons pour valeur initiale de M la matrice identité. Si, à une itération, la matrice calculée n'est pas définie positive, elle est réinitialisée à la matrice identité.
Une méthode "quasi newtonienne", n'est efficace que si elle est appliquée au voisinage d'un minimum. D'autre part, la règle du gradient simple est efficace lorsqu'on est loin du minimum et sa convergence ralentit considérablement lorsque la norme du gradient diminue (c'est-à-dire lorsque l'on s'approche du minimum). Ces deux techniques sont donc complémentaires. De ce fait, l'optimisation s'effectue en deux étapes : utilisation de la règle du gradient simple pour approcher un minimum, et de l'algorithme de BFGS pour l'atteindre.

II.6. Modélisation à l'aide des réseaux de neurones

Deux principales stratégies de modélisation qui emploient des réseaux de neurones peuvent être distinguées: la première appelée approche par "boite noire", quand le processus entier est représenté avec un réseau neuronal approprié, et approche "hybride" qui est une combinaison de la modélisation traditionnelle du processus avec un réseau neuronal qui représente les phénomènes les moins connus du processus.

II.6.1. Modèle "boîte noire"

Le terme de "boîte noire" s'oppose aux termes de "modèle de connaissance" ou "modèle de comportement interne" qui désignent un modèle mathématique établi à partir d'une analyse physique du processus que l'on étudie. Ce modèle peut contenir un nombre limité de paramètres ajustables, qui possèdent une signification physique. Les réseaux de neurones peuvent être utilisés pour l'élaboration de modèle "boîte grise", intermédiaire entre les modèles "boîtes noires" et les "modèles de connaissance".
Le modèle "boîte noire" constitue la forme la plus primitive de modèle mathématique (Fig. 8): il est réalisé uniquement à partir de données expérimentales ou d'observations. Il peut avoir une valeur prédictive, dans un certain domaine de validité, mais n'a aucune valeur explicative. Ainsi, le modèle de l'univers selon Ptolémée était un modèle "boîte noire" : il ne donnait aucune explication de la marche des astres, mais il permettait de la prédire avec toute la précision souhaitable au regard des instruments de mesure disponibles à l'époque.

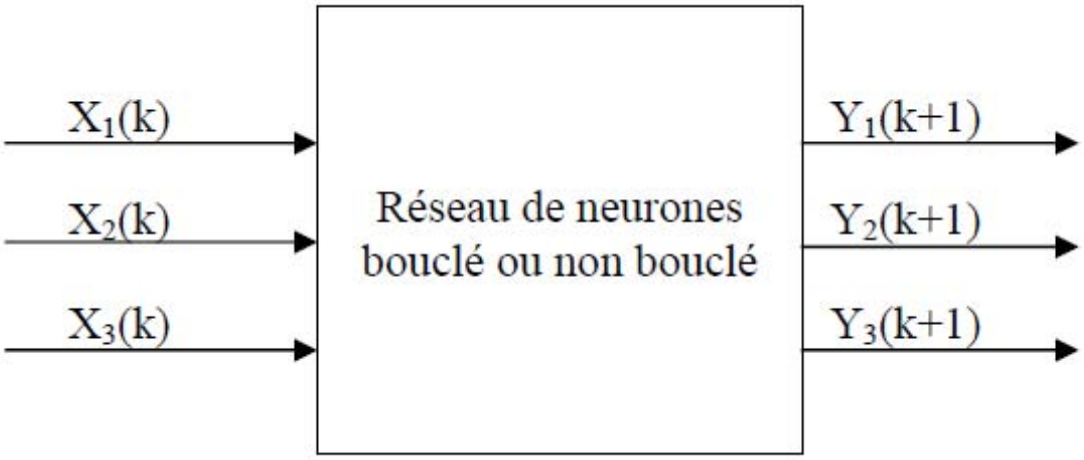


Fig.8. Diagramme schématique d'un modèle neuronal "boîte noire" (Ammar, 2007).

Entre la boîte noire et le modèle de connaissance se situe la modèle semi-physique, ou modèle "boîte grise" (appelé aussi modèle hybride), qui contient à la fois des équations résultant d'une théorie, et des équations purement empiriques, résultant d'une modélisation de type "boîte noire".

II.6.2. Modèle "boîte grise" ou hybride

Lorsque des connaissances, exprimables sous forme d'équations, sont disponibles, mais insuffisantes pour concevoir un modèle de connaissance satisfaisant, on peut avoir recours à une modélisation "boîte grise" (ou modélisation semi-physique) qui prend en considération à la fois les connaissances et les mesures. Une telle démarche peut concilier les avantages de l'intelligibilité d'un modèle de connaissance avec la souplesse d'un modèle comportant des paramètres ajustables.

Un modèle hybride peut être considéré comme un compromis entre un modèle de connaissance et un modèle "boite noire". Il peut prendre en considération toutes les connaissances que l'ingénieur possède sur le processus, à condition que celles-ci puissent être exprimées par des équations algébriques ou différentielles, et, de surcroît, ce modèle peut utiliser des fonctions paramétrées, dont les paramètres sont déterminés par apprentissage.
Dans la mesure où l'on met en œuvre davantage de connaissances expertes, les données expérimentales nécessaires pour estimer les paramètres d'une manière significative peuvent être en quantité plus réduites.

La conception d'un modèle hybride exige que l'on dispose d'un modèle de connaissance, qui se présente habituellement sous la forme d'un ensemble d'équations algébriques, différentielles, et aux dérivées partielles, non-linéaires couplées. Par la suite, on doit procéder à l'apprentissage de ce modèle (ou une partie de celui-ci) à partir de données obtenues par intégration numérique du modèle de connaissance, et de données expérimentales. Psichogios et Ungar (1992) ont introduit l'idée du modèle neuronal hybride (modèle "boîte grise"). Un tel modèle hybride utilise toute connaissance accessible et possible. Dans le modèle hybride, le réseau de neurones est utilisé pour représenter les éléments inconnus de cette modélisation.

II.7. Conception d'un réseau de neurones

Les réseaux de neurones réalisent des fonctions non-linéaires paramétrées. Leurs mises en œuvre nécessitent :

- La détermination des entrées et des sorties pertinentes, c'est-à-dire les grandeurs qui ont une influence significative sur le phénomène que l'on cherche à modéliser.
- La collecte des données nécessaires à l'apprentissage et à l'évaluation des performances du réseau de neurones.
- La détermination du nombre de neurones cachés nécessaires pour obtenir une approximation satisfaisante.
- La réalisation de l'apprentissage
- L'évaluation des performances du réseau de neurones à l'issue de l'apprentissage.

II.7.1. Détermination des entrées/sorties du réseau de neurones

Pour toute conception de modèle, la sélection des entrées doit prendre en compte deux points essentiels :

- Premièrement, la dimension intrinsèque du vecteur des entrées doit être aussi petite que possible, en d'autres termes, la représentation des entrées doit être la plus compacte possible, tout en conservant pour l'essentiel la même quantité d'information, et en gardant à l'esprit que les différentes entrées doivent être indépendantes.
- En second lieu, toutes les informations présentées dans les entrées doivent être pertinentes pour la grandeur que l'on cherche à modéliser : elles doivent donc avoir une influence réelle sur la valeur de la sortie.

II.7.2. Choix et préparation des échantillons

Le processus d'élaboration d'un réseau de neurones commence toujours par le choix et la préparation des échantillons de données. La façon dont se présente l'échantillon conditionne le type de réseau, le nombre de cellules d'entrée, le nombre de cellules de sortie et la façon dont il faudra mener l'apprentissage, les tests et la validation (Bishop, 1995). Il faut donc déterminer les grandeurs qui ont une influence significative sur le phénomène que l'on cherche à modéliser.

Lorsque la grandeur que l'on veut modéliser dépend de nombreux facteurs, c'est-à-dire lorsque le modèle possède de nombreuses entrées, il n'est pas possible de réaliser un " pavage " régulier dans tout le domaine de variation des entrées : il faut donc trouver une méthode permettant de réaliser uniquement des expériences qui apportent une information significative pour l'apprentissage du modèle. Cet objectif peut être obtenu en mettant en œuvre un plan d'expériences. Pour les modèles linéaires, l'élaboration de plans d'expériences est bien maîtrisée, par ailleurs, ce n'est pas le cas pour les modèles non linéaires.

Afin de développer une application à base de réseaux de neurones, il est nécessaire de disposer de deux bases de données, une pour effectuer l'apprentissage et l'autre pour tester le réseau obtenu et déterminer ses performances.

Notons qu'il n'y a pas de règle pour déterminer ce partage d'une manière quantitative, néanmoins chaque base doit satisfaire aux contraintes de représentativité de chaque classe de données et doit généralement refléter la distribution réelle, c'est-à-dire la probabilité d'occurrence des diverses classes (Arbib et al., 1995).

II.7.3. Elaboration de la structure du réseau

La structure du réseau dépend étroitement du type d'échantillons. Il faut d'abord choisir le type de réseau : un perceptron standard, un réseau de Hopfield, un réseau à décalage temporel (TDNN), un réseau de Kohonen, un ARTMAP, etc.

Par exemple, dans le cas du perceptron multicouche, il faudra aussi bien choisir le nombre de couches cachées que le nombre de neurones dans cette couche.

- ***Nombre de couches cachées :***

Mis à part les couches d'entrée et de sortie, il faut décider du nombre de couches intermédiaires ou cachées. Sans couche cachée, le réseau n'offre que de faibles possibilités d'adaptation. Néanmoins, il a été démontré qu'un perceptron multicouche avec une seule

couche cachée pourvue d'un nombre suffisant de neurones, peut approximer n'importe quelle fonction avec la précision souhaitée (Arbib et al., 1995).

Nombre de neurones cachés :

Il n'existe pas, à ce jour, de résultat théorique permettant de prévoir le nombre de neurones cachés nécessaires pour obtenir une performance spécifique du modèle, compte tenu des modèles disponibles. Il faut donc nécessairement mettre en œuvre une procédure numérique de conception de modèle.

II.7.4. Apprentissage

L'apprentissage est un problème numérique d'optimisation. Il consiste à calculer les pondérations optimales des différentes liaisons, en utilisant un échantillon. La méthode la plus Utilisée est la rétropropagation, qui est généralement plus économe que les autres en terme de nombres d'opérations arithmétiques à effectuer pour évaluer le gradient.

Pour rendre l'optimisation plus performante, on peut utiliser des méthodes de second ordre. Le calcul est très efficace, mais lourd. Elles ont de nombreuses limitations, quant aux conditions de convergence, sur les dérivées secondes. Des corrections sont proposées pour éviter ce problème, et sont prises en compte par les méthodes dites de Quasi-Newton ou de Newton modifiée.

Il a été observé que les poids calculés par la méthode de rétropropagation sont plus faibles que ceux obtenus par la technique de Quasi-Newton, ce qui semblerait montrer que la recherche d'un minimum par rétropropagation est restreinte à un voisinage immédiat des poids initiaux, d'où une dépendance plus forte de cette méthode par rapport à l'initialisation (Thibault, 1991).

II.7.5. Validation et Tests

Alors que les tests concernent la vérification des performances d'un réseau de neurones hors échantillon et sa capacité de généralisation, la validation est parfois utilisée lors de l'apprentissage. Une fois le réseau de neurones développé, des tests s'imposent afin de vérifier la qualité des prévisions du modèle neuronal (Fig. 9).

Cette dernière étape doit permettre d'estimer la qualité du réseau obtenu en lui présentant des exemples qui ne font pas partie de l'ensemble d'apprentissage. Une validation rigoureuse du modèle développé se traduit par une proportion importante de prédictions exactes sur l'ensemble de la validation.

Si les performances du réseau ne sont pas satisfaisantes, il faudra, soit modifier l'architecture du réseau, soit modifier la base d'apprentissage.

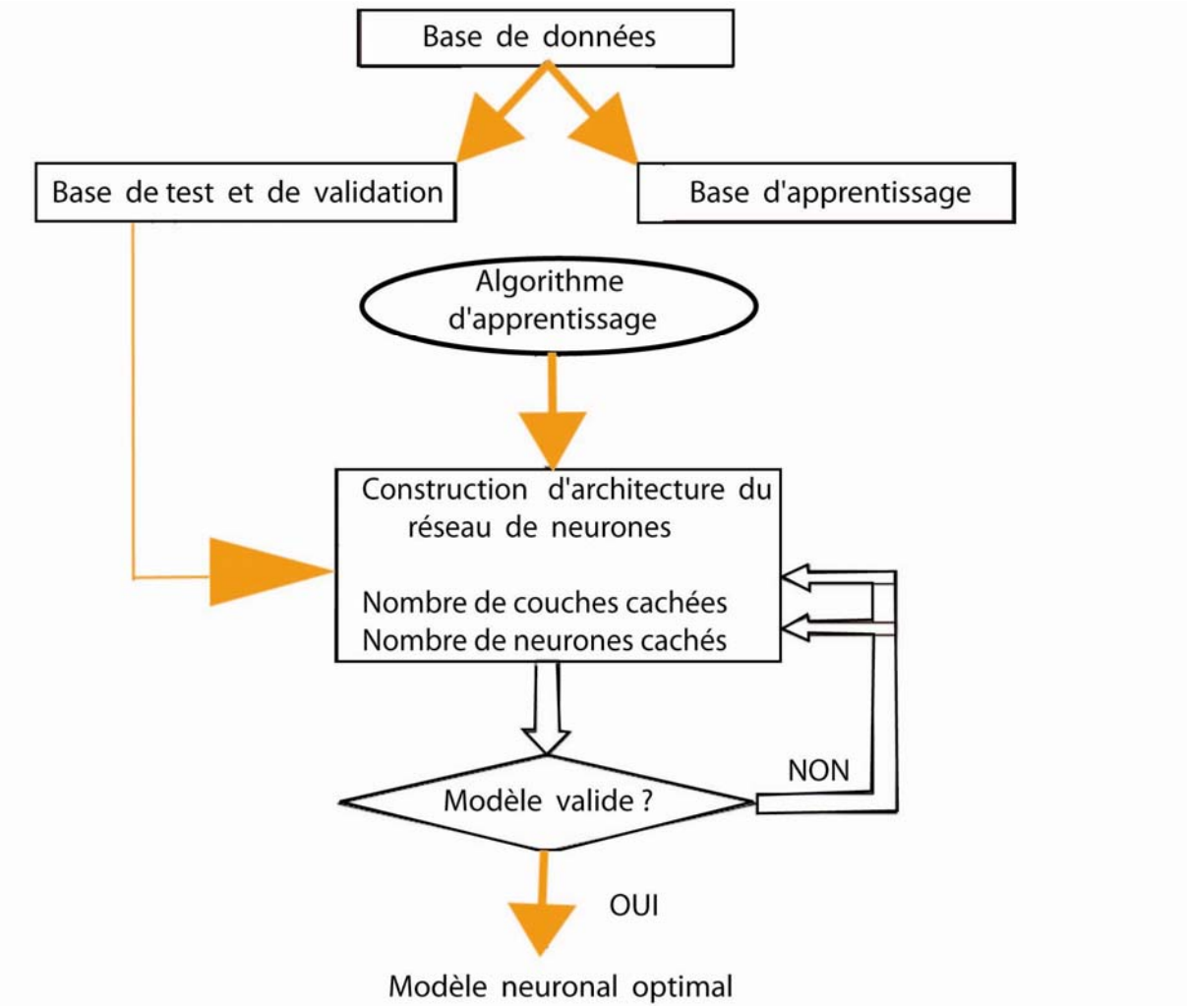


Fig.9. Organigramme de conception d'un réseau de neurones (Ammar, 2007).

II.8. Analogie de régression multilinéaire avec le réseau de neurones

Le procédé d'identification de la relation entre la réponse et la variable y, les variables causales x_1, x_2 et x_3 utilisant une régression multilinéaire ou "MultiLinear Regression" (MLR) est analogue à la procédure d'identification de la relation entre l'entrée et les variables de sortie à l'aide de réseaux neuronaux. Les paramètres η_1, η_2 et η_3 sont similaires aux poids du

réseau. Les paramètres de régression sont estimés par le principe de la méthode des moindres carrés, en minimisant la somme des carrés des écarts de même les poids et la polarisation dans le réseau neuronal sont ajustés en minimisant la fonction coût.

La différence entre un réseau de neurones et une MLR est que la relation d'entrée-sortie est linéaire dans la MLR où la relation est non-linéaire dans le réseau neuronal. L'application de la méthode des réseaux de neurones aux puits de Hassi R'Mel (well : HR-167) nous donne les résultats suivants pour la prédiction de tous les paramètres diagraphiques (Fig. 10**) :**

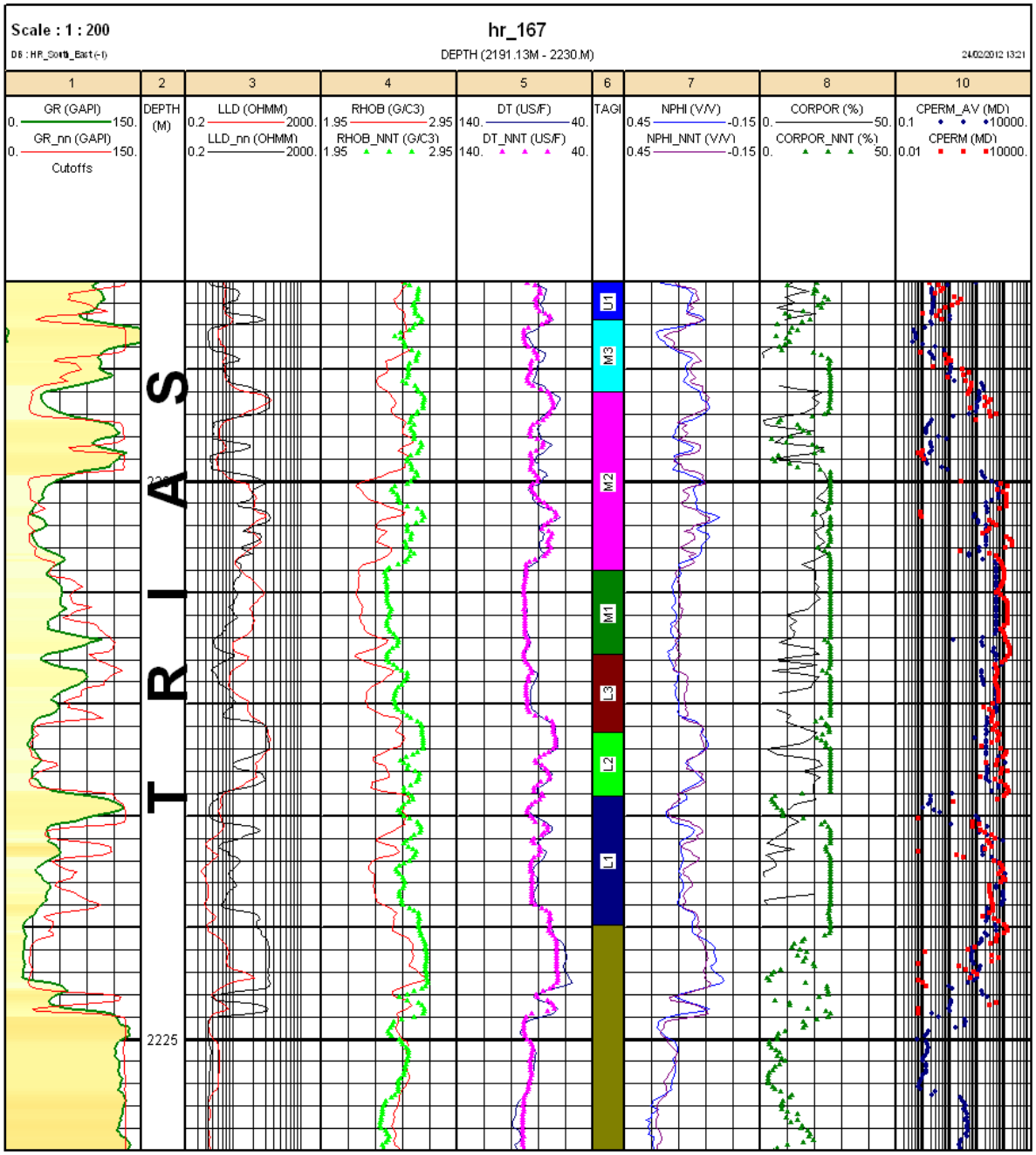


Fig.10. Modèle de logs synthétiques pour les données de réseaux de neurones (LLD_nn : résistivité; Rhob_nn : densité; Δt_nn : sonique; Nphi_nn : neutron; CorPor_nn : porosité et Cperm_nn : perméabilité).

IIIa.1. Rappel sur la logique floue

Ces dernières années ont connu une croissance rapide du nombre et de la variété des applications de la logique floue (**FL**). Les techniques FL ont été utilisées dans la compréhension et les applications d'image comme la détection et l'extraction de ses caractéristiques, la classification et le groupage (clustering). La logique floue s'intéresse à la capacité d'imiter l'esprit humain à utiliser efficacement les modes de raisonnement qui sont approximatifs plutôt que précis. En informatique traditionnelle, les décisions ou les actions sont basées sur la précision, la certitude et la rigueur. La précision et la certitude ont un coût. En informatique, la tolérance et l'impression sont explorées dans la prise de décision. L'exploration de la tolérance pour l'imprécision et l'incertitude sous-entend la capacité humaine remarquable à comprendre la parole déformée, déchiffrer l'écriture bâclée, comprendre les nuances de la langue naturelle, résumer le texte, et reconnaître et classer les images. Avec la FL, nous pouvons spécifier des règles de correspondance en termes de mots plutôt que de chiffres.

Le calcul avec les mots explore l'imprécision et la tolérance. Un autre concept de base est la règle floue If-Then. Bien que les systèmes à base de règles aient une longue histoire d'utilisation de l'intelligence artificielle, ce qui manque dans de tels systèmes sont les machines pour traiter les conséquents floues ou les antécédents flous. Dans la plupart des applications, une solution FL est une traduction d'une solution humaine.

D'autre part, la FL peut modéliser des fonctions non-linéaires de complexité arbitraire à un degré de précision désiré.

La FL est un moyen pratique pour cartographier un espace d'entrée d'un espace de sortie. La FL est l'un des outils utilisés pour modéliser une entrée multi, système à multiples sorties.

Le "Soft Computing" comprend la logique floue, les réseaux de neurones, le raisonnement probabiliste et les algorithmes génétiques. Aujourd'hui, les techniques ou une combinaison de techniques de tous ces domaines sont utilisées pour concevoir un système d'intelligence. Les réseaux de neurones fournissent des algorithmes d'apprentissage, la classification et l'optimisation, alors que la logique floue traite de questions telles que la formation d'impressions et de raisonnement sur un plan sémantique ou linguistique. Le raisonnement probabiliste traite de l'incertitude. Bien qu'il existe d'importantes zones de chevauchement entre les réseaux de neurones, FL, et raisonnement probabiliste, sont en général complémentaires plutôt que concurrents. Les systèmes flous récents contenant de nombreux

systèmes intelligents appelés " neuro " ont été utilisés. Il existe de nombreuses façons de combiner les réseaux de neurones et les autres techniques. Cependant avant de le faire, il est nécessaire de comprendre les idées de base de la conception des autres techniques. Dans ce chapitre, nous allons présenter les concepts tels que les ensembles flous et leurs propriétés, les opérateurs FL, haies, proposition floue et systèmes à base de règles, cartes floues et moteur d'inférence, méthodes de défuzzification, et conception d'un système de décision FL.

IIIa.2. Les principes de base de la logique floue

La logique floue est une méthode mathématique complexe qui permet de résoudre des problèmes difficiles simulées avec de nombreuses entrées et variables de sortie. La logique floue est capable de donner des résultats sous forme de recommandation pour un intervalle spécifique de l'état de sortie, il est donc essentiel que cette méthode mathématique soit strictement distincte des logiques plus familières, comme l'algèbre de Boole. Ce travail contient une vue d'ensemble des principes de la logique floue.

IIIa.3. Système de logique floue

Les systèmes de contrôle d'aujourd'hui sont généralement décrits par des modèles mathématiques qui suivent les lois de la physique des modèles ou des modèles stochastiques qui ont émergé de la logique mathématique. Une difficulté générale de ce modèle construit est de savoir comment passer d'un problème donné à un modèle mathématique approprié. Sans aucun doute d’après la technologie informatique de pointe d'aujourd'hui mais la gestion de ces systèmes est encore trop complexe.

Ces systèmes complexes peuvent être simplifiés en utilisant une marge de tolérance pour un montant raisonnable de l'imprécision et l'incertitude au cours de la phase de modélisation. Comme un résultat pas tout à fait parfait un système vient à l'existence. Néanmoins dans la plupart des cas il est capable de résoudre le problème d’une manière appropriée. Même les informations d'entrée manquantes sont déjà avérées satisfaisantes dans les systèmes à base de connaissances. La logique floue permet donc de réduire la complexité par l'utilisation de l'information imparfaite de manière sensible. Elle peut être mise en œuvre en matériel, logiciel, ou en une combinaison des deux.

Les analyses et les méthodes de contrôle de la logique floue (Fig. 11) peuvent être décrites comme suit:

1) La réception d’un grand nombre de mesures ou d'autres évaluations des conditions existantes dans un système qui sera analysé ou contrôlé.
2) Les traitements ont tous reçu des contributions basées, règles floues "**If-then**", qui peuvent être exprimées par des mots de la langue simples et combinés avec un traitement non floue traditionnelle.
3) la compensation et la pondération des résultats de toutes les règles individuelles en une seule décision de sortie unique ou un signal qui décide ce qu'il faut faire ou dit un système contrôlé. Le signal de sortie du résultat est une valeur précise "déffuzzifiée".
4) Le graphe est le schéma de la Logique floue de la méthode d'analyses de contrôle.

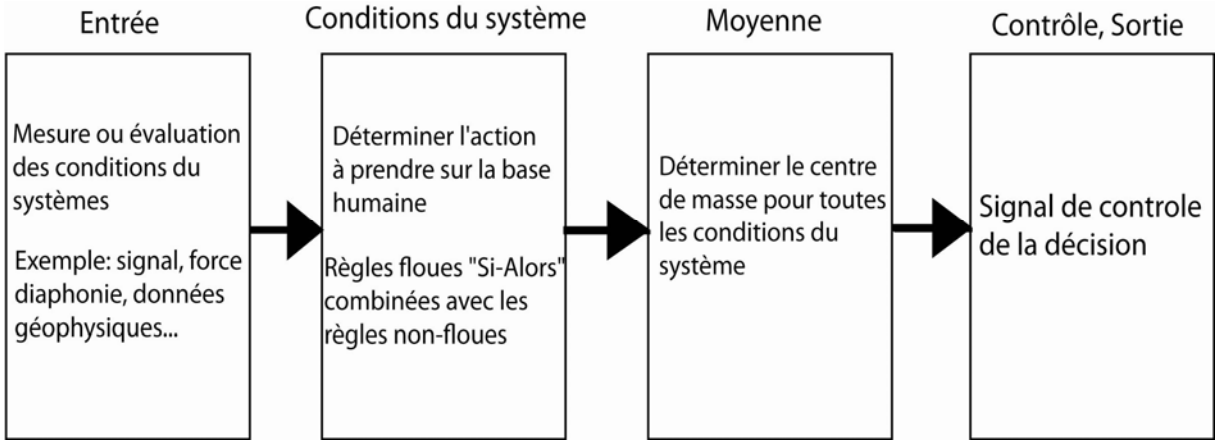


Fig.11. Méthode de contrôle et d’analyse de logique floue (Nauck et al., 1997)

Afin de la faire fonctionner, la logique floue a besoin d'être représentée par des nombres ou des descriptions. Par exemple, la vitesse peut être représentée par une valeur de "5 m/s" ou par une description "lente". Le terme "lente" peut avoir un sens différent s’il est utilisé par différentes personnes et doit être interprété à l'égard de l'environnement observé. Certaines valeurs sont faciles à classer, tandis que d'autres peuvent être difficiles à déterminer en raison de la compréhension humaine de situations différentes. On peut dire "lente", tandis que d'autres peuvent dire "pas rapide" pour décrire la même vitesse. Ces différences peuvent être distinguées à l'aide d'ensembles dits flous.

Habituellement le système flou de contrôle logique est créé à partir de quatre éléments majeurs présentés (Fig.12): l'interface de fuzzification, moteur d'inférence floue, la matrice de règles floues et l'interface de défuzzification.

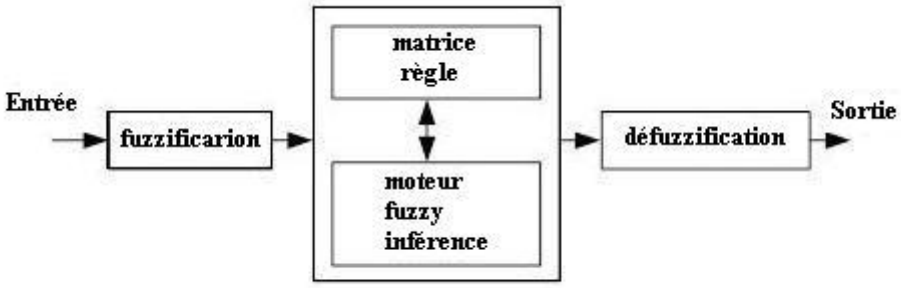


Fig. 12. Contrôleur en logique floue (Füller et al., 1996)

IIIa.4. Logique floue : Opérations de base

Les informations de base sur la logique floue seront présentées ainsi qu'une théorie complète de la logique floue tirée de la littérature.

Il s'agit d'un ensemble de valeurs possibles considérées comme entrée du système flou.

• Les ensembles flous :

Un ensemble flou : μ est une fonction de l'ensemble de référence X à l'intervalle unitaire, à savoir :

$$\mu : X \rightarrow [0,1] \qquad (1)$$

$\mu(X)$ représente l'ensemble de tous les sous-ensembles flous de X.

• Fonction d'appartenance

Il s'agit d'une représentation graphique d'ensembles flous, $\mu_F(x)$.

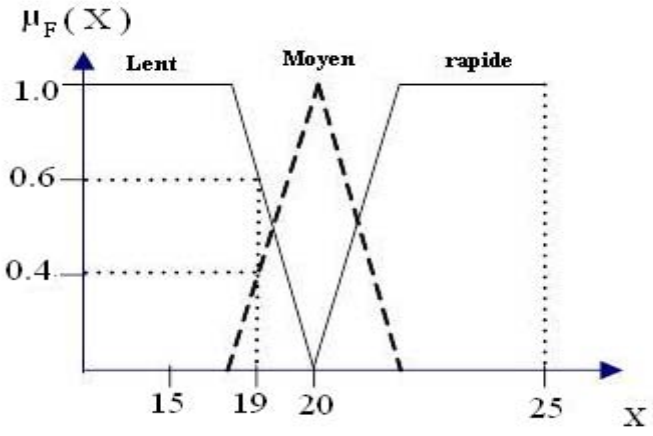


Fig. 13. Exemple de fonction d'adhésion de la logique floue (Füller et al., 1996)

La figure 13 montre les fonctions d'appartenance des trois ensembles flous, "lent", "moyenne", et "rapide", pour une vitesse variable floue. L'univers de discours crée toutes les valeurs possibles de la vitesse, c'est-à-dire, X = 19 Pour la valeur de la vitesse 19 km/h, l'ensemble flou "ralentir" a la valeur d'appartenance 0.6. Par conséquent, μ_{slow} (19) = 0.6. De même, $\mu_{average}$ (19) = 0.4, et μ_{fast} (19) = 0.

• Support

Le soutien d'un ensemble flou F est l'ensemble croustillant de tous les points de l'univers du discours U tel que la fonction d'appartenance de F n'est pas égale à zéro :

$$\mu_F(u) > 0 \qquad (2)$$

- **Point de croisement "Crossover point"**

Il s'agit d'un élément en U, où la fonction d'appartenance est égale à 0.5.

• Centre

Le centre d'un ensemble flou F est le point (ou points) à laquelle μ_F (u) atteint sa valeur maximale.

IIIa.5. Méthode de "Fuzzification"

La première phase de la procédure logique floue est de fournir des paramètres d'entrée pour le système flou donné sur la base de laquelle le résultat de sortie est calculé. Ces paramètres sont fuzzifiés avec l'utilisation des fonctions d'appartenance d'entrée prédéfinies, qui peuvent avoir des formes différentes. Les plus courantes sont: forme triangulaire, cependant cloche, trapézoïdale, sinusoïdale et exponentielle peut également être utilisées. Les fonctions les plus simples ne seront pas exigées informatiquement complexes et ne seront pas surchargées à la mise en œuvre. Le degré de la fonction d'appartenance est déterminé en plaçant une variable d'entrée choisie sur l'axe horizontal, tandis que l'axe vertical représente la quantification du degré d'appartenance de la grandeur d'entrée. La seule condition à une fonction d'appartenance à répondre est qu'elle doit varier entre zéro et un. La valeur zéro signifie que la variable d'entrée n'est pas un élément de l'ensemble flou, alors que la valeur de l'une des variables d'entrée est entièrement un élément de l'ensemble flou.

A chaque paramètre d'entrée, il y a une fonction d'appartenance unique associée. Les fonctions de membre associent un facteur de pondération des valeurs de chaque entrée et les règles en vigueur. Ces facteurs de pondération déterminent le degré d'influence ou degré d'appartenance (DOM) de chaque règle active. En calculant le produit logique des poids d'appartenance pour chaque règle active, un ensemble de grandeurs de sortie de réponse logique floue sont produites. Tout ce qui reste est de combiner et défuzzifier ces réponses de sortie.

- Matrice rule

La matrice de la règle est utilisée pour décrire des ensembles flous et des opérateurs flous sous forme de déclarations conditionnelles. Une seule règle floue "if-Then" peut s'exprimer comme suit :

Si x est A y est alors Z,

où A est un ensemble de conditions qui doivent être satisfaites et Z est un ensemble de conséquences qui peuvent en être déduites.

En règle avec plusieurs parties, les opérateurs flous sont utilisés pour combiner plus d'une entrée: ET = min, OU = max et NON = complément additif. La démonstration géométrique des opérateurs flous est illustrée en figure 14.

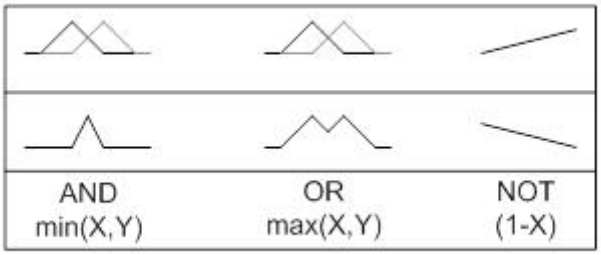


Fig. 14. Interprétation graphique des opérateurs flous (Fullér et al., 1996)

La matrice de règle est un outil graphique simple pour la cartographie des règles floues du système de contrôle de la logique. Il permet d'utiliser deux ou plusieurs variables d'entrée et exprime leur produit logique (ET ou OU) comme une variable de réponse de sortie. Le degré

d'appartenance pour la sortie de la matrice de règle peut prendre la valeur de maximum, et le minimum du degré du précédent. Il est souvent probable que, après évaluation de toutes les règles applicables à l'entrée, nous obtenions plus d'une valeur pour le degré d'appartenance. Dans ce cas, la simulation doit prendre en considération toutes les trois possibilités : le minimum, le maximum ou la moyenne des degrés des membres.

Mécanisme d'inférence : il permet de cartographier l'entrée donnée à une sortie en utilisant la logique floue. Il utilise toutes les pièces décrites dans les sections précédentes: fonctions d'appartenance, opérations logiques et règles "If-then". Les types les plus communs des systèmes d'inférence sont Mamdani et Sugeno (1977). Ils varient suivant les moyens à déterminer les sorties.

IIIa.6. Méthode de Mamdani

Ci-dessous, les exemples sont basés sur deux règles de commande floue, sous la forme de :

R_1: if *x* is A_1 and *y* is B_1 then *z* is C_1

R_2: if *x* is A_2 and *y* is B_2 then *z* is C_2

Résultat: z est C, où x est égal à x_0 et y est égal à y_0.

Les niveaux de déclenchement des règles, notés α_i, i = 1, 2 sont calculés par

$$\alpha_1 = A_1(x_0) \wedge B_1(y_0) \qquad (3)$$

$$\alpha_2 = A_2(x_0) \wedge B_2(y_0) \qquad (4)$$

Les sorties de règles individuelles sont obtenues par :

$$C_1^{'}(\omega) = (\alpha_1 \wedge C_1(\omega)) \qquad (5)$$

$$C_2^{'}(\omega) = (\alpha_2 \wedge C_2(\omega)) \qquad (6)$$

Ensuite, la sortie du système global est calculée en utilisant les sorties de règles individuelles :

$$C(\omega) = C_1^{'}(\omega) \vee C_2^{'}(\omega) = (\alpha_1 \wedge C_1(\omega)) \vee (\alpha_2 \wedge C_2(\omega)) \qquad (7)$$

Enfin, pour obtenir une action de contrôle déterministe, le mécanisme de défuzzification choisi doit être mis en œuvre.

IIIa.7. Méthode de Sugeno

Ci-dessous, les exemples sont basés sur deux règles de commande floue, sous la forme de :

R_1: if x is A_1 and y is B_1 then z is $z_1 = a_1x_1+b_1y_1$
R_2: if x is A_2 and y is B_2 then z is $z_2 = a_2x_2+b_2y_2$

Résultat: z_0, où x est équivalent x_0 et y est équivalent à y_0.

Les niveaux de la règle sont calculés de la même manière que dans la méthode de Mamdani, sur la base des équations (3) et (4). Les sorties de règles individuelles sont calculées à partir des relations ci-dessous :

$$z_1 = \alpha_1.x_0 + b_1.y_0 \qquad (8)$$

$$z_2 = \alpha_2.x_0 + b_2.y_0 \qquad (9)$$

Le résultat du contrôle est dérivé des équations suivantes :

$$z_0 = \frac{\sum_{i=1}^{n} \alpha_i z_i}{\sum_{i=1}^{n} \alpha_i} \qquad (10)$$

où α_i est un niveau de mise à feu de la $i^{ème}$ règle, et i = 1, …., n.

Le procédé Sugeno fonctionne bien avec les techniques linéaires, comme il est informatiquement efficace. Il convient d'appliquer l'optimisation et les techniques adaptatives. En outre, il garantit la continuité de la surface de sortie et il est bien adapté à l'analyse mathématique.

IIIa.8. Mécanismes de "Défuzzification"

La tâche " Défuzzification " consiste à trouver une valeur unique non floue qui résume l'ensemble flou. Il existe plusieurs techniques mathématiques disponibles: centre de gravité, bissectrice, moyenne maximale, moyenne maximale et pondéré. La figure 15 montre l’illustration de la façon dont les valeurs de chaque méthode sont choisies.

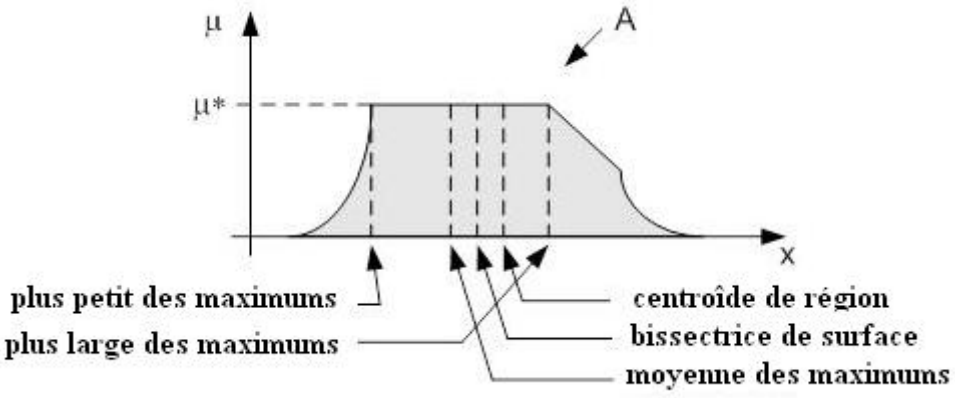


Fig. 15. Démonstration graphique des méthodes de "Défuzzification" (Füller et al., 1996)

La défuzzification centre de gravité est la méthode la plus couramment utilisée, car elle est très précise. Elle fournit le centre de l'aire sous la courbe de fonction d'appartenance. Pour les fonctions de membre complexes, elle met des exigences élevées sur le calcul. Elle peut être exprimée par la formule suivante

$$z_0 = \frac{\int \mu_i(x) x ds}{\int \mu_i(x) dx} \qquad (11)$$

où z_0 est la sortie défuzzifiée, u_i est une fonction d'appartenance et x est la variable de sortie.
La bissectrice défuzzification utilise la ligne verticale qui divise l'aire sous la courbe en deux zones égales.

$$\int_{\alpha}^{z} \mu_A(x) dx = \int_{z}^{\beta} \mu_A(x) dx \qquad (12)$$

Le moyen de la méthode de défuzzification maximale utilise la valeur moyenne des sorties de fonctions d'appartenance globales.

$$z_0 = \frac{\int_x x dx}{\int_x dx} \qquad (13)$$

où x' = $\{x; \mu_A(x) = \mu^*\}$.

La plus petite de la méthode de défuzzification maximale utilise la valeur minimale des sorties de fonctions d'appartenance globales (données globales).

$$z_0 \in \left\{ x \middle| \mu(x) = \min_{\omega} \mu(\omega) \right\} \qquad (14)$$

La plus grande de la méthode de défuzzification maximale utilise la valeur maximale des sorties de fonction d'appartenance agrégées (données fusionnées).

$$z_0 \in \left\{ x \middle| \mu(x) = \max_{\omega} \mu(\omega) \right\} \qquad (15)$$

La méthode de défuzzification moyenne pondérée, sur la base de la valeur de crête de chaque sous-ensemble flou, calcule la somme pondérée de ces valeurs de crête. D'après ces valeurs de poids et le degré d'appartenance floue pour la sortie, la valeur exacte de sortie est déterminée par la formule suivante :

$$z_0 = \frac{\sum \mu(x)_i \times w_i}{\sum \mu(x)_i} \qquad (16)$$

où μ_i est le degré d'appartenance de la production unique i, W_i est la valeur du poids de la sortie floue pour le singleton de sortie i.

IIIa.9. Conclusion sur la logique floue

La logique floue offre une approche complètement différente. On peut se concentrer sur la résolution du problème plutôt que d'essayer de modéliser le système mathématiquement, si c'est encore possible. Cela conduit presque toujours à des solutions moins coûteuses, plus rapides. Une fois compris, cette technologie n'est pas difficile à mettre en œuvre et les résultats sont généralement assez surprenant et plus que satisfaisant.

IIIb.1. Rappel sur les systèmes neuro-flous

Les techniques de l'intelligence artificielle à base de réseaux de neurones et logiques floues sont souvent appliquées ensemble. Les raisons de combiner ces deux paradigmes sont en

dehors des difficultés et des limites inhérentes à chacun des paradigmes séparément. Génériquement, quand ils sont utilisés de façon combinée, ils sont appelés *Systèmes neuro-flous*. Ce terme, cependant, est souvent utilisé pour attribuer un type de système qui intègre les deux techniques. Ce type de système est caractérisé par un système flou où les ensembles flous et les règles floues sont réglés à l'aide des modèles d'entrées-sorties. Il existe plusieurs implémentations différentes de systèmes neuro-flous, où chaque auteur définit son propre modèle. Ce travail, montre une application de cette méthode dans le domaine de la prédiction des paramètres pétrophysiques, tels que la porosité et la perméabilité des réservoirs.

Les techniques modernes de l'intelligence artificielle ont trouvé une application dans presque tous les domaines de la connaissance humaine. Cependant, une grande importance est accordée aux domaines des sciences exactes, peut-être la plus grande expression de la réussite de ces techniques est dans le domaine de l'ingénierie. Ces deux techniques de réseaux de neurones et la logique floue sont appliquées plusieurs fois ensemble pour résoudre les problèmes d'ingénierie, où les techniques classiques ne fournissent pas une solution facile et précise. Le terme de neuro-flou est né par la fusion de ces deux techniques. Comme chaque chercheur combine ces deux outils de façon différente, alors, une certaine confusion a été créée sur le sens exact de ce terme. Pourtant, il n'existe pas de consensus absolu, mais en général, le terme neuro-flou est un type de système caractérisé par une structure similaire d'un contrôleur flou où les décors et les règles floues sont ajustés à l'aide de techniques d'optimisation de réseaux de neurones de manière itérative avec des vecteurs de données (données d'entrée et de sortie du système).

Ces systèmes présentent deux façons distinctes de comportement :

Dans une première phase, dite phase d'apprentissage, le système se comporte comme un réseau de neurones qui apprend ses paramètres internes hors ligne. Plus tard, dans la phase d'exécution, il se comporte comme un système à logique floue. Par ailleurs, chacune de ces techniques présentent des avantages et des inconvénients qui, lorsqu'ils sont mélangés, leurs traitements fournit de meilleurs résultats comparés à chaque technique isolée.

IIIb.2. Systèmes flous

Les systèmes flous proposent un calcul mathématique pour traduire la connaissance humaine subjective des processus réels. C'est une façon de manipuler des connaissances pratiques avec un certain niveau d'incertitude. La théorie des ensembles flous a été initiée par Zadeh (1965).

Le comportement de ces systèmes est décrit par un ensemble de règles floues, comme :

IF <premise> THEN <consequent>

SI <prémisse> Alors <conséquente>

qui utilisent des variables linguistiques avec des termes symboliques. Chaque terme représente un ensemble flou. Les termes de l'espace d'entrée (généralement 5-7 pour chaque variable linguistique) composent la partition floue. Le mécanisme d'inférence floue est constitué de trois étapes: dans la première étape, les valeurs des entrées numériques sont représentées par une fonction relative à un degré de compatibilité des ensembles flous respectifs; cette opération peut être appelée fuzzification. Dans la deuxième étape, le système traite les règles de logique floue en conformité avec les forces des entrées. Dans la troisième étape, les valeurs floues obtenues sont transformées à nouveau en valeurs numériques; cette opération peut être appelée défuzzification. Cette procédure permet, essentiellement à des catégories floues, l'utilisation possible dans la représentation résumée des mots et des idées des êtres humains dans la description de la prise de décision.

Les avantages des systèmes flous sont:

- capacité de représenter les incertitudes inhérentes à la connaissance humaine avec des variables linguistiques;
- interaction simple de l'expert du domaine avec le concepteur de l'ingénieur du système;
- facilité de l'interprétation des résultats, en raison de la représentation des règles naturelles;
- extension facile de la base de connaissances grâce à l'ajout de nouvelles règles;
- robustesse en relation des perturbations possibles dans le système.

Leurs inconvénients sont les suivants:

- incapacité de généraliser, ou alors le système peut uniquement répondre à ce qui est écrit dans sa base de règles;
- pas robuste en ce qui concerne les modifications topologiques du système, ces changements pourraient exiger des modifications dans la base de règles;
- dépendance de l'existence d'un expert afin de déterminer les règles logiques d'inférence;

IIIb.3. Réseaux de Neurones (Neural Networks)

Les réseaux de neurones sont faits pour façonner les fonctions biologiques du cerveau humain. Cela conduit à l'idéalisation des neurones comme des unités individuelles de traitement distribué. Ses connexions locales ou globales à l'intérieur d'un réseau sont aussi idéalisées, ainsi conduisant à la capacité du système nerveux à assimiler l'apprentissage ou à prévoir des réactions ou des décisions à prendre. Mcculloch et Pitts (1943) ont décrit le premier modèle de réseau neuronal et Rosenblatt (Perceptron) (1958) et Widrow (Adaline) (1960) ont développé le premier algorithme d'apprentissage. La caractéristique principale des réseaux de neurones est le fait que ces structures peuvent apprendre des exemples (vecteurs d'apprentissage d'entrée, et des échantillons de sortie du système).

Les réseaux de neurones modifient la structure interne et les poids des connexions entre les neurones artificiels pour rendre la correspondance avec un niveau d'erreur acceptable pour l'application de la relation d'entrée/sortie qui représente le comportement du système modélisé.

Les avantages des réseaux de neurones sont:

- capacité d'apprentissage;
- capacité de généralisation;
- robustesse par rapport aux perturbations.

Leurs inconvénients sont les suivants:

- interprétation impossible de la fonctionnalité;
- difficulté à déterminer le nombre de couches et le nombre de neurones.

IIIb.4. Les systèmes "Neuro-Fuzzy"

Depuis le moment où les systèmes flous deviennent populaires dans les applications industrielles, la communauté perçoit que le développement d'un système flou avec une bonne performance n'est pas une tâche facile. Le problème de trouver des fonctions d'appartenance et des règles appropriées est souvent un processus fatigant de tentatives et d’erreurs. Cela a conduit à l'idée d'appliquer des algorithmes d'apprentissage pour les systèmes flous. Les réseaux de neurones qui ont des algorithmes d'apprentissage efficaces ont été présentés comme une alternative à automatiser ou à soutenir le développement de réglage de systèmes flous.

Les premières études des systèmes neuro-flous datent du début des années 90, avec par exemple Lin et Lee (1991), Berenji et Khedkar (1992), Jang (1992) et Nauck (1994). La majorité des premières applications étaient un contrôle de processus. Peu à peu, son application s'est propagée à tous les domaines de la connaissance comme l'analyse des données, la classification des données, la détection des imperfections et aide à la prise de décision, etc.

Les réseaux de neurones et systèmes flous peuvent être combinés pour rejoindre les avantages par exemple dans le domaine de la médecine (guérison de maladie humaine). Les réseaux de neurones présentent les caractéristiques de calcul de l'apprentissage dans les systèmes flous et reçoivent l'interprétation et la clarté de la représentation des systèmes. Ainsi, les inconvénients des systèmes à logique floue sont compensés par les capacités des réseaux de neurones. Ces techniques sont donc complémentaires, ce qui justifie leur utilisation conjointe.

IIIb.5. Types de systèmes Neuro-Fuzzy

En général, toutes les combinaisons de techniques basées sur les réseaux de neurones et la logique floue peuvent être appelées *systèmes neuro-flous*. Les différentes combinaisons de ces techniques peuvent être divisées (Nauck et al., 1997), suivant les classes suivantes:

- ✓ ***Système neuro-flou coopératif "Cooperative Neuro-Fuzzy"*** **:** Dans les systèmes coopératifs il y a une phase de pré-traitement où les réseaux de neurones consistent à déterminer les mécanismes de l'apprentissage des sous-blocs du système flou. Par exemple, il s'agit de déterminer les ensembles flous et / ou règles floues (mémoires associatives floues ou l'utilisation d'algorithmes de classification pour déterminer la position des règles et des ensembles flous). Une fois les sous-blocs flous sont calculés les méthodes d'apprentissage du réseau de neurones sont enlevées, l'exécution concerne seulement le système flou.
- ✓ ***Système neuro-flou concurrent "Concurrent Neuro-Fuzzy"*** : Dans les systèmes concurrents du réseau de neurones, le système fonctionne en ensemble flou continu. D'une manière générale, les réseaux de neurones pré-traitent les entrées (ou les procédés post-sorties) du système flou. En informatique, "*Système Concurrent*" est une propriété de plusieurs systèmes dans lesquels les calculs sont exécutés simultanément, et potentiellement en interaction les uns avec les autres.

- ✓ ***Système neuro-flou hybride "Hybrid Neuro-Fuzzy" :*** Dans cette catégorie, un réseau neuronal est utilisé pour tirer des paramètres du système flou (paramètres des ensembles flous, et des règles de logique floue : le poids des règles) d'un système à logique floue de façon itérative. La majorité des chercheurs utilisent le terme de neuro-flou pour désigner le système *hybride.*

IIIb.6. Systèmes "Cooperative Neuro-Fuzzy"

Dans un système coopératif les réseaux neuronaux sont utilisés que dans une phase initiale. Dans ce cas, le réseau de neurones détermine des sous-blocs du système de logique floue à l'aide de données d'apprentissage, après cela, les réseaux de neurones sont éliminés et le système flou est exécuté. Dans les systèmes neuro-flous de coopération, la structure n'est pas complètement interprétable ce qui peut être considéré comme un inconvénient (Fig.16).

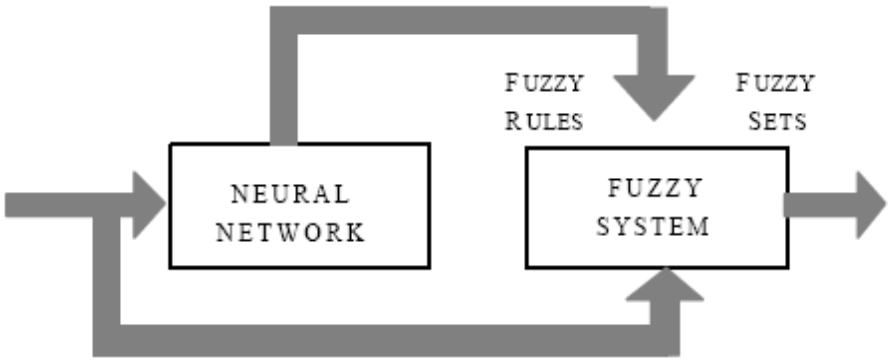


Fig. 16. Système neuro-flou de Coopération (Czogala, 2000)

IIIb.7. Systèmes "Concurrent Neuro-Fuzzy"

Un système concurrent n'est pas un système neuro-flou au sens strict, car le réseau de neurones fonctionne en même temps que le système flou. Cela signifie que les entrées pénétrant dans le système flou sont pré-traitées, puis le réseau de neurones traite les sorties du système concurrent ou dans le sens inverse. Dans les systèmes neuro-flous simultanées, les résultats ne sont pas complètement interprétables, ce qui peut être considéré comme un inconvénient (Fig. 17).

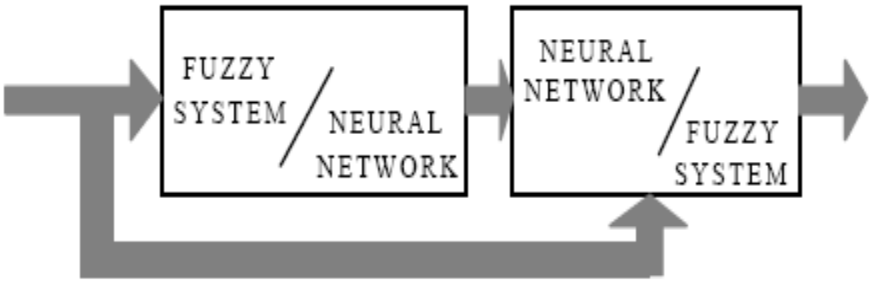


Fig. 17. Systèmes neuro-flou concurrent (Nauck, 1997)

IIIb.8. Systèmes "Hybrid Neuro-Fuzzy"

Dans la définition de Nauck (1997) : "Un système neuro-flou hybride est un système flou qui utilise un algorithme d'apprentissage basé sur les gradients ou inspiré par la théorie des réseaux de neurones (stratégies d'apprentissage heuristiques) pour déterminer ses paramètres (ensembles flous et règles floues) par le traitement des modèles (entrée et sortie)".

Un système neuro-flou peut être interprété comme un ensemble de règles de logique floue. Ce système peut être créé totalement de données d'entrée-sortie ou initialisé avec la connaissance a priori de la même façon que celles des règles floues. Le système résultant de la fusion de systèmes flous et des réseaux de neurones a pour avantages l'apprentissage à travers des modèles et l'interprétation facile de ses fonctionnalités.

Il y a plusieurs façons de développer des systèmes neuro-flous hybrides. Chaque utilisateur peut définir ses propres modèles qui sont similaires dans leur sens, mais peuvent présenter des différences fondamentales.

De nombreux types de systèmes neuro-flous sont représentés par des réseaux de neurones qui mettent en œuvre des fonctions logiques. Ce qui n'est pas nécessaire pour l'application d'un algorithme d'apprentissage dans un système flou. Cependant, la représentation d'un réseau de neurones est plus pratique car elle permet de visualiser la circulation des données à travers le système et les signaux d'erreur qui sont utilisées pour mettre à jour ses paramètres. L'avantage est de permettre la comparaison des différents modèles et de visualiser les différences structurelles. Il existe plusieurs architectures neuro-floues comme:

a) *Fuzzy Adaptive Learning Control Network* (FALCON) (Lin et Lee, 1991);
b) *Adaptive Network based Fuzzy Inference System* (ANFIS) (Jang, 1992);
c) *Generalized Approximate Reasoning based Intelligence Control* (GARIC) (Berenji et Khedkar, 1992);

d) *Neuronal Fuzzy Controller* (NEFCON) (Nauck et Kruse, 1997);
e) *Fuzzy Inference and Neural Network in Fuzzy Inference Software* (FINEST) (Tano et al., 1996);
f) *Fuzzy Net* (FUN) (Sulzberger et al., 1993);
g) *Self Constructing Neural Fuzzy Inference Network* (SONFIN) (Juang et Lin, 1998);
h) *Fuzzy Neural Network* (NFN) (Figueiredo et Gomide, 1999);
i) *Dynamic/Evolving Fuzzy Neural Network* (EFuNN and dmEFuNN) (Kasabov et Song, 1999);

Une description sommaire des cinq architectures neuro-flous les plus populaires est faite dans la section suivante.

IIIb.9. Architecture FALCON

Le Fuzzy Adaptive Learning Network Control (FALCON) est une architecture de cinq couches comme il est montré (Fig. 18). Il existe deux nœuds linguistiques pour chaque sortie. La première couche est pour les motifs et l'autre pour la production réelle du FALCON. La première couche cachée est responsable de l'application des variables d'entrée par rapport à chacune des fonctions d'appartenance. La deuxième couche cachée définit les antécédents des règles suivies par les conséquences de la troisième couche cachée. FALCON utilise un algorithme d'apprentissage hybride composé d'un apprentissage non supervisé pour définir les fonctions initiales et la base de la règle initiale en utilisant un algorithme d'apprentissage basé sur le gradient conjugué pour optimiser/régler les paramètres définitifs des fonctions d'appartenance afin de produire la sortie désirée.

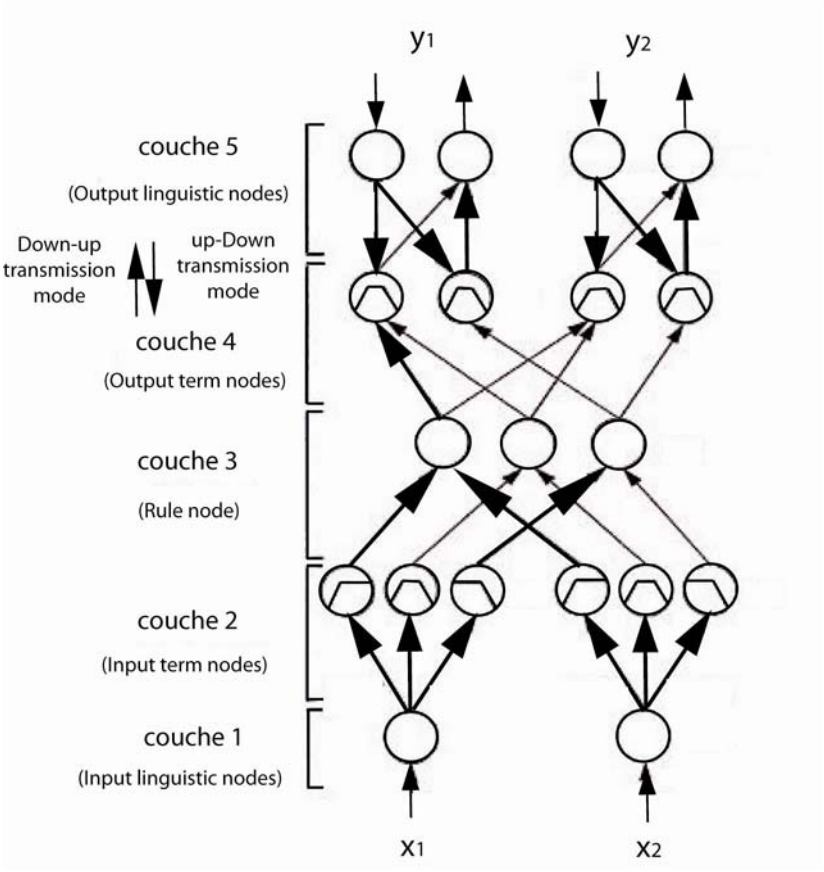


Fig. 18. Architecture FALCON (Lin et al., 1991).

Couche 1 (layer 1) : les nœuds d'entrée linguistiques (variables linguistiques d'entrée)

Couche 2 (layer 2) : nœuds à terme d'entrée

Couche 3 (layer 3) : nœuds de règles

Couche 4 (layer 4) : nœuds à terme de sortie

Couche 5 (layer 5) : nœuds linguistiques de sortie

IIIb.10. Architecture "ANFIS"

Le système d'inférence floue ANFIS basé sur les réseaux adaptatifs met en œuvre un système d'inférence flou à cinq couches (Fig.19) (Takagi et Sugeno, 1985). La première couche cachée est responsable de la cartographie de la variable d'entrée relativement à chacune des fonctions d'appartenance. L'opérateur T-norme est appliqué dans la deuxième couche cachée pour calculer les antécédents des règles. La troisième couche cachée normalise les forces règles suivies par la quatrième couche cachée où les conséquences des règles sont déterminées. La couche de sortie calcule la sortie globale en tant que somme de tous les signaux qui arrivent à cette couche.

L’ANFIS utilise le retour apprentissage de propagation pour déterminer les entrées des fonctions d'appartenance des paramètres et la méthode des moindres carrés moyens pour déterminer les paramètres de conséquences. Chaque étape de l'algorithme d'apprentissage itératif comporte deux parties. Dans la première partie, les formes d'entrée sont propagées et les paramètres conséquents sont calculés en utilisant la méthode des moindres carrés d’un algorithme itératif (Levenberg-Marquardt, 1963). Il permet d'obtenir une solution numérique au problème de minimisation d'une fonction, souvent non linéaire et dépendant de plusieurs variables. L'algorithme interpole l'algorithme de Gauss-Newton et l'algorithme du gradient. Plus stable que celui de Gauss-Newton, il trouve une solution même s'il est démarré très loin d'un minimum. Cependant, pour certaines fonctions très régulières, il peut converger légèrement moins vite. L'algorithme fut développé par Levenberg (1944), puis publié par Marquardt (1963).

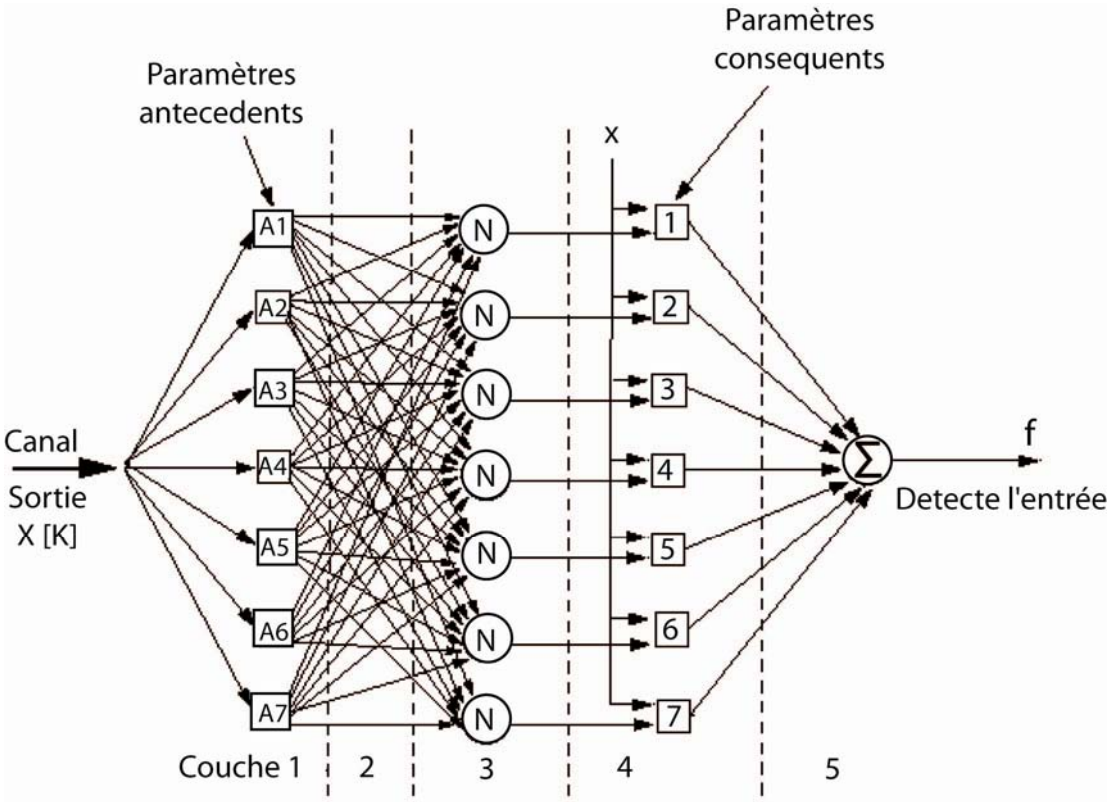


Fig. 19. Architecture ANFIS (Jang et al., 1992

IIIb.11. Architecture "GARIC"

Le raisonnement approximatif basé sur le contrôle de l’lntelligence généralisée (GARIC) met en œuvre un système neuro-flou en utilisant deux modules de réseaux neuronaux : le réseau

d'action de sélection (ASN) et le réseau d'action d'évaluation d'état (AEN). L'AEN est un évaluateur adaptatif des actions de l'ASN. L'ASN du GARIC est un réseau avancé de cinq couches. La structure GARIC-ASN est caractérisée par des connexions non pondérées entre couches (Fig.20).

La première couche cachée stocke les valeurs de la linguistique de toutes les variables d'entrée. Chaque entrée peut se connecter à la première couche qui représente ses valeurs linguistiques associées. La deuxième couche cachée représente les nœuds des règles floues qui déterminent le degré de chaque règle en utilisant un opérateur de compatibilité "softmin" (mécanisme d'activation pour résoudre les interactions additives). La troisième couche cachée représente les valeurs de la linguistique des variables de sortie. Les conclusions de chaque règle sont calculées en fonction de la force de règles antécédentes calculées dans les ganglions de la règle. GARIC utilise la moyenne de la méthode du maximum pour calculer la sortie des règles. Cette méthode nécessite une valeur numérique à la sortie de chaque règle. Ainsi, les conclusions doivent être transformées à partir de valeurs floues pour les valeurs numériques avant d'être accumulées dans la valeur de sortie finale du système. GARIC utilise un mélange de descente de la pente et l'apprentissage par renforcement pour un réglage fin de ses paramètres internes.

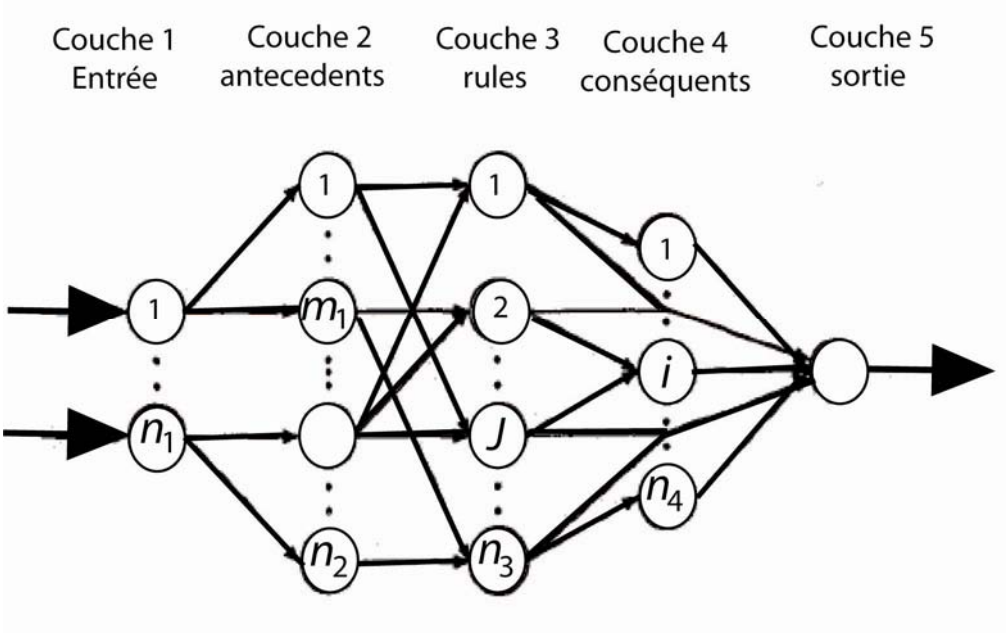


Fig. 20. Architecture GARIC (Bherenji et al., 1992)

IIIb.12. Architecture "NEFCON"

Le contrôleur réseau flou ou Neural Fuzzy (NEFCON) a été élaboré pour mettre en œuvre un système flou d'inférence de type Mamdani (Fig. 21). Les connexions dans cette architecture sont pondérés avec des jeux et des règles floues en utilisant les mêmes antécédents (appelés poids partagés), qui sont représentés par les ellipses tirées. Elles assurent l'intégrité de la base de règles. Les unités d'entrée assument la fonction interface de fuzzification, l'interface logique est représentée par la fonction de propagation et l'unité de sortie est responsable de l'interface de défuzzification. Le processus d'apprentissage de l'architecture NEFCON est basé sur un mélange de l'apprentissage par renforcement avec l'algorithme de propagation. Cette architecture peut être utilisée pour apprendre la base de règles depuis le début, s'il n'y a pas de connaissance a priori du système, ou pour optimiser une base de règles définies initialement manuellement. NEFCON a deux variantes: pour l'approximation de fonctions (NEFPROX) et pour les tâches de classification (NEFCLASS).

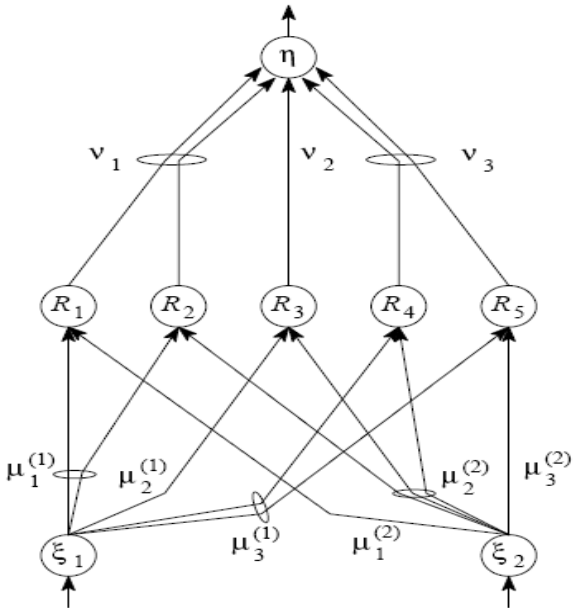


Fig. 21. Architecture NEFCON (Nauck et al., 1997).

IIIb.13. Architecture EFuNN

Dans l'architecture en évolution Neural Fuzzy (EFuNN) tous les nœuds sont créés pendant la phase d'apprentissage. La première couche de données passe à la seconde couche qui calcule le degré de compatibilité en fonction des fonctions d'appartenance prédéterminées. La troisième couche contient des nœuds de règles floues représentant les prototypes de données

d'entrée-sortie en tant qu'association de hyper-sphères de l'entrée floue et espaces de sortie floue. Chaque nœud de la règle est défini par deux vecteurs de poids de connexion qui sont ajustés au moyen d'une technique d'apprentissage hybride. La quatrième couche calcule la mesure dans laquelle les fonctions d'appartenance de sortie sont mises en correspondance. Les données d'entrée et la cinquième couche effectuent la défuzzification et calculent la valeur numérique de la grandeur de sortie. Le réseau dynamique en évolution Neural Fuzzy (dmEFuNN) est une version modifiée de l'EFuNN (Fig.22). L'idée est non seulement l'activation de la règle nœud qui se propage mais un groupe de nœuds de règle qui est dynamiquement sélectionné pour chaque nouveau vecteur d'entrée. Leurs valeurs d'activation sont utilisées pour le calcul des paramètres dynamiques de la fonction de sortie. Alors qu'EFuNN met en œuvre des règles floues de type Mamdani, dmEFuNN utilise des règles floues de type Takagi-Sugeno.

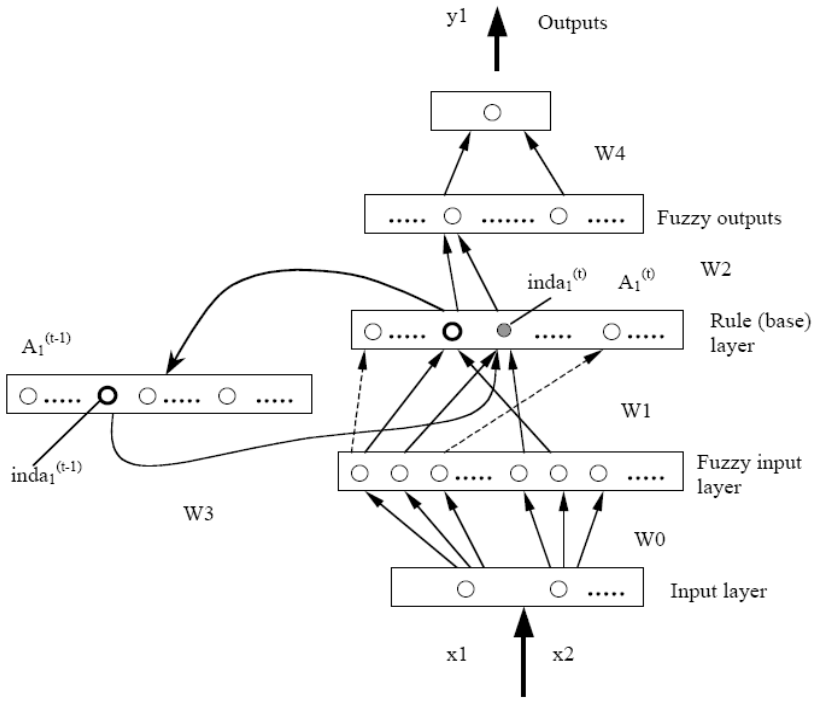


Fig. 22. Architecture EFuNN (Kasabov et al., 1999)

Pour obtenir une description détaillée de plusieurs de ces architectures, au-delà des références pointues spécifiques formulées dans le présent document, une étude détaillée a été faite par Abraham et Nath (2000), où il peut être trouvé une description détaillée de plusieurs architectures neuro-flous bien connus et leur algorithme d'apprentissage respectif.

IIIb.14. Discussion et Applications

Les systèmes neuro-flous hybrides présentent un modèle interprétable qui a des capacités d'apprentissage supervisée. Dans FALCON, GARIC, ANFIS, NEFCON, SONFIN et FINEST le processus d'apprentissage ne concerne que l'adaptation des paramètres internes d'une structure fixe du système. Pour les problèmes complexes, le calcul sera exigeant pour déterminer tous les paramètres (paramètres locaux, paramètres de conséquences, nombre de règles, etc.) car les paramètres vont croître de façon exponentielle.

Une caractéristique importante de l'architecture dmEFuNN et EFuNN est de faire de la formation dans une seule itération. Cette caractéristique permet à l'application de l'adaptation en ligne d'une manière simple.

Abraham et Nath (2002) ont proposé une approche évolutive basée sur des algorithmes génétiques pour l'optimisation de tous les paramètres de la structure d'un système neuro-flou (type de système flou, nombre de règles, paramètres, opérateurs d'inférence, règles et fonctions d'appartenance).

Dans le domaine industriel, ces architectures ont été d'abord appliquées à la modélisation de systèmes non-linéaires et à la technique de commande. Cependant, ces architectures sont utilisées dans presque tous les domaines de connaissances où une fonction non-linéaire doit être approchée. Les domaines d'application des systèmes neuro-flous sont la médecine, l'économie, le contrôle, la mécanique, la physique, la chimie, etc.

IIIb.15. Synthèse des résultats

Les articles cités plus haut résultent, d'une manière générale, de la dernière décennie de l'enquête dans le domaine des fonctions de modélisation non-linéaire à travers les systèmes neuro-flous.

Un duo pour le grand nombre d'outils communs : il est toujours difficile de comparer conceptuellement les différentes architectures et à évaluer comparativement leurs performances. En termes génériques les points de bibliographie que les neuro-systèmes flous mettent en œuvre sont des systèmes d'inférence floue de type Takagi-Sugeno. Ceux-ci arrivent à obtenir des résultats plus précis que les approches mettant en œuvre des systèmes d'inférence neuro-floue de type de Mamdani, bien que la complexité du calcul reste grande. La ligne directrice pour mettre en œuvre des systèmes neuro-flous très efficaces doit passer par les caractéristiques suivantes: apprentissage rapide, adaptabilité en ligne, auto-ajustement

dans le but d'obtenir la petite erreur globale possible, petite complexité de calcul. L'acquisition des données et le pré-traitement de ces données d'entraînement d'entrée sont également très importants pour le succès de l'application des architectures neuro-floues.

Toutes les architectures neuro-floues utilisent les techniques de l’algorithme du gradient (steepest descent) pour l'apprentissage de ses paramètres internes. Pour accélérer la convergence du calcul de ces paramètres, il serait intéressant d'explorer d'autres algorithmes efficaces de l'apprentissage des réseaux neuronaux comme le gradient conjugué ou recherche de Levenberg-Marquardt, en dépit de l'algorithme de rétropropagation.

IV.1. Rappel sur les réservoirs triasiques de *Hassi R'Mel Est*

Dans le domaine pétrolier, les méthodes de caractérisation des réservoirs sont précieuses, car elles fournissent une meilleure description des capacités de stockage et d'écoulement d'un réservoir pétrolier (Aggoun et al., 2006). Les réservoirs de sable schisteux représentent un défi pour les ingénieurs et les géologues qui s'efforcent de les caractériser, car ils ont tendance à être très hétérogènes en raison du processus de déposition et diagénétique (Elgaghah et al., 2001). La forte hétérogénéité pétrophysique au niveau des sables argileux est démontrée par la grande variabilité observée, en particulier dans la relation porosité-perméabilité rencontrée dans l'analyse des données de base (Carlos, 2004).

La caractérisation des réservoirs de sable schisteux en terme de perméabilité (Abbaszadeh et al., 1995) est un moyen pratique de les définir. La présence d'unités distinctes ayant des caractéristiques pétrophysiques particulières telles que la porosité, la perméabilité, la saturation en eau, le rayon des pores, le stockage et la capacité d'écoulement contribuent à créer une forte caractérisation de réservoir (Bear, 1972). Le plus tôt dans la vie d'un réservoir la détermination de l'unité de débit (HFU) est réalisée, le mieux est la compréhension de sa performance future (Hambalek et Gonzalez, 2003). Au fil des ans, il y a eu plusieurs tentatives pour corréler la perméabilité avec d'autres données de diagraphies pétrophysiques telles que la porosité, la saturation en eau irréductible, la teneur en argile, etc. (Ebanks et al., 1992). Cependant, ces dernières années, de nombreux auteurs ont tenté d'établir une corrélation entre la perméabilité de base et les données de carottes. La connaissance de la perméabilité est essentielle pour développer une description efficace du réservoir. La perméabilité de la formation contrôle les stratégies impliquant les complétions de puits, la stimulation et la gestion des réservoirs (Lim, 2005). Les données de perméabilité peuvent être obtenues à partir d'essai de puits, d'analyse des données de carottes et des diagraphies de puits. Les puits ne sont tous carottés, en raison de problèmes qui peuvent survenir au cours du carottage et de leurs coûts élevés. En général, l'estimation du log de perméabilité (par diagraphie) est considérée comme la méthode la moins coûteuse, dans la mesure où l'on peut utiliser les valeurs de porosité et les paramètres correspondants, pour l'estimation des saturations en eau et donc en hydrocarbures. Mais la prédiction de la perméabilité dans les formations de sable schisteux à partir de données de diagraphie est difficile et complexe. Une corrélation de base entre la perméabilité et la porosité ne peut pas être établie, en raison de l'effet d'autres paramètres de forage qui doivent être ancrées dans la corrélation. Outre tous

ces défis pour l'estimation de la perméabilité de diagraphies de puits, leur utilisation fournit un profil de perméabilité continue tout au long de l'intervalle particulier qui peut être décrit comme une unité de débit hydraulique (Al-Ajmi et Holditch, 2000). Bon nombre de problèmes rencontrés dans l'ingénierie et en géosciences peuvent effectivement être modélisés mathématiquement. Cependant, lors de la construction de ces modèles, plusieurs hypothèses doivent être faites qui peuvent être fausses en réalité. Une approche développée par Zadeh (1965) est basée sur un concept connu sous le nom d'*ensembles flous*. La théorie des sous-ensembles flous est une théorie mathématique du domaine de l'algèbre abstraite. Elle a été développée par Zadeh (1965) afin de représenter mathématiquement l'imprécision relative à certaines classes d'objets et sert de fondement à la logique floue. Le développement de cette idée a conduit à de nombreuses implémentations, une réussite de systèmes flous ou de systèmes d'inférence floue (Mohaghegh, 2000). *Un système d'inférence floue (FIS) est un système qui utilise des ensembles flous pour prendre des décisions ou de tirer des conclusions.*

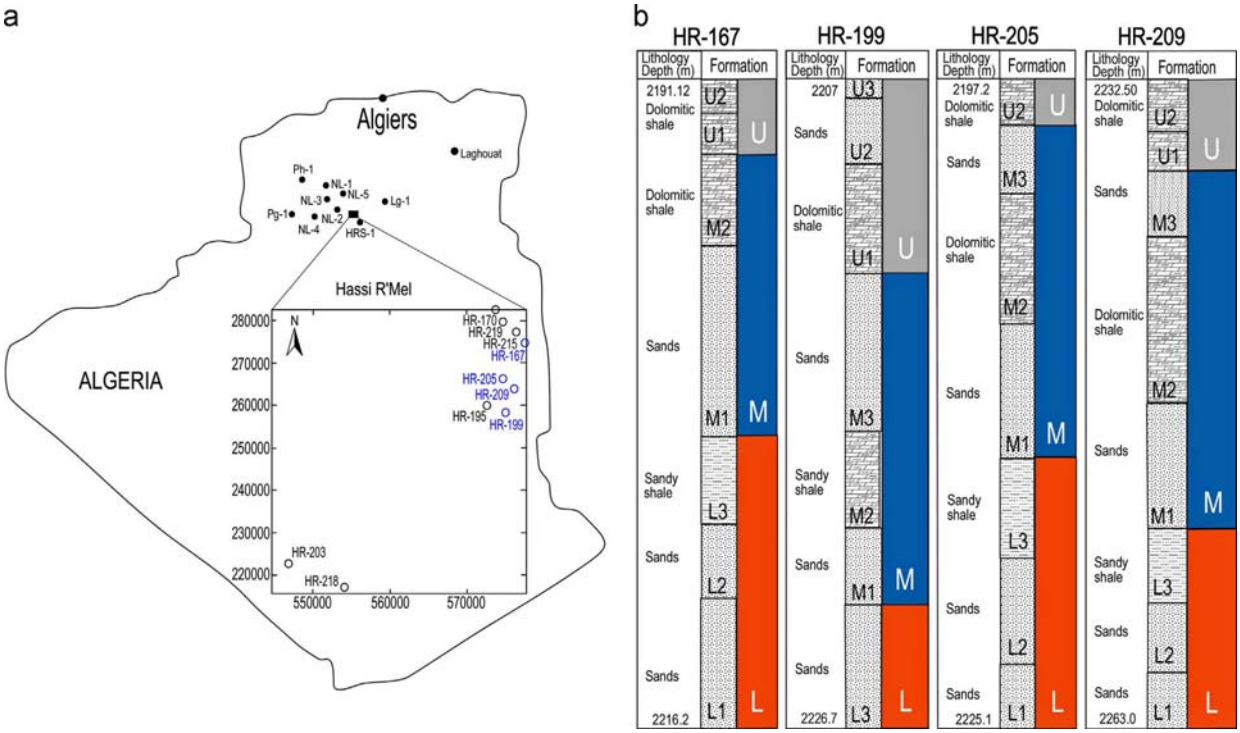


Fig. 23. (a) Partie du champ montrant les principaux puits de Hassi R'Mel-Est (HRE). (b) logs lithostratigraphiques des formations de la région HRE dans les puits HR-167, HR-199, HR-205 et HR-209.

IV.2. Zones d'intérêt à huile

Le présent travail traite de plusieurs puits situés entre deux failles parallèles dans la partie huile du champ de Hassi R'Mel. Nous allons nous concentrer sur quatre puits documentés: HR-195, HR-199, HR-205 et HR-209 (Fig. 23) ayant fait objet d'enregistrements diagraphiques et les données de carottes de tous les puits qui étaient disponibles. Le puits HR-195 est partiellement carotté. La zone d'intérêt du réservoir se trouve dans les grès du Trias. La zone la plus importante de ce réservoir est l'intervalle A qui a une épaisseur moyenne de 31 m. Elle a été subdivisée en trois principales unités géologiques U, M et L. La zone A est subdivisé en 7 sous-unités ici de L1 à U3. Les unités U2 et U3 sont très argileuses et leur épaisseur nette est proche de zéro, elles n'ont pas été considérées dans cette étude. Ainsi, le réservoir d'intérêt, appelé réservoir gréseux contient les unités de couches suivantes: L1, L2, L3, M1, M2, M3, et U1. U1 fait défaut dans le puits HR-205 qui est liée à l'érosion où probablement à une discordance.

IV. 4. Back-propagation neural networks (BPNN)

Un réseau de neurones artificiels (ANN) est un système basé sur l'exploitation de réseaux de neurones biologiques, et peut donc être défini comme une émulation de systèmes de neurones biologiques (Matlab User's Guide, 2012a). Le réseau de rétro-propagation de neurones (BPNN), utilisé dans cette étude, est une technique d'apprentissage supervisé qui calcule la différence entre la sortie ANN-calculé et le résultat souhaité correspondant, à partir de l'ensemble des données d'apprentissage (Bhatt et al., 2002).

Les réseaux de neurones ont été utilisés avec succès dans une variété d'applications d'ingénierie du pétrole telles que la caractérisation de réservoir, la conception optimale des traitements de stimulation, et l'optimisation des opérations de terrain (Mohaghegh, 2000 ; Tamhane et al., 2000). L'élément de traitement de base d'un réseau de neurones est un *neurone*. Fondamentalement, un neurone biologique reçoit des entrées provenant d'autres sources, les combine d'une certaine façon, effectue une opération non-linéaire en général sur le résultat, puis délivre en sortie le résultat. Un neurone typique contient un corps cellulaire, des dendrites, et un axone (Mohaghegh, 2000). Le déroulement des opérations standard de la méthode de rétropropagation utilisée dans ce travail se fait en cinq étapes:

1) initialisation des poids synaptiques du réseau pour les petites valeurs aléatoires ;

2) l'ensemble de formation de paires d'entrée / sortie qui présente une configuration d'entrée et calcule la réponse du réseau ;
3) la réponse du réseau est comparée aux données réelles en utilisant les équations et les erreurs locales calculées ;
4) les poids du réseau sont mis à jour ;
5) les étapes 2 à 4 sont répétées jusqu'à ce qu'une erreur globale minimale soit obtenue.

En général, il est supposé un nombre "n" d'une entrée et une sortie, la sortie (y) peut être présentée selon la relation suivante (x_1, x_2, x_3, ..., x_n.)

$$Y = f\left(\sum_{i=1}^{n} W_i X_i\right) \qquad (5)$$

et $f(net) = W^T X = W_1X_1 + W_2X_2 \ldots W_nX_n$ (6)

où f (net) fait référence à la fonction d'activation ou de transfert et w_i est le vecteur de pondération.

La méthode présentée dans cet article utilise une méthodologie générale pour prédire la perméabilité à partir des données de diagraphie. L'analyse typologique est constamment utilisée pour identifier et classer les groupes en fonction des caractéristiques uniques et les mesures de l'information reflètent les faciès du réservoir, alors un modèle de logique floue est appliquée à chaque groupe pour déterminer la perméabilité de la formation dans l'intervalle du Trias de Hassi R'Mel. Les données réelles obtenues dans le domaine des formations des réservoirs gréseux sont utilisées pour démontrer l'applicabilité et la compétence de la technique proposée. Ce qui fournit un profil continu de la perméabilité à travers un horizon de formation particulière, par opposition à la perméabilité présentée par l'analyse obtenue à partir des données d'essai de puits, ce qui donne une valeur moyenne de la zone de drainage de l'assemblage (Amaefule et al., 1993).

La caractérisation de réservoir est un processus permettant d'attribuer quantitativement les propriétés des réservoirs, en reconnaissant l'information et les incertitudes géologiques de la variabilité spatiale (Lake et Carroll, 1986). La perméabilité est souvent évaluée par deux méthodes différentes : (i) par mesure directe, au moyen de fiches de base ou (ii) par estimation afin de bien identifier les paramètres, en utilisant soit des relations empiriques ou une certaine forme paramétrique/non paramétrique de l'analyse statistique de régression (Nashawi et Malallah, 2010). La perméabilité est contrôlée par deux caractéristiques

sédimentaires : (i) la taille de grain et le tri ainsi que (ii) les caractéristiques diagénétiques (Nashawi et Malallah, 2010).

IV. 5. Modèle de logique floue

Un système d'inférence floue (FIS) est un processus de formulation de la cartographie à partir d'une entrée donnée à une sortie en utilisant la logique floue (Matlab User's Guide, 2012a). La cartographie fournit alors une base à partir de laquelle les décisions peuvent être prises. Dans la plupart des problèmes résolus en géosciences en utilisant ANFIS et ANN ce sont les prédictions de comportement basées sur des résultats expérimentaux donnés qui sont utilisées pour les données de formation et de test.

Le processus d'inférence floue implique toutes les étapes qui sont décrites dans les sections précédentes: fonctions d'appartenance, opérations logiques, et modèle "If-Then" (Fig. 24). Nous pouvons mettre en œuvre deux types de FIS, dans la boîte à outils: type Mamdani (Mamdani, 1976) et type Sugeno (Sugeno, 1985a ; Sugeno, 1985b). Ils varient quelque peu de la manière dont les sorties sont déterminés (Mamdani et Assilian 1975 ; Mamdani ,1976 ; Sugeno, 1985a ; Sugeno, 1985b).

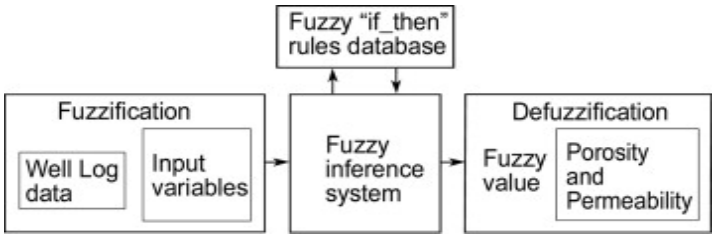


Fig.24. Modèle de règle "if-Then" (Matlab User's Guide, 2012a).

Les incertitudes peuvent se produire lorsque les données géologiques sont utilisées principalement dans la prédiction des paramètres pétrophysiques où les données de diagraphies de puits présentent des erreurs (Nikravesh et Aminzadeh, 2003).

Les FIS ont été appliqués avec succès dans des domaines tels que le contrôle automatique, la classification des données, l'analyse de la décision et les systèmes experts. En raison de leur caractère multidisciplinaire, ils sont associés à un certain nombre de noms, tels que les systèmes à base de règle de logique floue, les systèmes experts flous, la modélisation floue, la mémoire associative floue et les contrôleurs de logique floue. La méthode d'inférence floue de Mamdani est une méthode floue la plus fréquemment observée. La méthode de Mamdani a été parmi les premiers systèmes de contrôles construits en utilisant la théorie des ensembles flous.

Le travail de Mamdani et Assilian (1975) était fondé sur l'article de Zadeh (1973) et donc sur des algorithmes flous pour les systèmes complexes et les processus de décision. Toutefois, le processus d'inférence décrit dans les prochaines sections montre quelque chose que nous faisons pour créer le modèle, afin d'évaluer l'efficacité du modèle et exécuter le modèle dans un sujet bien spécifique (étude des puits par exemple). Il est alors possible de mettre en œuvre deux types de FIS (Mamdani et Sugeno) dans la boîte à outils.

Ces deux types de systèmes d'inférence varient quelque peu à la façon dont les sorties sont déterminées.

(i) Le type d'inférence Mamdani, tel que défini dans la boîte à outils, s'attend à ce que les fonctions d'appartenance de sortie doivent être des ensembles flous. Après le processus d'agrégation, il existe un sous-ensemble flou pour chaque variable de sortie qui doit faire la "défuzzification ". Il est possible, et dans de nombreux cas beaucoup plus efficaces, d'utiliser un seul pic pour les fonctions d'appartenance de sortie plutôt que d'un ensemble flou distribué. Elle améliore l'efficacité du processus de défuzzification parce qu'elle simplifie considérablement les calculs requis par le procédé de Mamdani plus général. Ce procédé trouve le centre de gravité d'une fonction bidimensionnelle, plutôt que d'intégrer la fonction pour ajuster le centre de gravité, et utilise la moyenne pondérée de quelques points de données.

(ii) Les systèmes de type Sugeno peuvent être utilisées pour modéliser un système d'inférence dans lequel les fonctions d'appartenance de sortie sont soit linéaires, soit constantes (Fig. 25).

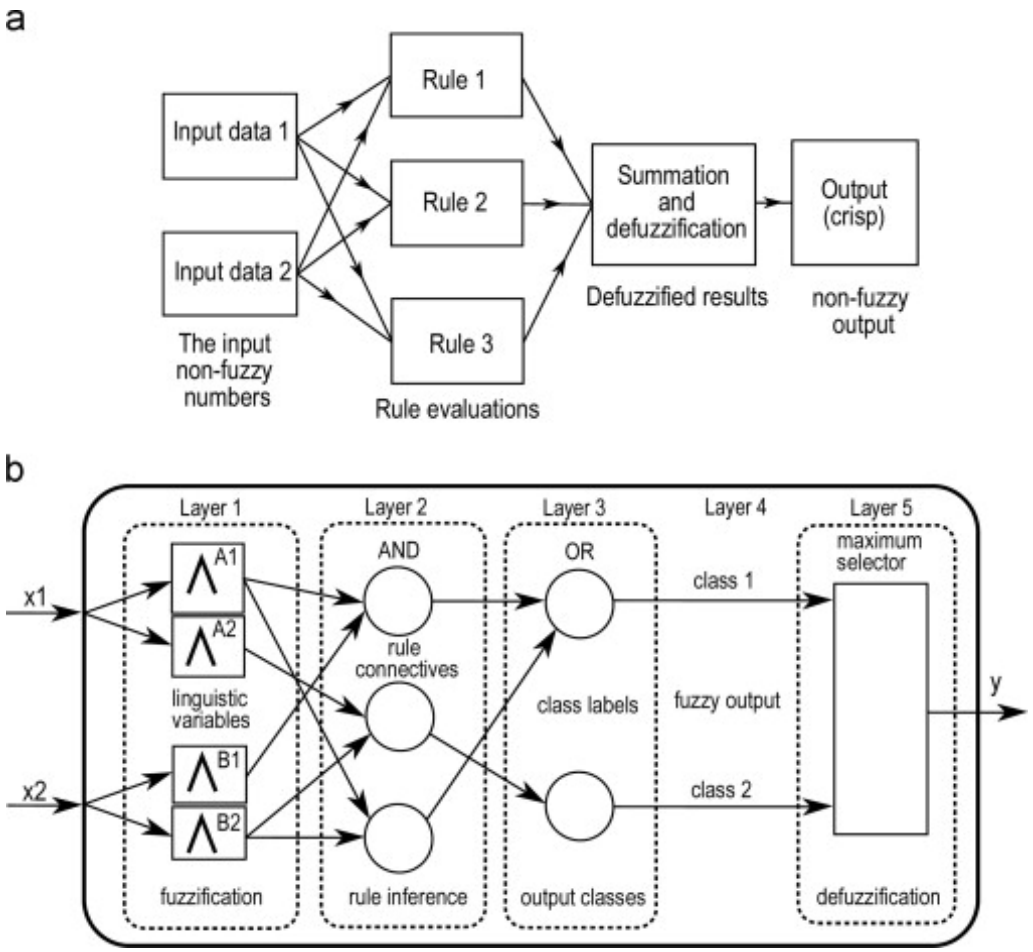


Fig.25. (a) Architecture neuronale du classificateur NF; (b) schéma de l'architecture neuro-floue.

Couche 3. Combinaison : Si plusieurs règles floues ont la même classe de conséquence, cette couche combine les points forts. Généralement, le maximum conjonctif (ou opération) est utilisé.

Couche 4. Sorties floues: Dans cette couche, les valeurs floues des classes sont disponibles et décrivent de combien l'entrée du système correspond aux classes.

Couche 5. Défuzzification: Si la classification de sortie est nécessaire, la meilleure classe pour l'entrée est choisie comme classe de sortie.

Avec l'utilisation des éditeurs et des afficheurs dans la boîte à outils de la logique floue, nous pouvons construire des règles, définir des fonctions d'appartenance, et analyser le comportement d'un FIS. Le Système d'architecture neuro-floue à l'aide des règles floues est représenté en (Fig. 25a)**.**

Rule 1: If x1 is A1 and x2 is B1, then class is 1.

Rule 2: If x1 is A2 and x2 is B2, then class is 2.

Rule 3: If x1 is A1 and x2 is B3, then class is 1.

La fonction "*subclust*" utilisée dans Matlab trouve le point de données optimal à définir un centre de cluster en fonction des rayons et a une valeur entre 0 et 1. Elle détermine la taille du groupe dans chaque dimension, lorsqu'il s'agit d'un vecteur; et a une entrée pour chaque dimension des données. Les valeurs des rayons varient généralement entre 0.2 et 0. 5.

Les rayons "*radii*" utilisés pour tous les puits avec "If-Then", règles pour les données d'entrée de paramètres pétrophysiques pour *la porosité* sont:

If (GR is in1mf1) and (ΔT is in2mf1) and (Rhob is in3mf1) and (RT is in4mf1) and (Nphi is in5mf1) and (S_W is in6mf1) then (CorPor is out1mf1).

If (GR is in1mf2) and (ΔT is in2mf2) and (Rhob is in3mf2) and (RT is in4mf2) and (Nphi is in5mf2) and (S_W is in6mf2) then (CorPor is out1mf2).

If (GR is in1mf3) and (ΔT is in2mf3) and (Rhob is in3mf3) and (RT is in4mf3) and (Nphi is in5mf3) and (S_W is in6mf3) then (CorPor is out1mf3).

If (GR is in1mf4) and (ΔT is in2mf4) and (Rhob is in3mf4) and (RT is in4mf4) and (Nphi is in5mf4) and (S_W is in6mf4) then (CorPor is out1mf4).

If (GR is in1mf5) and (ΔT is in2mf5) and (Rhob is in3mf5) and (RT is in4mf5) and (Nphi is in5mf5) and (S_W is in6mf5) then (CorPor is out1mf5).

Les rayons "*radii*" utilisés pour tous les puits avec "If-Then" règles pour les données d'entrée de paramètres pétrophysiques pour la *perméabilité* sont:

If (GR is in1mf1) and (ΔT is in2mf1) and (Rhob is in3mf1) and (RT is in4mf1) and (Nphi is in5mf1) and (S_W is in6mf1) then (CPerm is out1mf1).

If (GR is in1mf2) and (ΔT is in2mf2) and (Rhob is in3mf2) and (RT is in4mf2) and (Nphi is in5mf2) and (S_W is in6mf2) then (CPerm is out1mf2).

If (GR is in1mf3) and (ΔT is in2mf3) and (Rhob is in3mf3) and (RT is in4mf3) and (Nphi is in5mf3) and (S_W is in6mf3) then (CPerm is out1mf3).

If (GR is in1mf4) and (ΔT is in2mf4) and (Rhob is in3mf4) and (RT is in4mf4) and (Nphi is in5mf4) and (S_W is in6mf4) then (CPerm is out1mf4).

If (GR is in1mf5) and (ΔT is in2mf5) and (Rhob is in3mf5) and (RT is in4mf5) and (Nphi is in5mf5) and (S_W is in6mf5) then (CPerm is out1mf5).

Les valeurs de in1mf1, in2mf1, ... représentent la fonction d'appartenance floue de chaque valeur des données de diagraphies dans l'éditeur ANFIS de Matlab, où toutes les fonctions d'appartenance associées à toutes les variables d'entrée/sortie pour l'ensemble de la FIS sont éditées.

L'architecture du classificateur neuro-flou (Fig. 25b) est légèrement différente de la structure utilisée dans les approximations faites par Tommi (1994). Le tableau de la procédure de formation dans un réseau de neurones supervisé (Fig. 26) est développé une fois avec succès et fournit une base de connaissances permanente.

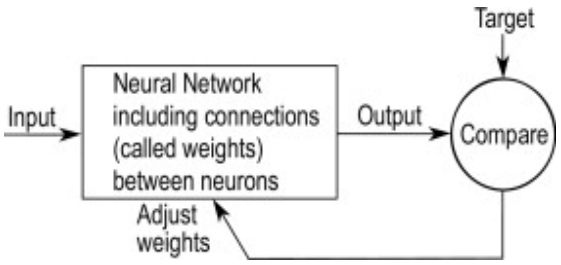


Fig.26. Diagramme de la procédure de formation dans un réseau de neurones supervisé (Matlab User's Guide, 2012a).

Le modèle de Takagi-Sugeno a été appliqué aux données de test et de formation des puits. Les fonctions d'appartenance et les règles "If-Then" sont générées par le programme fuzzy sous Matlab par la méthode de classification soustractive. Par conséquent, les résultats de la porosité et de la perméabilité ont été comparés à la porosité et perméabilité mesurées sur carottes. Pour tester les performances du système d'inférence, la prédiction (R^2), qui représente le coefficient de corrélation au carré est utilisé dans la plage de 0 à 1 :

Après construction d'un modèle de régression, il est souvent nécessaire de vérifier dans quelle mesure le modèle correspond aux données et, par conséquent, sa capacité à prédire avec précision les observations futures utilisant des statistiques de régression:

SS_{Error}: Somme des carrés des erreurs due à la variabilité inexpliquée dans les données est :

$$SS_{Error} = \sum_i \left(Y_i - \hat{Y}_i \right)^2 \qquad (1)$$

$SS_{Régression}$: La somme des carrés des erreurs en régression est exprimée par :

$$SS_{Regressionr} = \sum_i \left(Y_i - \hat{Y}_i \right)^2 \quad (2)$$

SS_{Total}: Somme totale des erreurs moindres carrés

R^2: Coefficient défini comme le rapport de la somme des carrés expliquée par le modèle de régression, et la somme totale des carrés.

$$R^2 = \frac{SS_{Regression}}{SS_{Total}} \quad (3)$$

Le coefficient R^2 peut aussi être exprimé en termes de SS_{Error} :

$$R^2 = 1 - \frac{SS_{Error}}{SS_{Total}} \quad (4)$$

Ainsi, meilleure est la régression correspondant aux données, R^2 est proche de l'unité (formation et tests),

Pour la porosité (de formation): $R^2 = 0.8263$

Pour la perméabilité prédite (test): $R^2 = 0.7302$

IV.6. Modèle neuro-flou

Les systèmes NF hybrides combinent les avantages des systèmes flous (qui traitent de la connaissance explicite) avec ceux des NN (qui traitent de la connaissance implicite). D'autre part, la logique floue (FL) augmente la capacité de généralisation d'un système de réseau neuronal (NN) en fournissant une sortie plus fiable et nécessaire lorsque l'extrapolation est au-delà de la limite des données d'apprentissage. Un diagramme schématique des flux d'information dans un système NF est représenté en (Fig. 27).

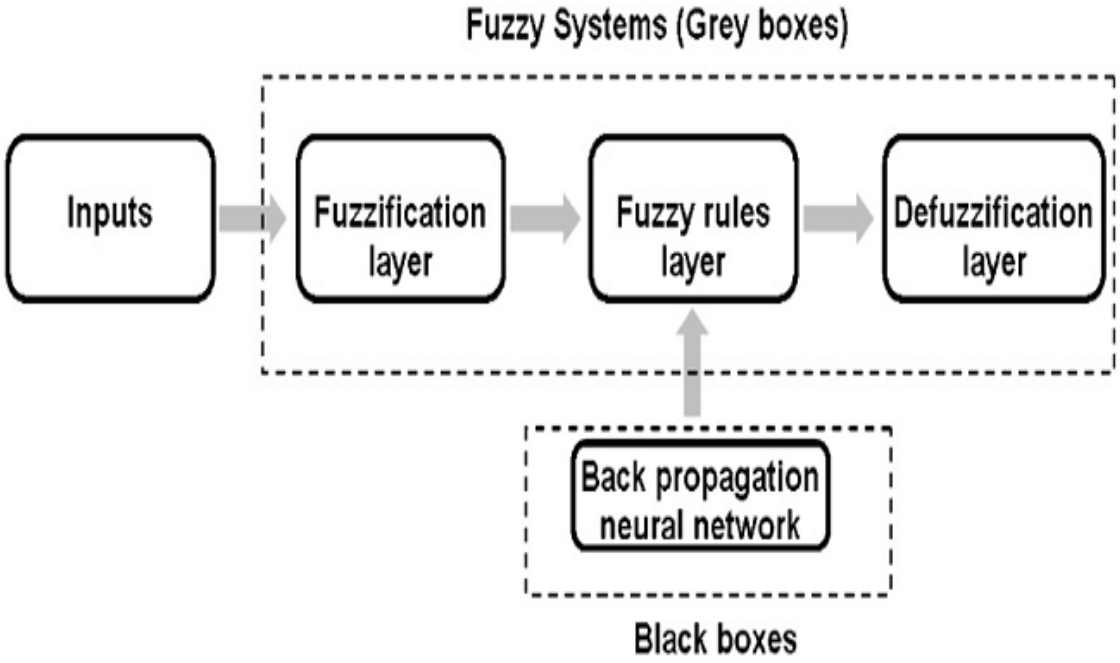


Fig.27. Représentation schématique du flux de l'information dans un système NF (Tommi, 1994)

Le Procédé d'apprentissage neuro-adaptatif fonctionne de façon similaire à celui des réseaux de neurones. Les techniques d'apprentissage neuro-adaptatif fournissent un procédé pour la procédure de modélisation floue afin d'avoir des informations sur un ensemble de données. La boite à outils (ToolBox) de la logique floue calcule les fonctions d'appartenance des paramètres qui permettent d'extraire le meilleur système d'inférence floue associée à suivre les données d'entrée/sortie. La fonction ToolBox de la logique floue qui accomplit cette adhésion d'ajustement des paramètres de fonctionnement est appelé ANFIS. Elle a été proposée à l'origine par Jang (1993). C'est un système flou formé par un algorithme dérivé de la théorie ANN. En outre, il devient un nouvel outil amélioré et une approche de modélisation fondée sur les données pour déterminer le comportement des systèmes dynamiques complexes (Jang et Sun, 1995 ; Jang et al., 1997). L'approche de modélisation utilisée par ANFIS est similaire à de nombreuses techniques d'identification du système. Tout d'abord, nous faisons l'hypothèse d'une structure paramétrée de modèle (entrées relatives à des fonctions d'appartenance à des règles aux sorties de fonctions d'appartenance, et ainsi de suite). Ensuite, nous recueillons des données d'entrée/sortie sous une forme qui sera utilisable par ANFIS pour former le modèle de la FIS pour émuler les données de formation qui lui sont présentées en modifiant les paramètres de fonction d'appartenance, selon un critère d'erreur choisi.

La création d'un système flou ou Fuzzy Logic (FL) se compose de quatre étapes de base:

1) Pour chaque variable d'entrée ou une variable de résultat, un ensemble de fonctions d'appartenance (PDE) doit être définis. Un "membership function" (MF) définit la mesure dans laquelle la valeur d'une variable appartient au groupe et est habituellement un terme linguistique, comme élevé ou faible.
2) Les déclarations, ou les règles sont définies pour les MF de chaque variable au résultat, normalement à travers une série de déclarations "If-Then". Par exemple, une règle serait: si le manteau neigeux (condition) est faible, les eaux de ruissellement (conclusion) sont faibles.
3) Les règles sont mathématiquement évaluées et les résultats sont combinés, c'est-à-dire chaque règle est évaluée grâce à un processus "d'implication", et les résultats de toutes les règles sont combinés dans un processus «d'agrégation».
4) La fonction résultante est évaluée comme un nombre à travers un processus de "défuzzification". Fondamentalement, un système FL typique est composé de logique floue, "defuzzifier", base de règles, MF et une procédure d'inférence (Fig. 25).

IV.7. Données utilisées et méthodologie

Le réservoir du Trias de Hassi R'Mel Est (HRE) a été enregistré avec une suite de diagraphies différées. De la même façon que d'autres puits les enregistrements ont été acquis par Schlumberger (1988).

Toutes les données obtenues à partir des carottes et des diagraphies de puits sont analysées de façon à modéliser une zonation du réservoir du Trias, avec des techniques statistiques dont les graphes et les histogrammes sont traités par les méthodes de régression linéaire et de régression multiple (Lim, 2005). L'analyse de base conventionnelle, y compris la densité des grains, porosité et perméabilité à l'air des mesures, a été réalisée pour tous les puits par Sonatrach (CRD) à Boumerdès. Jusqu'à ce jour, aucune donnée n'a été acquise afin de déterminer l'impact de la pression de surcharge net sur la perméabilité. Par conséquent, les données de perméabilité utilisées dans la caractérisation de réservoir est celle diffusée par le laboratoire.

Plusieurs puits de production de pétrole, à savoir HR-167, HR-199, HR-205, HR-209, et RH-170, HR-203, HR-215, HR-218, HR-219 ont été étudiés afin de caractériser la formation en unités du Trias dont seulement 4 ont fait l'objet de notre étude pour des raisons de disponibilité des carottes. Le puits HR-195 est partiellement carotté.

Les enregistrements des puits HR-167, HR-199, HR-205, et HR-209 sont le gamma Ray classique (*GR*), le temps de transit (ΔT), la densité apparente (δ_b), la résistivité (R_t), la porosité neutron (NPhi), et la saturation en eau (Sw) comme paramètre interprété à partir des données de diagraphies.

Ainsi les diagraphies pour les 10 puits sont disponibles sous des formes classiques. Les enregistrements ont été lus avec 1 m d'incrément. Les interprétations incluent seulement la formation du Trias. Les paramètres des données de diagraphie pour les puits étudiés sont donnés en annexe A.

L'estimation de la lithologie est le résultat des premières interprétations : calculs du volume d'argile, détermination de la porosité sonique, neutron et densité. Des crossplots nécessaires à la détermination de la porosité et des correctifs ont été construits.

Les étapes nécessaires à l'établissement de ce processus sont les suivantes :

1) **Elimination de données** erronées et aberrantes à partir des données de diagraphies brutes des puits de sondage.
2) **Organiser les données** en ensembles de données d'entrée, y compris GR, ΔT, Rohb, RT, Phin, Sw et les ensembles de données de sortie comprenant la perméabilité et porosité carotte.
3) **Normalisatio**n de l'entrée et de la sortie des ensembles de données (entre les intervalles 0 et 1) pour dimensionner les données et supprimer l'effet de mise à l'échelle.
4) **Diviser les données** en formation, contrôle et essai d'ensembles de données (Partitionnement)
5) le **regroupement des ensembles** d'entrée et de sortie des données à l'aide de C-moyennes floues (Fuzzy C-Means), k-means floues (Fuzzy k-means) ou des méthodes de classification soustractive.
6) **fuzzification** impliquant la conversion des données numériques dans le domaine du monde réel à nombres flous dans le domaine flou, donnant lieu à la construction d'un système d'inférence floue (FIS) qui consiste à fixer les fonctions d'appartenance et établir de règles floues.
7) **Deffuzzification**, facultative, consistant à convertir la logique floue dérivée des données numériques dans le domaine du monde réel.

IV.7.1 Organisation des données : Les données pour le modèle neuro-flou viennent des puits de Hassi R'Mel. Le choix de ce champ est fondé sur des considérations géologiques avec la présence d'une bonne couverture de base de la formation Triasique. La calibration des données de diagraphie a été soigneusement réalisée pour compenser les différences de profondeur. Le tableau 3 représente les statistiques des ensembles de données d'entrée et de sortie utilisées dans la modélisation NF.

Descriptive Statistics (Hassi R'Mell Wells)				
	Mean	**Minimum**	**Maximum**	**Std.Dev.**
Cperm	467.2778	0.12000	1473.200	330.1228
Corpor	17.6821	2.59000	25.930	5.5492
GR	49.9431	25.12500	106.063	17.5520
ΔT	80.0319	64.00000	91.188	8.2897
Rhob	2.3685	2.20300	2.570	0.1063
RLLD	21.9661	3.57600	66.426	15.8215
Nphi	0.1633	0.03100	0.258	0.0606
Sw	43.3738	15.14000	90.610	17.3844

Tab.3. Descriptives Statistiques (Puits de Hassi R'Mel)

IV.7.2 Normalisation des données. Lors du traitement des données réelles, en raison des différentes dimensions du paramètre d'évaluation des roches réservoirs, le niveau de données réelles de volume varie considérablement. Si l'on calcule en utilisant directement les données brutes, les données ayant un volume plus important seraient de plus en plus remarquables, alors que l'indicateur avec un volume plus faible et une plus grande sensibilité, sera sous-estimée. Ainsi, nous devrions pré-traiter et normaliser les données brutes. Dans ce travail, la normalisation des données a lieu en utilisant les valeurs maximales et minimales des données.

IV.7.3 Fuzzy clustering: Durant la classification floue (Fuzzy classification), il est nécessaire de classer les ensembles de données d'entrée et de sortie en groupes en utilisant des méthodes de classes (clustering). Dans cette étude, une méthode de classification soustractive, qui est un moyen utile et efficace pour la modélisation, est utilisée pour l'extraction des groupes et des règles floues "If-then". Les détails de regroupement soustractif pourraient être trouvés dans Chiu (1994), Chen et Wang (1999), et Jarrah et Halawani (2001). Le paramètre important dans le regroupement soustractif qui contrôle le nombre de groupes et les règles floues "If-Then" est le rayon des clusters (Radius). Ce paramètre peut prendre des valeurs de

l’intervalle [0,1]. Généralement, le plus petit rayon (par exemple 0.1) est choisi avec les règles du fuzzy inférence système. En revanche, un grand rayon des groupes (par exemple 0.9) est donné quelquefois dans les grands groupes.

L'efficacité d'un modèle flou est liée à la recherche d'un rayon de groupement optimal (radii), qui est un paramètre de contrôle pour déterminer le nombre de règles flous "If-Then". Certaines règles ne peuvent couvrir la totalité des domaines d’études, et compliquent le comportement du système ce qui peut conduire à une faible performance du modèle. En ce qui concerne le modèle de perméabilité, cinq centres résultent de regroupement où le modèle flou a été créé par cinq règles floues "If-Then" et les quatre fonctions d'appartenance pour les données d'entrée et de sortie. Le modèle de porosité contient cinq centres (clusters) caractérisés par cinq règles et cinq fonctions d'appartenance. Les figures 28 et 29 montrent les groupes de soustraction de données de perméabilité et de porosité.

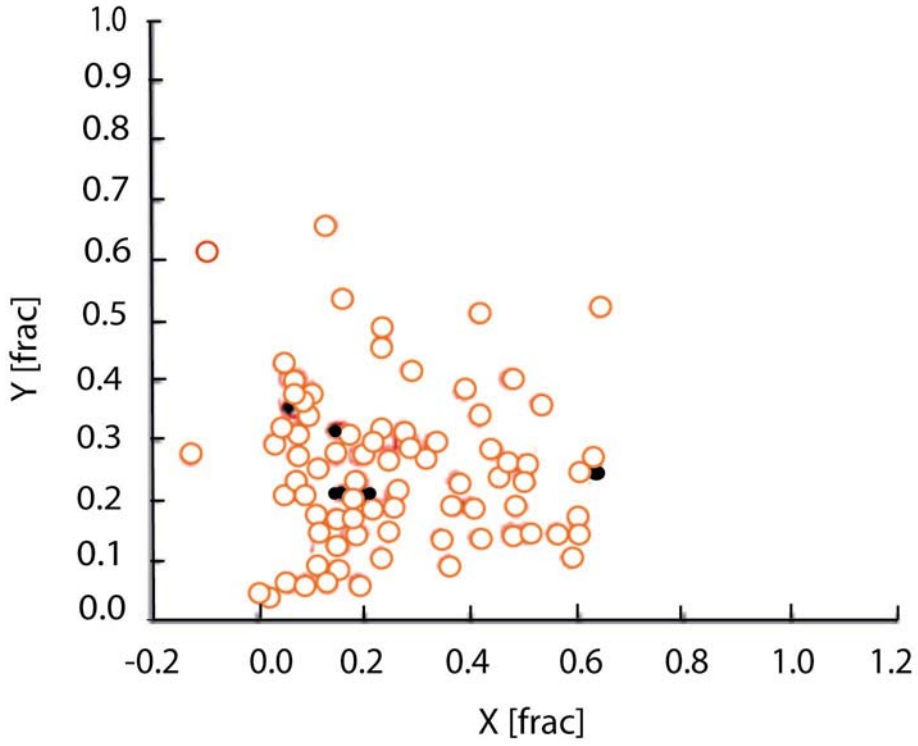


Fig.28. Regroupement soustractive de modèle flou de perméabilité. X : index ; Y : Perméabilité normalisée.

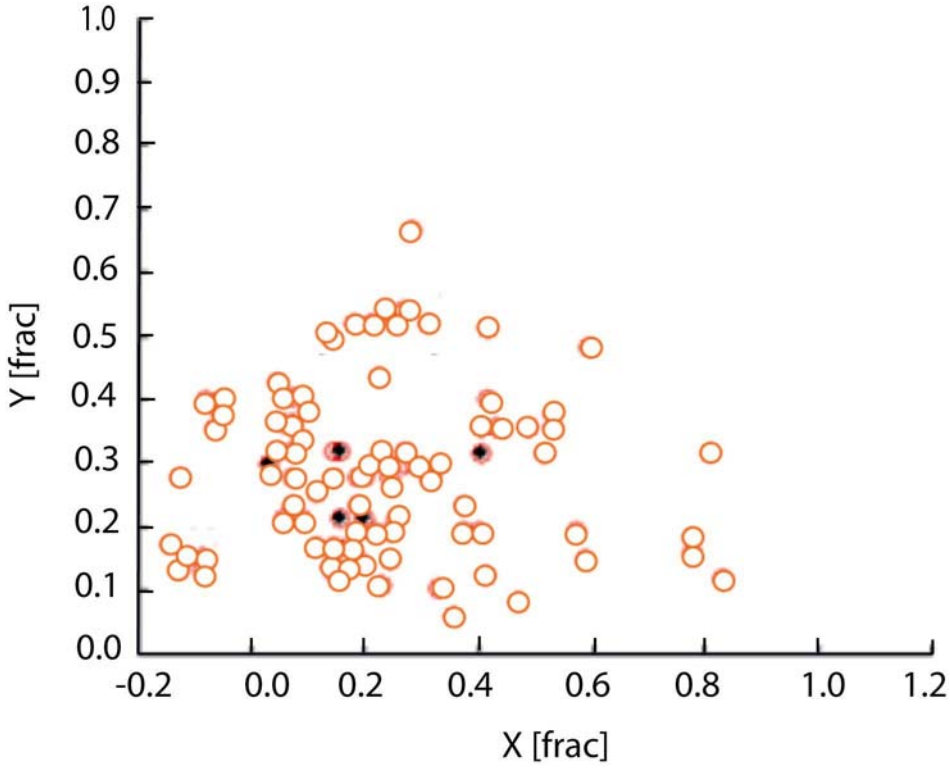


Fig.29. Regroupement soustractive de modèle flou de porosité. X : index ; Y : Porosité normalisée.

IV. 7.4 Construction du Système Fuzzy Inférence (FIS)

L'inférence Fuzzy est le processus de formulation d'une cartographie à partir d'une entrée donnée à une sortie en utilisant la logique floue (Kaufmann et al., 1985). La cartographie fournit alors une base à partir de laquelle les décisions peuvent être prises, ou des modèles qu'on en juge. Le processus d'inférence Fuzzy consiste à mettre les fonctions d'appartenance et l'établissement de règles floues (Matlab User's Guide, 2012a).

1. **Définition des fonctions d'appartenance (MF)**. Une fonction d'appartenance (MF) est une courbe qui définit la façon dont chaque point de l'espace d'entrée est cartographiée avec une valeur d'appartenance (ou degré d'appartenance) entre 0 et 1. L'espace d'entrée est parfois appelé "univers". La seule condition à laquelle une fonction d'appartenance doit répondre c'est qu'elle doit varier entre 0 et 1. La fonction elle-même peut être une courbe arbitraire dont la forme peut être définie comme une fonction qui nous convient du point de vue de la simplicité, commodité, vitesse et efficacité. Il existe de nombreux types de fonctions d'appartenance construites à partir de plusieurs fonctions de base :

- fonctions linéaires " *Piece-wise* "
- fonction de distribution *Gaussienne*
- courbe *sigmoid*
- courbe *quadratique et* courbe *polynomiale*

Dans cette étude, une fonction de distribution de Gauss est utilisée pour définir les pôles d'entrées extraites. Une fonction gaussienne f (x) représente la distribution normale des données (x) :

$$f(x)=\frac{e^{-(x-\mu)^2/\sigma^2}}{\sigma\sqrt{2\pi}}$$

où μ et σ sont des paramètres de la distribution normale montrant respectivement la moyenne et l'écart-type des données. Ces fonctions d'appartenance gaussiennes sont construites à partir des valeurs moyennes et des clusters. La moyenne représente les centres de classes et σ est la dérivée:

σ = (rayon (maximum de données - données minimum))/sqrt (2π). Les paramètres d'entrée de la fonction d'appartenance gaussienne pour la perméabilité et la porosité sont présentés dans les tableaux 4a et 5a.

(a)

Inputs	GR (API)		ΔT (μs/ft)		Rhob (g/cc)		Rlld (Ωm)		Nphi (%)		Sw (%)	
Paramètres	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ
Input MF No	0.19	0.15	0.18	0.21	0.17	0.83	0.18	0.03	0.17	0.11	0.18	0.19
MF1	0.19	0.16	0.18	0.25	0.17	0.79	0.18	0.06	0.17	0.12	0.18	0.25
MF2	0.19	0.62	0.18	0.32	0.17	0.87	0.18	0.07	0.17	0.11	0.18	0.26
MF3	0.19	0.24	0.18	0.23	0.17	0.42	0.18	0.37	0.17	0.13	0.18	0.22
MF4	0.19	0.34	0.18	0.27	0.17	0.58	0.18	0.42	0.17	0.11	0.18	0.22
MF5	0.19	0.27	0.18	0.29	0.17	0.47	0.18	0.33	0.17	0.15	0.18	0.21

(b)

Output	Corpor (%)						
	C1	C2	C3	C4	C5	C6	C7
Output MF No	0.0	0.0	0.0	0.0	0.0	0.0	-0.27
MF1	-0.42	-0.45	-4.20	18.52	0.74	2.55	0.82
MF2	0.31	0.48	-3.50	-0.14	0.14	2.58	-0.35
MF3	0.19	0.58	-0.47	0.37	0.17	-0.87	0.58
MF4	0.22	0.62	0.29	0.44	0.57	-0.57	0.41
MF5	0.35	0.87	0.42	0.48	0.98	-0.41	0.47

Tab. 4. Entrée (a) et sortie (b) des paramètres de la fonction des MF obtenus par TS-FIS pour la porosité.

(a)												
Inputs	GR (API)		ΔT (µs/ft)		Rhob (g/cc)		Rlld (Ωm)		Nphi (%)		Sw (%)	
Paramètres	σ	µ	σ	µ	σ	µ	σ	µ	σ	µ	σ	µ
Input MF No	0.18	0.17	0.17	0.15	0.15	0.51	0.18	0.14	0.19	0.11	0.21	0.14
MF1	0.18	0.14	0.17	0.11	0.15	0.42	0.18	0.20	0.19	0.14	0.18	0.20
MF2	0.18	0.20	0.17	0.20	0.15	0.22	0.18	0.22	0.19	0.48	0.31	0.22
MF3	0.18	0.22	0.17	0.21	0.15	0.44	0.18	0.29	0.19	0.55	0.42	0.29
MF4	0.18	0.29	0.17	0.27	0.15	0.45	0.18	0.38	0.19	0.60	0.55	0.38
MF5	0.18	0.38	0.17	0.33	0.15	0.69	0.18	0.17	0.19	0.75	0.60	0.17

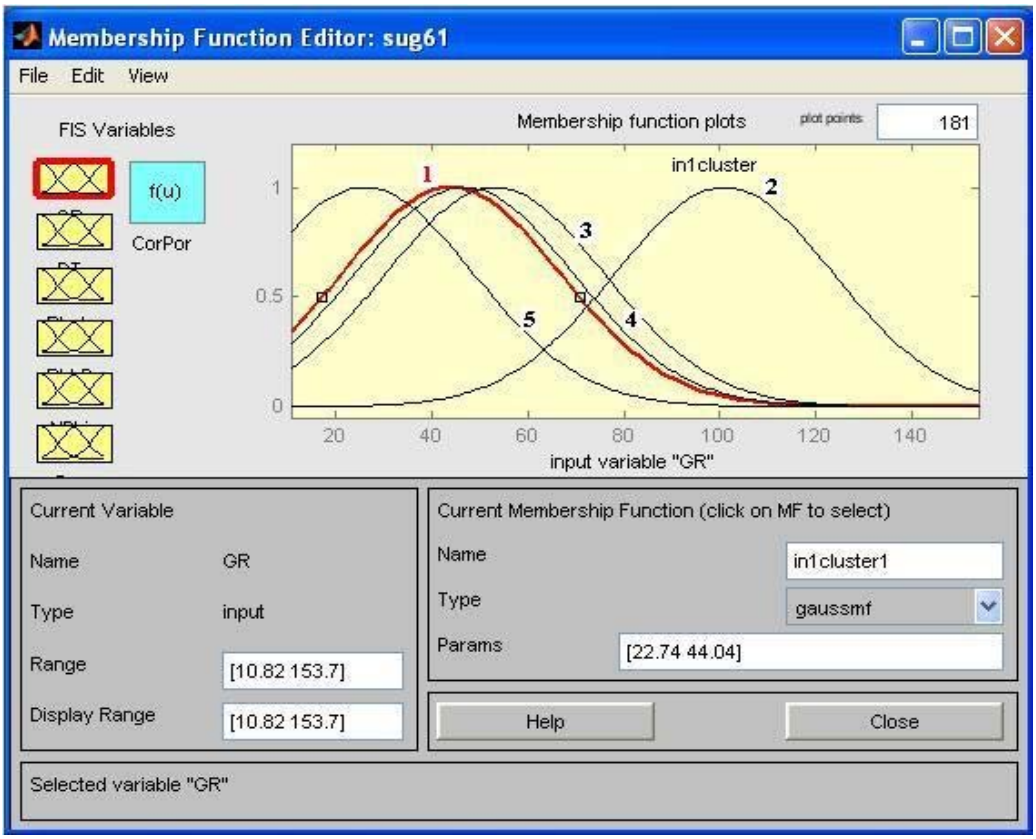
(b)							
Output	CPerm (mD)						
	C1	C2	C3	C4	C5	C6	C7
Output MF No							
MF1	-0.35	-0.55	-0.08	0.60	1.25	0.75	-0.55
MF2	-0.44	-0.19	-0.05	1.53	1.89	-0.15	0.27
MF3	0.15	0.11	0.18	0.22	0.27	0.27	0.33
MF4	0.20	0.22	0.41	0.32	0.41	0.38	0.42
MF5	0.30	0.33	0.44	0.45	0.48	0.45	0.49

Tab. 5. Entrée (a) et sortie (b) des paramètres de la fonction des MF obtenus par TS-FIS pour la perméabilité.

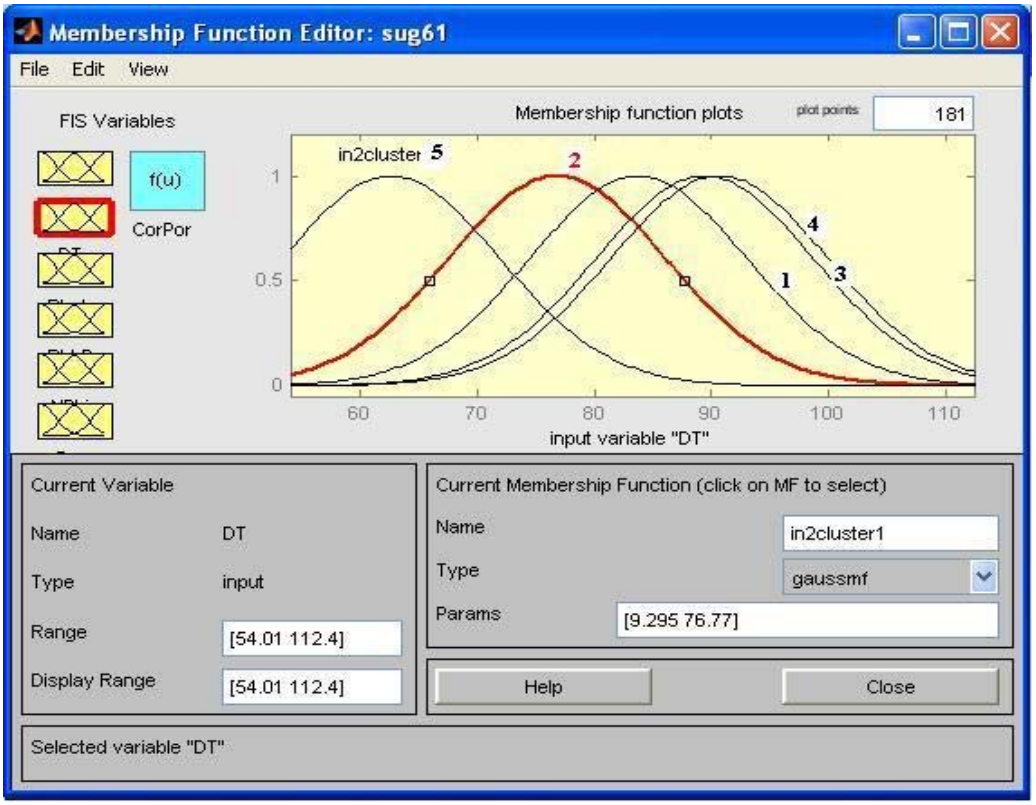
L'Inférence Fuzzy (FIS) et les fonctions d'appartenance de sortie sont des équations linéaires qu'il a construites à partir des paramètres d'entrée. Par exemple, l'adhésion de sortie de la fonction nommée (MF1), qui est la conséquence de la règle no.1, est construite à partir de six entrées pétrophysiques de la manière suivante :

$$\text{Output MF1} = C_1 \text{ x GR} + C_2 \text{ x DT} + C_3 \text{ x Rhob} + C_4 \text{ x RD} + C_5 \text{ x Phin} + C_6 \text{ x SW} + C_7$$

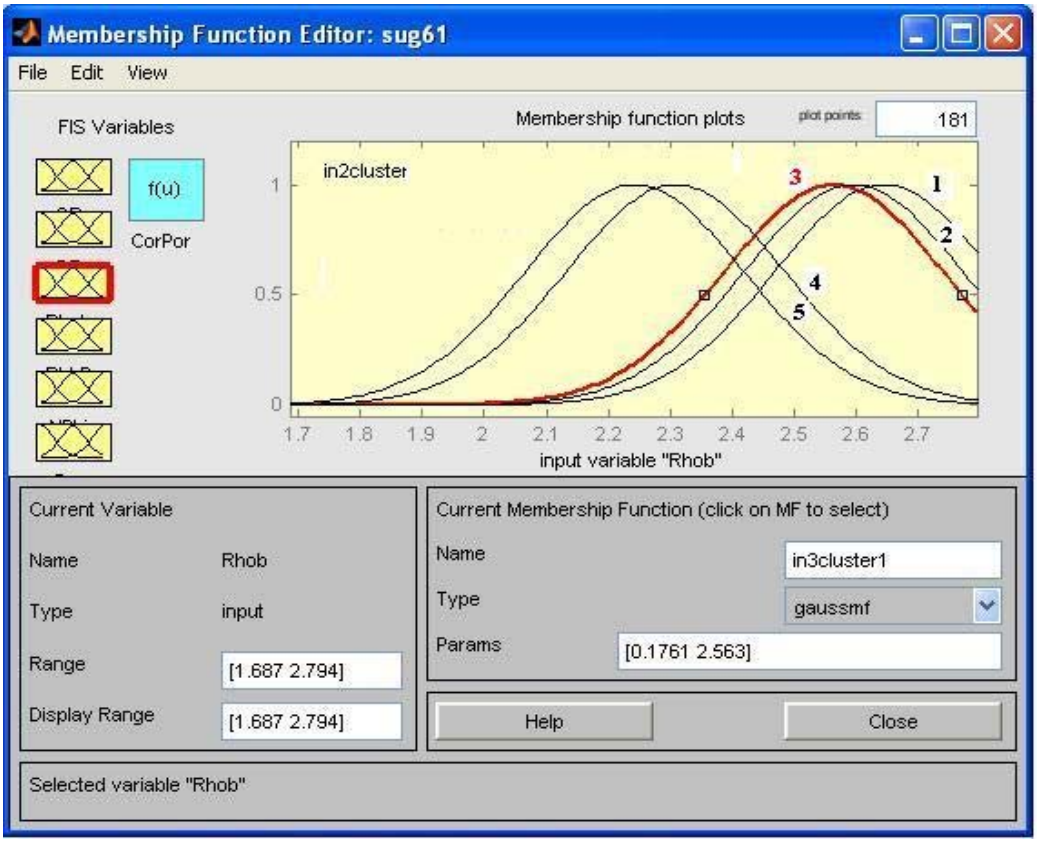
Dans cette équation, les paramètres C_1, C_2, C_3, C_4, C_5 et C_6 sont respectivement les coefficients correspondant à GR, ΔT, Rhob, RD, Phin et l'entrée SW. Ces paramètres sont obtenus par estimation linéaire des moindres carrés. Avec ces explications, il y aura sept paramètres pour chaque fonction d'appartenance de sortie, qui sont représentés dans les tableaux 4b et 5b respectivement pour la perméabilité et de la porosité. Les figures 30 et 31 représentent le FIS qui a généré les fonctions d'appartenance gaussiennes des données d'entrée respectivement pour la porosité et de la perméabilité modèle.



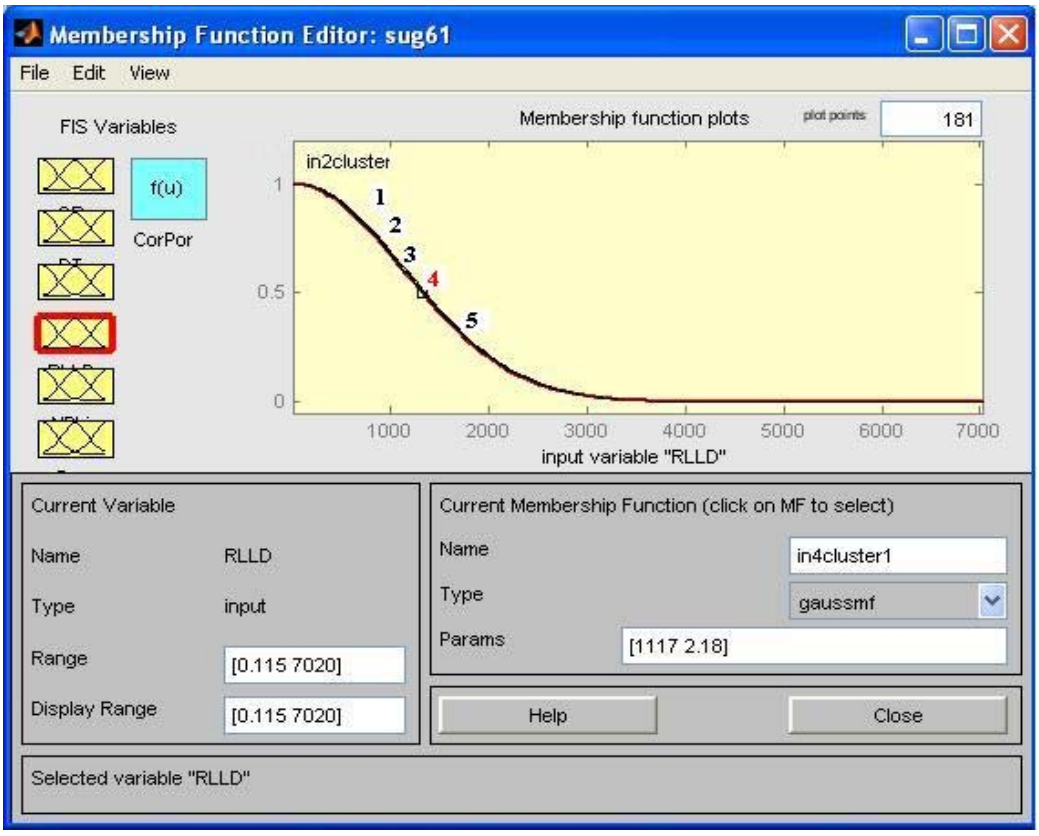
a) "Membership function" pour la variable : Radioactivité GR ; sortie: Porosité CorPor



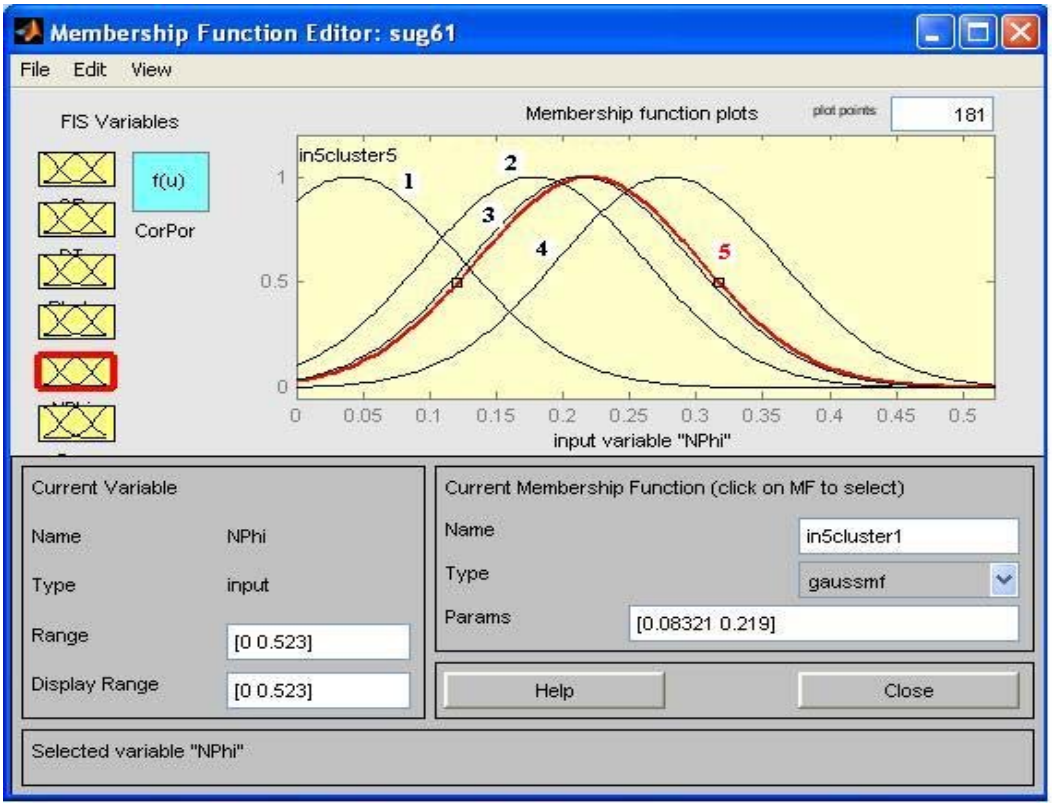
b) "Membership function" pour la variable : temps de transit ΔT ; sortie: Porosité CorPor



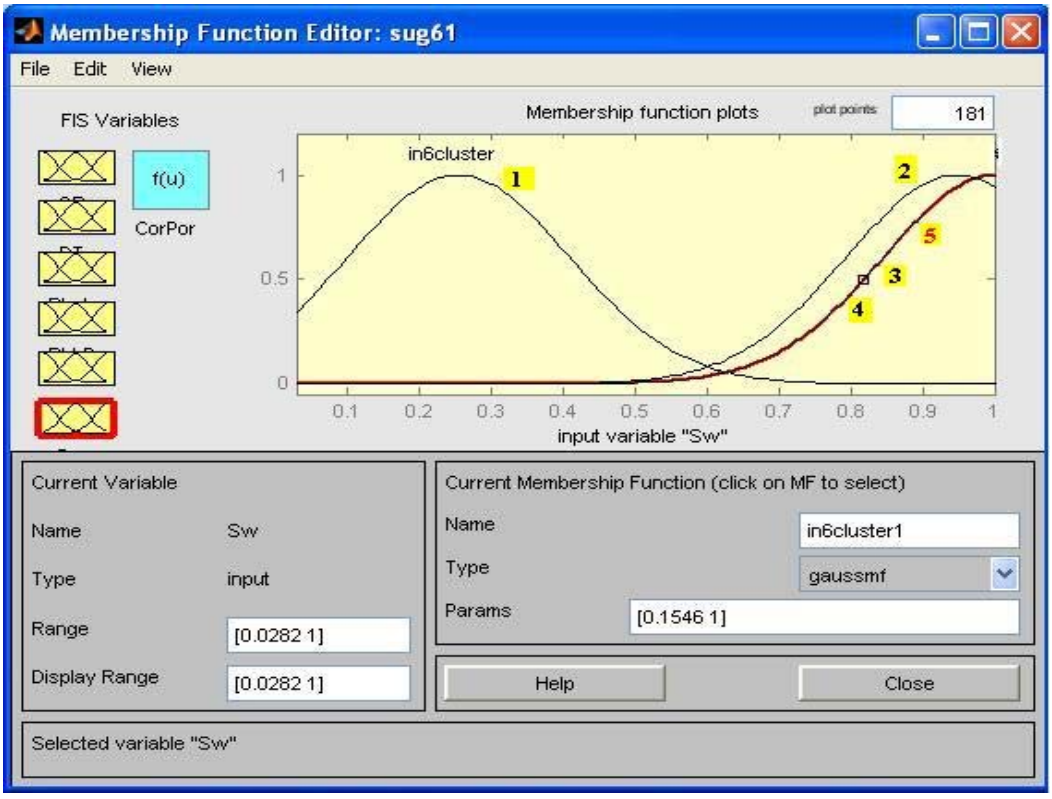
c) "Membership function" pour la variable : densité Rhob ; sortie: Porosité CorPor



d) "Membership function" pour la variable : résistivité RLLD ; sortie: Porosité CorPor

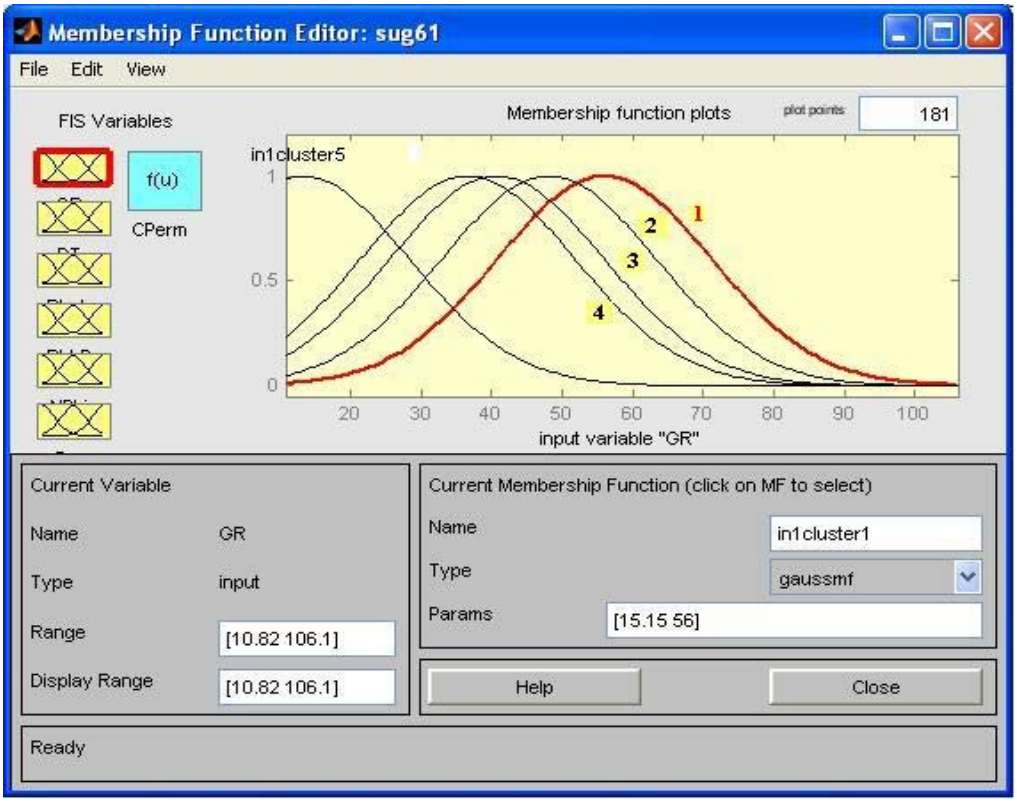


e) "Membership function" pour la variable: porosité neutron NPhi; sortie: Porosité CorPor

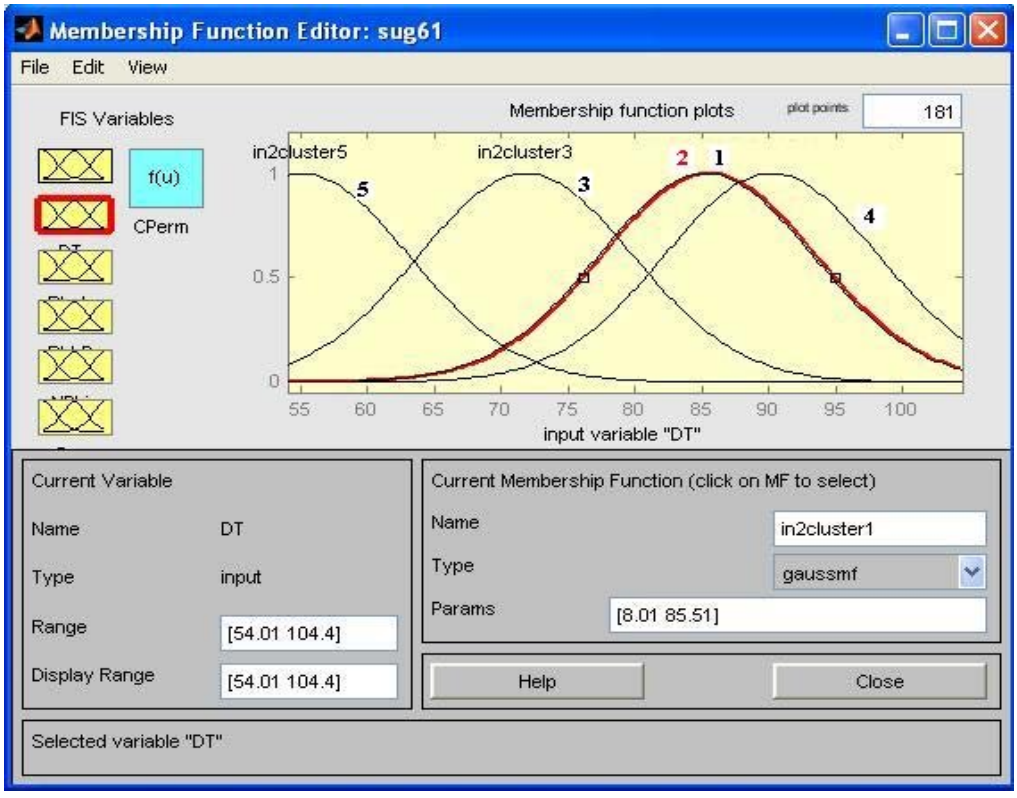


f) "Membership function" pour la variable: Saturation Sw ; sortie : Porosité CorPor

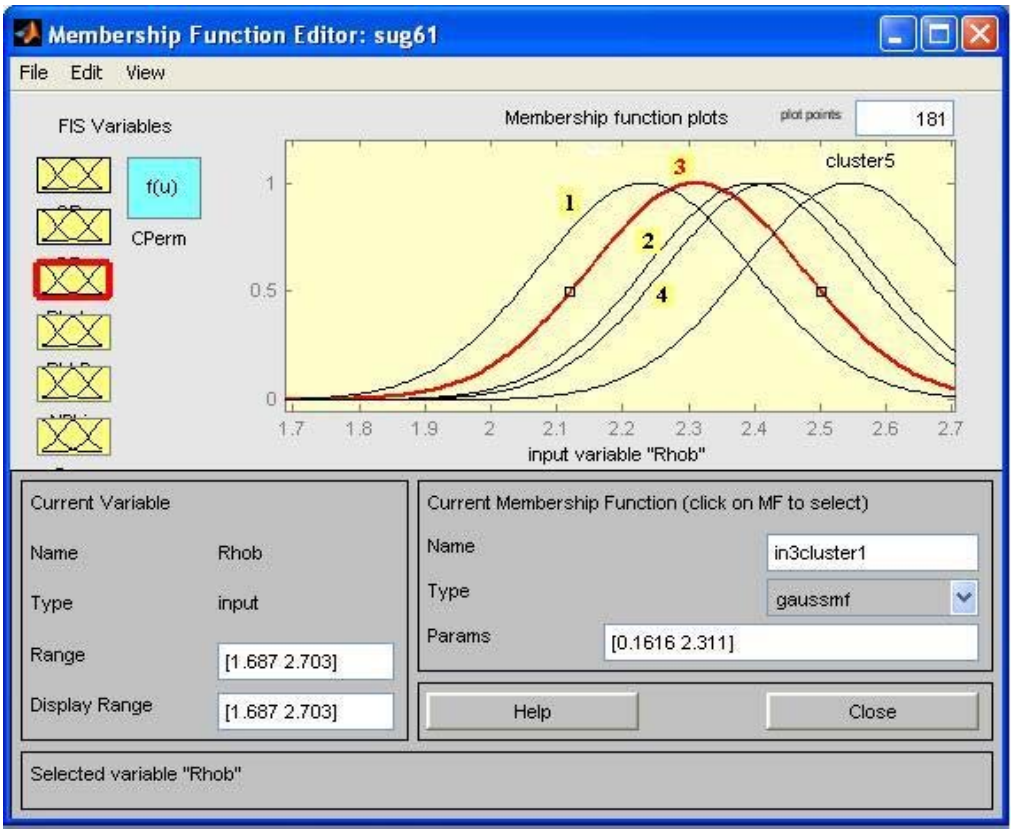
Fig.30 (a-f). Fonctions FIS "Membership Gaussian", générées par le Système Fuzzy Inference pour le modèle de données d'entrée pour la porosité



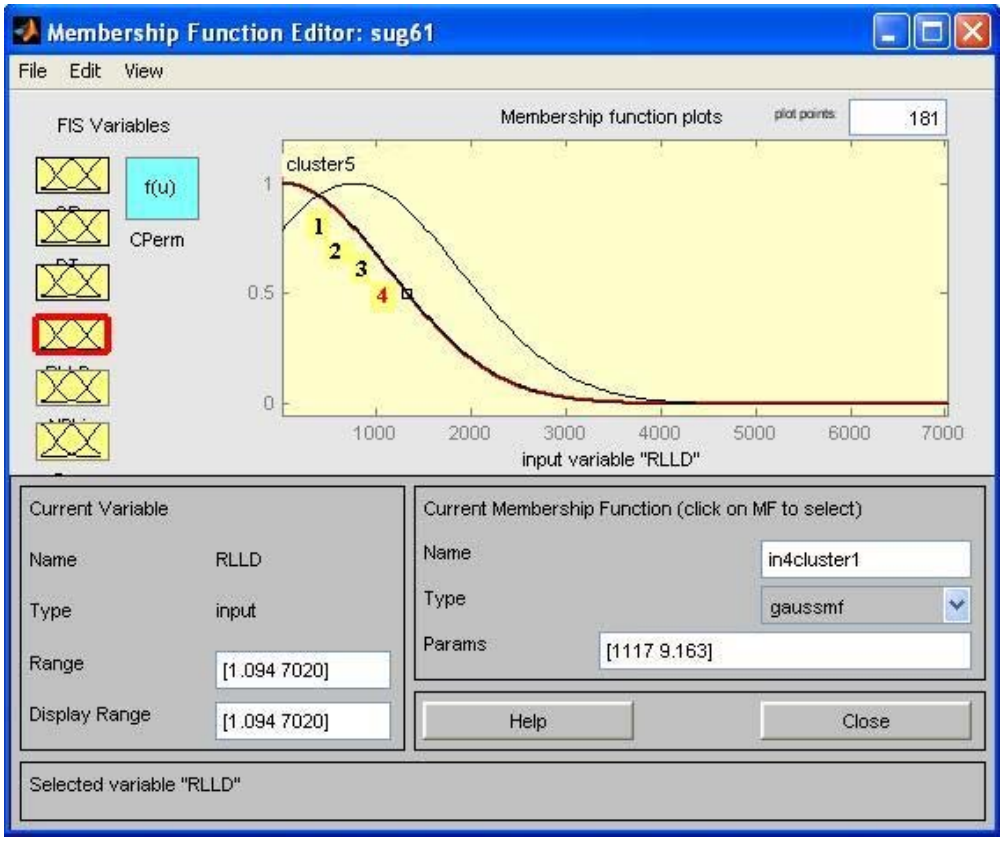
a) "Membership function" pour la variable GR ; sortie : Perméabilité : CPerm



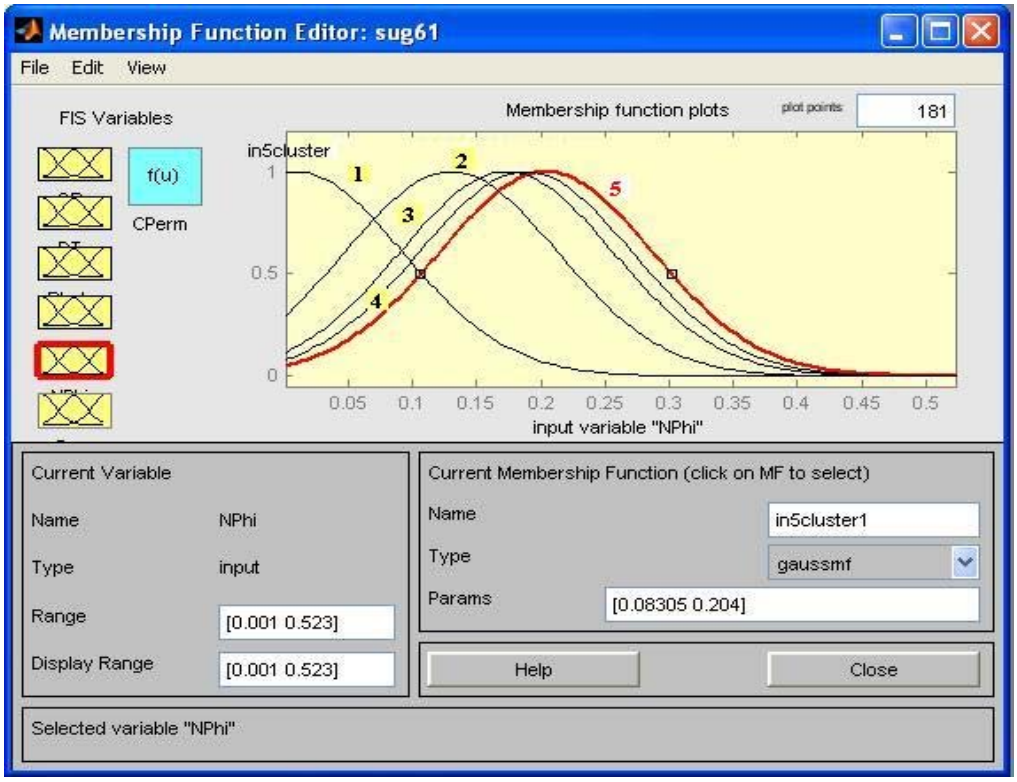
b) "Membership function" pour la variable ΔT ; sortie : Perméabilité CPerm



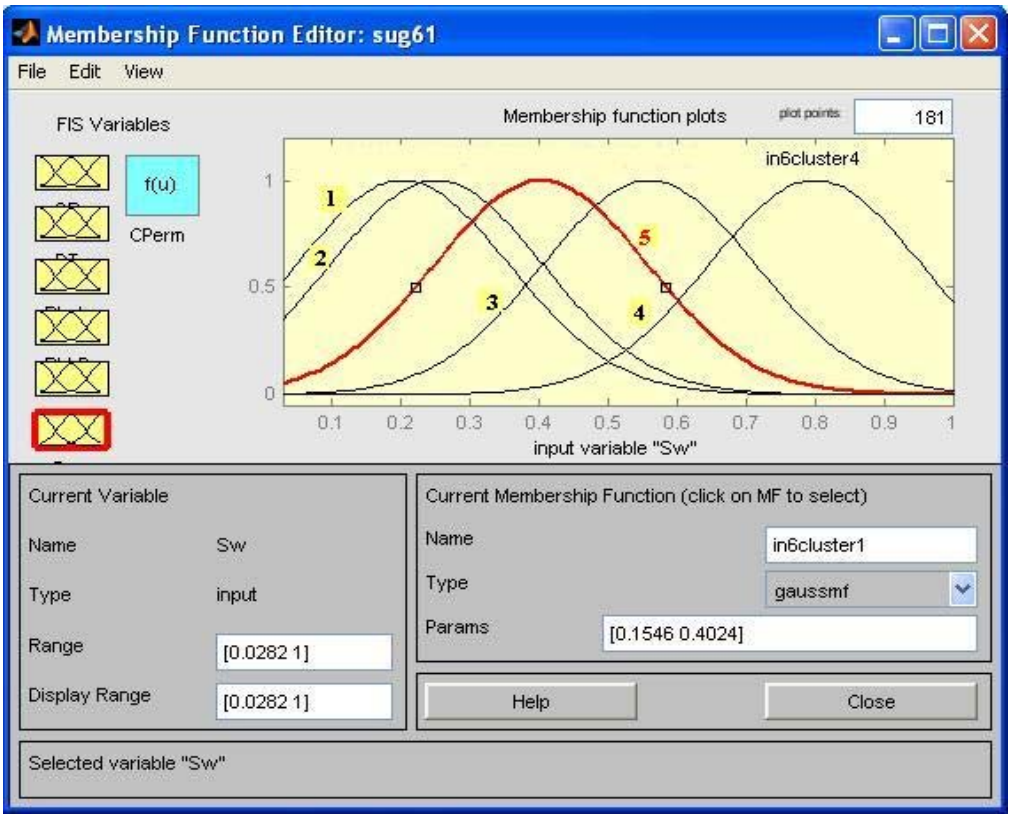
c) "Membership function" pour la variable Rhob ; sortie : Perméabilité CPerm



d) "Membership function" pour la variable RLLD ; sortie : Perméabilité CPerm



e) "Membership function" pour la variable NPhi ; sortie : Perméabilité CPerm



f) "Membership function" pour la variable Sw ; sortie : Perméabilité CPerm

Fig.31 (a-f). Fonctions FIS "Membership Gaussian", générées par le Système Fuzzy Inference pour le modèle de données d'entrée pour la perméabilité.

2. **Mise en place de règles floues.** Les déclarations de règles floues sont utilisées pour formuler les instructions conditionnelles qui composent la logique floue. Une seule règle floue "If-Then" prend la forme si x est A, y est alors B où A et B sont des valeurs linguistiques définies par des ensembles flous respectivement sur les plages (univers du discours) X et Y. Le " if-partie " de la règle "x est A" est appelé l'antécédent ou prémisse, alors que « Alors-partie » de la règle "y est B" est appelé conséquent ou conclusion.

Les Fuzzy générés " If-Then " : règles de formulation des données d'entrée pétrophysiques de perméabilité (CPerm ou K) sont (Fig.32):

1. If (GR is in1mf1) and (ΔT is in2mf1) and (Rhob is in3mf1) and (RD is in4mf1) and (Nphi is in5mf1) and (SW is in6mf1) then (K is out1mf1).
2. If (GR is in1mf2) and (ΔT is in2mf2) and (Rhob is in3mf2) and (RD is in4mf2) and (Nphi is in5mf2) and (SW is in6mf2) then (K is out1mf2).
3. If (GR is in1mf3) and (ΔT is in2mf3) and (Rhob is in3mf3) and (RD is in4mf3) and (Nphi is in5mf3) and (SW is in6mf3) then (K is out1mf3).
4. If (GR is in1mf4) and (ΔT is in2mf4) and (Rhob is in3mf4) and (RD is in4mf4) and (Nphi is in5mf4) and (SW is in6mf4) then (K is out1mf4).
5. If (GR is in1mf5) and (ΔT is in2mf5) and (Rhob is in3mf5) and (RT is in4mf5) and (Nphi is in5mf5) and (SW is in6mf5) then (CPerm is out1mf5).

Les règles fuzzy " If-Then " de formulation des données d'entrée pétrophysiques de porosité (CorPor) sont (Fig.33):

1. If (GR is in1mf1) and (ΔT is in2mf1) and (Rhob is in3mf1) and (RD is in4mf1) and (Nphi is in5mf1) and (SW is in6mf1) then (PHI is out1mf1).
2. If (GR is in1mf2) and (ΔT is in2mf2) and (Rhob is in3mf2) and (RD is in4mf2) and (Nphi is in5mf2) and (SW is in6mf2) then (PHI is out1mf2).
3. If (GR is in1mf3) and (ΔT is in2mf3) and (Rhob is in3mf3) and (RD is in4mf3) and (Nphi is in5mf3) and (SW is in6mf3) then (PHI is out1mf3).
4. If (GR is in1mf4) and (ΔT is in2mf4) and (Rhob is in3mf4) and (RD is in4mf4) and (Nphi is in5mf4) and (SW is in6mf4) then (PHI is out1mf4).

5. If (GR is in1mf5) and (ΔT is in2mf5) and (Rhob is in3mf5) and (RD is in4mf5) and (Nphi is in5mf5) and (SW is in6mf5) then (CorPor is out1mf5).

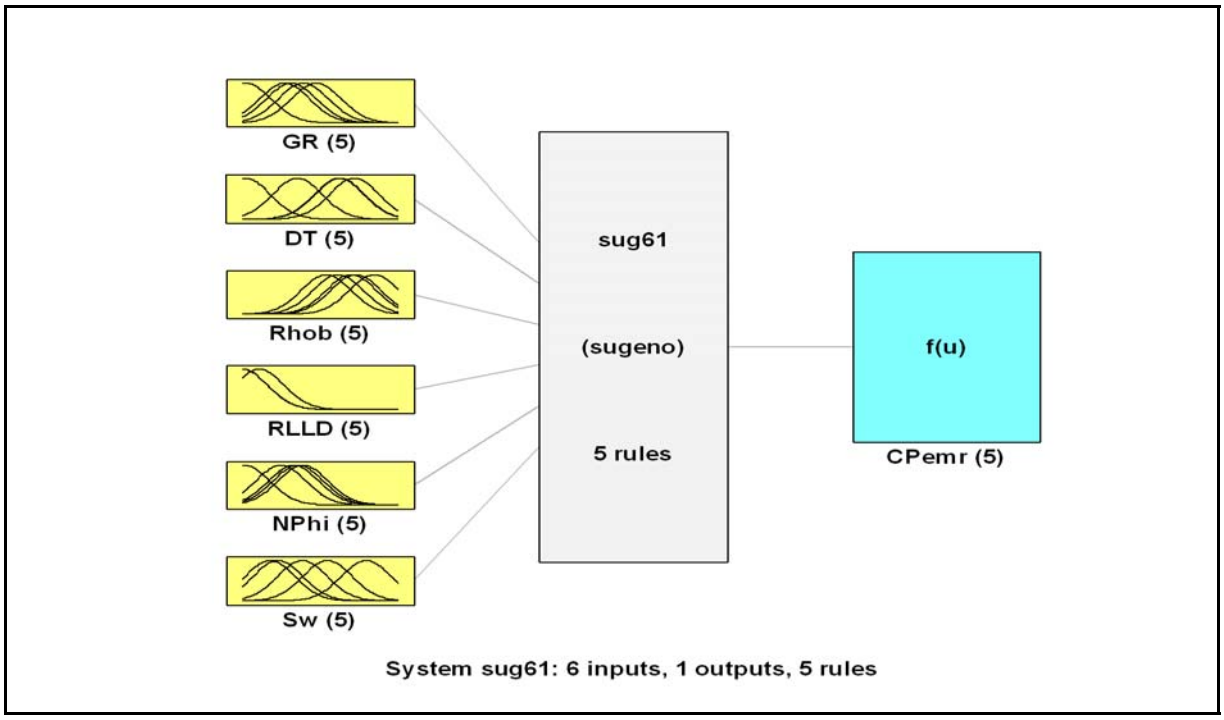


Fig.32. Système d'inférence flou de tous les puits de Hassi R'Mel pour la perméabilité.

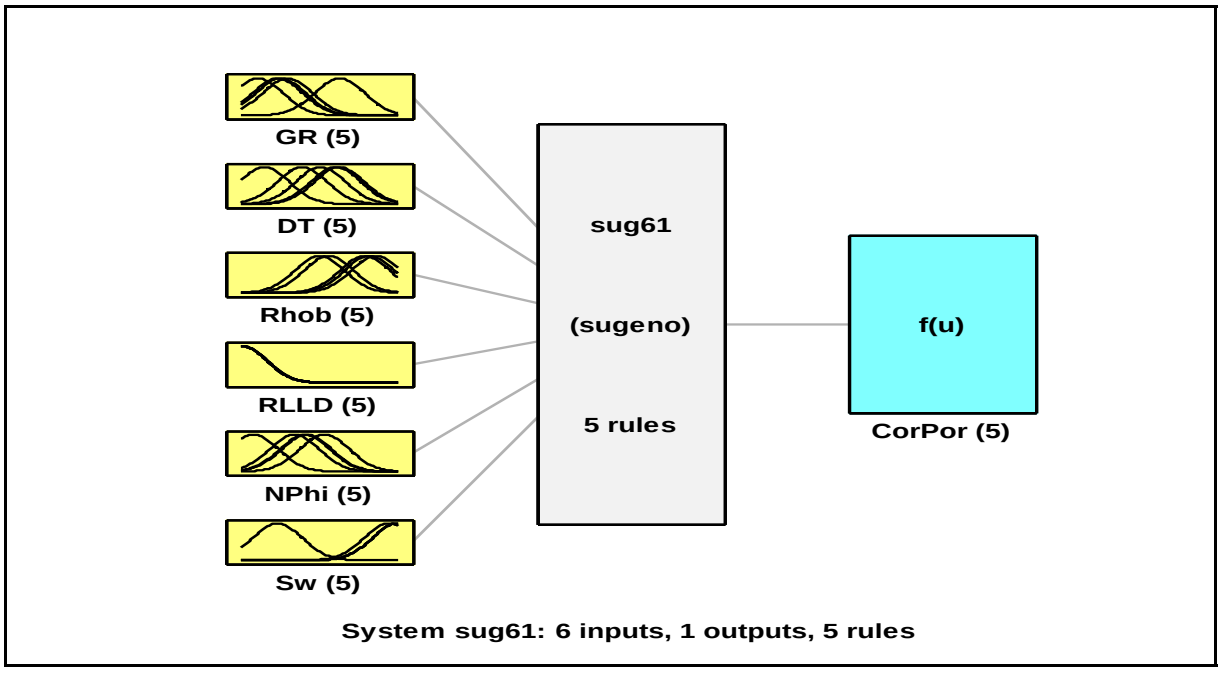


Fig.33. Système d'inférence flou de tous les puits de Hassi R'Mel pour la porosité.

Une illustration graphique montrant les étapes de formulation d'entrée des données pétrophysiques de la perméabilité en utilisant quatre règles " If-Then " flous générés par la FIS, est représenté en figures 34 et 35. Chaque figure affiche une feuille de route de

l'ensemble du processus d'inférence floue. Les sept parcelles dans la partie supérieure de la figure représentent l'antécédent et le conséquent de la première règle. Chaque règle est une rangée de parcelles, et chaque colonne est une variable. Les numéros de règles sont affichés à gauche de chaque ligne.

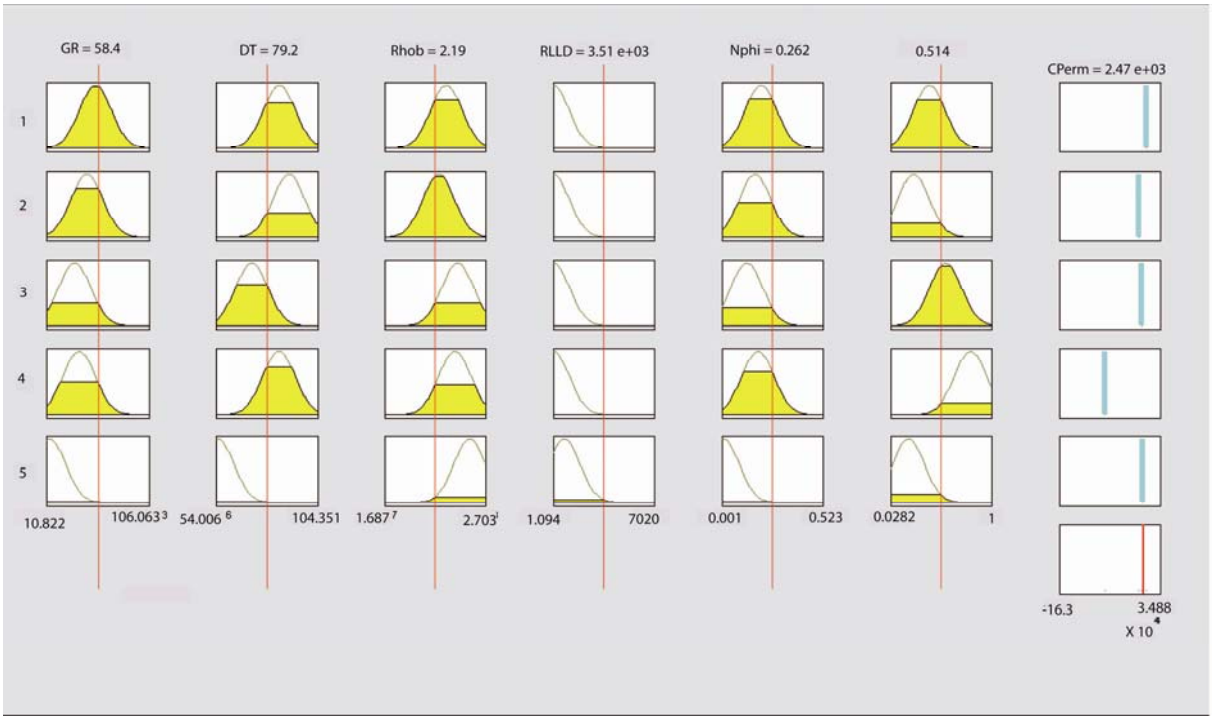


Fig.34. Fonctions d'appartenance (Membership functions) pour la perméabilité (CPerm)

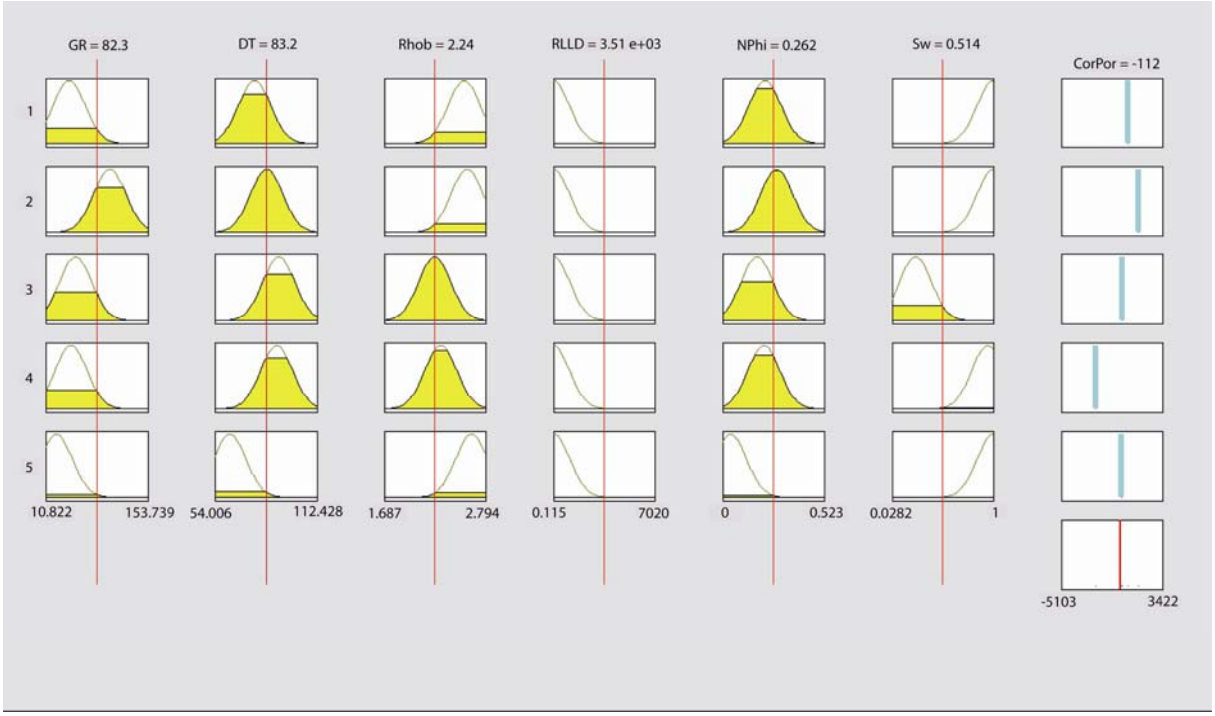


Fig.35. Fonctions d'appartenance (Membership functions) pour la porosité (CorPor)

La structure du modèle NF est maintenant générée pour la perméabilité (Fig.36) et la porosité (Fig.37). L'entrée est représentée par le nœud le plus à gauche et la sortie par le nœud le plus à droite. Le nœud représente un facteur de normalisation pour les règles.

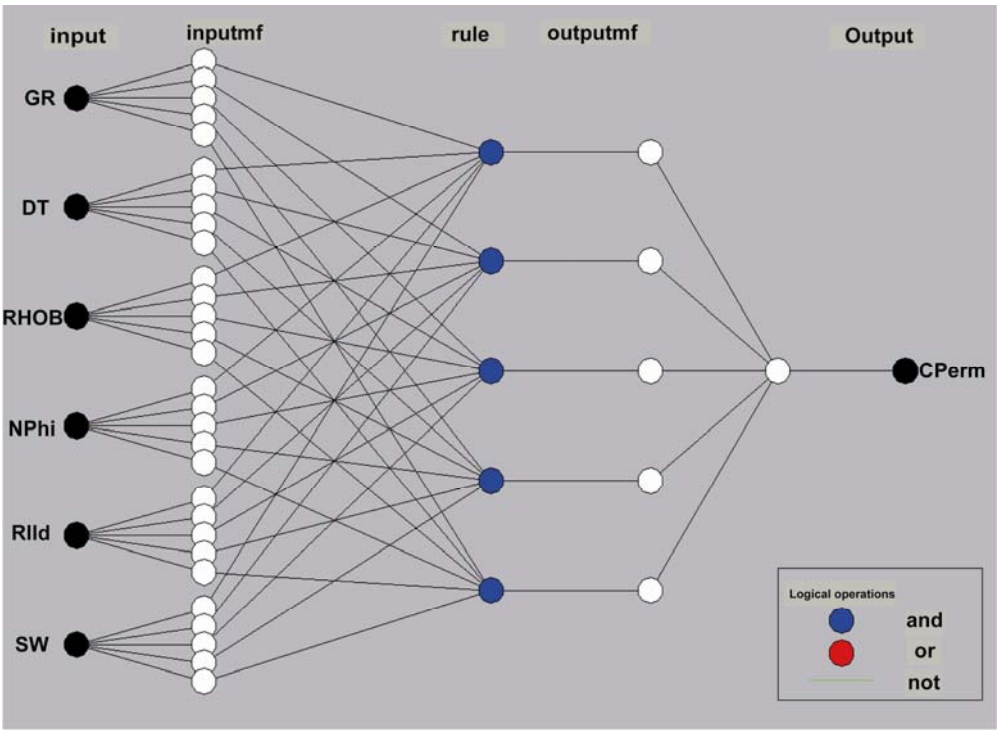


Fig.36. Structure du modèle NF pour la perméabilité (CPerm)

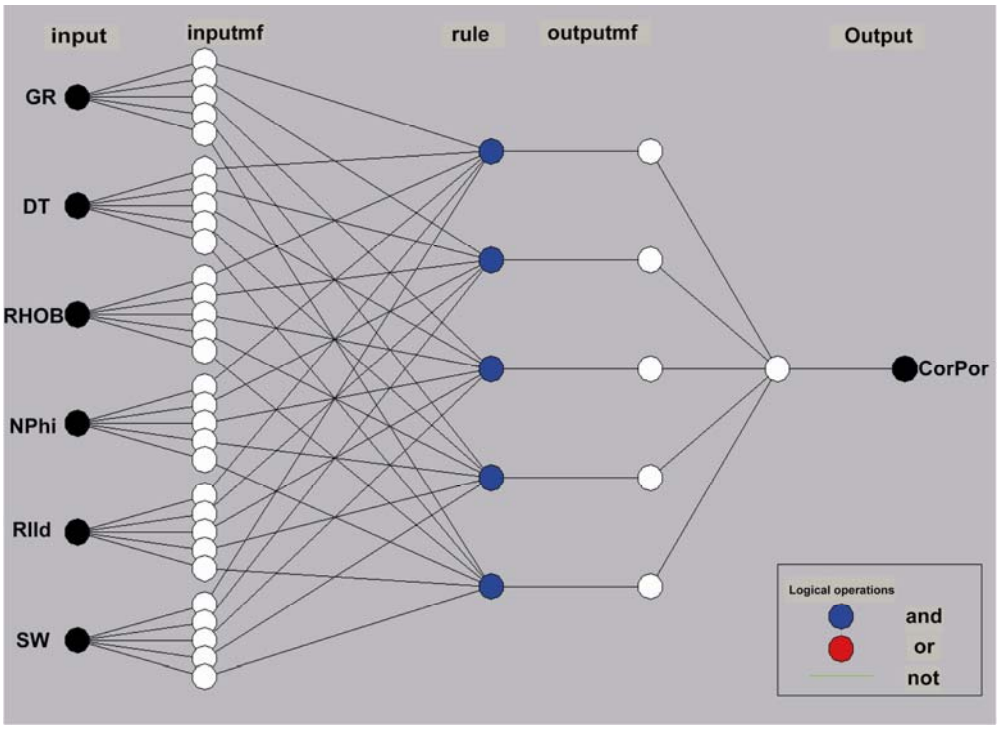


Fig.37. Structure du modèle NF pour la porosité (CorPor)

IV. 7. Résultats et discussion

Les résultats obtenus par les méthodes décrites précédemment sont expliqués individuellement, mais la caractérisation des réservoirs par les diagraphies nécessite la combinaison et l'application de plusieurs méthodes. Les résultats obtenus ne sont bons que pour une meilleure explication des résultats finaux ayant abouti aux objectifs souhaités. Cette étude a nécessité principalement deux approches : l'analyse des données des modules de base et l'interprétation des données de diagraphie.

Dix puits ont été utilisés dans le champ de Hassi R'Mel pour cette étude, et chaque puits a été étudié individuellement par les méthodes habituelles d'interprétation. Seuls 4 puits sont traités dans cet article pour raison de disponibilité des carottes : HR-167, HR-199, HR-205 et HR-209 pour l'intérêt d'huile dans le pourtour du réservoir. Les estimations de la lithologie sont basées sur les logs de porosité. Comme mentionné précédemment, les calculs de volume d'argiles nécessaires pour l'évaluation qualitative et quantitative des réservoirs ont été faits sur la base du diagramme de densité en fonction du neutron (Fig. 38) pour les puits carottés. Les résultats ont montré que les lithologies peuvent être décrites comme des sables argilo-gréseux. La partie à haute teneur en argile de chaque puits observé à la partie gréseuse (Fig. 38) est la limite pour la formation du Trias. Une augmentation typique de contenu GR est aussi suivie dans les premières sections de dolomies qui sont l'argile dolomitique pouvant agir comme joint.

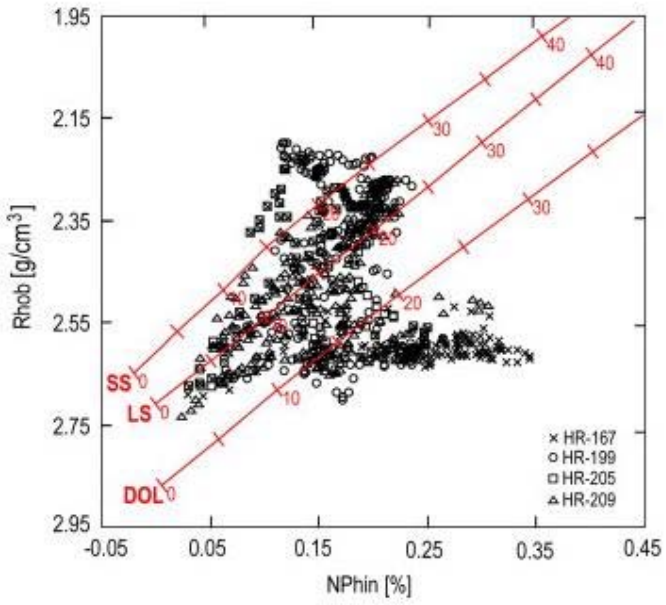


Fig. 38. Diagramme de densité en fonction du Neutron établi au niveau du réservoir de Hassi R'Mel.

Dans la partie juste en-dessous, les zones à faible teneur en argile et en dolomie montrent des porosités relativement élevées et des logs de résistivité indiquant la présence d'hydrocarbures. Ces zones poreuses et perméables sont également indiquées par la formation de "dépôt de boue".

Les calculs de la porosité et en particulier les saturations prouvent effectivement des sections à hydrocarbures du réservoir. Les dolomies peuvent être considérées comme argileuses. Comme on le voit dans la partie de la lithologie, certaines dolomies calcaireuses sont à prédominance dolomitique. Après correction de l'effet d'argile, la formation a été considérée comme propre. La porosité des formations a été déterminée par plusieurs méthodes: porosité sonique, porosité de densité et de porosité neutron. Dans tous les puits, les valeurs de la porosité les plus fiables sont considérées comme des porosités de densité neutrons. Pour les formations de porosité secondaire, la taille des pores, en raison du tri des grains et la dolomitisation peuvent affecter aussi l'écoulement du fluide dans la formation. Cette porosité est caractéristique de l'effet principal de la limitation de l'écoulement du fluide. Ces paramètres peuvent évidemment affecter la perméabilité en raison des faibles valeurs de porosité et de perméabilité déterminées à partir des données de base. La porosité maximale pour les dolomies est observée dans le puits HR-167 avec une porosité moyenne de 40%, et 10% pour le sable argileux. En général, les gammes de porosité pour les dolomies varient entre 10% et 40% (Fig. 38).

La qualité du réservoir est contrôlée par le stockage des hydrocarbures et la capacité d'écoulement. Ceux-ci permettent de définir des intervalles caractéristiques d'écoulement semblables et prévisibles, qui sont les unités de débit. Le stockage des hydrocarbures est fonction de la porosité et la capacité de débit est fonction de la perméabilité. Les unités de débit peuvent être identifiées à partir d'une série de cross plots pétrophysiques et du calcul des rayons de pores (R35, taille des pores) au volume poreux de 35% en utilisant les équations de débit (Aguilera et Aguilera, 2001).

La logique floue est utilisée pour prédire les courbes de diagraphies manquantes (ou faciès) et de procéder à une étude multi-puits. Elle utilise des courbes de diagraphies en continu ou discrétisées et les informations de faciès de base ou d'autres sources dans un puits.

Les courbes sont sélectionnées dans un jeu complet de données diagraphiques (Fig. 39) qui doit être estimée dans d'autres puits (Figs. 40-42). L'ensemble de la formation est traité en

utilisant la logique floue pour déterminer la valeur la plus probable pour diverses combinaisons de valeurs de log qui représentent les données de diagraphies manquantes où tous les logs contiennent sept pistes. De gauche à droite, nous pouvons voir sur ces mêmes figures: la formations en profondeur, le gamma naturel GR [API], la profondeur [m] à l'échelle 1/200, la résistivité vraie RT [Ω.m], la densité globale Rhob [g /cm^3], porosité neutron [%], sonique [µs/ft], porosité carotte CorPor [%], la prévision de la porosité carottes CorPor_nn [%], la perméabilité carottes CPerm [mD] et la prédiction de la perméabilité carotte Cperm_nn [mD].

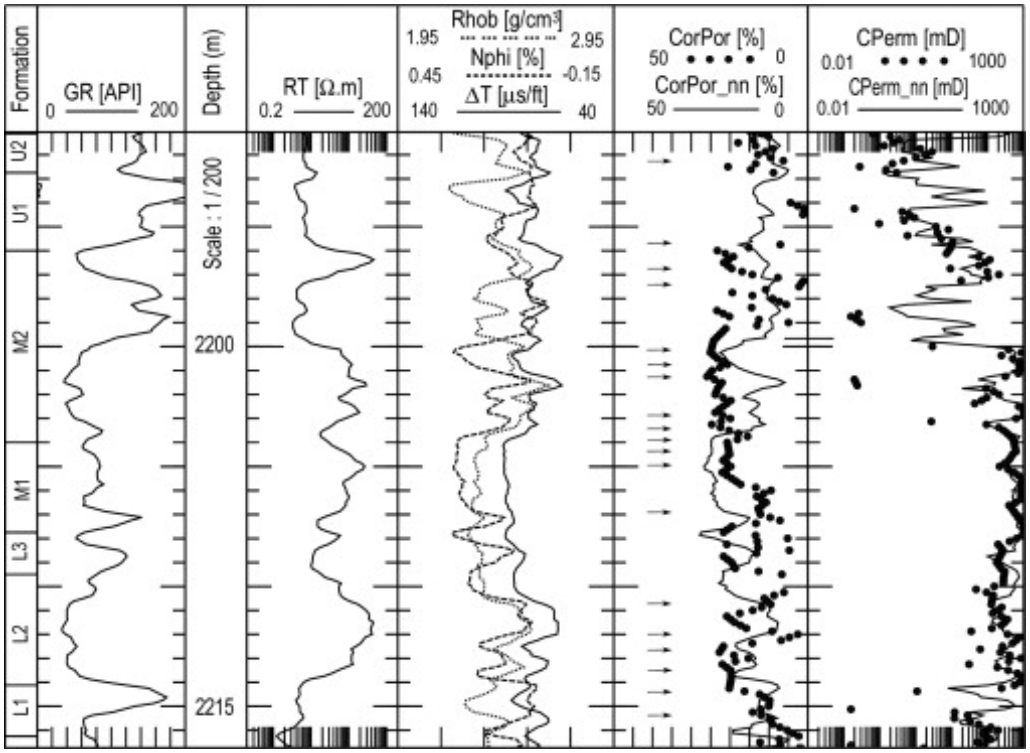


Fig. 39. Log composites ainsi que les prévisions de diagraphies pour le puits HR-167. U1, U2, M1, M2, L1, L2 et L3: formations in situ; GR [API]: rayon gamma; Rhob [g / cm^3]: densité; NPhi [%]: porosité neutron; Δt [µs/ft]: sonique; RT [Ω.m]: résistivité vraie, CPerm [MD]: perméabilité carotte; CPerm_nn [MD] : valeurs prédites du réseau de neurones. Les flèches horizontales représentent les niveaux des échantillons utilisés dans la formation.

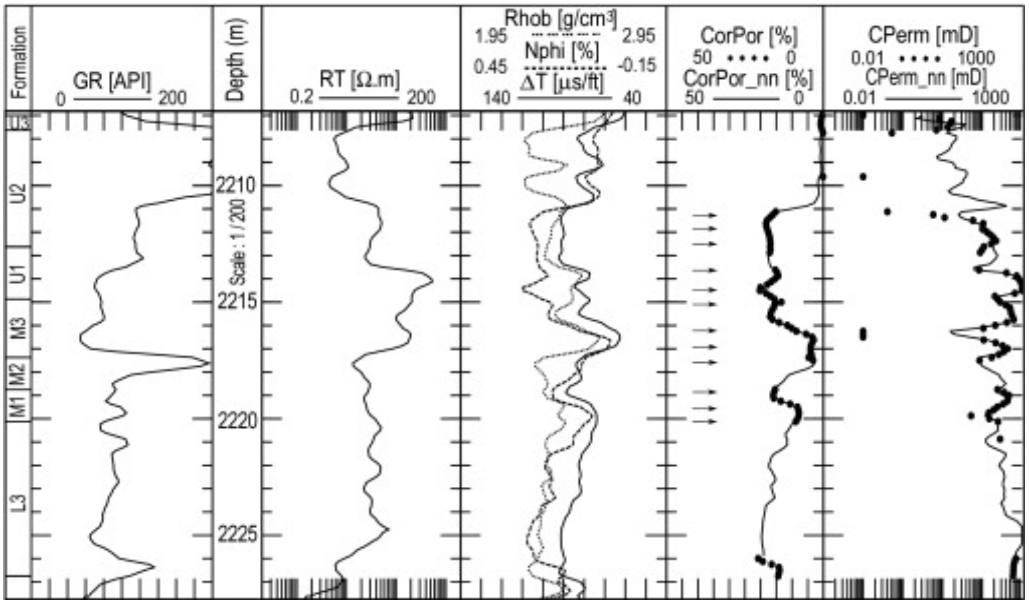


Fig. 40. Log composites ainsi que les prévisions de diagraphies pour le puits HR-205. Même notations et symboles qu'en Fig. 39.

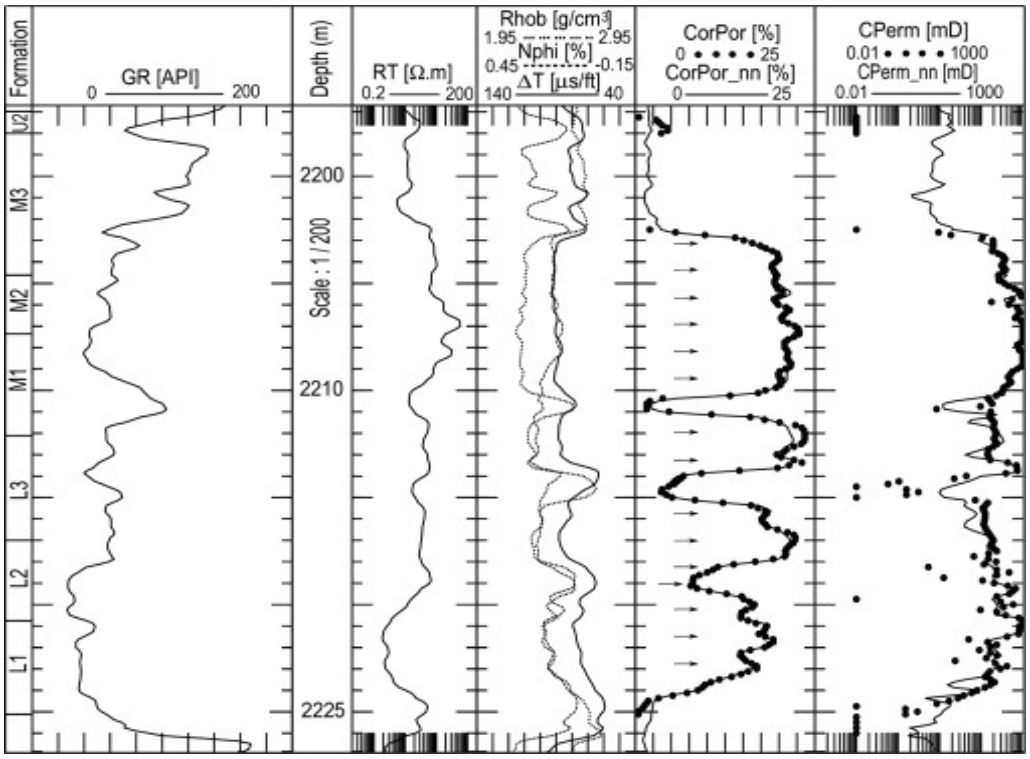


Fig.41. Log composites ainsi que les prévisions de diagraphies pour le puits HR-209. Même notations et symboles qu'en Fig. 39.

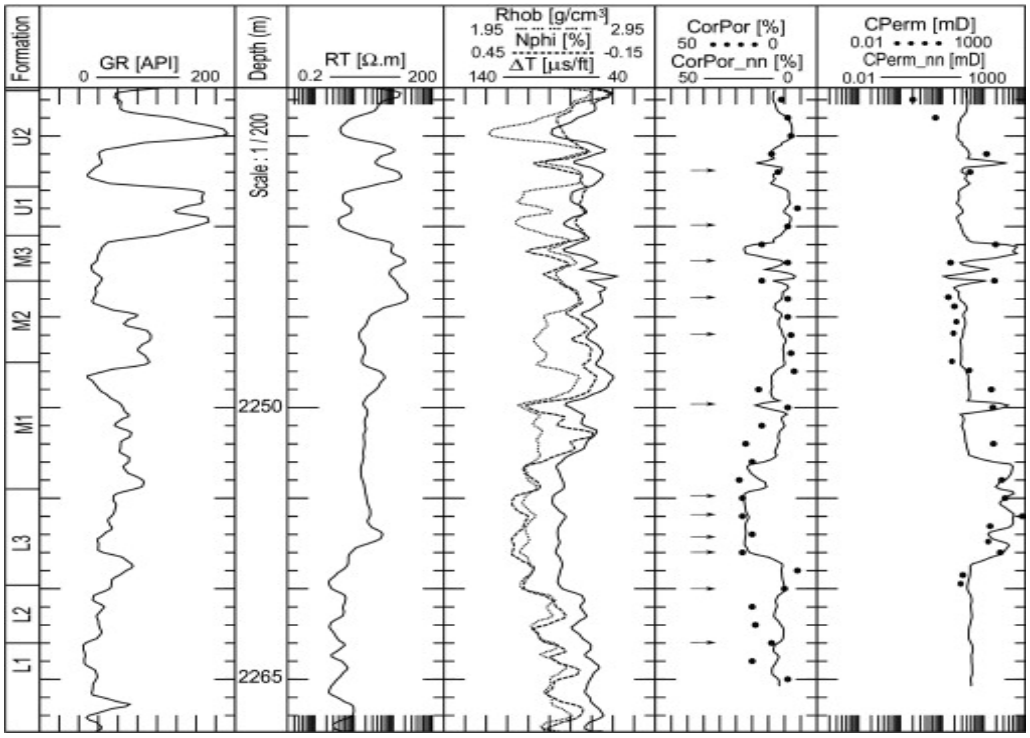


Fig. 42. Log composites ainsi que les prévisions de diagraphies pour le puits HR-199. Même notations et symboles qu'en Fig. 39.

En utilisant les propriétés des réservoirs choisis au hasard, les résultats de chaque analyse sont accumulés et affichés dans des distributions appelés "Bin" (Senergy, 2010) pour prédire la porosité (Fig. 43a) et la perméabilité (Fig. 43b) carotte. Les données de la courbe d'entrée sont divisées en un nombre d'intervalles de données à utiliser dans le modèle. Deux types de Bins peuvent être appliqués: les Bins de taille variable et les bins échantillonnés. Le poids Bin par nombre d'échantillons est utilisé en mode de prédiction lorsque l'option de taille variable des Bins est choisie suivant la règle de probabilité des données appartenant à un Bin par le nombre d'échantillons. Les résultats montrent la valeur moyenne du Bin par rapport au nombre de Bin pour chaque courbe d'entrée : temps de transit (ΔT), gamma ray (GR), porosité neutron (Nphi), densité (Rhob), saturation en eau (SW) et résistivité vraie (RT). Les barres vertes représentent l'écart-type de chaque côté de la valeur moyenne (la barre horizontale rouge, Fig. 43a, b). Les données de la courbe d'entrée seront divisées en un certain nombre de données pour utilisation dans le modèle (Fig. 43a, b). Le nombre de Bins doit être compris entre 2 et 100 (Fig. 43a, b). Le type de Bin peut être appliqué comme échantillon égal: le programme fera un passage préalable par les données d'entrée pour calculer les données de maxima et minima pour toutes les courbes. L’interactif petrophysique (Senergy, 2010) alors

définira l'espacement Bin de telle sorte qu'une proportion égale de données sera placée dans chaque groupe. Ce ne sera pas toujours pour fournir un nombre exactement égal à l'échantillon de chaque Bin. Des problèmes peuvent survenir lorsque l'on a des valeurs de données identiques au nombre d'échantillons qui doivent être dans un groupe.

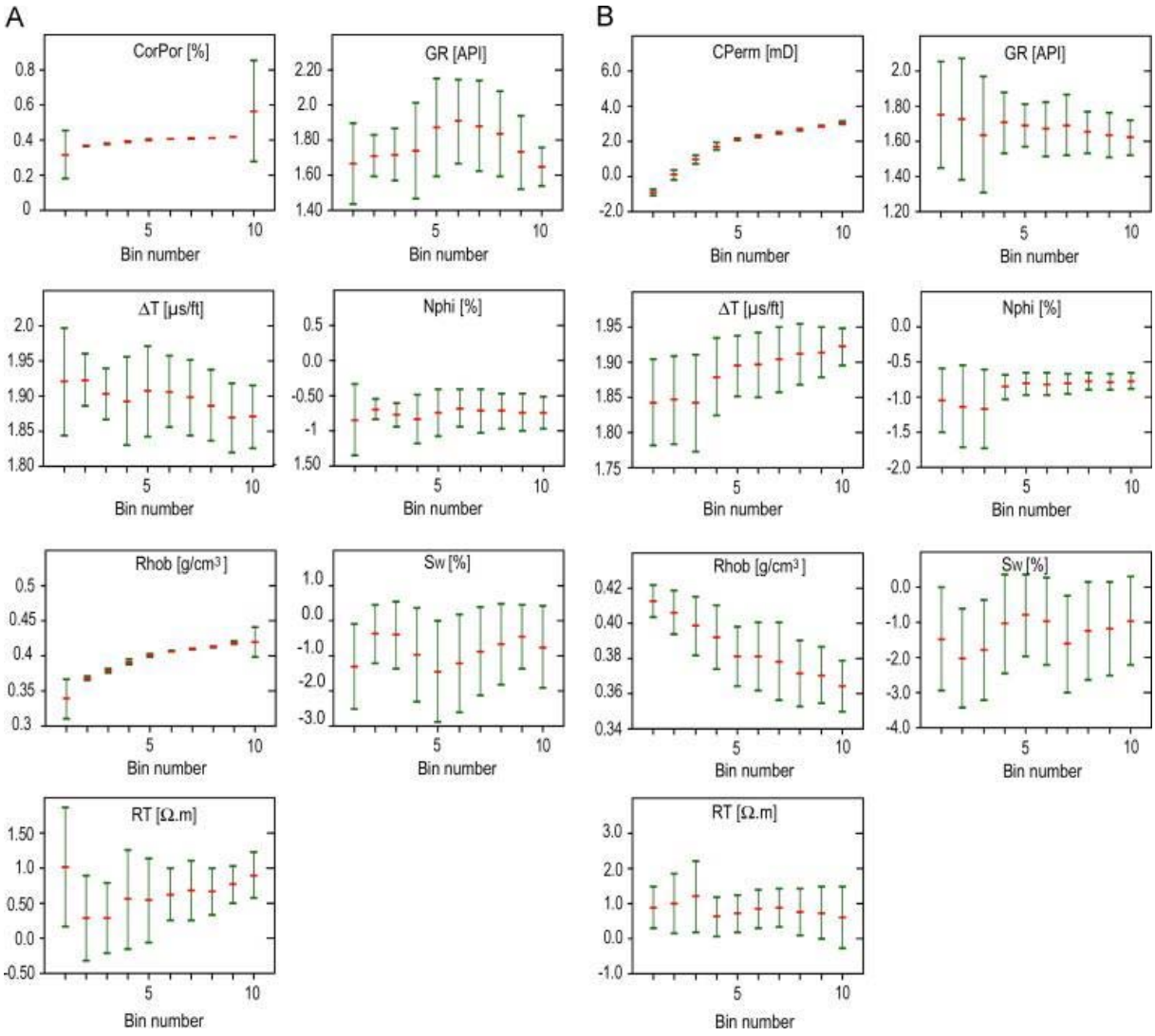


Fig. 43. Distributions du nombre de bins de la logique floue dans le réservoir du Trias de Hassi R'Mel pour (a) la porosité carotte (CorPor) et (b) la perméabilité carotte (CPerm); gamma ray (GR), sonique (ΔT), porosité neutron (Nphi), densité (Rhob), saturation en eau (Sw) et résistivité vraie (RT); Nombre de relations: valeurs prédites.

La Figure 43a montre toutes les données utilisées dans la prédiction de porosité (CorPor): les fonctions obtenues après transformation floue montrent que les données de diagraphies de puits telles que GR, ΔT, Nphi, Rhob, Sw et Rt sont fonction du nombre de bins. Les valeurs de porosité neutron peuvent être négatives par rapport à celles du calcaire dont la valeur est égale à 0.

La Figure 43b montre les mêmes fonctions obtenues par le logiciel interactif pétrophysique (Senergy, 2010) qui est utilisé pour prédire la perméabilité de carottes (CPerm) comme : GR, ΔT, Nphi, Rhob, Sw et RT.

450 points de données de porosité et de perméabilité de la formation ont été utilisées (Fig. 44a, b), 120 points de données pour vérification de l'erreur (Fig. 45a, b) et 120 points de données sont utilisées pour tester les modèles neuro-flous (FN) de la perméabilité et de la porosité. Le FIS est ainsi formé en utilisant les données d'apprentissage, puis vérifié et testé à l'aide du contrôle des ensembles de données et des essais. L'ensemble de données de test est ensuite utilisé pour vérifier la généralisation. Par conséquent, l'erreur du modèle pour l'ensemble des données de contrôle tend à diminuer à mesure que la formation a lieu au point où le sur-ajustement (over fitting) commence.

La méthode des moindres carrés peut être utilisée pour déterminer les valeurs optimales. L'erreur quadratique moyenne (RMSE) du système généré par les données d'apprentissage est respectivement 0.1575 et 0.4556 pour la porosité et la perméabilité. Les données d'essai ont été appliquées à la FIS pour valider le modèle. Dans ce travail, les données de contrôle sont utilisées à la fois pour vérifier et tester les paramètres de la FIS.

Lors de la distribution des erreurs possibles dans les paramètres d'interprétation et des courbes d'entrée, le programme sélectionne de manière aléatoire la porosité des paramètres d'entrée, la saturation, le volume d’argile à partir des séquences sélectionnées par l'utilisateur. De nombreux passages, des centaines de milliers, peuvent être définis par les modules d'analyse IP3.6. (Senergy, 2010).

La logique floue améliore la capacité de généralisation d'un réseau de neurones (NN) portant sur la connaissance implicite, le système nerveux en fournit une sortie plus fiable quand l’extrapolation est nécessaire au-delà des limites des données de formation. Dans les systèmes hybrides (FN), la combinaison des deux systèmes peut montrer les avantages des systèmes flous (Fig. 43a, b), où l'information est traitée explicitement par les réseaux de neurones. Cette modélisation a pour avantage en ce qu'il ne nécessite aucune hypothèse a priori sur la base de la complexité des réservoirs physiques ou expérimentaux ou la construction d'un modèle raisonnable à partir d'un ensemble spécifique de données de mesures. De bons coefficients de corrélation ont été obtenus pour la porosité (R^2 = 0.9879) et la perméabilité (R^2 = 0.9689) en utilisant une combinaison des logiques floue et neuro-floue. Le tableau 5 montre les résultats

des coefficients de corrélation et des taux d'erreur obtenus pour chaque technique de prédiction utilisée dans ce travail.

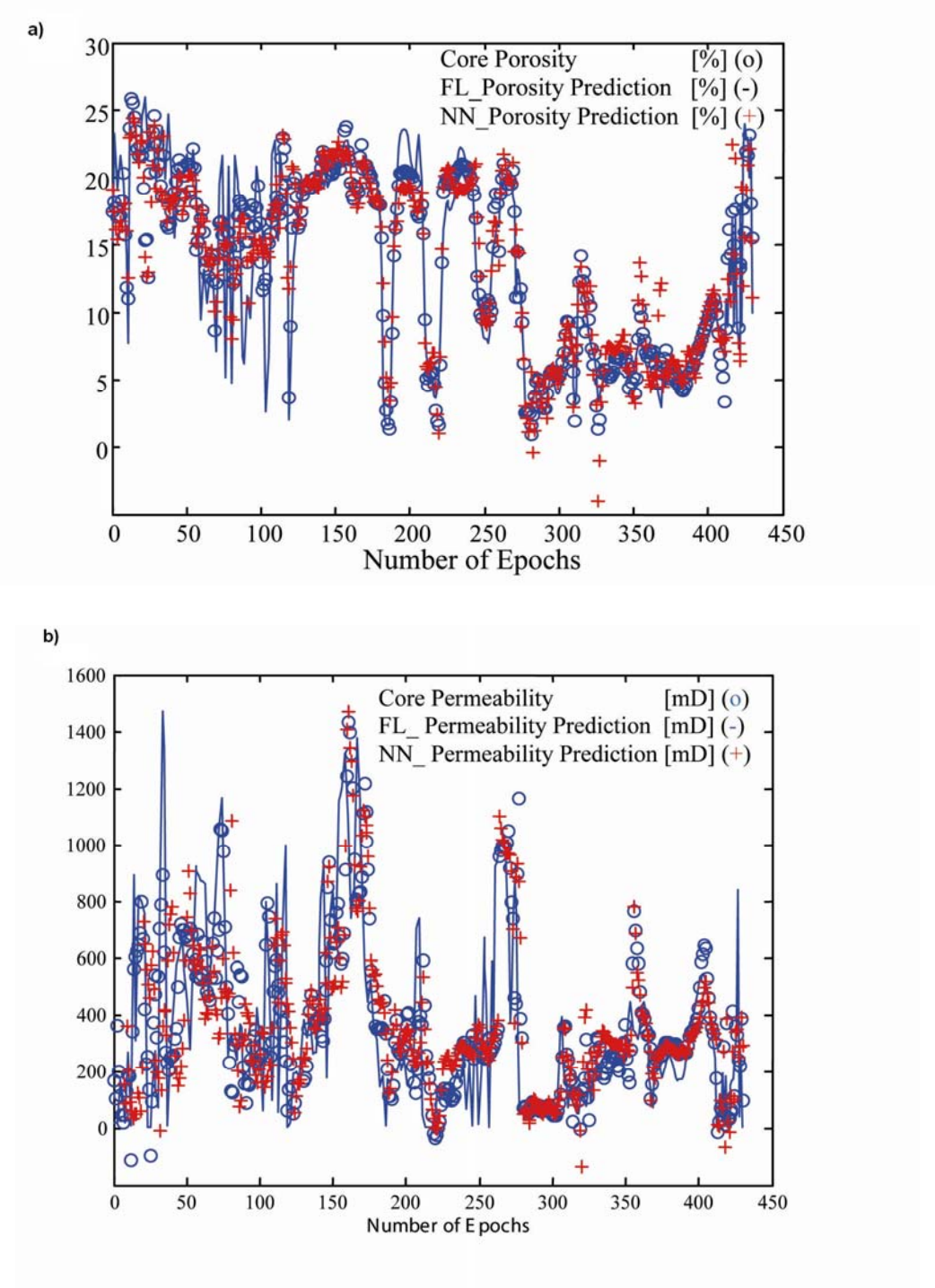


Fig. 44. Comparaison de la porosité (a) et la perméabilité (b) et des prédictions pour NN et NF, respectivement pour la porosité et la perméabilité carotte.

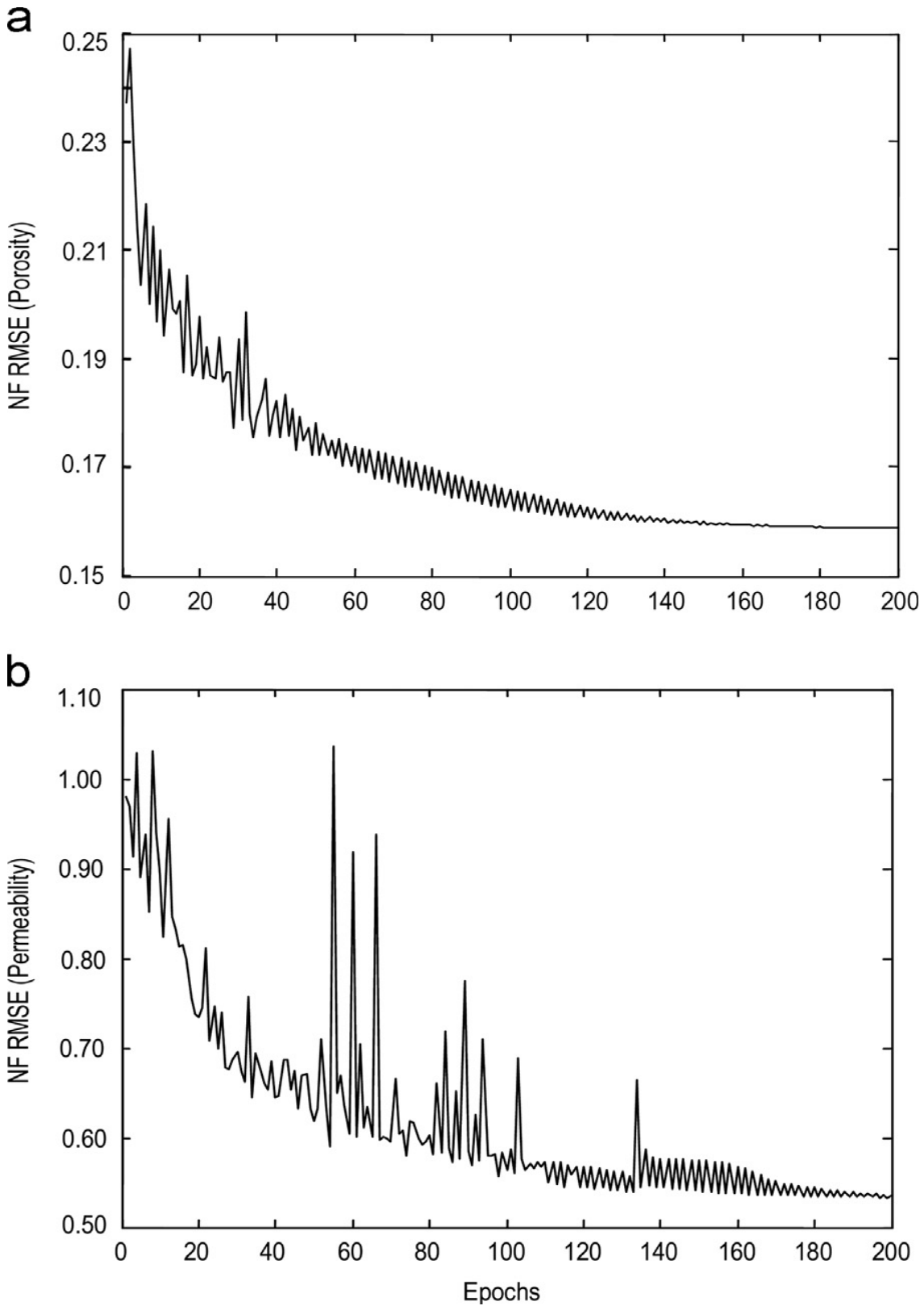


Fig. 45. Erreur quadratique moyenne pour la porosité (a) et la perméabilité (b) en utilisant la méthode NF.

	Technique de prédiction	Coefficient de corrélation	RMSE
Prédiction de porosité	Neuro Fuzzy	**0.9879**	0.1575
	Neural Network	0.9655	0.3340
	Fuzzy Logic	0.9275	0.3915
Prédiction de perméabilité	Neuro Fuzzy	**0.9689**	0.4556
	Neural Network	0.9431	0.5984
	Fuzzy Logic	0.9125	0.6571

Tab. 6. Comparaison des coefficients de corrélation et RMSE pour les trois méthodes utilisées dans le champ de Hassi R'Mel.

IV.8. Analyse et Interprétation des résultats obtenus

Des systèmes intelligents, y compris la logique floue, ont été utilisés avec succès pour bien prédire les données enregistrées en fonction de la profondeur par rapport à l'analyse des carottes. La perméabilité prédite traitée en utilisant la logique floue a montré de bons résultats avec un bon coefficient de corrélation : ($R^2 = 0.96$).

Une comparaison entre les valeurs mesurées et prédites de la perméabilité et de la porosité en fonction de la profondeur a montré un bon accord pour la technique Fuzzy C-means.

La comparaison entre la porosité et la perméabilité prédites dans le Trias de Hassi R'Mel, vérifiées par les données de carottes, a été utilisée pour tester la capacité prédictive de la technique. Pour tester la méthode de prédiction fuzzy, cette technique a été calibrée dans un puits carotté et bien testé sur un maximum de valeurs diagraphiques et de données de carottes (HR-167). Un bon ajustement entre les données de carottes (mesurées) et les données de porosité et de perméabilité prédites peuvent être remarquées (Fig. 44a, b), ce qui suggère une bonne estimation pour la prédiction de ces deux paramètres. Par conséquent, à partir des résultats de logs composites, la méthode de la logique floue a produit les courbes prévues de diagraphies de puits et de la perméabilité carotte avec un bon facteur de corrélation ($R^2 = 0.97$).

La fonction servant à générer un modèle de données à savoir " genfis2 " (Matlab User's Guide, 2012a) a été appliquée pour spécifier un rayon des groupes (0.9). Le rayon de cluster indique la zone d'influence d'un groupe quand l'on considère l'espace des données comme une unité hypercube. La spécification d'un petit rayon de cluster donne habituellement de nombreux

petits groupes dans les données, et les résultats dans beaucoup de règles et une structure de la FIS est renvoyée. Le type de modèle pour la structure de la FIS est une première pour le modèle Sugeno avec quatre règles.

Nous remarquons que l'erreur de la formation est à l'index 200 et de la même manière la plus petite valeur d'erreur de vérification de base et à l'index 140 (Fig. 45a). De là, elle commence à augmenter de façon à minimiser l'erreur de la base des données de la formation du réservoir à la 200^{e} fois (Fig. 45a, b). En fonction de la tolérance d'erreur déterminée, la courbe montre la capacité du même modèle de généraliser les données d'essai. Ces travaux ont fait l'objet d'études par utilisation de NN et NF avec FL, pour la prévision de la porosité et de la perméabilité pour un réservoir gréseux. Les valeurs de RMSE de 0.1575 et 0.4556 et les valeurs R^2 de 0.9879 et 0.9689 respectivement pour la porosité et la perméabilité, indiquent que la méthode NF est le meilleur outil pour la prédiction par rapport à un réseau de neurones classique.

Cependant, il faut tenir en compte que les résultats obtenus dans cette étude pour l'ensemble des données ont le même effet que pour d'autres types de données. Il est donc recommandé d'identifier et d'évaluer les effets des autres enregistrements de paramètres classiques des puits pour les travaux futurs.

V.1. Application de l'analyse multivariée dans la classification des Faciès

Ce travail examine l'utilité des concepts mathématiques et les capacités prédictives de l'analyse multivariée (Wendt et al., 1985), en vue d'une meilleure estimation des faciès en utilisant les méthodes de régression et l'amélioration des estimations de perméabilité pour les formations triasiques du champ de Hassi R'Mel. Dans ce contexte, nous utilisons la méthode de l'algorithme de régression linéaire pas à pas afin de mieux améliorer les estimations de perméabilité. Nous démontrons d'abord la disponibilité de l'algorithme pas à pas en appliquant la régression par étape (Hastie et al., 1990) aux formations triasiques de Hassi R'Mel. Une méthode combinant l'algorithme pas à pas et "attentes conditionnelles alternatives" (ACE) est proposée et appliquée avec des résultats qui sont comparés à ceux de la régression sans sélection de variables.

Afin d'améliorer l'estimation de la perméabilité des réservoirs pétroliers, plusieurs techniques de régression statistique linéaires ont déjà été testées dans des travaux antérieurs (Wendt et al., 1986 ; Yumei Li et al., 2006 ; Bagheripour et al., 2013) et ont corrélé la perméabilité à différents rapports de forage. Tous ces travaux ont montré que la régression statistique pour la corrélation des données est très prometteuse :

Le modèle de régression linéaire a de nombreuses applications pratiques. Il permet notamment de faire des analyses de prédiction. Après avoir estimé un modèle de régression linéaire, on peut prédire quel serait le niveau de y pour des valeurs particulières de x.

Il permet également d'estimer l'effet d'une variable sur une autre en contrôlant par d'autres facteurs. Un certain nombre de phénomènes : physiques, biologiques, économiques,…peuvent se modéliser par une loi affine, de type :

$$y = fa_0, a_1, ..., a_n\left(x_1, ..., x_n\right) = a_0 + a_1 * x_1 + ... + a_n x_n$$

Les paramètres de cette loi, c'est-à-dire les coefficients a_i, permettent de caractériser le phénomène. On effectue donc des mesures, c'est-à-dire que l'on détermine des n+1-applets $\left(x_1, ..., x_n, y\right)$.

Une mesure est nécessairement entachée d'erreur. C'est cette erreur qui "crée" le résidu r : chaque n +1 applet j fournit une équation :

$$y_j = fa_0, a_1, ...a_n\left(x_{1.j}, ..., x_{n.j}\right) = a_0 + a_1 * x_{1.j} + ... + a_n x_{n.j} + r_j$$

La régression linéaire permet de déterminer les paramètres du modèle, en réduisant l'influence de l'erreur. Par exemple, en électricité, un dipôle passif (résistance) suit la loi d'Ohm : $U = R*I$ ou, pour reprendre la notation précédente : $Y = U$, $x = I$, $f_R(x) = 0 + R * x$

En mesurant plusieurs valeurs de couple (U, I), on peut déterminer la résistance R par régression.

Nous proposons une approche en deux étapes pour la prédiction de la perméabilité qui utilise la régression non paramétrique en conjonction avec l'analyse statistique multivariée. D'abord, nous classons les données des diagraphies de puits en types d'électrofaciès. Une combinaison de l'analyse en composantes principales, l'analyse de cluster basée sur un modèle et l'analyse discriminante est utilisée pour caractériser et identifier les types d'électrofaciès. Nous appliquons également les techniques de régression non paramétrique pour prédire la perméabilité à l'aide des diagraphies de puits dans chaque électrofaciès. Trois approches non paramétriques sont examinées par la méthode ACE, le modèle généralisé additif (GAM) et les réseaux de neurones (Nnet), de même que les avantages et les inconvénients sont explorés. Les résultats sont comparés à trois autres approches à la perméabilité de prédiction qui utilisent le partitionnement des données en fonction des couches réservoirs, les informations lithofaciès et les unités de débit hydraulique (HFU). L'examen des pourcentages d'erreurs associés à l'analyse discriminante pour les puits non carottés indique que la classification des données pour une caractérisation faite sur la base d'électrofaciès est plus robuste par rapport aux autres approches. Les résultats obtenus de ce travail ont montré que le modèle de l'ACE semble être le meilleur parmi les trois approches non paramétriques.

V.2. Cas de la formation *Hassi R'Mel Sud*

Les principales études menées dans cette formation sont principalement focalisées dans la caractérisation des électrofaciès. Les techniques de l'ACE de régression non paramétrique (Gupta et al., 1999) se sont révélés être positives pour la prédiction de la perméabilité des réservoirs argilo-gréseux d'âge triasique. En effet, La prédiction de la perméabilité absolue est un élément clé dans la description des réservoirs et a un impact direct sur les autres paramètres (Jensen, 1985).

Les applications des résultats obtenus lors de l'analyse de faciès, à partir des données de diagraphies disponibles pour les formations du Trias de Hassi R'Mel Sud ont permis, en premier lieu, de définir dix électrofaciès dans les puits d'études (grès, argile, dolomite, évaporite, andésite, etc.). Une évaluation manuelle des faciès par analyse des données de

diagraphie des formations du Trias a été réalisée dans dix puits (Baouche et al., 2012). Cela nous a permis, dans un premier temps, de définir dix lithologies. Après élaboration d'un modèle sur un puits en utilisant le logiciel "Petrolog", et la vérification sur d'autres puits, un traitement semi-automatique des données a été effectué sur sept autres puits. L'analyse des faciès dans le champ de Hassi R'Mel peut être illustrée en figure 46.

Cette approche a été développée en deux étapes pour la prédiction de la perméabilité à l'aide de la régression non paramétrique en conjonction avec une analyse statistique à plusieurs variables (Lee et al., 2002). Dans une première étape, une classification des données de diagraphies dans de nombreux types d'électrofaciès est faite en conformité avec les caractéristiques uniques de données et de mesures reflétant les minéraux et les lithofaciès dans l'intervalle étudié.

V.3. Méthodologie utilisée

Les techniques de régression linéaire simple pour le calcul de la porosité et de la perméabilité faisant appel à des formules empiriques utilisant diverses réponses diagraphiques ont été utilisées pour prédire la perméabilité. Ces méthodes empiriques ne s'appliquent que localement et ignorent le fait qu'il n'y ait aucune base théorique d'une relation entre la porosité et la perméabilité. En outre, la dispersion des données sur la ligne de régression est ignorée explicitement et implicitement attribuée à des erreurs de mesure ou de commande inférieure à la variabilité des caractéristiques du réservoir (Amaefule et al., 1993). La méthodologie du "Flow Zone Index" été proposée pour l'identification des unités de débit hydraulique basé sur l'équation de Kozeny-Carman modifiée (Carman, 1956 ; Amyx et al., 1960 ; Timur, 1968 ; Hearst et al., 2000). Ce faisant, il faut un usage intensif des données de base pour décrire les variations complexes de la géométrie des pores et l'identification des indicateurs de l'Index de la Zone d'écoulement de fluide (FZI) des unités hydrauliques qui sont, à leur tour, implicitement liées à l'amélioration des estimations de la perméabilité (Xue et al, 1997).

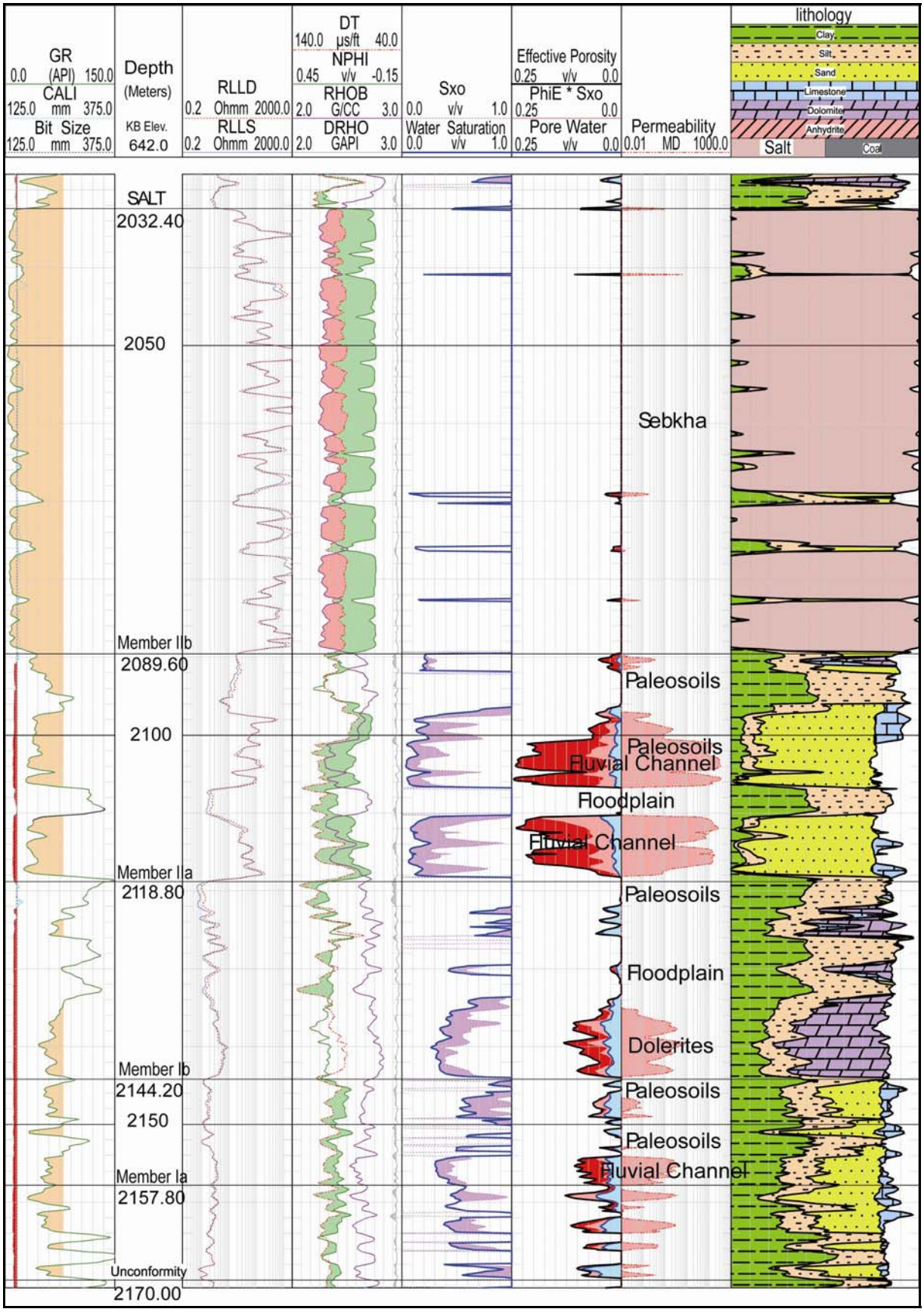


Fig.46. Analyse des faciès dans le champ de Hassi R'Mel Sud. (Baouche et al., 2010)

V.4. Données disponibles

Les données présentées dans cette analyse pour vérifier nos corrélations à l'aide de tests concernent deux puits parmi sept (HRS-7 et HRS-8). La prise en compte de la qualité des données et la disponibilité de carottes à l'échelle d'une série de diagraphies de puits est sélectionnée pour l'analyse. Dans ce domaine, nous avons 8 diagraphies de puits: gamma ray (GR), trois résistivités différentes (LLD, LLS et MSFL), temps de transit (ΔT), neutron (Nphi), et densité (Rhob), et diagraphies photoélectriques (PEF). Parmi les 8 données de diagraphies, seulement 6 paramètres (GR, RLLD, ΔT, Nphi, Rhob, et Sw) ont été choisis pour caractériser les groupes d'électrofaciès.

Une fois le partitionnement des réponses diagraphiques établi dans les groupes des électrofaciès, des techniques de régression statistique ont été appliquées pour modéliser la corrélation entre la perméabilité et attribuer les réponses au sein des groupes cloisonnés. Dans cette étude, trois techniques non paramétriques sont examinées à l'aide de l'ACE, le GAM, et le Nnet et leurs performances prédictives relatives sont évaluées. Dans la modélisation des réseaux de neurones, les données des échantillons (927) prévues à partir de 8 puits ont été divisées en deux sous-ensembles pour la **formation** et pour la **supervision**. L'ensemble des données de supervision est utilisé pour tester si le réseau de neurones peut généraliser à partir de l'ensemble de données d'apprentissage. Pour chaque groupe d'électrofaciès, nous avons modélisé les réseaux contenant le nombre optimal de nœuds dans la couche cachée produisant l'erreur moyenne des moindres carrés pour l'ensemble de données de surveillance. Les Tableaux 7a et 7b comparent les erreurs de régression pour les trois modèles utilisés pour développer les corrélations. Ces erreurs sont résumées en termes d'erreur quadratique moyenne (MSE) et d'erreur absolue moyenne (MAE) durant la régression. Le GAM semble correspondre le mieux pour les données par rapport aux autres modèles.

	GR (API)	**ΔT (µs/ft)**	**Rhob (cc)**	**RLLD (Ohmm)**	**NPhi (%)**	**SW (%)**
MSE	1420.377	4485.003	177.4659	349071.5	230.4	224.171
MAE	33.329	66.021	11.2626	109.9	13.4	13.136
MRSE	0.496	0.696	34.1510	2.5	630535.9	4185.298
MRAE	0.681	0.829	4.8987	0.9	212.8	45.430
R^2	0.275	0.343	0.4992	-0.1	0.5	0.494

Tab. 7a. Erreurs quadratiques moyennes (MSE) et erreurs absolues moyennes (MAE) pour la porosité. MSE : Erreur quadratique moyenne (Mean square error) ; MAE : Erreur absolue moyenne (Mean absolute error) ;

MRSE : Erreur quadratique moyenne relative (Mean relative square error) ; MRAE : Erreur absolue moyenne relative (Mean relative absolute error) ; R^2 : Coefficient de corrélation.

	GR (API)	**ΔT (µs/ft)**	**Rhob (cc)**	**RLLD (Ωm)**	**NPhi (%)**	**SW (%)**
MSE	193258.4	171866.0	223655.5	514328.3	225150	225010
MAE	317.8	289.0	355.7	415.0	358	358
MRSE	110.7	24.7	45558.6	5268.2	474475956	4928654
MRAE	7.5	3.5	156.8	32.1	4785	1295
R^2	0.1	0.4	-0.1	-0.1	-0	-0

Tab. 7b. Erreurs quadratiques moyennes (MSE) et erreurs absolues moyennes (MAE) pour la perméabilité. Mêmes notations qu'au Tab. 7a.

V.5. Analyses des résultats obtenus

V.5.1. Analyse des électrofaciès

Des méthodes récentes utilisées dans la classification des électrofaciès sont principalement basées sur les techniques d'identification des groupes à partir des données de réponses diagraphiques avec les mêmes caractéristiques. (Fig. 47). A cet effet, les résultats du processus de l'analyse en composantes principales sont indiqués aux tableaux 8 et 9. La première composante principale (PC1) semble indiquer la porosité (Nphi) de la formation alors que la deuxième composante principale (PC2) montre une forte corrélation avec des densités (Rhob) de 0.802908 et 0.415180 respectivement pour PC1 et PC2. Les vecteurs propres de la matrice de covariance (de Σ) fournissent les coefficients de la transformation de composants principaux.

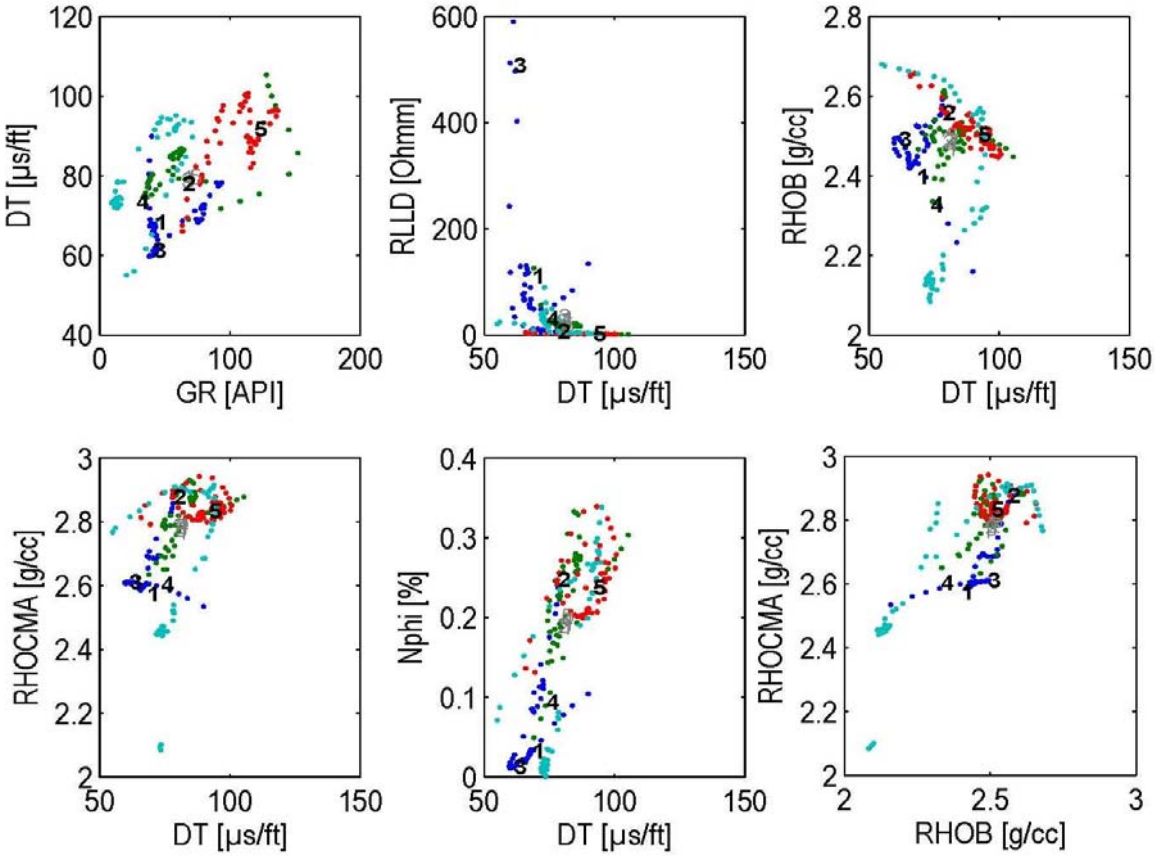


Fig.47. Electrofacies des Puits du Trias de Hassi R'Mel Hassi R'Mel Sud. RHOCMA : densité corrigée de la matrice [g/cc]; Sebkha (évaporite) : bleu; Chenaux fluviatiles (Fluvial Channel): rouge ; Paléosols (Paléosols) : Vert foncé; Plaine d'inondation (Flood Plain) : Rose; Dolérite (éruptive) : Noir.

Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7
GR	0.810222	-0.221781	0.112247	-0.326747	0.254331	0.332007	0.008781
ΔT	0.928652	-0.085377	-0.060521	0.064263	0.149098	-0.298393	0.106090
Rhob	0.802908	0.415180	-0.075180	0.373197	0.141656	-0.053781	-0.122791
RLLD	-0.771005	0.191981	-0.209594	0.454785	0.309703	0.135997	0.059369
Nphi	-0.017782	-0.613081	0.664005	0.427526	0.008350	-0.002590	-0.007632
Sw	0.431749	0.779061	0.310636	0.177623	-0.216682	0.163703	0.069734
CPor	0.525146	-0.521029	-0.491840	0.356037	-0.239513	0.162650	0.016158
Valeur propre	3.220466	3.220466	3.220466	3.220466	3.220466	3.220466	3.220466
Contribution (%)	46.0067	46.0067	46.0067	46.0067	46.0067	46.0067	46.0067
variance Totale (%)	46.00666	46.00666	46.00666	46.00666	46.00666	46.00666	46.00666

Tab. 8. Corrélations des diagraphies et principaux composants pour la porosité Cpor.

VARIABLES	PC1	PC2	PC3	PC4	PC5	PC6	PC7
GR	0.780222	-0.1117	0.112247	-0.28674	0.224331	0.312007	0.015781
ΔT	0.888652	-0.0853	-0.06052	0.064263	0.149098	-0.29839	0.106090
Rhob	0.892908	0.4151	-0.07518	0.373197	0.141656	-0.05378	-0.122791
RLLD	-0.68100	0.1919	-0.20959	0.454785	0.309703	0.135997	0.059369
Nphi	-0.05778	-0.6130	0.664005	0.427526	0.008350	-0.00259	-0.007632
Sw	0.481749	0.7790	0.310636	0.177623	-0.21668	0.163703	0.069734
CPerm	0.457146	-0.6210	-0.52184	0.456037	-0.33951	0.262650	0.026158
Valeur propre	3.125478	3.1254	3.125478	3.125478	3.125478	3.125478	3.125478
Contribution (%)	54.0057	54.00	54.0057	54.0057	54.0057	54.0057	54.0057
variance Totale (%)	55.00576	55.005	55.00576	55.00576	55.00576	55.00576	55.00576

Tab. 9. Corrélations des diagraphies et principaux composants pour la perméabilité Cperm.

V.5.1.1. Analyse en Composantes Principales

L'analyse en composantes principales (ACP) est probablement la plus ancienne et la plus connue des techniques d'analyse multivariée. Elle a été introduite par Pearson (1901), et développée indépendamment par Hotelling (1933). Comme beaucoup de méthodes multivariées, elle n'a pas été largement utilisée jusqu'à l'avènement des ordinateurs électroniques, mais elle est maintenant bien ancrée dans pratiquement tous les progiciels statistiques.

Les composantes principales constituent une autre forme d'affichage des données (Wang, 2004), permettant ainsi une meilleure connaissance de sa structure sans modifier les informations. En outre, parce que la variance totale de l'ensemble des données peut être définie comme la somme des variances associées à chaque composante principale, les quelques premières composantes principales qui expliquent plus la variation des variables initiales sont souvent utiles pour révéler la structure dans les données. Cela peut réduire la dimension du problème et de la complexité de l'analyse des clusters et de l'analyse discriminante. Dans le nuage de points (Fig. 48), la relation entre les propriétés

des réservoirs et les 3 principales composantes principales générées à partir des données de diagraphies des puits d'étude peut être explorée. La première composante principale (PC1) s'affiche pour indiquer la porosité de la formation tandis que la deuxième composante principale (PC2) montre une forte corrélation avec la lecture du rayon gamma.

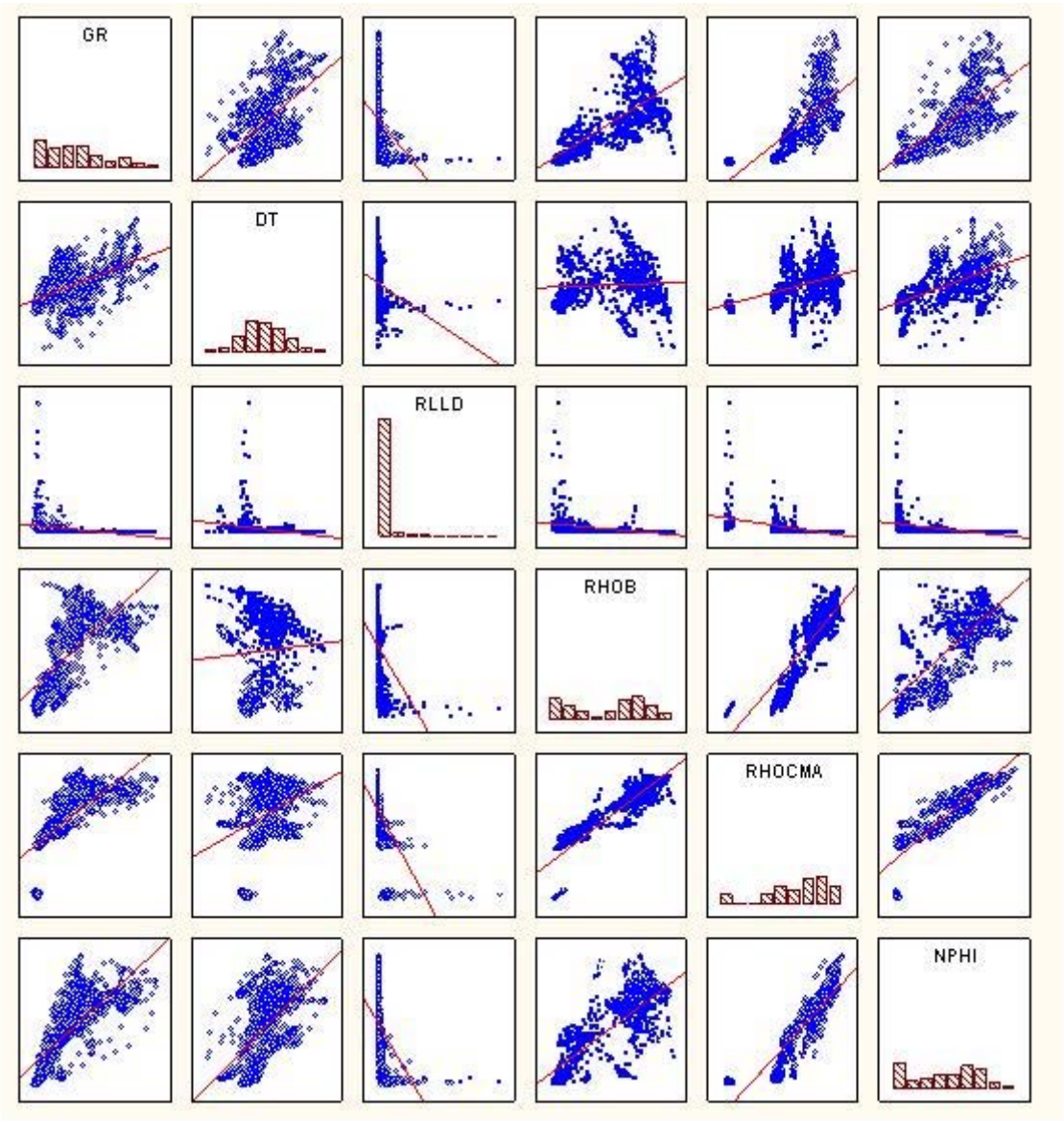


Fig. 48. Scatter plot de: GR, ΔT, RLLD, Rhob, RhocMa et Nphi, Sw.

La Figure 49a montre la projection des variables diagraphiques sur le plan des facteurs (1x2), où la contribution du facteur 1 est de 46.01% et de 21.71% pour le deuxième facteur.

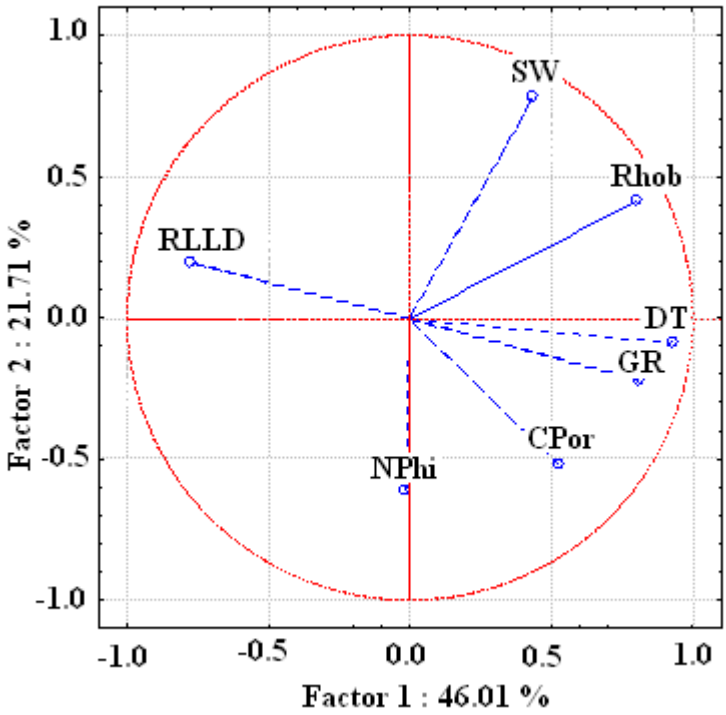


Fig.49a. Projection des variables sur le facteur plan (1x2) avec porosité CPor

Pour le PC1 et PC2 (Tab.10a), avec la porosité, les composants sont donnés par :

PC1 = 0,3994GR - 0,2456RLLD + 0,4688NPhi- 0,4329Rhob - 0.5153DT - 0,1417Sw+ 0,3328 CPor

PC2 = 0.3426GR - 0.4122RLLD + 0.1544NPHI + 0.3402Rhob + 0.0746DT + 0.6803Sw- 0.3156 Cpor

	Numéro de la Variable	Composant 1	Composant 2
GR	1	0.399453	0.342616
DT	2	0.491891	0.074640
Rhob	3	-0.432947	0.340221
RLLD	4	-0.245687	-0.418269
Nphi	5	0.468838	0.154481
Sw	6	-0.141704	0.680331
CPor	7	0.332846	-0.315642

Tab.10a. Composants des vecteurs propres vs variables diagraphiques pour la variable porosité (Cpor).

La Figure 49b montre la projection des variables diagraphiques sur le plan des facteurs (1x2), où la contribution du facteur 1 est de 48.22% et 35.61% pour le

deuxième facteur.

Pour le PC1 et PC2 (Tab.10b), avec une perméabilité CPerm, les composants sont données par:

PC1 = 0,4818GR -0,2642RLLD + 0,3283NPhi + 0,4406Rhob - 0.1455DT + 0,5331Sw+ 0,3009CPerm

PC2 = 0.0930GR - 0.4067RLLD + 0.4906NPhi- 0.3862Rhob + 0.6197DT + 0.0280Sw - 0.2259Cperm

	Numéro de la Variable	**Composant 1**	**Composant 2**
DT	1	-0.145534	0.619764
GR	2	0.481838	0.093089
RLLD	3	-0.264245	-0.406760
NPhi	4	0.328384	0.490641
Rhob	5	0.440670	-0.386256
Sw	6	0.533144	0.028047
Cperm	7	-0.300927	0.225976

Tab.10b. Composants des vecteurs propres vs variables diagraphiques pour la variable perméabilité (CPerm).

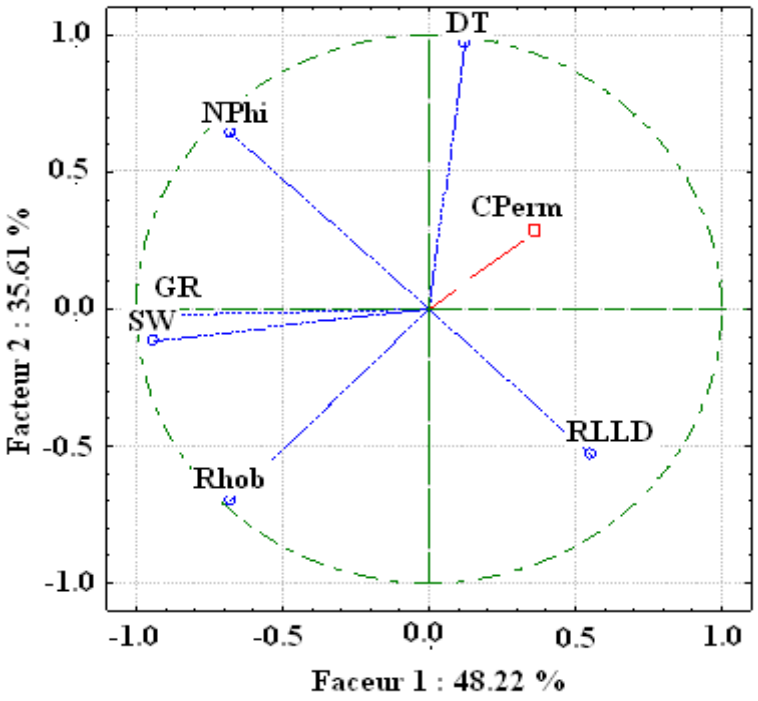


Fig.49b. Projection des variables sur le facteur plan (1x2) avec perméabilité (Cperm).

V.5.1.2. Analyse des clusters

Le but de l'analyse de classification consiste à établir des catégories de groupes de données homogènes, et isolés de l'extérieur sur la base d'une mesure de ressemblance ou de dissemblance entre les groupes (Fig. 50a, b). Dans cette étude, le groupement basé sur un modèle d'une technique de regroupement hiérarchique par agglomération est utilisé. Cette méthode peut fournir de meilleures performances par rapport aux méthodes traditionnelles telles que les voisins les plus proches et méthodes k-Means de classification, qui ne sont souvent pas en mesure d'identifier les groupes (Sebkha, chenaux Fluviatils, paléosols, plaines d'inondation et dolérites).

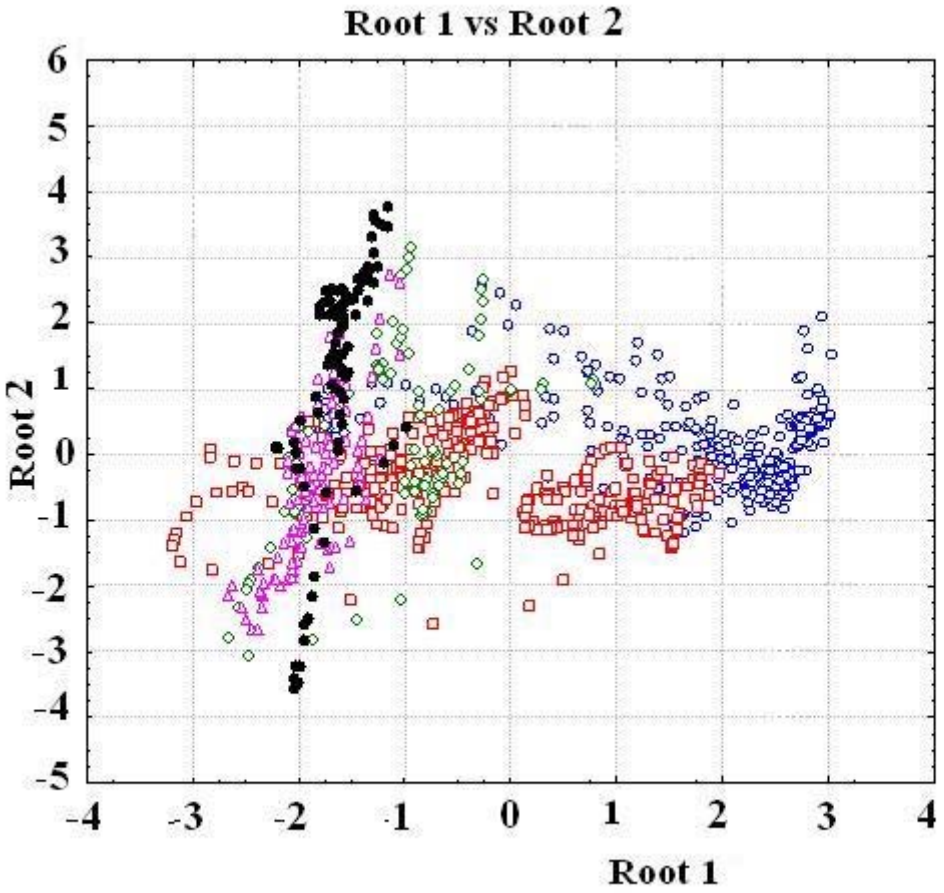


Fig. 50a. Analyse des clusters des environnements de facies (Root 1 vs Root 2). Sebkha : bleu; Chenaux fluviatils (Fluvial Channel): rouge ; Paléosols (Paleosoils) : Vert ; Plaine d'inondation (Flood Plain) : Rose; Dolérite (Dolerite). Noir.

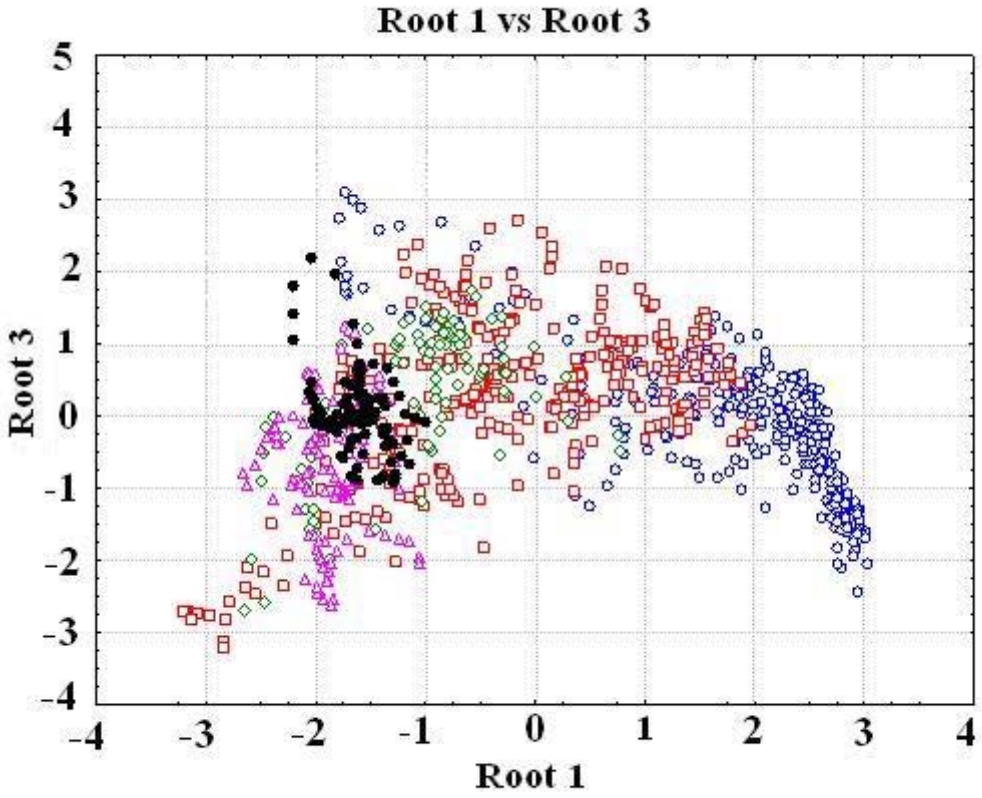


Fig. 50b. Analyse des clusters des environnements de facies (Root 1 vs Root 3). Mêmes notations qu'en Fig. 50a.

Un autre avantage de l'approche basée sur un modèle, c'est qu'il y a un critère bayésien associé pour évaluer le modèle (Wang et al., 2004) lorsque le nombre de groupes nécessaires dans d'autres techniques d'analyse de groupe conventionnel peut être estimé.

V.5.1.3. Analyse discriminante

C'est une analyse statistique pour prédire une variable dépendante (appelée variable de regroupement) par une ou plusieurs variables continues ou binaires indépendantes (appelées variables explicatives). L'analyse discriminante est utilisée lorsque les groupes sont connus a priori (contrairement à l'analyse de cluster). Chaque cas doit avoir un score sur une ou plusieurs des mesures quantitatives de prédiction, et un score sur une mesure de groupe. En termes

simples, l'analyse discriminante est la classification du fait de distribuer les choses en groupes, classes ou catégories de même type.

Cette technique est basée sur l'hypothèse selon laquelle un échantillon individuel est pris à partir de l'une des populations ou des groupes G ($\Pi_1,...\Pi_g$, **g** > 2) (Friedman et al., 1991). Si chaque groupe se caractérise par une fonction de densité de probabilité fc spécifique au groupe (x) et la probabilité a priori du groupe c est connue, d'après le théorème de Bayes, la distribution a posteriori des classes compte tenu de l'observation x est (Eq.1):

$$p(c / x) = \frac{\pi_c\, p(x / c)}{p(x)} = \frac{\pi_c\, f_c(x)}{p(x)} \alpha\ \pi_c\, f_c(x) \qquad (1)$$

Les observations doivent être attribuées dans le groupe avec la probabilité à posteriori maximale p (c / x). L'analyse discriminante nécessite une classification préalable (Tab. 11) des données en sous-groupes relativement homogènes dont les caractéristiques peuvent être décrites par les distributions statistiques des variables de regroupement associés à chaque sous-groupe (Friedman et al., 1995)

	Percent	Sebkha	chenaux Fluviatils	Paléosols	Plaine Innondation	Dolerites
Sebkha	87.15278	251	17	13	0	7
Chenaux Fluviales	73.47561	46	241	0	41	0
Paléosols	3.09278	2	46	3	23	23
Plaine Inondation	61.81818	0	18	4	68	20
Dolerites	72.11539	0	7	7	15	75
Total	68.82417	299	329	27	147	125

Tab.11. Classification préalable des données en sous-groupes relativement homogènes.

V.5.2. Corrélation de la perméabilité

Les techniques de régression non paramétrique qui ne nécessitent pas d'hypothèses a priori sur les formes fonctionnelles pour modéliser les données peuvent être appliquées si les électro

faciès sont connus. Dans ce cas, les corrélations entre la perméabilité et les réponses logarithmiques peuvent être établies pour chaque type d'électrofaciès. Il y a généralement trois approches non paramétriques différentes qui traitent de cette question: GAM, ACE, et Nnet.

V.5.2.1. Les modèles (GAM)

Les modèles de régression additive généralisée ont la forme générale (Eq.2):

$$E(y/x_1,x_2,...,x_p)=\alpha+\sum_{i=1}^{p}\phi_1(x_l)+\varepsilon \qquad (2)$$

où χ_L sont les variables prédictives et φ_L des fonctions de prédiction. Ainsi, les modèles additifs remplacent le problème de l'estimation d'une fonction d'une variable p-dimensionnelle χ par une estimation de p fonctions unidimensionnelles séparées, φ_L. Ces modèles sont attrayants s'ils correspondent aux données, car ils sont beaucoup plus faciles à interpréter que pour une surface multivariée de dimension p.

Ajustement des modèles additifs généralisés : Les ouvrages de Hastie et Tibshirani (1990), et de Schimek (2000) présentent en détail la manière dont les Modèles Additifs Généralisés sont ajustés aux données. D'une manière générale, cet algorithme utilise deux opérations itératives distinctes, qui sont souvent appelées boucle externe (*outer*) et boucle interne (*inner*). L'objectif de la boucle externe (*outer*) consiste à maximiser l'ajustement global du modèle, en minimisant la vraisemblance globale des données compte tenu du modèle (c'est l'équivalent des procédures d'estimation par le maximum de vraisemblance telles qu'elles sont décrites, par exemple, dans le module (*Estimation Non-Linéaire*). L'objectif de la boucle interne (*inner*) consiste à affiner le lissage du nuage de points, à l'aide d'un lissage par splines cubiques. Le lissage s'effectue par rapport aux résidus partiels; c'est-à-dire que pour chaque prédicteur k, il faut trouver l'ajustement spline cubique pondéré qui représente le mieux la relation entre la variable k et les résidus (partiels) qui sont calculés en éliminant l'effet de tous les autres j prédicteurs ($j \neq k$). La procédure itérative d'estimation prend fin lorsque la vraisemblance des données, compte tenu du modèle, ne peut plus être améliorée. Les modèles ont la forme générale suivante (Eq.3):

$$\theta(y)=\alpha+\sum_{l=1}^{p}\phi_l(x)+\varepsilon \qquad (3)$$

L'algorithme de l'ACE, initialement proposé par Breiman et Friedman (1985), fournit une méthode pour estimer les transformations optimales pour des régressions multiples qui se traduisent par une corrélation maximale entre une variable dépendante (réponse), et multiple indépendante variables explicatives.

De telles transformations optimales peuvent être obtenues par minimisation de la variance d'une relation linéaire entre la variable de réponse transformée et la somme des variables prédictives transformées.

Les transformations optimales sont exclusivement dérivées sur les ensembles de données et peuvent entraîner une corrélation maximale dans l'espace transformé. Les transformations ne nécessitent pas d'hypothèses a priori de toute forme fonctionnelle pour les variables de réponse ou prédictives et constituent donc un outil puissant pour l'analyse exploratoire des données et leur corrélation. Les corrélations de vecteurs dans une matrice représentant les valeurs prédictives et la valeur prédite de la perméabilité carotte (CPerm) et la porosité carotte (CPor) sont résumés dans les tableaux 12a et 12b.

	ΔT	GR	RLLD	NPhi	Rhob	SW	CPerm
ΔT	1.000000	-0.089607	-0.339964	0.571517	-0.758730	-0.224298	0.309288
GR	-0.089607	1.000000	-0.330645	0.580730	0.563389	0.731145	-0.256745
RLLD	-0.339964	-0.330645	1.000000	-0.530421	-0.011256	-0.457333	0.047765
NPHI	0.571517	0.580730	-0.530421	1.000000	0.046107	0.554921	-0.090472
RHOB	-0.758730	0.563389	-0.011256	0.046107	1.000000	0.708240	-0.495770
SW	-0.224298	0.731145	-0.457333	0.554921	0.708240	1.000000	-0.380541
CPerm	0.309288	-0.256745	0.047765	-0.090472	-0.495770	-0.380541	1.000000

Tab. 12a. Corrélations de vecteurs dans une matrice représentant les variables de prédiction et la valeur prédictive (Perméabilité).

	GR	ΔT	RHOB	RLLD	NPhi	SW	CPor
GR	1.000000	0.619490	0.138185	-0.313039	0.153794	0.150801	0.275141
ΔT	0.619490	1.000000	0.075433	-0.327866	0.102853	0.084975	0.342660
RHOB	0.138185	0.075433	1.000000	0.002034	0.998923	0.999061	0.499198
RLLD	-0.313039	-0.327866	0.002034	1.000000	-0.006388	-0.014186	-0.120998
NPHI	0.153794	0.102853	0.998923	-0.006388	1.000000	0.998692	0.515081
SW	0.150801	0.084975	0.999061	-0.014186	0.998692	1.000000	0.493719
CPor	0.275141	0.342660	0.499198	-0.120998	0.515081	0.493719	1.000000

Tab. 12b. Corrélations de vecteurs dans une matrice représentant les variables de prédiction et la valeur prédictive (Porosité).

V.5.2.2. Réseau de neurones feed-forward (Nnet).

Le réseau de neurones feed-forward (Nnet) est un réseau neuronal artificiel où les connexions entre les unités ne forment pas un cycle dirigé. Ce qui est différent du réseau de neurones récurrent. Dans ce réseau, l'information se déplace dans une seule direction, vers l'avant, à partir de nœuds d'entrée, par l'intermédiaire des nœuds cachés (le cas échéant) et les nœuds de sortie.

Habituellement, le réseau neuronal "feed forward" est disposé en couches multiples (Fig.51): Six couches entrées, deux couches de sortie et les couches cachées. Chaque couche contient un certain nombre de nœuds (également appelées unités de traitement de neurones) qui sont connectées à chaque nœud de la couche précédente pondérée par des liaisons simples. A l'exception des nœuds dans la couche d'entrée, chaque nœud multiplie la valeur d'entrée déterminée par le poids correspondant, puis additionne toutes les entrées pondérées.

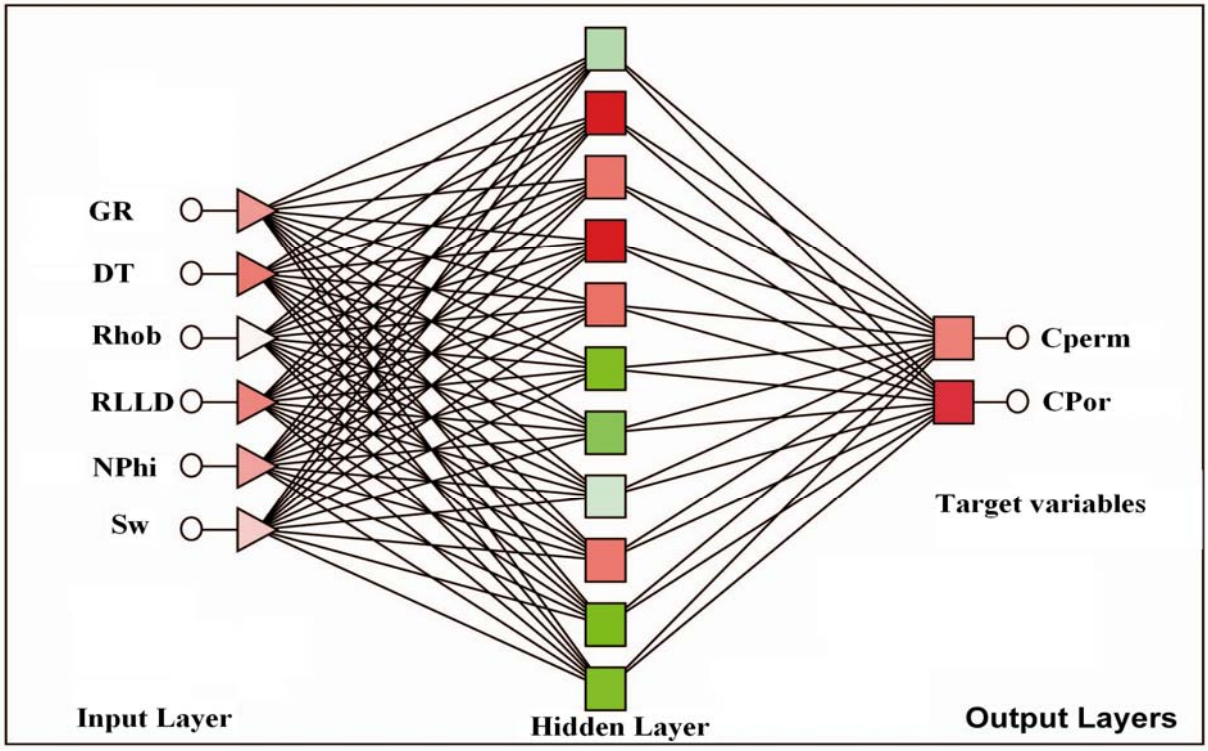


Fig. 51. Feed Forward Neural Net en couches multiples. Input Layer: Couche d'entrée; Hidden Layer: Couches cachées; Output: Couche de sortie. Target variables: Variables cibles: CPerm et CPor.

Parfois, une constante (terme de «partialité») peut être impliquée dans la sommation. La sortie finale à partir du nœud est calculée en appliquant une fonction d'activation (fonction de transfert) à la somme des entrées pondérées comme représenté ici (Fig.51, Eq.4) :

$$y_j = f_o\left(\alpha_j + \sum_{h=1}^{nh} w_{hj}\, f_h\left(\alpha_h + \sum_{i=1}^{ni} w_{ih}\, x_i\right)\right) \qquad (4)$$

où y est le vecteur des variables de sortie, x est le vecteur des variables d'entrée, w_{ij} sont les poids de connexion sur la liaison à partir de i au nœud j, *nh* est le nombre de nœuds cachés, et n_i est le nombre des variables d'entrée.

Dans cette étude, la fonction logistique (Sigmoïde) est utilisée comme fonction d'activation, ce qui donne des valeurs allant de 0 à 1 en tant que (Eq.5):

$$f(x) = \frac{\exp(x)}{1 + \exp(x)} \qquad (5)$$

Dans cette étude, nous avons utilisé un modèle d'alimentation typique qui consiste en une couche d'entrée, une couche cachée et une couche de sortie. En

outre, le nombre optimal de nœuds cachés est déterminée par essai et erreur. Cette technique est appliquée à des sables réservoirs très argileux dans les formations du Trias de Hassi R'Mel. Les régressions multiples classiques n'ont pas fourni de perméabilité et de porosité des prédictions satisfaisantes dans ce domaine. La capacité prédictive supérieure de notre approche proposée est vérifiée à l'aide de tests. Nous avons également comparé nos résultats avec la méthode de réseau de neurones pour la perméabilité et la prédiction de la porosité basée sur les lithofaciès et les résultats sont meilleurs : la perméabilité n'existe que dans les milieux réservoirs or dans un système pétrolier ce sont les chenaux fluviatiles qui sont les plus favorables aux accumulations d'hydrocarbures ayant une certaine marge de perméabilité. Contrairement aux évaporites, à la plaine d'inondation qui ne renferment pas d'hydrocarbures et donc absence ou très faible perméabilité.

V.5.2.3. Perméabilité dans les puits et HFU (Hydraulic Flux Unit)

Beaucoup de modèles pétrophysiques utilisés pour les études de simulation sont basés sur l'approche classique de relations reportant les valeurs logarithmiques de perméabilité par rapport à la porosité (Fig. 52), puis, en ajustant une droite de régression sur cette partie, la prédiction de la perméabilité à travers la roche réservoir a donné un coefficient de corrélation (R^2 = 0.029). Cette approche est essentielle lorsqu'elle est utilisée pour modéliser les roches perméables, car elle implique deux concepts. Tout d'abord, on ne considère que la relation entre les logarithmes de base par rapport à la perméabilité et porosité linéaire. Deuxièmement, en utilisant des porosités de diagraphies sur cette partie de prédiction, les perméabilités impliqueraient un accord de mise à l'échelle. La discrétisation du réservoir en unités telles que des couches et des blocs, et l'attribution de valeurs de toutes les propriétés physiques pertinentes à ces blocs donnera un meilleur point de vue du réservoir.

Certaines unités d'écoulement dans le puits HR-7, la relation de la porosité et la perméabilité construites pour les HFU uniques sont illustrées en figure 53. Les échantillons dans le nuage de points sont regroupés par HFU après avoir calculé

le FZI en utilisant les équations : Eq.6 à Eq.11 plus bas, à partir des données de diagraphie, la perméabilité peut être déterminée pour chaque HFU (valeur moyenne FZI) en utilisant l'équation Eq.12.

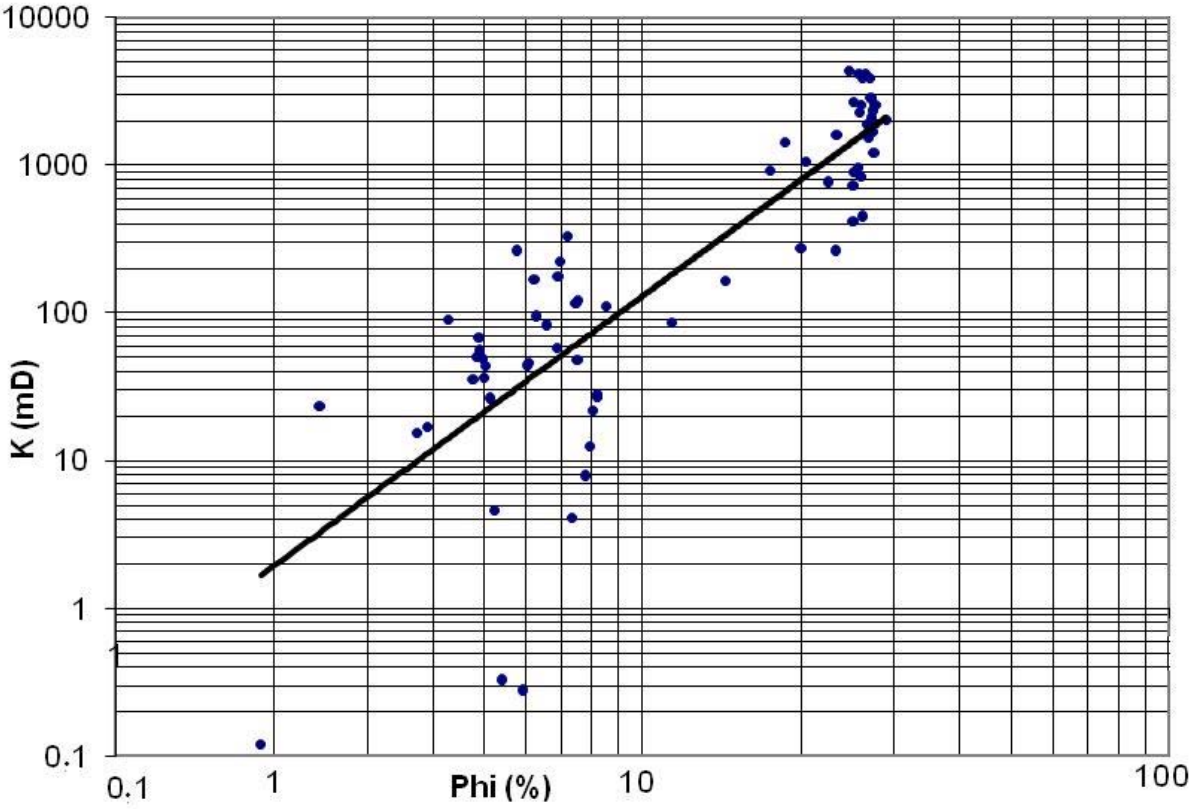


Fig. 52. Perméabilité vs Porosité au puits HRS-7. $Y = 0.314\,X^{2.612}$; CPerm [mD] ; $R^2 = 0.717$

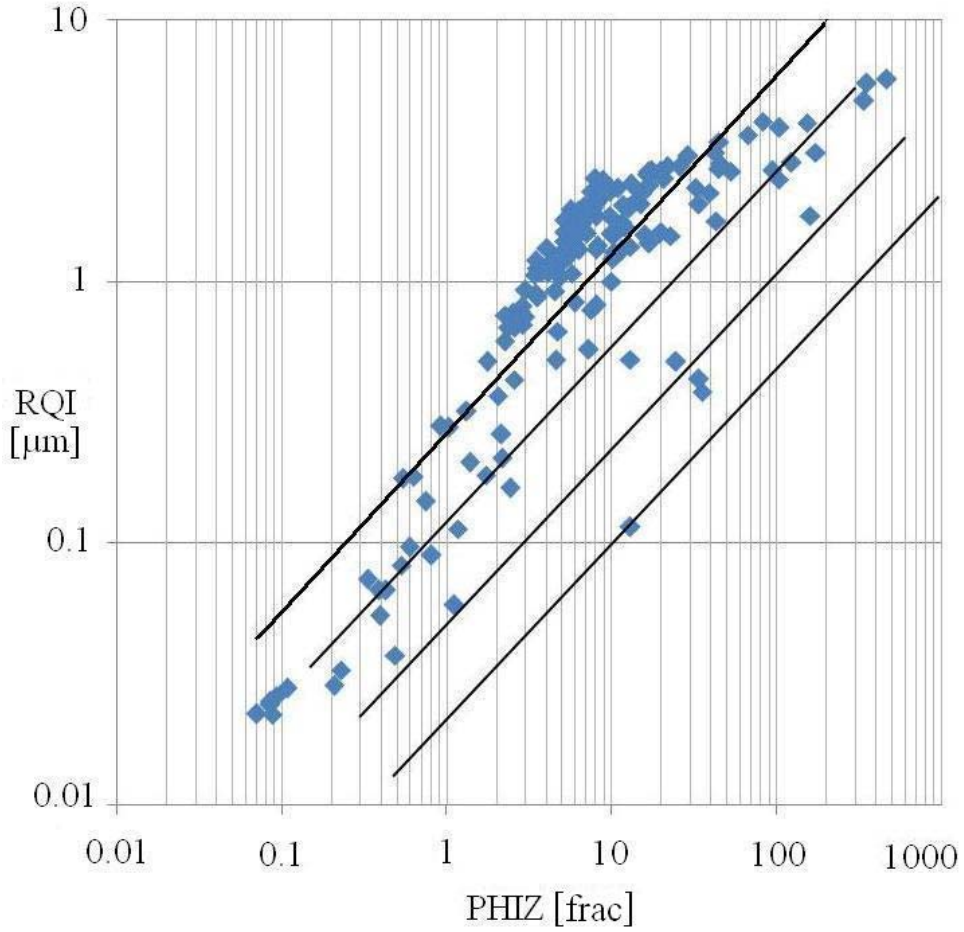


Fig. 53. Scatter plot : log RQI vs log PHIZ où HFU données au puits HRS-7.

$Y = 0.0.262\,X^{0.684}$; CPerm [mD] ; $R^2 = 0.738$

La perméabilité calculée (K_H) par rapport à la perméabilité de carotte (Permc) en fonction de la profondeur est représentée en figure 54 et la figure 55. Les profils de perméabilité établis sont comparés avec les valeurs prédites de la perméabilité (KH-NN) et les résultats montrent une très bonne corrélation (0.98824) et pour la porosité CPhi (0.9736). Cela a été fait seulement pour vérifier la précision de la méthode HFU pour la prédiction de la perméabilité dans ces puits.

La formation III se caractérise principalement par un faciès évaporitique de type Sebkha halitique (Fig. 47). A cet effet, l'estimation de la perméabilité par application de l'équation de corrélation établie dans chaque zone peut être déterminée directement sur les réponses des diagraphies correspondantes à la zone prédéfinie, sans l'utilisation de l'analyse discriminante. Toutes les

prédictions de perméabilité et de porosité en fonction de la classification de zone pour le puits HRS-7 (perméabilité carotte) sont représentées en figure 54: KH-NN (Neural Network perméabilité) et Permc (FHU perméabilité). Les mêmes résultats ont été obtenus pour le puits HRS-8 (Fig. 55).

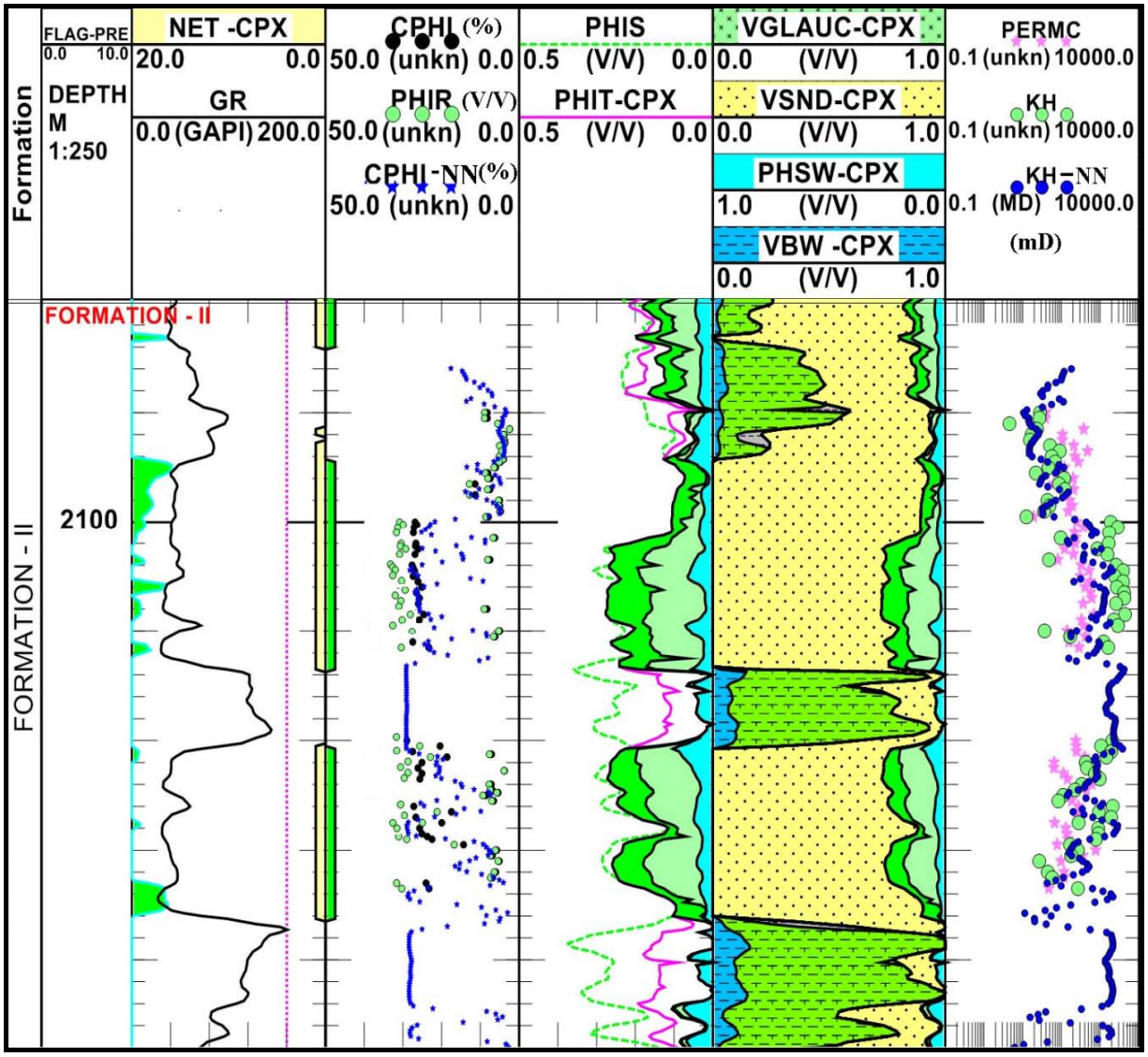


Fig. 54. Prédiction de la perméabilité et de la porosité en fonction de la classification zonale au puits HRS-7. KH : perméabilité carotte ; Porosités prédites : CPhi (Porosité carotte) ; CPhi_NN (Porosité neural network). Coefficient de corrélation : 0.9824.

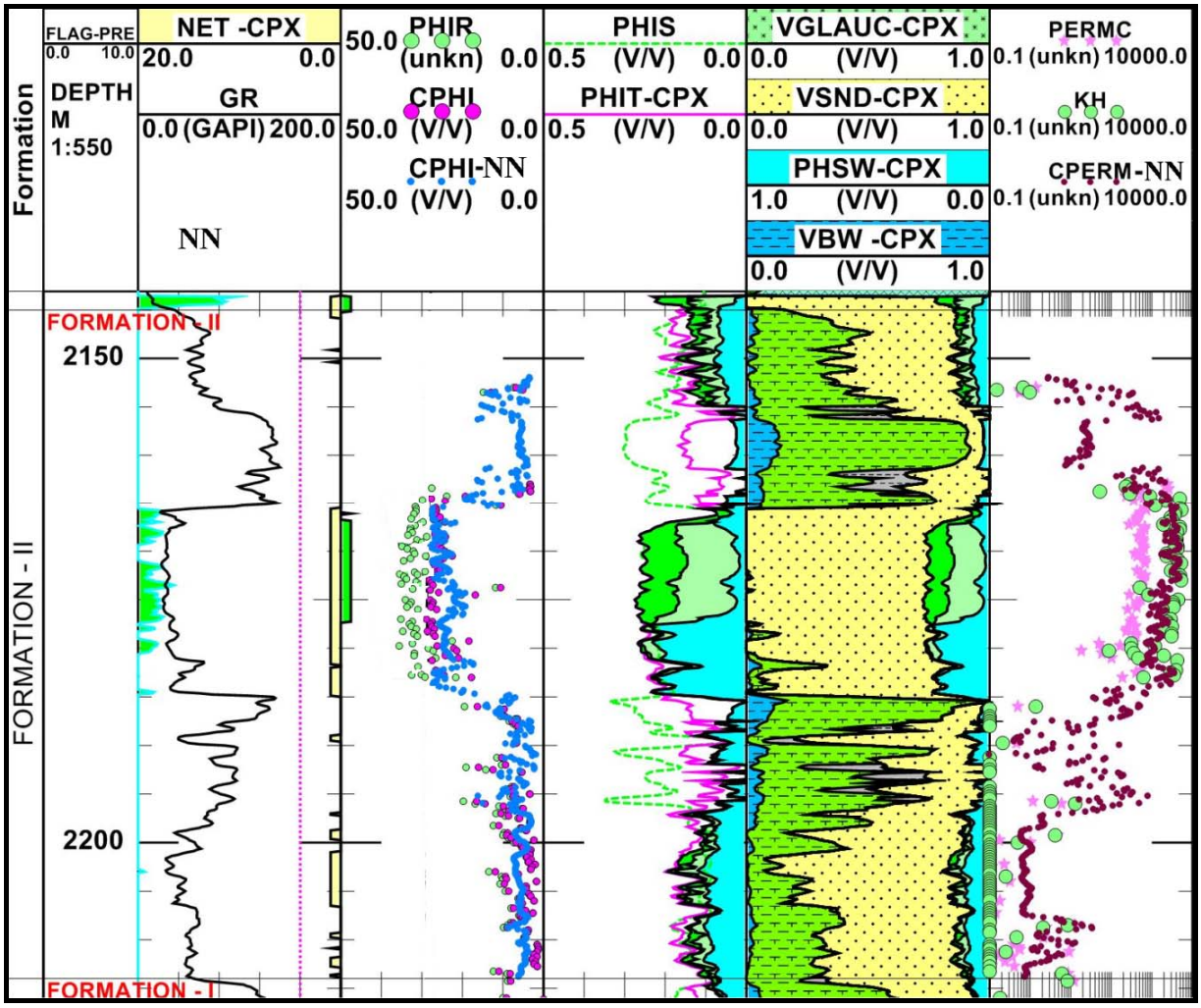


Fig.55. Prédictions de la perméabilité et de la porosité en fonction de la classification zonale au puits HRS-8. Coefficient de corrélation : 0.9845 ; mêmes notations qu'en Fig.54.

V.5.3. Classification basée sur les Lithofaciès: Les études de faciès ont commencé par une analyse de base, sur les formations du Trias du sud de Hassi R'Mel. Cela a nécessité une analyse de base au laboratoire de Sonatrach (CRD, Boumerdès), ainsi que la combinaison des différentes données de diagraphies de puits: Gamma Ray, neutron et sonique, densité et résistivité ainsi que la saturation en eau. L'étude des figures sédimentaires, qui représente les conditions hydrodynamiques et physiques de l'environnement de sédimentation, a également été prise en compte.

Le partitionnement a été fait en utilisant les résultats lithofaciès qui sont habituellement déterminés par des descriptions de base et analyse de lames

minces pour tous les puits (HRS-2, 4, 7, 8, 9,11 et 12). Le champ de Hassi R'Mel peut être décrit par 10 différents types de lithofaciès comme indiqué au Tableau 13.

Nombre de faciès (Facino)	Description de Faciès
1.	Halite
2.	Halite argileuse
3.	Grès
4.	Grès faiblement argileux
5.	Grès argileux
6.	Argile dolomitique
7.	Grès argileux dolomitique
8.	Grès faiblement argileux dolomitique
9.	Argile gréseuse
10	Dolérite

Tab. 13. Description des différents types de lithofaciès.

Les moyennes de variation des porosités et perméabilités de base, associées aux erreurs en fonction de la lithologie avec un intervalle de confidence qui sont donnés en figure 56a, b.

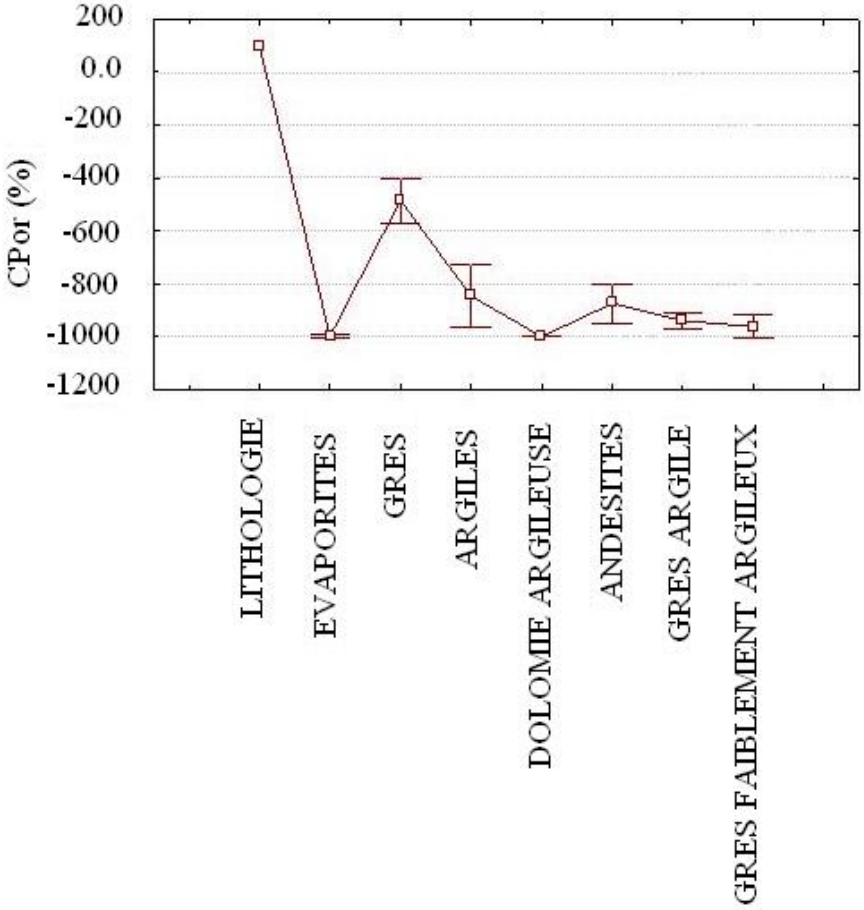


Fig. 56a. Moyennes avec erreurs de porosité en fonction de la lithologie.

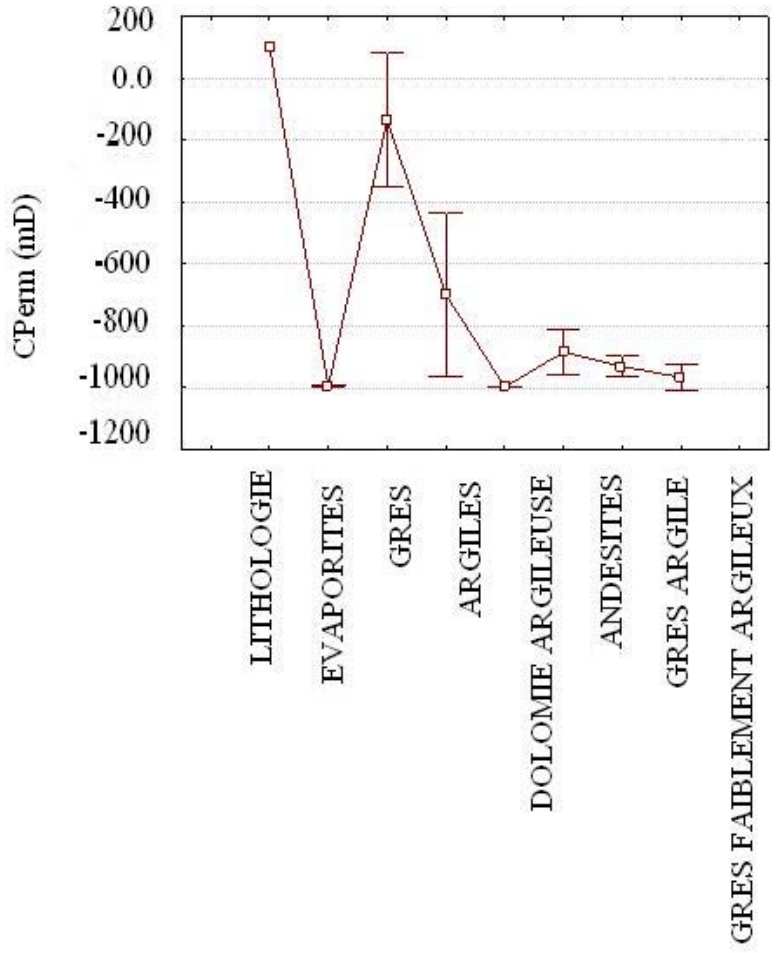


Fig. 56b. Moyennes avec erreurs de perméabilité en fonction de la lithologie.

L'examen de cette partie montre que les variables évaporite, argile, dolomie argileuse, grès et de grès à faible argilosité sont le moins influent dans la détermination de la première composante principale tandis que les sables jouent le rôle le plus important expliquer mieux ainsi que les barres d'erreurs pour les figures 56a,b. La performance prédictive des différents classificateurs est définie comme un taux d'erreur qui représente la fraction de formation ou un ensemble de test qui a été mal classé. Les taux d'erreur valorisés pour ce contexte sont résumés dans le tableau 7 pour les puits HRS-7 et HRS-8.

La fonction crossval (Matlab User's Guide, 2012a) permet d'estimer l'erreur de classification pour l'analyse discriminante linéaire (lda) et l'analyse discriminante quadratique (qda) en utilisant la partition de données donnée pour les deux puits HRS-7 et HRS-8, où l'adq a une plus grande erreur de validation

croisée de lda. Elle montre qu'un modèle plus simple peut obtenir une performance comparable ou supérieure à celle d'un modèle plus compliqué d'autres classifications (Tab. 14). Les résultats obtenus (Tab.14) montrent que les valeurs de la performance obtenues sont les plus siginificatives pour la lithologie (0.7535 et 0.7515).

		ldaCVerr	qdaCVerr	Perf	Tr_Perf	Val_Perf	Test_Perf
HRS-7	HFU	0.4052	0.4125	0.7425	0.5266	1.0002	0.9255
	LITHO	0.4444	0.3919	0.7535	0.6496	1.0285	0.9622
	Facies	0.4450	0.3871	0.7337	0.6894	0.9048	0.7693
HRS-8	HFU	0.4129	0.4135	0.7140	0.4828	0.7797	0.8358
	LITHO	0.4401	0.3910	0.7515	0.6396	1.0200	0.9522
	Facies	0.4189	0.3221	0.1954	0.2181	0.1415	0.1439

Tab. 14. Erreurs de classification pour analyse discriminante linéaire (lda) et analyse discriminante quadratique (qda). ldaCVerr: analyse discriminante linéaire (cross validation error) ; qdaCVerr : analyse discriminante quadratique (cross validation error) ; Perf : Performance ; Tr_Perf : Train Performance ; Val_Perf : Validation de la performance; Test_Perf : Test de performance.

Le taux d'erreur associé à la classification des HFU est le plus élevé parmi toutes les techniques considérées dans cette étude. Les données utilisées pour l'analyse discriminante quadratique (QDA) montre que les erreurs calculés par la fonction cross-validation "ldaCVerr" et "qdaCVerr" représentent l'erreur de prédiction prévue sur un ensemble indépendant. A cet effet, l'analyse discriminante sur la base des données de lithofaciès est déterminée à partir de l'information des électrofaciès (Mathisen, 2003).

Parmi les trois méthodes non paramétriques, le modèle de l'ACE qui semble être le meilleur en termes de modèles de prédiction de la porosité et de la perméabilité. Le modèle du réseau neuronal a tendance à surestimer ou sous-estimer la valeur de prédiction dans certains intervalles de la formation. En particulier, la performance prédictive dans les puits HRS-8 indique qu'un réseau de neurones ne fournira pas nécessairement des prévisions précises pour les données qui n'ont pas été utilisées. Par conséquent, le modèle GAM donne de bons résultats pour la plupart des intervalles dans les puits non testés.

V.6. Comparaison à d'autres techniques de perméabilité prévisionnelle

La notion d'unité hydraulique (Amaefule et al., 1993) a été choisie pour subdiviser le réservoir en deux types pétrophysiques distincts. Chaque type de réservoir distinct a une valeur unique d'indicateur de zone de débit (FZI).

Selon Tiab (2000), une unité d'écoulement hydraulique (HFU) est un corps continu sur un volume de réservoir spécifique qui possède pratiquement une uniformité. Il est caractérisé par les propriétés pétrophysiques du fluide "caractérisant" de façon unique la communication statique et dynamique avec le puits de forage.

Cette technique est basée sur celle de Kozeny et Carman (1927) modifié, (cité dans Amaefule, 1993) et le concept de rayon hydraulique moyen (Eq.6):

$$k = \left\{ \frac{1}{2\tau^2 S_{gv}^2} \right\} \left(\frac{\varphi_e^3}{(1-\varphi_e)^2} \right) \qquad (6)$$

S_{gv} peut également définir l'aire de surface des grains par rapport à l'unité du volume de fluide.

L'indicateur de zone d'écoulement est fonction des caractéristiques géologiques de la roche et de la géométrie de pores. Par conséquent, il représente un bon paramètre pour déterminer les unités de débit hydraulique (HFU). Il est également une fonction de l'indice de la qualité du réservoir et de l'indice des vides.

Amaefule et al. (1993) ont proposé une méthode de relation pétrophysique en zonation en utilisant un indicateur de zone d'écoulement (FZI) sur un cross plot de porosité en fonction perméabilité comme indiqué dans l'équation 7.

Cette technique de caractérisation des unités de flux est basée sur le concept de rayon hydraulique moyen de la relation de Kozeny-Carman (1956).

Définition de la zone d'indicateur d'écoulement FZI (µm) est définie comme:

$$FZI = \frac{1}{S_{gv}\tau\sqrt{F_s}} \qquad (7)$$

l'indice de qualité du réservoir RQI (pm) est donné par:

$$RQI = 0.0314\sqrt{\frac{K}{\varphi_e}} \qquad (8)$$

et la porosité normalisée (fraction) s'écrit :

$$\varphi_Z = \left(\frac{\varphi_e}{1-\varphi_e} \right) \qquad (9)$$

A cet effet, l'équation 8 devient:

$$RQI = FZI \times \varphi_Z \qquad (10)$$

En prenant le logarithme des deux termes de l'équation 10 on aura :

$$Log(RQI) = Log(FZI) + Log(\varphi_Z) \qquad (11)$$

Sur une échelle log-log de RQI en fonction de φ_z, tous les échantillons avec des valeurs FZI similaires se situeront sur une ligne droite (Fig.54). La valeur de la constante FZI peut être déterminée à partir de la relation PHIZ (%) vs RQI (Fig.54). Les points qui se trouvent sur la même ligne droite ont des caractéristiques similaires, et de ce fait, constituent l'unité hydraulique. La perméabilité d'un point de l'échantillon est ensuite calculée à partir d'une HFU pertinent en utilisant la valeur moyenne de FZI et la porosité de l'échantillon correspondant à l'aide de la formule d'Amaefule et al. (1993):

$$K = 1014 \times FZI^{2} \frac{\varphi_e}{(1-\varphi_e)^2} \qquad (12)$$

Toutefois, il est important de mentionner que, compte tenu de la porosité réelle et une vraie valeur de HFU (basée sur les données de base), la perméabilité prédite montre un accord presque parfait avec la vraie perméabilité. Ainsi, la principale difficulté semble être l'identification des unités de débit hydraulique dans les puits non carottés. Les performances prédictives des quatre approches sont comparées en utilisant les erreurs linéaires et quadratiques pour les deux puits HRS-7 et HRS-8 (Tab. 7).

V.7. Analyse des méthodes utilisées

L'utilisation des méthodes de régression non paramétrique dans l'étude de la caractérisation des réservoirs permet d'avoir des résultats d'électrofaciès plus significatifs dans le modèle de réservoir à huile argilo-gréseux. Par conséquent, tous les résultats de prédiction de la porosité et de la perméabilité en utilisant les trois méthodes montrent une très bonne corrélation des données de diagraphies de puits avec la prédiction des trois types de porosité et de perméabilité calculée à partir de la HFU (Permc), à partir des réseaux de neurones (K_{H_NN}) corrélées avec les données de carottes (K_H et CPhi). A cet effet, sur la base de ce travail, on peut constater que :

1. La sélection des variables en conjonction avec les méthodes de régression montre un potentiel important pour les prédictions de perméabilité dans les réservoirs argilo-gréseux.
2. Un examen des taux d'erreur pour les puits HRS-7 et HRS-8 de Hassi R'Mel (formation triasique), indique que les étapes effectuées avec les méthodes ACE sont des méthodes très favorables pour la prédiction de la perméabilité et de la porosité.
3. Dans le cas de la prédiction de la perméabilité, en utilisant une combinaison de corrélation de données et de variables, le succès de la méthode dépend fortement de la puissance discriminatoire de la technique de classification des données. Les résultats obtenus montrent que la difficulté d'identification des électrofaciès dans les puits non carottés peut entraîner une corrélation incorrecte et, par conséquent, les prévisions de la perméabilité seraient de mauvaises qualités.
4. En comparant la performance prédictive relative des trois méthodes de régression utilisées, le pas à pas avec la méthode ACE semble surpasser les deux autres méthodes.

CONCLUSION GENERALE

La caractérisation des réservoirs argilo-gréseux par les données de diagraphies est une façon habituelle de décrire les réservoirs dans les champs de gaz.

Notre étude menée dans le domaine de l'ingénierie pétrolière a montré une très grande importance de l'intelligence artificielle dans les applications basées sur les données pétrophysiques utilisées conjointement avec les enregistrements des diagraphies de puits et des données de modules de base, dans un réservoir de sable argileux de la formation du Trias du champ de Hassi R'Mel.

1. Dans ce travail, la technique intelligente NF a été utilisée pour estimer la porosité des réservoirs productifs ainsi que la perméabilité à partir d'enregistrements de diagraphies de puits conventionnelles. L'analyse des courbes de prédiction de la porosité et perméabilité basées sur la logique floue peuvent être utilisées pour mieux sélectionner les paramètres qui sont en excellent rapport avec les propriétés des réservoirs.
2. L'approche de la modélisation NF présentée dans ce travail a été appliquée avec succès pour la prédiction des paramètres de réservoir pétrophysiques. Cette approche de modélisation a l'avantage significatif en ce qu'elle ne nécessite aucune hypothèse précédente fondée sur des considérations physiques ou expérimentales sur les complexités de réservoirs de construire un modèle raisonnable et précis d'un ensemble de données mesurées.
3. Une approche de logique floue est utilisée pour calibrer la perméabilité calculée et la perméabilité carotte; et un réseau de neurones a été développé dans ce modèle, sur la base des données disponibles. Dans ce travail, la technique intelligente NF a été utilisée pour estimer la porosité des réservoirs productifs ainsi que la perméabilité à partir d'enregistrements de diagraphies de puits conventionnelles. L'analyse des courbes de prédiction de la porosité et perméabilité basées sur la logique floue peuvent être utilisées pour mieux sélectionner les paramètres qui sont en excellent rapport avec les propriétés des réservoirs.
4. Un réseau de neurones est utilisé comme une méthode de régression non linéaire afin d'effectuer des estimations prédictives des paramètres

pétrophysiques obtenus à partir des diagraphies de puits choisis et des mesures de carottes.

5. Les résultats obtenus ont montré d'excellents coefficients de corrélation pour la porosité 0.957 et la perméabilité 0.825 en utilisant les modèles neuro-flous, comparés aux réseaux de neurones. Ces techniques peuvent faire des estimations de propriétés réservoirs d'une manière plus précise et plus fiables et par conséquent peuvent être utilisés comme un puissant outil réservoir dans l'industrie pétrolière, avec une application aux différents puits et champs pétroliers.

- D'autres approches, utilisant des informations de lithofaciès identifiés à partir des électrofaciès et des données de carottes, ont montré de bons résultats.

Afin de mieux améliorer les mesures prévisionnelles de la perméabilité et de la porosité, il faut s'assurer que :

1. les résultats du traitement des données soient soigneusement vérifiés,
2. le nombre des unités de débit hydraulique (HFU) doit être optimal,
3. les calculs des paramètres FZI et HFU soient d'abord faits,
4. une meilleure qualité de perméabilités soit pratiquement obtenue (en laboratoire) et des corrections aient été introduites dans la phase d'analyse pétrophysique.

Cette méthode hybride devient un outil puissant pour l'estimation des propriétés des réservoirs pétroliers dans les projets de développement du pétrole et du gaz naturel.

PERSPECTIVES DE RECHERCHE

La porosité et la perméabilité sont deux paramètres les plus importants des réservoirs d'hydrocarbures. Leur investigation dans le domaine pétrolier, par le biais d'outils adéquats, permet d'évaluer et de minimiser le risque et l'incertitude dans l'exploration et la production de gisements de pétrole et de gaz.

Les différentes méthodes directes et indirectes sont utilisées pour mesurer ces deux paramètres, dont la plupart (ex. analyses de carottes) nécessitent beaucoup de temps

d’extraction ainsi que des coûts élevés pour leurs analyses. L’utilisation des méthodes efficaces dans le but d’apporter de meilleures solutions à ces problèmes sont les bienvenues.

L’application des méthodes de classification ("pattern matching"), ainsi que les méthodes d’optimisation et l’exploration des données dans l’intelligence artificielle, s’avèrent très robustes dans la résolution des problèmes d'ingénierie de pétrole et gas.

Les dernières problématiques abordées et traitées dans le contexte de la prédiction des paramètres pétrophysiques les plus importants du domaine de réservoir, ont concerné essentiellement l'intégration entre la théorie des ondelettes et l’intelligence artificielle (ANN) et un réseau d'ondelettes (Saljooghi et al., 2014). Dans ce domaine d’étude, différentes ondelettes sont appliquées en tant que fonctions d'activation pour prédire la porosité à partir des données de diagraphies de puits. L’efficacité de ce type de réseau dans l'apprentissage de la fonction et l'estimation sont comparées à l’ANN. L’analyse des résultats obtenus, montrent bien une diminution des valeurs d'erreur d'estimation que représente sa capacité à améliorer l’efficacité de la fonction d'approximation et présente une excellente performance par rapport à l'apprentissage du réseau de neurones classique avec la fonction sigmoïde ou d'autres fonctions d'activation.
La deuxième méthode est celle de l’Estimation de la porosité des fractures par l'intelligence (http://hdl.handle.net/123456789/135). Dans ce travail, la caractérisation et la modélisation d'un réservoir fracturé ont été traitées à partir des données de diagraphies. Durant la phase d'apprentissage, le coefficient de corrélation obtenu entre la porosité de fractures estimée par RNA et celle calculée par diagraphie en phase d'apprentissage est 0.96 et dans la phase de test, il est de 0.88. Ce qui montre bien l’efficacité de la méthode utilisée ainsi que l’avancement des travaux dans l’intelligence artificielle.

Les travaux envisagés à l'avenir concernant l'intelligence artificielle et les géosciences :

L'un des problèmes numéro un de l'industrie pétrolière est la modélisation des réservoirs pétroliers. En effet, la connaissance des paramètres fondamentaux en pétrophysique d'un gisement pétrolier permet de mieux appréhender ses paramètres de stockage et les calculs de réserves récupérables afin de mieux assurer une meilleure rentabilité pendant le régime de production.

Le deuxième problème lié à cette modélisation est la prédiction des paramètres pétrophysiques par les nouvelles méthodes de diagraphie (résonance magnétique nucléaire : RMN), dans laquelle la perméabilité est très bien estimée par comparaison à celle de la carotte. A cet effet, La récupération des carottes au niveau des forages pétroliers n'est pas toujours intégrale, pour des raisons techniques et géologiques rencontrés lors des opérations du carottage.

L'approche des facies lithologiques dans les intervalles non carottés ainsi que les environnements de dépôts aux niveaux des formations géologiques, par l'utilisation de l'intelligence artificielle et des méthodes statistiques par les variables diagraphiques permet de palier ces inconvénients.

Toujours dans le cadre de ce travail, deux méthodes de modélisations ont été appliquées sur la base des données de diagraphies ; la première basée sur la technique des Réseaux de Neurones Artificiels et la deuxième est une technique statistique dite " Analyse des Clusters ", appliquée dans l'analyse des faciès géologiques. Dans ce contexte, des variables ont été utilisées comme paramètres d'entrée. Il s'agit de la radioactivité naturelle (Gamma Ray), de la densité de la roche (Rhob), de la résistivité (R_t), de la porosité (Φ) et de la saturation en eau de la roche (S_w). Tous ces paramètres ont été utilisés pour la prédiction des faciès ainsi que la porosité et perméabilité, représentant la sortie des modèles.

Deux articles soumis en Septembre 2014:

a) Les systèmes intelligents pour la prédiction de la porosité et la perméabilité à partir des données de la résonance magnétique nucléaire et des diagraphies conventionnelles enregistrées au niveau des puits de Hassi Messaoud (Journal of African Earth Sciences) ;
b) Analyse des Faciès et prédiction de la perméabilité et porosité à partir des données de diagraphie conventionnelles en utilisant les régressions non paramétriques en conjonction avec l'analyse multivariée et les réseaux de Neurone dans les réservoirs du Champ de Hassi R'Mel Sud (Arabian Journal of Geosciences).

Références

Abbaszadeh, M., Fujii, H., Fujimoto, F., (1995). Permeability Prediction by Hydraulic Flow Units-Theory and Applications, paper SPE 30158 prepared for presentation at the SPE Petrovietnam Conference held in Hochiminh, Vietnam, 1-3 March, 1995.

Abraham, A., and Baikunth Nath, (2000). Hybrid Intelligent Systems: A Review of a decade of Research. School of Computing and Information Technology, Faculty of Information Technology, Monash University, Autralia, Technical Report Series, 5/2000, pp. 1-55.

Aguilera, R., Aguilera, M.S., (2001). The Integration of Capillary Pressures and Pickett Plots for Determination of Flow Units and Reservoir Containers, paper SPE 71725 prepared for presentation at the 2001 SPE Annual and Technical Conference and Exhibition held in New Orleans, Louisiana, 30 September- 3 October.

Aggoun R.C., Djebbar T., Jalal, F., (2006). Characterization of flow units in shaly sand reservoirs-Hassi R'Mel Oil Rim, *v.*50 *fasc.*16 p., 211-226, 2006, Algeria.

Al-Ajmi, F., Holditch, S.A., (2000). Permeability Estimation Using Hydraulic Flow Units in a Central Arabia Reservoir, paper SPE 63254 prepared for presentation at the 2000 SPE Annual Technical Conference and Exhibition held in Dallas, Texas, 1-4 October, 2000.

Amaefule, J.O., Altunbay M., Tiab, D., Kersey, D.G., Keelan, D.K., 1993. Enhanced Reservoir Description: Using Core and Log Data to Identify Hydraulic (Flow) Units and Predict Permeability in Uncored Intervals/Wells, paper SPE 26436 prepared for presentation at the 68th Annual Technical Conference and Exhibiton of SPE held in Houston, Texas, 3-6 October.

Ammar, M.Y., (2007). Transfert, Dynamique des Fluides, Energétique & Procédés. Thèse Doctorat**.** Spécialité: Génie des procédés et de l'Environnement. ENSCI, Limoges.

Arbib, M. (ed), (1995). The Handbook of Brain Theory and Neural Networks, The MIT Press.

Baouche, R., Baddari K., Djeddi M., (2012). Analyse faciologiques des formations triasiques des puits de Hassi R'Mel à partir des diagraphies différées: reconnaissances des paléosols. Africa Géosciences Review. Vol. N°3, 191-213. France.

Bhatt, A., Helle, H.B., (2002). Committee neural networks for porosity and permeability prediction from well logs. Geophysical Prospecting 50, pp. 645–660.

Bear J., (1972). Dynamics of Fluids in Porous Media, Elsevier, New York. *Israel J. Technol.*, 10, 391-403.

Berenji H. R. and P. Khedkar, (1992). Learning and Tuning Fuzzy Logic Controllers through Reinforcements. IEEE Transactions on Neural Networks, Vol. 3, pp. 724-740.

Bishop, C.M., (1995).Training with noise is equivalent to Tikhonov regularization. Neural Computation 7 (1), 108-116.

Breiman, L., Friedman, J., Pregibon, D., Vardi, Y., Buja, A., Kass, R., Fowlkes, E., and Kettenring, J., (1985). Estimating Optimal Transformations for Multiple Regression and Correlation. Comments and Rejoinder," *Journal of the American Statistical Association*" (1985), 80, No.391, 580-619.

Carman, P. (1937), Fluid flow through a granular bed, Trans. Inst. Chem. Eng., 15, 150–167.

Carlos, F.H., (2004). The perfect permeability transforms using logs and cores. Proceedings of the SPE 89516 presented at the 2004 SPE Annual Technical Conference and Exhibition in Houston, Sept. 26-29, Texas, USA, 17-17.

Courel L., Ait Salem H., Ben Ismail H., El Mostaine M., Fekirine B., Kamoun F., Mami L., Oujidi M., Soussi M., (2000). An overview of the epicontinental, Triassic series of the Maghreb (NW Africa). *In: G. H. Bachmann and Ian Lerche (eds.).* An overview of the epicontinental Triassic series of Maghreb (N-W Africa). In Bachmann G.H & Lerche I. (eds.): Epicontinental Triassic. Zbl. Geol. Päont., tome 2, 9-10 (1998), 1145-1166.

Czogala E., and J. Leski., (2000). Neuro-Fuzzy Intelligent Systems, Studies in Fuzziness and Soft Computing, Springer Verlag, Germany.

D.W. Marquardt, 1963. An algorithm for least squares estimation of nonlinear inequalities SIAM J. Appl. Math., pp. 431–441.

Dorronsoro, J. Ginel, F. Sanchez, C. & C Cruz. (1997). 'Neural Fraud Detection in Credit Card Operations'. IEEE Transactions on Neural Networks, 8; 827-834.

Ebanks, W. J. Jr., Scheihing, M. H., and Atkinson, C. D. (1992). Flow units for reservoir characterization. In: *Development Geology Manual* . AAPG Methods in Exploration Series No. 10, Tulsa, 282-285.

Elgaghah, D. Tiab and S.O. Osisanya., (2001). Influence of stress on the characteristic of flow units in shaly formations, *Journal of Canadian Petroleum Technology*, 40 (4), 25-35.

Friedman, J. H., (1991). Rejoinder: Multivariate Adaptive Regression Splines. *Annual of Statistics*, 19, No.1, 123-141.

Friedman, J. and Roosen, C., (1995). An Introduction to Multivariate Adaptive Regression Splines. *Statistical Methods in Medical Research*, 4, No.3, 197-217.

Fullér, R., Hassanein H., Ali A.N., (1996). Neural fuzzy systems towards IMT-advanced networks. Åbo: Åbo akademi, xxvii, 275 p. ISBN 95-165-0624-0.

Hambalek, N., R. Gonzalez, (2003). Fuzzy logic applied to lithofacies and permeability forecasting. Proceedings of the SPE-81078 Latin American and Caribbean Petroleum Engineering Conference, Apr. 27-30, Trinidad, West Indies, 1-10.

Hamel A., Mania J., Perriaux J., (1988). Etude géologique des grès triasiques du gisement pétrolier de Hassi R'Mel (Algérie). Caractérisation, extension et milieu de dépôt. Volume 40, No. 4. *3.*

Hastie, T. and Tibshirani, R., (1990). Generalized Additive Models. Chapman & Hall/CRC, New York.

Hassoum, M. H. (1995). Fundamentals of artificial neural networks. Cambridge: MIT Press.

Hebb, D.O. (1949). The Organization of Behavior. New York: Wiley & Sons.

Hinton, D.E., The, S. J., Okihiro, M. S., Cooke, J.B., and Parker, L.M., (1992) Phenotypically altered hepatocyte populations in dietylnitrosamine-induced medaka liver carcinogenesis: Resistance, gouwth, and fate. Marine Environment. Res. 34, 1-5.

Hopfield J., (1992). Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences, vol.79, pp. 2554-2558, (1982).

Jang, R. (1992). Neuro-Fuzzy Modelling, Architectures, Analysis and Applications. PhD Thesis, University of California, Berkley.

Jensen, J. L. and Lake, L. W., (1985). Optimization of Regression-Based Porosity-Permeability Predictions," *CWLS 10th Symposium, Calgary, Alberta, Canada, September.*

Juang, F. C., T. Chin Lin., (1998). An On-Line Self Constructing Neural Fuzzy Inference Network and its applications. IEEE Transactions on Fuzzy Systems, Vol. 6, pp. 12-32.

D.W. Marquardt, 1963. An algorithm for least squares estimation of nonlinear inequalities SIAM J. Appl. Math., pp. 431–441

Kaufmann, A. and Gupta, M. M., (1985). Introduction to Fuzzy Arithmetic. New York: Van Nostrand.

Kasabov N and Qun Song (1999). Dynamic Evolving Fuzzy Neural Networks with 'm-out-of-n' Activation Nodes for On-line Adaptive Systems, Technical Report TR99/04, Department of information science, University of Otago.

Kohonen T., (1982). Self organized formation of topologically correct feature maps, Biol Cybernetics, Vol. 43, pp. 59-69.

Kozeny, J. (1927), Uber kapillare Leitung der Wasser in Boden, Sitzungs-ber. Akad. Wiss. Wien, 136, 271–306

L. A. Zadeh., (1965). Fuzzy Sets", Information and Control. Vol. 8, pp. 338-353.

Lake, L.W., Carroll, H.B., JR. (eds) 1986. *Reservoir Characterization.* Proceedings of a conference held in Dallas, Texas, 29 April-1 May 1985. Academic Press, 659p.

Lee, S. H., Arun, K., and Datta-Gupta, A. (2002). Electrofacies Characterization and 75 Permeability Predictions in Complex Reservoirs," *SPE Reservoir Evaluation & Engineering*. 5, No.3, 237-248

Le Cun Y., (1985). Une procédure d'apprentissage pour réseau à seuil asymétrique, Cognitiva 85, Paris.

Lim, J.-S., (2005). Reservoir properties determination using fuzzy logic and artificial neural network from well data in offshore Korea. *J. Petrol. Sci. Eng.*, 49, 182-192.

M. Figueiredo and F. Gomide; "Design of Fuzzy Systems Using Neuro-Fuzzy Networks", IEEE Transactions on Neural Networks, (1999), Vol. 10, no. 4, pp.815-827.

McCulloch, W. S., Pitts, W., (1943) *A Logical Calculus of the Ideas Immanent in Nervous Activity*, Bulletin of Mathematical Biophysics, vol. 5, pp. 115-133.

Mamdani, E. H., (1977). Applications of fuzzy logic to approximate reasoning using linguistic synthesis. *IEEE Transactions on Computers*, vol. 26, no. 12, pp. 1182–1191.

Matlab User's Guide, (2012a). Version a, Fuzzy Logic Toolbox. Math Works, USA, pp 235.

Mathisen, T., Lee, S. H., and Datta-Gupta, A. (2003). Improved Permeability Estimates in Carbonate Reservoirs Using Electrofacies Characterization: A Case Study of the North Robertson Unit, West Texas," *SPE Reservoir Evaluation & Engineering* (2003), 6, No.3, 176-184.

Mohaghegh, S., (2000). Virtual-intelligence applications in petroleum engineering: Part I. Artificial neural networks. *J. Pet. Technol.*, 52, 64-73.

Minsky M. et Papert S., (1988). Perceptrons: an introduction to computational geometry, MIT Press, expanded edition,

Nauck D., F. Klawon., 1997. R. Kruse, "Foundations of Neuro-Fuzzy Systems", J. Wiley & Sons.

Nauck D., R, Kurse., 1997. Neuro-Fuzzy Systems for Function Approximation, 4th International Workshop Fuzzy-Neuro Systems.

Nauck D., 1995. "Beyond Neuro-Fuzzy Systems: Perspectives and Directions". Proc. of the Third European Congress on Intelligent Techniques and Soft Computing (EUFIT'95), Aachen.

Nauck D., (1994). A Fuzzy Perceptron as a Generic Model for Neuro-Fuzzy Approaches. Proc. Fuzzy-Systems, 2nd GI-Workshop, Munich.

Nashawi, I.S., Malallah, A., (2010). Permeability Prediction from Wireline Well Logs Using Fuzzy Logic and Discriminant Analysis. SPE Asia Pacific Oil and Gas Conf and Exhibition, 8-20 October, Brisbane, Queensland, Australia.

Nerrand O., P. Roussel-Ragot, L. Personnaz, G. Dreyfus (1993). Neural Networks and Non-linear Adaptive Filtering: Unifying Concepts and New Algorithms. *Neural Computation* **5**, 165-197.

Nikravesh, M., Aminzadeh, F., (2003). Soft Computing and Intelligent Data Analysis in Oil Exploration. Part1: Introduction: Fundamentals of Soft Computing. Elsevier, Berkeley, USA, pp 744.

Petrolog, "*Advanced Log Analysis Software*", v .10, (2009). Crocker data processing. Petroleum House. WA 6102, Australia.

Pollard HB, Guy HR, Arispe N, (1992). De la Fuente M, Lee G, Rojas EM, Pollard JR, Srivastava M, Zhang-Keck ZY, Merezhinskaya N, et al. Calcium channel and membrane fusion activity of synexin and other members of the Annexin gene family. Biophys J. Apr; 62(1):15–18

Psichogios, D.C. and L.H. Ungar (1990). Nonlinear Internal Model Control Using. Neural. Networks, *Proceedings of the IEEE Fifth Int'l. Symposium on Intelligent Control,* September.

Refenes A.N., Zapranis. A.S., and Francis, G., (1994), "Stock Performance Modeling Using Neural Networks Comparative study with Regressive Models", Neural Networks, 7(2), 375-712.

Rosenblatt F**.,** The perceptron, (1958). A probalistic model for information storage and organization in the brain, Psycological Review, vol.65, 386-408,.

Rumelhart, G.E., Hinton, G.E. and Williams, R.J. (1986) "Learning Internal Representations by error propagation." In McClelland and Rumelhart.

Rumelhart D. E. et Mc Clelland. J. L., (1986). Parallel Distributed Processing: Exploration in the MicroStructure of Cognition, MIT Press, Cambridge,

Schimek, M. G. (2000). *Smoothing and regression: Approaches, computations, and application.* New York: Wiley

Sulzberger S., N. Tschichold e S. Vestli, (1993). FUN: Optimization of Fuzzy Rule Based Systems Using Neural Networks", Proceedings of IEEE Conference on Neural Networks, San Francisco, pp. 312-316.

Sejnowski T. J., and C. R. Rosenberg, (1986) 'Connectionist Models of Learning", in Perspectives in Memory Research and training, edited by M. S. Gazzaniga, (MIT Press).

Sugeno, M., (1977). Fuzzy measures and fuzzy integrals: A survey. In *Fuzzy Automata and Decision Processes*, M. M. Gupta, et al. eds., New York: North-Holland.

Sugeno, M., (1985a). An introductory survey of fuzzy control. Inf. Sci. 36, 59–83.

Sugeno, M., (1985b). Industrial Applications of Fuzzy Control. Elsevier Science Pub. Co, North Holland, Amsterdam.

Senergy, (2010). Interactive Petrophysics, Trial software. Version 3.6. Www.senergyltd.com. Schlumberger. Scotland.

Schlumberger, (1988). Basic Log Interpretation, Schlumberger Educational Services, U.S.A.

Sulzberger, S., N. Tschichold e S. Vestli, (1993). FUN: Optimization of Fuzzy Rule Based Systems Using Neural Networks", Proceedings of IEEE Conference on Neural Networks, San Francisco, pp. 312-316.

Takagi, T., Sugeno, M., (1985). Identification of systems and its application to modeling and control. IEEE Transaction on Systems, Man and Cybernetics 15, pp. 116–132.

Tano S., T. Oyama, T. Arnould, (1996). Deep Combination of Fuzzy Inference and Neural Network in Fuzzy Inference. Fuzzy Sets and Systems, 1996, Vol. 82(2), pp. 151-160.

Thibault, P.J. (1991). *Social semiotics as praxis: Text, social meaning making, and Nabokov's Ada*. Minneapolis: University of Minnesota Press.

Psichogios D.C. et Ungar L.H., (1991). Direct and indirect model based control using artificial neural networks, Ind. Eng. Chem. Res. 30 pp. 2564-2573.

Psichogios D.C. et Ungar L.H., (1992). A hybrid neural network-first principles approach to process modelling, AIChE J. 38, pp. 2269-2276.

Wang, Li-Xin and Mendel, J. M., (1991). Generating fuzzy rules by learning from examples. *Proceedings of the IEEE International Symposium on Intelligent Control*, Arlington, VA, pp. 263–268.

Wang, D. and Murphy, M., (2004). Estimating Optimal Transformations for Multiple Regressions Using the ACE Algorithm. *Journal of Data Science* (2004), 2, No.4, 329-346.

Wang L-X., HASSANEIN H., ALI A.N., (1994). Adaptive fuzzy systems and control: design and stability analysis. Englewood Cliffs, N.J: Prentice Hall, 1994, xxvii, 275 p. ISBN 978-013-1471-092.

Weiss, S. M. & Kulikowski, C. A. (1991). Computer systems that learn. San Mateo, CA: Morgan Kaufmann

Wendt, W. A., Sakurai, S., and Nelson, P. H., (1985). *Permeability Prediction from Well Logs Using Multiple Regression*, in *Reservoir Characterization*, L.W. Lake and H.B.J. Caroll (eds), Academic Press, New York.

Werbos, P.J. (1986) Generalized Information requirements of intelligence decision-making systems. SUGI 11 *Proceedings,* SAS Institute, Cary, NC. A version updated later in 1986 is available from the author.

Widrow B. and Hoff D.E., (1960). Adaptative switching circuits, IRE Western Electric Show and convention Record", vol.4, pp. 96-104.

Xue, G., Datta-Gupta, A., Valko, P., and Blasingame, T. (1997). Optimal Transformations for Multiple Regression: Application to Permeability Estimation from Well Logs," *SPE Reservoir Evaluation & Engineering*, 12, No.2, 85-94.

Zadeh, L. A., (1965). Fuzzy sets. *Information and Control*, vol, 8, pp, 338–353.

Zadeh, L. A., (1973). Outline of a new approach to analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, pp. 28–44.

Appendix A

SS: Grès (sandstones) ;

LS: Calcaire (limestones) ;

Dol: Dolomie (dolomites) ;

Y_i: Y valeur à $N^{ème}$ échantillon;

$\dot{Y}_i$: Déviation de l'écart quadratique moyen des échantillons ;

Y_{mean}: Valeur moyenne des échantillons;

SS_{Total}: Somme totale des erreurs quadratiques;

CorPor: Porosité carotte;

CPerm : permeabilité carotte;

Nomenclature

HRE: Hassi R'Mel Est ;

X, Y: coordonnées [UTM] ;

K: perméabilité [mD] ;

Φ : porosité [%] ;

Nphi: porosité neutron [%] ;

ΔT: temps de transit [μs/ft] ;

Sw: saturation en eau [%] ;

Rhob: densité globale [g/cm3] ;

GR: Rayons gamma [API] ;

RT: resistivity vraie [Ω m] ;

Nn: Réseaux de neurons (neural network) ;

Nf: Neuro-flou (neuro-Fuzzy) ;

Mlr: Régression multi linéaire (multi-linear regression) ;

in1mf1, in2mf1,…, in6mf6: les fonctions d'appartenance pour les données d'entrée: GR, ΔT, Rhob, RT, Nphi, Sw ;

out1mf1, out1mf2,…, out1mf5: les fonctions d'appartenance pour les données de sortie: Porosité carotte et perméabilité carotte ;

U: Unité supérieure du Trias avec les subdivisions U1, U2 et U3

M: Unité moyenne du Trias avec les subdivisions M1, M2 et M3

L: Unité inférieure du Trias avec les subdivisions L1, L2 et L3

ANNEXE 1

Article publié dans "Journal of Petroleum Science and engineering» 123 (2014), p.217-219, Elsevier. **"Neuro-Fuzzy system to predict Permeability and Porosity from well log data: a case study of Hassi R'Mel Gas Field, Algeria"**

Abstract.....140

1. Introduction.....140

2. Geological setting of the Hassi R'Mel field 143

3. Data sets and methodology 143

1.1 Fuzzy Logic.....144

1.2 Back propagation neural networks 144

1.3 Neuro Fuzzy Model.....144

4. Results and discussions 145

5. Conclusion.....145

6. References.....151

Liens pour quelques publidations

1. Tahar Aïfa, Rafik Baouche, Kamel Baddari. Neuro-fuzzy system to predict permeability and porosity from well log data: A case study of Hassi R'Mel gas field, Algeria**.** *Journal of Petroleum Science and Engineering, Volume 123, November 2014, Pages 217-229.* http://www.sciencedirect.com/science?_ob=ArticleListURL&_method=list&_ArticleListID=-705109804&_sort=r&_st=13&view=c&md5=debfb04f3507c2271a986b65d20a287b&searchtype=a

2. R. Baouche, A. Nedjari, S. El Aadj, R. Chaouchi, 2009. Facies Analysis of Triassic Formations of the Hassi R'Mel in Southern Algeria Using Well Logs: Recognition of Paleosols Using Log Analysis. *The Open Geology Journal*, 3, 39-57. 1874-2629/09 2009 Bentham Open**.** http: //creativecommons.org/licenses/by-nc/

3. R. Baouche, A. Nedjari, S. El Aadj, R. Chaouchi, 2009. Analysis and interpretation of environment sequence models of the Hassi R'Mel Field in Algeria. https://www.google.dz/?gws_rd=cr&ei=j2qgVOLSHaWjyAPV8YKgDA#q=SPE-117134-