République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Yahia Fares de Médéa
Faculté des Sciences et de la Technologie
Département de Génie des Procédés et Environnement
Laboratoire de Biomatériaux et Phénomènes de Transport (LBMPT)



THESE

Présentée par :
AMMI Yamina

En vue de l'obtention du diplôme de Doctorat en Génie des Procédés

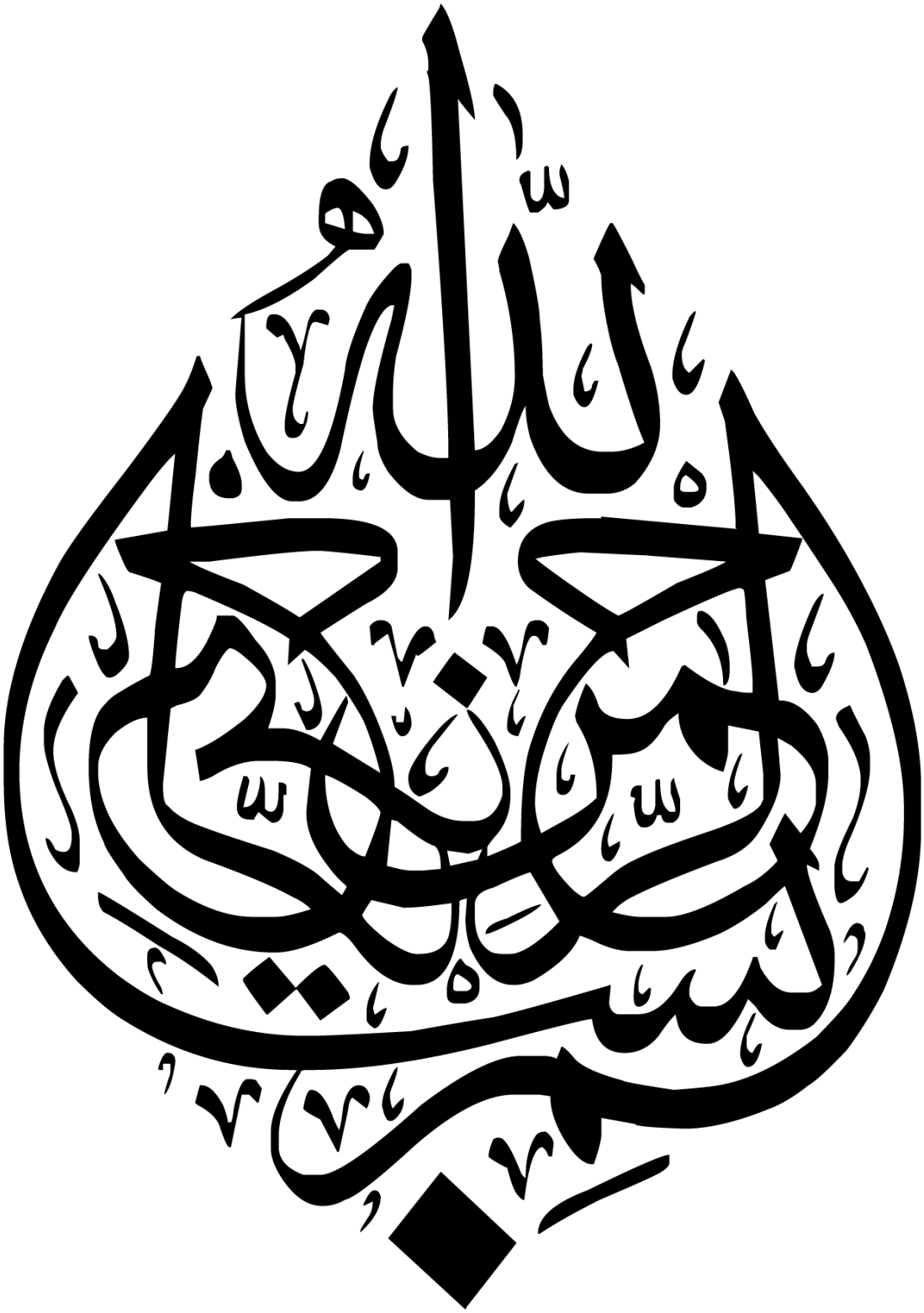
Option : Modélisation des phénomènes de transport en sciences de l'eau et d'environnement

Thème

MODELISATION DES PHENOMENES DE TRANSPORT LORS D'OSMOSE INVERSE

Soutenue le 21 mai 2016, devant le jury composé de :

M. Othmane BENKORTBI	Maître de Conférences (A), Université de Médéa	Président
M^{me} Rachida RIHANI	Maître de Conférences (A), USTHB, Alger	Examinatrice
M. Mounir MELLAL	Maître de Conférences (A), USTHB, Alger	Examinateur
M. Salah HANINI	Professeur, Université de Médéa	Directeur de thèse
M^{elle} Latifa KHAOUANE	Maître de Conférences (B), Université de Médéa	Invitée



Dédicaces

À ma grand-mère,
À mes chers parents,

À ma sœur Nadjet,
À mes frères Redoiane, Mohamed El Amine, Abd El Reouf, louanes, Soufiane et Abd El Bassite,

À toutes la famille AMMI

À toutes mes amies

À tous ceux qui m'ont soutenu de près ou de loin.

Yamina.

Remerciements et gratitude

Tout d'abord, je tiens à remercier ***ALLAH*** *Tout Puissant pour toutes les merveilles qu'il fait dans ma vie et pour m'avoir donné le courage, la force et la patience d'achever ce modeste travail.*

Je tiens à remercier mon directeur de thèse, Monsieur ***Salah HANINI****, Professeur à l'UYFM, qui m'a guidé avec grande patience tout au long de l'élaboration de ce travail, pour ses aides précieuses qui ont judicieusement éclairés mon chemin vers l'aboutissement et la réussite de la concrétisation de ma thèse de Doctorat.*

J'adresse également mes plus grands remerciements au, Docteur ***Latifa KHAOUANE****, Maître de Conférences (B) à l'UYFM, pour son aide, son accompagnement et son éclairage technique. Je lui dois les remerciements les plus sincères. Je ne peux conclure sans un mot de remerciements à ma Co-directrice de thèse de doctorat.*

Je profite de l'occasion de la présentation de ce travail pour exprimer mes vifs remerciements à Monsieur ***Othmane BENKORTBI****, Maître de Conférences (A) à Université de Médéa, pour avoir accepté de présider le jury de soutenance.*

Je tiens également à présenter ma profonde gratitude à Madame ***Rachida RIHANI****, Maître de Conférences (A) à l'USTHB et Monsieur* ***Mounir MELLAL****, Maître de Conférences (A) à l'USTHB, d'avoir accepté de juger ce modeste travail.*

Merci à tous ceux qui m'ont aidé de près ou de loin durant ces années de recherches et d'études.

Résumé

Modélisation des phénomènes de transport lors d'osmose inverse/nanofiltration

L'objectif assigné à ce travail est d'élaborer des modèles QSAR (réseaux de neurones et régressions linéaires multiples) pour la modélisation du processus de la rétention des composés organiques (neutres et ioniques) par les membranes d'osmose inverse et de nanofiltration.

Trois modèles neuronaux de type feedforward (QSAR-RN) caractérisés par une structure similaire (onze neurones pour QSAR-RN1et QSAR-RN2 et douze neurones pour QSAR-RN3 dans la couche d'entrée, une couche cachée et un neurone dans la couche de sortie) sont construits dans le but de prédire la rétention des composés organiques (neutres et ioniques). Un ensemble de 956 points de données pour QSAR-RN1 et 701 points de données pour QSAR-RN2 et QSAR-RN3 ont été utilisés pour tester les réseaux de neurones. 80%, 10% et 10% de l'ensemble des données ont été utilisées respectivement pour l'apprentissage, la validation et le test des trois modèles. Pour les modèles les plus prometteurs de réseau neuronal, les valeurs de la rétention prédite de la base de données totale ont été comparées aux valeurs de la rétention mesurées, de bonnes corrélations ont été trouvées (R = 0.9128 pour QSAR- RN1, R = 0.9419 pour QSAR-RN2 et R = 0.9527 pour QSAR-RN3). Les racines des erreurs quadratiques moyennes pour la base de données totale étaient de 11.2430% pour QSAR-RN1, de 9.0742% pour QSAR-RN2 et de 8.2047% pour QSAR-RN3. Les valeurs prédites par QSAR-RLM pour la rétention étaient en accord avec les valeurs expérimentales (les racines des erreurs quadratiques moyennes pour la base de données totale étaient de 19.7134 pour QSAR-RLM1, de 19.3437 pour QSAR- RLM2 et de 18.8702 pour QSAR-RLM3). En outre, la comparaison des résultats entre les modèles de QSAR-RN et les modèles QSAR-RLM a montrée la supériorité des modèles QSAR-RN. Une étude comparative entre les modèles QSAR développés dans ce travail et les modèles QSAR de Yangali-Quintanilla montre la priorité de nos modèles QSAR.

Mots-clés : Nanofiltration ; Osmose inverse ; Composés Organiques ; Rétention ; Réseaux de Neurones ; Régressions Linéaires Multiples ; QSAR ; Modélisation.

Summary

Modeling the phenomena of transport in reverse osmosis/nanofiltration

The objective set for this work is to develop QSAR models (neural networks and Multiple Linear Regressions) in modeling the rejection processes of organic compounds (neutral and ionic) by nanofiltration and reverse osmosis membranes.

Three feed-forward neural network (QSAR-NN) models characterized by a similar structure (eleven neurons for QSAR-NN1 and QSAR-NN2 and twelve neurons for QSAR-NN3 in the input layer, one hidden layer and one neuron in the output layer) are constructed with the aim of predicting the rejection of organic compounds (neutral and ionic). A set of 956 data points for QSAR-NN1 and 701 data points for QSAR-NN2 and QSAR-NN3 were used to test the neural networks. 80%, 10%, and 10% of the total data were used respectively for the training, the validation, and the test of the three models. For the most promising neural network models, the predicted rejection values of the total dataset were compared to measured rejections values, and good correlations were found (R =0.9128 for QSAR-NN1, R =0.9419 for QSAR-NN2, and R =0.9527 for QSAR-NN3). The root mean squared errors for the total dataset were 11.2430% for QSAR-NN1, 9.0742% for QSAR-NN2, and 8.2047% for QSAR-NN3. The predicted values by QSAR-MLR for the rejection were in agreement with the observed values (The root Mean Squared Error for the total database were 19.7134 for QSAR-RLM1, 19.3437 for QSAR-RLM2, and 18.8702 for QSAR-RLM3). Furthermore, the comparison of results between the QSAR-RN and QSAR-RLM models showed the superiority of the QSAR-RN models. A comparative stady between the QSAR models developed in this work and models QSAR of Yangali Quintanilla shows the priority of our QSAR models.

Keywords : Nanofiltration ; Reverse osmosis ; Organic compounds ; Rejection ; Neural networks ; Multiple Linear Regressions ; QSAR ; Modeling.

ملخص

نمذجة ظواهر النقل في التناظح العكسي/الترشيح من فئة النانو

الهدف المعين لهذا العمل هو تطوير نماذج (العلاقات الكمية البنية-النشاط : QSAR) "الشبكات العصبية والانحدار الخطي المتعدد : RLM" لنمذجة عمليات استبقاء المركبات العضوية (محايدة و أيونية) من خلال أغشية التناضح العكسي والترشيح من فئة النانو.

ثلاثة نماذج عصبية من نوع feedforward (QSAR-RN) تتميز ببنية مماثلة (إحدى عشر عصبون ل RN1 -QSAR و QSAR-RN2 و اثني عشر عصبون لQSAR-RN3 في طبقة المدخلات، طبقة مخفية واحدة و عصبون في طبقة الإخراج) شيدت لغرض التنبؤ باستبقاء المركبات العضوية (محايدة و أيونية) من خلال التناضح العكسي والترشيح من فئة النانو. و استخدمت مجموعة من 956 نقاط البيانات ل QSAR-RN1 و701 نقاط البيانات ل -QSAR RN2 و3 QSAR-RN لاختبار الشبكات العصبية. واستخدم 80٪، 10٪ و 10٪ من إجمالي البيانات على التوالي للتدريب، والتحقق من صحة، واختبار النماذج الثلاثة . لنماذج الشبكات العصبية الواعدة، تمت مقارنة قيم الاستبقاء المتوقعة بقيم الاستبقاء التجريبية لمجموع نقاط البيانات ووجدت ارتباطات جيدة ($R = 0.9128$ لQSAR-RN1 ، $R = 0.9419$ ل RN2 -QSAR و $R = 0.9527$ ل QSAR-RN3). جذر الأخطاء التربيعية المتوسطة لمجموعة الاختبار كانت 11.2430٪ ل QSAR-RN1، 9.0742٪ لQSAR-RN1 و8.2047٪ ل QSAR-RN3. القيم التي تنبأ بها QSAR-RLM من أجل الاحتفاظ في اتفاق مع القيم التجريبية (جذر الأخطاء التربيعية المتوسطة لقاعدة البيانات الإجمالية كانت 19.7134 ل QSAR-RLM1، 19.3437 ل QSAR-RLM2 و 18.8702 ل QSAR-RLM3). وعلاوة على ذلك، مقارنة النتائج بين نماذج QSAR-RN ونماذج QSAR-RLM أظهرت تفوق نماذج QSAR-RN. دراسة مقارنة بين نماذج QSAR المتقدمة في هذا العمل ونماذج QSAR ل Yangali Quintanilla تظهر أولوية نماذج QSAR لدينا.

كلمات مفاتيح : التصفية من درجة النانو ، التناضح العكسي ، الجزيئات العضوية ، قدرة الاستبقاء ، الشبكة العصبية ، (الانحدار الخطي المتعدد : RLM) ، (العلاقات الكمية البنية-النشاط : QSAR) ، نمذجة .

Table des matières

Abréviations i
Nomenclature et symboles ii
Liste des tableaux v
Liste des figures vii

Introduction générale 1

Chapitre I. Description de la technologie membranaire

I.1. Introduction 05
I.2. Généralités 05
I.2.1. Définition de la membrane 05
I.2.2. Classification des membranes 06
I.2.2.1. Classification d'après la morphologie 06
I.2.2.2. Classification d'après la nature chimique 07
I.2.2.3. Classification d'après la géométrie de préparation des membranes 08
I.2.2.4. Classification d'après le mécanisme de séparation 09
I.2.3. Osmose inverse et nanofiltration 11
I.2.4. Sélectivité d'une membrane 12
I.2.5. Influence des paramètres opératoires 12
I.2.5.1. Pression 12
I.2.5.2. Température 13
I.2.5.3. Taux de récupération "Taux de conversion" 13
I.3. Interactions et propriétés de séparation 13
I.3.1. Encombrement stériques (effet du tamisage) 14
I.3.1.1. État physique de la membrane 15
I.3.1.2. Taille moléculaire 17
I.3.2. Répulsion électrostatique (effet de charge) 19
I.3.2.1. Charge de la surface 19
I.3.2.2. Charge de soluté 20
I.3.2.3. Influence du pH 20
I.3.3. Interaction hydrophobe/adsorption 21
I.3.3.1. Hydrophobicité de la membrane 21
I.3.3.2. Hydrophobicité des solutés 21
I.4. Applications des procédés d'OI et de NF 22
I.5. Avantages et inconvénients des procédés d'OI et de NF 24
I.6. Conclusion 24

Chapitre II. Revue critique sur la modélisation de la rétention des MO lors d'OI/NF par le modèles QSAR

II.1. Introduction 25
II.2. Méthodes d'apprentissages statistiques 25
II.2.1. Régression linéaire multiple (RLM) 26
II.2.2. Réseaux de neurones artificiels (RNA) 26
II.2.2.1. Neurone biologique 26
II.2.2.2. Neurone formel 27
II.2.2.3. Architecture d'un réseau de neurones 28
II.2.2.4. Perceptron multicouches 29
II.3. Revue des travaux de la rétention des MO lors d'OI/NF 30
II.4. Revue des travaux de la prédiction de la rétention par les modèles QSAR 32
II.5. Conclusion 37

Chapitre III. Matériels et méthodes

III.1. Introduction 39
III.2. Collecte des données, prétraitement et analyse 39
III.3. Développement des modèles QSAR 45
III.3.1. Réseaux de neurones artificiels 45
III.3.2. Régression linéaire multiple 47
III.4. Analyse de sensibilité par la méthode de "poids" 47
III.5. Conclusion 50

Chapitre IV. Résultats et discussion

IV.1. Introduction 51
IV.2. Modèles QSAR-RN 51
IV.2.1. Performances des modèles QSAR-RN 51
IV.2.2. Expressions mathématiques des modèles neuronaux optimisés 63
IV.2.3. Analyse de sensibilité par la méthode de "poids" 68
IV.2.4. Comparaison entre la rétention expérimentale et calculée par QSAR-RN3 71
IV.3. Modèles QSAR- RLM 82
IV.4. Comparaison entre les modèles QSAR 94
IV.4.1. Comparaison entre les modèles QSAR-RN et QSAR-RLM 94
IV.4.2. Comparaison entre nos modèles QSAR et les modèles QSAR de la littérature 96
IV.5. Rétention des composés organiques en fonction du pH 100
IV.6. Conclusion 105

Conclusion générale et perspectives 107
Références bibliographiques 110
Annexe A 117
Annexe B 126
Annexe C 129

Abréviations

Symbole	Signification
ART	Adaptive Resonance Theory
BD	Base de Données
BFGS	Broyden–Fletcher–Goldfarb–Shanno
DS	Déviation Standard
FFNN	Feed Forward Neuronal Network
HB	Hydrophobe
HL	Hydrophile
IR	Importance Relative
LM	Levenberg–Marquardt
MF	Microfiltration
MO	Molécule Organique
NF	Nanofiltration
OI	Osmose Inverse
PES	Polyethersulfone
PMC	Perceptron Multi-Couches
PS	Polysulfone
QSAR	Quantitative Structure Activity Relationship
QSPR	Quantitative Structure Property Relationship
RBF	Radial Basis Function Network
RLM	Régression Linéaire Multiple
RN	Réseaux de Neurones
RNA	Réseaux de Neurones Artificiels
RNFF	Réseaux de Neurones Feedforward
SCG	Scaled Conjugate Gradient
TFC	Thin-Film Composite
3D	Trois Dimension
UF	Ultrafiltration

Nomenclatures et symboles

Lecture	Désignation	Unité
$\mathbf{C_{SP}}$	Concentration d'un soluté dans le perméat	(kg/m^3)
$\mathbf{C_{SA}}$	Concentration d'un soluté dans l'alimentation	(kg/m^3)
C	Concentration de la solution	$(mol.\,m^{-3})$
$\mathbf{C_0}$	Concentrations des molécules dans la couche de n-octanol	$(mol.\,m^{-3})$
$\mathbf{C_w}$	Concentrations des molécules dans la couche de l'eau	$(mol.\,m^{-3})$
DCO	Demande Chimique en Oxygène	$(mg.L^{-1})$
Depth	Profondeur moléculaire	(nm)
EAM	Erreur Absolue Moyenne	(-)
EQM	Erreur Quadratique Moyenne	(-)
EqWidth	Largeur équivalente moléculaire	(nm)
ESP	Erreur Standard de Prédiction	(%)
Length	Longueur moléculaire	(nm)
LogD	Logarithme du coefficient de distribution	(-)
$\mathbf{Logk_{ow}}$ ou **LogP**	Logarithme du coefficient de partage octanol-eau	(-)
MV	Volume molaire	$cm^3/mole$
MW	Masse molaire	$(g.mol^{-1})$
Ptm	Pression transmembranaire	(Pa)
$\mathbf{Q_A}$	Débit volumique d'Alimentation	(m^3/h)
$\mathbf{Q_P}$	Débit volumique de Permeat	(m^3/h)
REQM	Racine de l'Erreur Quadratique Moyenne	(-)
S	Surface	(nm^2)
$\mathbf{S_{max}}$	Surface maximale	(nm^2)
SR	Rejet de Sel	(-)
T	Température	(K)
$\mathbf{T_r}$	Taux de rétention	(%)
Width	Largeur	(nm)
y	Taux de récupération (Taux de conversion)	(%)
Z	Potentiel zêta	(Mv)

Lettres Grecques

Lecture	Désignation	Unité
ΔΠ	Différence de pression osmotique de part et d'autre de la membrane	(Pa)
α	Pente de la droite du meilleur ajustement linéaire	(-)
β	Ordonnée à l'origine de la droite du meilleur ajustement linéaire	(-)
Π	Pression osmotique	(Pa)

Exposants

Lecture	Désignation
cal	Calculée
exp	Expérimentale
H	Cachée (Hidden)
I	Entrée (Input)
O	Sortie (Output)

Indices

Lecture	**Désignation**
I	Élément i d'un vecteur
i,exp	Élément i d'un vecteur de valeurs expérimentales
i,cal	Élément i d'un vecteur de valeurs calculées
Ij	Croisement de la ligne i avec la colonne j d'une matrice
j	Élément j d'un vecteur
max	Maximum
min	Minimum
$\bar{x}$	Moyenne

Liste de tableaux

Chapitre I. Description de la technologie membranaire

Tableau I.1. Applications des procédés l'OI et NF. 23
Tableau I.2. Avantages et inconvénients de procédés d'OI et NF. 24

Chapitre II. Revue critique de la littérature sur la modélisation de la rétention des MO lors d'OI/NF par RNA

Tableau II.1. Fonctions de transfert usuelles, *f(x)*. 28
Tableau II.2. Vue d'ensemble des travaux antérieurs sur les modèles de prédiction de la rétention des molécules organiques lors d'osmose inverse et de nanofiltration. 36

Chapitre III. Matériels et méthodes

Tableau III.1. Taille de la base de données et le nombre de composés organiques utilisés dans ce travail. 39
Tableau III.2. Logiciels utilisés pour calculer $LogK_{ow}$, *LogD* et le moment dipolaire. 40
Tableau III.3. Entrées et sortie utilisées pour le développement des modèles QSAR (QSAR1, QSAR 2 et QSAR 3). 41
Tableau III.4. Analyse statistique des données d'entrée et de sortie (BD1). 42
Tableau III.5. Analyse statistique des données d'entrée et de sortie (BD2). 42
Tableau III.6. Analyse statistique des données d'entrée et de sortie (BD3). 43

Chapitre IV. Résultats et discussion

Tableau IV.1. Structure des modèles RN optimisés. 51
Tableau IV.2. Vecteurs de régression linéaire issue de l'équation linéaire : [équation linéaire : $y^{cal} = \alpha y^{exp} + \beta$, $\boldsymbol{\alpha}$(pente), $\boldsymbol{\beta}$(ordonnée à l'origine), $\boldsymbol{R}$(coefficient de corrélation)] et la racine de l'erreur quadratique moyenne (***REQM***). 60
Tableau IV.3. Poids et biais du QSAR-RN1 optimal. 65
Tableau IV.4. Poids et biais du QSAR-RN2 optimal. 66
Tableau IV.5. Poids et biais du QSAR-RN3 optimal. 67
Tableau IV.6. Vecteurs de régression linéaire issue de l'équation linéaire : [équation linéaire : $y^{cal} = \alpha y^{exp} + \beta$, $\boldsymbol{\alpha}$(pente), $\boldsymbol{\beta}$(ordonnée à l'origine), $\boldsymbol{R}$(coefficient de corrélation)], l'erreur absolue moyenne ***(EAM)***, la racine de l'erreur quadratique moyenne (***REQM***) et l'erreur standard de prédiction ***(ESP)***. 91
Tableau IV.7. Coefficients du modèle estimés par analyse de régression linéaire multiple (QSAR-RLM1). 92
Tableau IV.8. Coefficients du modèle estimés par analyse de régression linéaire multiple (QSAR-RLM2). 92
Tableau IV.9. Coefficients du modèle estimés par analyse de régression linéaire 93

multiple (QSAR-RLM3).

Annexe A

Tableau 1 Liste de 82 molécules organiques utilisées pour développer les modèles QSAR (RN et RLM). 117

Annexe C

Tableau 1 Matrice des coefficients de corrélation pour la première base de données (BD1). 129

Tableau 2 Matrice des coefficients de corrélation pour la deuxième base de données (BD2). 130

Tableau 3 Matrice des coefficients de corrélation pour la troisième base de données (BD3). 131

Liste de figures

Chapitre I. Description de la technologie membranaire

Figure I.1.	Membrane sélective.	06
Figure I.2.	Représentation schématique des différentes structures membranaires.	06
Figure I.3.	Représentation schématique d'une membrane multicouche.	07
Figure I.4.	Représentation schématique d'un module tubulaire.	08
Figure I.5.	Représentation schématique d'un module à fibres creuses.	09
Figure I.6.	Représentation schématique d'un module spiralé.	09
Figure I.7.	Gamme de rétention de différentes particules par les procédés membranaires.	10
Figure I.8.	Principe des phénomènes d'osmose inverse.	11
Figure I.9.	Molécule de polyamide et ses équilibres acido-basiques.	14
Figure I.10.	Illustration du principe de détermination du seuil de coupure d'une membrane, la rétention en fonction de la masse molaire (défini ici pour un taux de rejet de 0.90).	16
Figure I.11.	Calcul des dimensions de l'octanol (length, width et depth) de la structure 3D.	19
Figure I.12.	Mesure d'angle de contact par la méthode de la goutte posée.	21

Chapitre II. Revue critique de la littérature sur la modélisation de la rétention des MO lors d'OI/NF par RNA

Figure II.1.	Neurone biologique.	27
Figure II.2.	Neurone formel.	27
Figure II.3.	Structure d'un réseau de neurones non bouclé.	30
Figure II.4.	Nombre de publications traitant la rétention des molécules organiques par (a) : nanofiltration (b) osmose inverse, (ScienceDirect le 05/04/2016).	31
Figure II.5.	Nombre de publications traitant la modélisation de la rétention des molécules organiques lors d'osmose inverse et de nanofiltration.	37

Chapitre III. Matériels et méthodes

Figure III.1.	Rétentions expérimentales en fonction de poids moléculaires : (a) BD1, (b) BD2, (c) BD3.	44
Figure III.2.	Représentation des pourcentages de division de la base de données.	45

Chapitre IV. Résultats et discussion

Figure IV.1.	Comparaison entre les valeurs expérimentales et les valeurs calculées pour QSAR-RN1 : (a) la phase d'apprentissage, (b) la phase de validation, (c) la phase de test, (d) la phase totale.	54
Figure IV.2.	Comparaison entre les valeurs expérimentales et les valeurs calculées pour QSAR-RN2 : (a) la phase d'apprentissage, (b) la phase de	56

validation, (c) la phase de test, (d) la phase totale.

Figure IV.3. Comparaison entre les valeurs expérimentales et les valeurs calculées pour QSAR-RN3 : (a) la phase d'apprentissage, (b) la phase de validation, (c) la phase de test, (d) la phase totale. 59

Figure IV.4. Performances des modèles neuronaux (RN*1*, RN*2* et RN*3*) pour : (a) la phase d'apprentissage, (b) la phase de validation, (c) la phase de test, (d) la phase totale. 62

Figure IV.5. Représentation schématique du QSAR-RN1 optimisés. 63

Figure IV.6. Architecture finale du réseau de neurone optimal de type feedforward pour la prédiction de la rétention des molécules organiques (neutres et ioniques) par les membranes d'OI/NF. 68

Figure IV.7. Présentation des pourcentages de l'importance relative (*IR*) de chaque variable d'entrée sur la rétention : (a) QSAR-RN1, (b) QSAR-RN2, (c) QSAR-RN3. 70

Figure IV.8. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de la masse molaire des MO. 71

Figure IV.9. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de *logD* des MO. 72

Figure IV.10. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de moment dipolaire des MO. 73

Figure IV.11. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de la longueur des MO. 74

Figure IV.12. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de la largeur équivalente des MO. 75

Figure IV.13. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de MWCO des MO. 76

Figure IV.14. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de SR($MgSO_4$) des MO. 77

Figure IV.15. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de potentiel zêta des MO. 78

Figure IV.16. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction d'angle de contact des MO. 79

Figure IV.17. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de pH des MO. 80

Figure IV.18. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de pression des MO. 81

Figure IV.19. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de taux de récupération des MO. 82

Figure IV.20. Modèle QSAR-RLM1 : les valeurs calculées vs. Les valeurs expérimentales pour les ensembles de données : (a) phase d'apprentissage, (b) phase de validation, (c) phase de test, (d) totale. 85

Figure IV.21. Modèle QSAR-RLM2 : les valeurs calculées vs. Les valeurs expérimentales pour les ensembles de données : (a) phase d'apprentissage, (b) phase de validation, (c) phase de test, (d) totale. 87

Figure IV.22. Modèle QSAR-RLM3 : les valeurs calculées vs. Les valeurs expérimentales pour les ensembles de données : (a) phase d'apprentissage, (b) phase de validation, (c) phase de test, (d) totale. 90

Figure IV.23. Comparaison des résultats de la base totale e entre le modèle de QSAR-RN et le modèle QSAR-RLM : (a) EAM, (b) REQM, (c) ESP. 96

Figure IV.24. Comparaison entre nos modèles QSAR et les modèles QSAR de la 98

littérature "MWCO" est indicateur de porosité" : (a) EAM, (b) REQM, (c) ESP.

Figure IV.25. Comparaison entre nos modèles QSAR et les modèles QSAR de la littérature "SR($MgSO_4$)" est un indicateur de porosité" : (a) EAM, (b) REQM, (c) ESP. 100

Figure IV.26. Comparaison entre la rétention expérimentale et calculée des composés organiques chargés positivement en fonction du pH de l'eau d'alimentation avec Trisep TS80 (Clenbuterol, (a)) et Desal HL (Salbutamol, (b)) pour chaque modèle (RN1, RN2 et RN3). 101

Figure IV.27. Comparaison entre la rétention expérimentale et calculée des composés organiques neutres en fonction du pH de l'eau d'alimentation avec Trisep TS80 (Cyclophosphamide, (a)) et Desal HL (Pentoxifylline, (b)) pour chaque modèle (RN1, RN2 et RN3). 103

Figure IV.28. Comparaison entre la rétention expérimentale et calculée des composés organiques chargés négativement en fonction du pH de l'eau d'alimentation avec Trisep TS80 (Acide clofibrique, (a)) et Desal HL (Fenoprofen, (b)) pour chaque modèle (RN1, RN2 et RN3) 104

Annexe B

Figure 1. Capture d'écran du site intitulé Kowwin, epa. 126

Figure 2. Capture d'écran du site intitulé Chemspider. 126

Figure 3. Capture d'écran des logiciels : (a) ChemBioDraw et (b) ChemBio3D. 127

Figure 4. Capture d'écran de logiciel STATISTICA. 128

Figure 5. Capture d'écran de logiciel MATLAB 2010. 128

Introduction générale

L'eau est un élément vital, précieux et rare. Son rôle économique et social est très important (Metaiche 2006). La quantité d'eau est abondante sur notre planète, elle représente 1380 millions de km^3. Elle est essentiellement constituée d'eau de mer (97.2%) et de glace (2.15%) inutilisables directement. Sur les 0.65% d'eau douce restante, l'eau douce, facilement disponible (lacs, fleuves, certaines eaux souterraines), ne représente que 0.07% de la ressource totale soit environ un million de km^3, mais, la répartition de cette eau est très inégale d'un point de vue géographique. En effet, dix pays se partagent 60% des réserves d'eau douce dont le Brésil qui à lui seul en possède 12% et vingt-huit autres principalement en Afrique et au Moyen-Orient sont au contraire confrontés à une pénurie chronique d'eau douce (Triki 2014). Les ressources naturelles (conventionnelles) en eau douce sont limitées (et épuisées pour certaines régions surtout dans les pays à climat aride et semi-aride). Elles sont inexistantes ou insuffisantes vis-à-vis de la croissance démographique ou de la production industrielle, et le recours vers les ressources non conventionnelles est primordial, dont la réutilisation d'eau et le dessalement de l'eau de mer et de l'eau saumâtre constituent une solution adéquate du problème de l'insuffisance des ressources naturelles. Le recyclage des eaux usées apparaît dés à présent comme une nécessité (Roorda 2004 ; Petrinic et al. 2015). La réutilisation des eaux usées est donc « *une action volontaire et planifiée qui permet de fournir une source supplémentaire d'eau pour différents usages comblant les déficits hydriques tout en protégeant le milieu environnant* » (Ecosse 2001). Dans la plupart des pays développés, la majeure partie des eaux usées et des effluents industriels est traitée afin de protéger la qualité du milieu récepteur.

Les procédés de séparation membranaires regroupent un grand nombre de techniques permettant de réaliser des séparations en phase liquide ou en phase gazeuse sous l'action de diverses forces de transfert (mécanique, électrique ou chimique). Dans le cas de la filtration en phase liquide réalisée au moyen de procédés baro-membranaires, c'est-à-dire utilisant un gradient de pression comme force motrice, on distingue quatre techniques répertoriées sous les noms de microfiltration (MF), ultrafiltration (UF), nanofiltration (NF) et osmose inverse (OI). La micro-, l'ultra- et la nanofiltration mettent en jeu des membranes possédant une structure poreuse tandis que les membranes d'osmose inverse sont des matériaux denses. La microfiltration est utilisée pour séparer des espèces dont la taille est de l'ordre du micromètre (cellule animale, levures, bactéries…), l'ultrafiltration pour séparer des espèces de taille comprise entre ~ 2 et 100 nm (peptides, protéines, virus…) et la nanofiltration pour traiter les espèces de taille inférieure à 2 nm (sucres, colorants, sels…). Tout comme la nanofiltration,

l'osmose inverse, permet de filtrer des solutions contenant des espèces de très faibles masses molaires. Elle est généralement utilisée lorsqu'une rétention totale des solutés est souhaitée (Bouranene 2008).

La littérature met en lumière plusieurs points. D'une part, la nanofiltration et l'osmose inverse sont des procédés de plus en plus répandues et de nombreuses études indiquent qu'il est possible de traiter différents types d'effluents en retenant à la fois les molécules inorganiques et organiques. La compréhension fondamentale et quantitative des mécanismes de transfert et des performances des membranes d'osmose inverse et de nanofiltration sont nécessaires à la sélection, au dimensionnement, au fonctionnement et au contrôle des systèmes industriels. D'autre part, le rôle complexe et encore mal cerné de nombreuses interactions qui peuvent s'établir entre les solutés, l'eau et la membrane. De nombreux modèles de transferts existent (thermodynamique irréversible, solubilisation-diffusion ou de pores), mais ils ne décrivent ou n'expliquent pas toujours correctement le procédé et les résultats obtenus dans le cas du transfert de molécules organiques nécessitent des modifications importantes. Le comportement du procédé d'osmose inverse et de nanofiltration pour les solutions complexe plus difficile à décrire et prédire, du fait du grand nombre de mécanismes mis en jeu. De plus, peu de travaux font le lien entre les mécanismes révélés par l'étude de solutions-modèles et les performances de traitement de rejets industriels. Ils ne se concentrent que sur l'un ou l'autre de ces aspects.

Pour surmonter ces problèmes, des tentatives ont été faites pour proposer des approches de modélisation permettant de diminuer le nombre d'expériences à effectuer pour déterminer les paramètres optimaux pour le traitement d'un effluent industriel. Si elle représente bien le procédé, les influences des paramètres opératoires pourront être évaluées sans aucune expérience et n'auront qu'à être validées sur le site. De plus, la modélisation peut également permettre de valider les hypothèses simplificatrices émises lors de la formulation mathématique des phénomènes de transfert. Pour cela, il faut chercher d'autres méthodes de modélisation telles que les sciences cognitives.

La révolution de l'informatique et le développement de nouvelles méthodes de calcul numérique basé sur l'intelligence artificielle et les sciences cognitives ont donné naissance durant les dernières décennies à des nouvelles méthodes de calcul et de modélisation telles que les réseaux de neurones artificiels "RNA", les algorithmes génétiques et la logique floue. L'efficacité, la fiabilité et la robustesse de ces méthodes par rapport aux difficultés de

l'utilisation des approches classiques, ont été démontrées par leurs nombreuses applications dans le domaine du génie des procédés (Bouamar et Ladjal 2007). Le principal avantage des RNA, comparé aux modèles de connaissance classique, est leur capacité à modéliser ces procédés sans avoir une connaissance approfondie de la phénoménologie mise en jeu. Un autre avantage important des RNA, comparé à d'autres approximateurs universels tels que les régressions linéaires ou polynomiales, est qu'ils sont des approximateurs non-linéaires parcimonieux, permettant une modélisation de procédés non-linéaires quel que soit le nombre de paramètres indépendants utilisés (Dreyfus et al. 2002).

Bien que les réseaux de neurones artificiels (RNA) et les régressions linéaires multiples (RLM) soient très appliqués dans le domaine du génie des procédés en raison de la précision des résultats qu'ils produisent et de la simplicité relative de la mise en œuvre. Peu de travaux de recherche utilisent les RNA et RLM comme moyen de calcul des phénomènes de transports par d'osmose inverse et de nanofiltration lors de la rétention des molécules organiques (Libotean et al. 2008 ; Yangali-Quintanilla et al. 2008, 2009b et 2010b ; Shahmansouri et al. 2013 ; Sadmani et al. 2014a et 2014b).

L'objectif de ce travail est l'application des RN de type feedforward et RLM pour modéliser la rétention des molécules organiques (neutres et ioniques) par les membranes d'OI et de NF. Ainsi, nous avons organisé cette thèse en quatre chapitres :

- Le premier chapitre est consacré à la description de la technologie membranaire, on donnera dans un premier temps un aperçu sur les généralités sur les membranes, ensuite les interactions et les propriétés de séparation, puis les applications et les avantages d'OI/NF.
- Le deuxième chapitre est dédié à la présentation de la revue critique sur la modélisation de la rétention des molécules organiques lors d'osmose inverse et de nanofiltration par les modèles QSAR (RNA et RLM).
- Nous présentons dans le troisième chapitre les matériels et les différentes procédures de modélisations utilisées. Il comptera trois axes : dans le premier axe, on examinera les données collectées, prétraitement et analyse. Le second axe sera consacré aux techniques les plus courantes pour établir des modèles QSAR : la régression linéaire multiple et les réseaux de neurones artificiels (RNA) et le développement des modèles QSAR. Le troisième et dernier axe traitera l'analyse de sensibilité basée la méthode de "poids".

- Dans le quatrième chapitre, les différents résultats prometteurs obtenus sont discutés, il est organisé en quatre parties : dans une première partie, on présentera les performances des modèles QSAR-RN, ensuite l'extraction de formules mathématiques à partir de ces modèles, par la suite, on présentera les résultats de l'analyse de sensibilité par la méthode de "poids", une comparaison entre la rétention expérimentale et calculée par QSAR-RN3 est ensuite présentée. Puis on y discutera sur les modèles QSAR- RLM dans la deuxième partie. Dans la troisième partie, on comparera les modèles QSAR-RN avec QSAR- RLM et avec les modèles QSAR de la littérature. La quatrième partie est consacrée à l'étude de la rétention de composés organiques en fonction du pH.

Les conclusions tirées de ce travail ainsi que les études qui peuvent être envisagées à l'avenir sont présentées à la fin de cette thèse.

Chapitre I

Description de la technologie membranaire

I.1. Introduction

La filtration membranaire est de plus en plus utilisée comme procédé de séparation dans de nombreux domaines notamment dans le cycle de l'eau (potabilisation de l'eau, traitement des effluents, réutilisation de l'eau, adoucissement, dessalement,…). La filtration membranaire est basée sur l'application d'une différence de pression qui permet le transfert du solvant à travers une membrane dont la taille des pores assure la rétention de solutés. Ces opérations, classées selon une taille des pores décroissantes, sont : la microfiltration (MF) l'ultrafiltration (UF), la nanofiltration (NF) et l'osmose inverse (OI).

Les procédés de séparation membranaire ont pris leur essor dans les années 70 en particulier grâce au développement des matériaux et des techniques de fabrication des membranes qui offrent une meilleure rentabilité dans tous les termes de consommation énergétique par apport à des techniques de séparation conventionnelles telle que la distillation, des conditions de fonctionnement tolérables, un traitement souhaitable, et qu'elles favorisent le développement dans toutes les branches de l'industrie (Bouranene et al. 2008). Parmi eux, l'osmose inverse et nanofiltration sont utilisées en traitement de séparation des différentes molécules organiques et inorganiques qui permettent une séparation conforme aux normes actuellement en vigueur.

Ce chapitre s'attache à faire les points sur les généralités des membranes d'osmose inverse et de nanofiltration, les interactions et les propriétés de séparation des molécules organiques, les applications des procédés d'OI/NF, les avantages et les inconvénients des procédés d'OI/NF.

I.2. Généralités

I.2.1. Définition de la membrane

Une membrane peut être définie comme étant une couche mince de matière, permettant l'arrêt ou le passage sélectif des substances dissoutes ou non, sous l'action d'une force motrice de transfert (Figure I.1). Les critères de séparation des particules, des molécules et/ou des ions peuvent être (Annane 2015) :

- La dimension et la forme ;
- La nature chimique ;
- L'état physique ;
- La charge électrique.

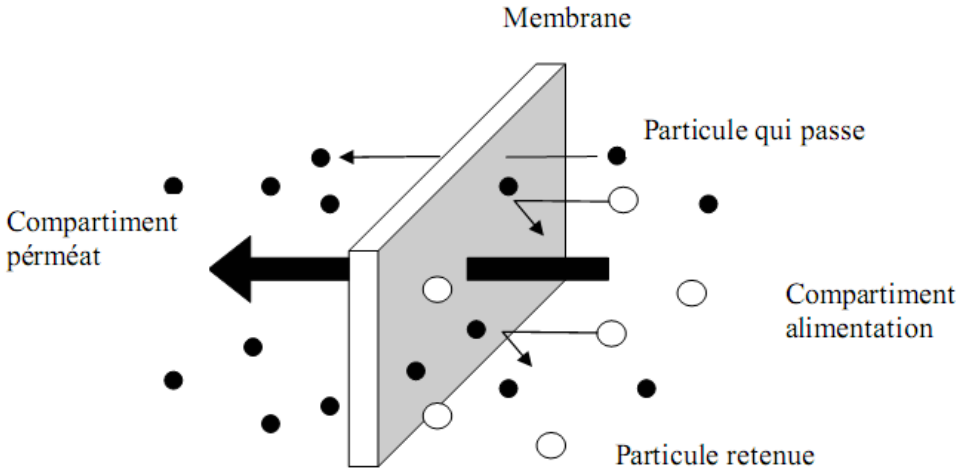


Figure I.1. Membrane sélective (Annane 2015).

I.2.2. Classification des membranes

I.2.2.1. Classification d'après la morphologie

Les membranes peuvent être également classées en fonction de la structure des différents matériaux, on distingue :

- **Membranes à structure symétrique "isotrope" :** ce sont des membranes denses ou poreuses ayant la même structure (structure homogène) sur toute leur épaisseur (Chikhi 2008) (Figure I.2 "a").
- **Membranes à structure asymétrique "anisotrope" :** la structure de la membrane varie d'une couche à une autre (Figure I.2 "b"). Celles-ci présentent une structure hétérogène.

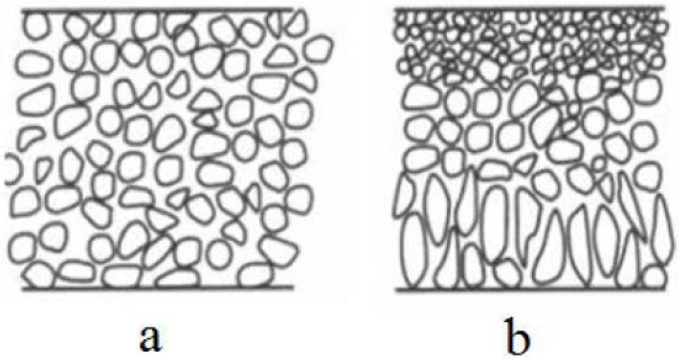


Figure I.2. Représentation schématique des différentes structures membranaires (Baker 2004 ; Chabane 2015).

Deux sous-types de membranes asymétriques peuvent été distingués (Annane 2015) **:**

- Membranes préparées à partir du même matériau ;
- Membranes composites : ce sont des membranes constituées principalement de deux couches :
    - **Peau :** une couche de très faible épaisseur en contact direct avec l'eau à traiter. Cette couche donne à la membrane sa sélectivité ;
    - **Couche support :** une couche plus épaisse, d'une perméabilité beaucoup plus grande, qui supporte la peau et confère à la membrane sa résistance mécanique.

La plupart des membranes commerciales de MF, UF et de NF sont des membranes à structure asymétrique, car ce type de structure permet d'avoir des perméabilités plus élevées. Une membrane asymétrique est en général constituée d'une couche support macroporeuse (assurant une bonne résistance mécanique) et d'une couche active (régissant les propriétés de rétention entre lesquelles s'insèrent une ou plusieurs couches intermédiaires dont le rôle est de permettre le dépôt de la couche active et sa tenue en pression (Figure I.3) (Chikhi 2008).

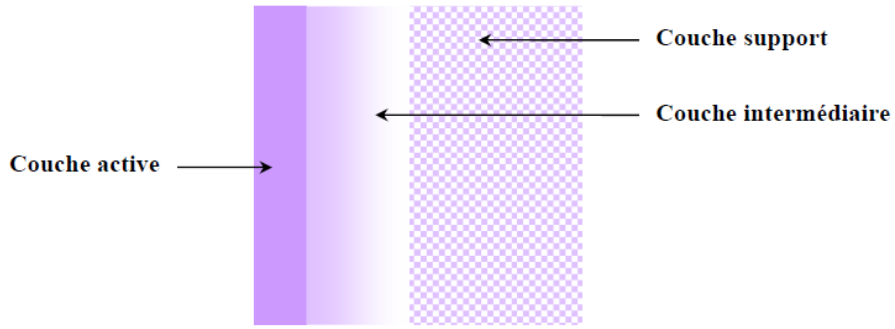


Figure I.3. Représentation schématique d'une membrane multicouche (Bouranene 2008).

I.2.2.2. Classification d'après la nature chimique

Les membranes sont fabriquées à partir des polymères organiques et de matières inorganiques. Les principaux polymères utilisés pour la fabrication des membranes sont (Annane 2015) :

a. **Dérivés de la cellulose** : ces polymères considérés généralement comme plus hydrophiles, peu coûteux et possédant une faible tendance à l'adsorption.
b. **Polysulfone (PS) et polyethersulfone (PES)**, qui ont une tendance à l'adsorption plus élevée que les membranes ci-dessus, mais avec une meilleure stabilité chimique, mécanique et thermique (Aptel et Buckley 1996).

c. **Les membranes inorganiques** sont fabriquées à base de céramique (oxydes d'aluminium, de zirconium ou de titane). Ces membranes ont une stabilité chimique, mécanique et thermique supérieure que les membranes organiques (Hamlaoui 2002).

I.2.2.3. Classification d'après la géométrie de préparation des membranes

Les membranes peuvent être classées en fonction de la forme sous laquelle elles sont fabriquées (Chikhi 2008) :

- ✓ Membranes planes ;
- ✓ Membranes spirales ;
- ✓ Membranes cylindriques ;
- ✓ Membranes tubulaires (ayant le diamètre interne plus grand que 3 mm) ;
- ✓ Membranes creuses (ayant le diamètre plus petit que 3 mm).

La forme des membranes conditionne la manière dont elles sont assemblées pour fabriquer des modules.

- **Modules membranaire**

La séparation se déroule dans un module de filtration pouvant prendre plusieurs formes (Lanteri 2009) :

- ✓ **Module plan :** le système utilise des membranes sous forme de feuilles. Il s'agit du type de configuration la plus simple. La surface filtrante reste toutefois faible par rapport à l'encombrement engendré.
- ✓ **Module tubulaire :** Il est constitué par des faisceaux de tubes de 4 à 25mm de diamètre intérieur. Ils sont basés sur une technologie simple, une consommation élevée en énergie est nécessaire pour une très faible surface d'échange par unité de volume (compacité réduite) (Figure I.4).

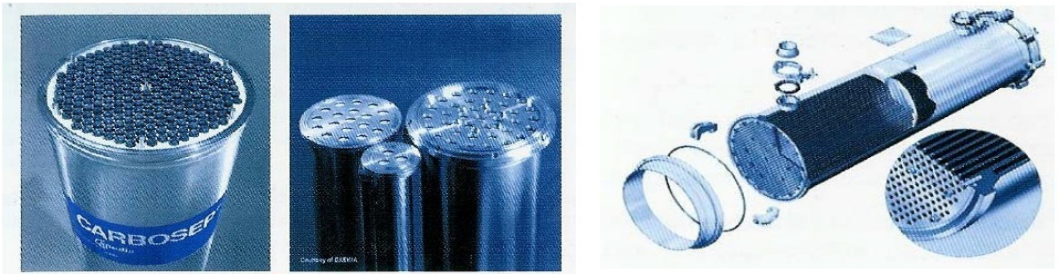


Figure I.4. Représentation schématique d'un module tubulaire (Chabane 2015).

- ✓ **Fibre creuse :** Elle rassemble un ensemble de fibres creuses de diamètre externe inférieur à 1 millimètre en un faisceau. Cette configuration leur procure la plus forte densité d'écoulement par module (Figure I.5).

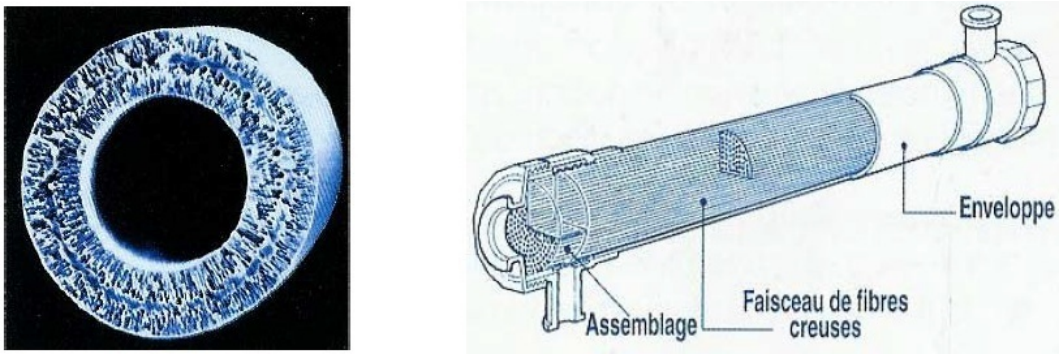


Figure I.5. Représentation schématique d'un module à fibres creuses (Lanteri 2009 ; Chabane 2015).

- ✓ **Module spiralé :** une ou plusieurs membranes planes sont enroulées autour d'un tube poreux et creux destiné à collecter le perméat. Ce type de module permet d'obtenir des surfaces filtrantes importantes pour un faible encombrement et un coût peu élevé. C'est une solution typiquement choisie dans l'industrie agro-alimentaire pour le traitement des fluides clarifiés (Figure I.6) (Antoine 2004).

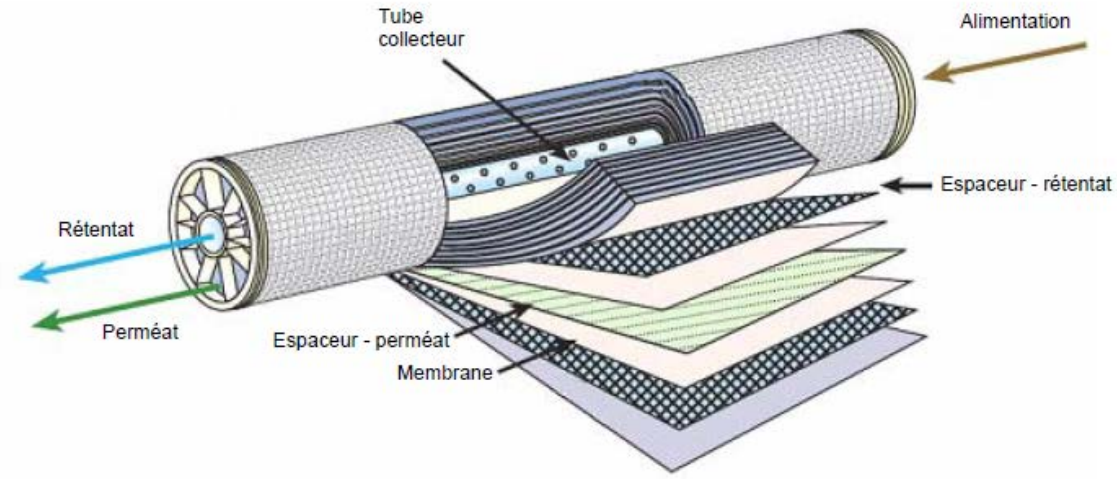


Figure I.6. Représentation schématique d'un module spiralé (Antoine 2004).

I.2.2.4. Classification d'après le mécanisme de séparation

La classification la plus répandue des membranes d'après le mécanisme de séparation est la suivante (Annane 2015) :

a. **Membranes poreuses (MF, UF, NF)**, où les effets de tamisage et de friction jouent des rôles importants, comme en NF, les forces de surface peuvent jouer un rôle important. Les pores peuvent être subdivisés en plusieurs catégories :
   - ✓ Macropores-plus grands que 50 nm (MF) ;
   - ✓ Mésopores-2 à 50 nm (UF) ;
   - ✓ Micropores-plus petits que 2 nm (NF).

b. **Membranes denses** (perméation de gaz, pervaporation, osmose inverse (OI)) ; ces membranes peuvent être considérées comme des milieux denses où la diffusion des espèces a lieu dans les volumes libres situés entre les chaînes moléculaires du matériel de la membrane ;

c. **Membranes échangeuses des ions** : Ces membranes sont considérées comme un type spécial de membranes non poreuses, consistant dans des gels denses, ayant une charge positive (échangeuses des anions) ou négative (échangeuses des cations).

Il a noté que la communauté scientifique ne s'entende pas sur cette classification. Aptel et Buckley 1996 considèrent que les membranes d'OI sont des membranes poreuses. La figure I.7 montre schématiquement la correspondance entre les opérations de filtration par membrane, la taille des pores et les composés retenus.

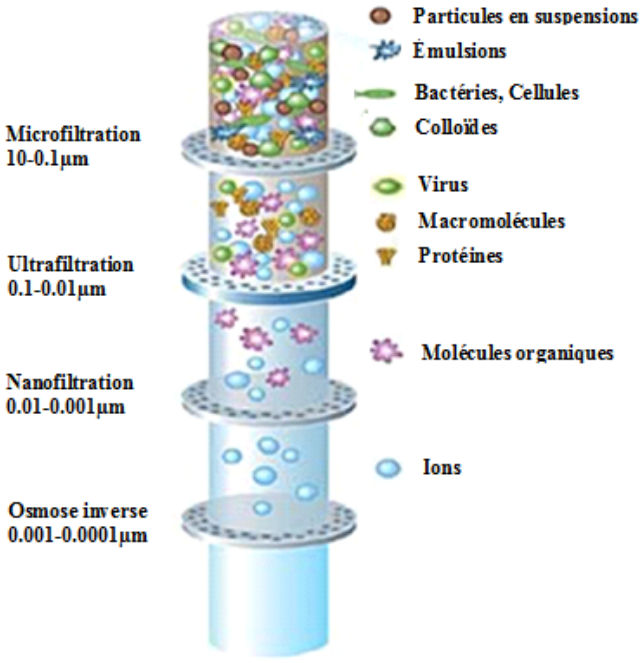


Figure I.7. Gamme de rétention de différentes particules par les procédés membranaires (Metaiche 2014).

I.2.3. Osmose inverse et nanofiltration

L'**osmose inverse** est un procédé de filtration tangentielle qui permet l'extraction d'un solvant, le plus souvent l'eau par perméation sélective à travers une membrane dense sous l'action d'un gradient de pression (Fell 1995).

- Elle s'oppose au phénomène naturel d'osmose qui tend à transférer le solvant d'une solution diluée vers une solution concentrée mise en contact par une membrane sélective sous l'action du gradient de concentration.
- Lorsqu'une pression est appliquée sur le compartiment le plus concentré, le flux de solvant diminue et tend vers une valeur nulle pour une pression égale à la pression osmotique de la solution.
- Lorsque la pression appliquée est supérieure à cette pression osmotique, le flux s'inverse : c'est le phénomène d'osmose inverse (Figures I.8).

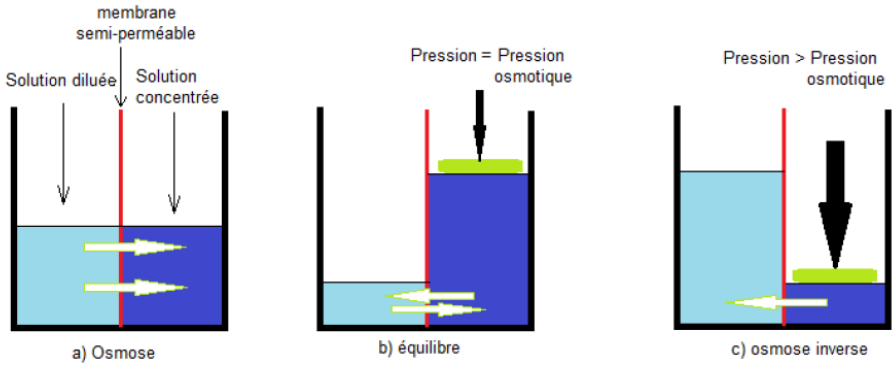


Figure I.8. Principe des phénomènes d'osmose inverse (Triki 2014).

Pour les solutions diluées, considérées comme thermodynamiquement idéales, la pression osmotique est calculée à partir de l'équation de Van't Hoff (Sagne 2008) :

$$\pi = CRT \qquad \text{(I.1)}$$

π : pression osmotique de la solution (Pa)

C : concentration de la solution ($mol.m^{-3}$)

R : constante des gaz parfait (8.31 J $mol^{-1}K^{-1}$)

T : température (K).

La **nanofiltration** est une technique intermédiaire entre l'osmose inverse et l'ultrafiltration, elle permet de retenir des molécules de faible poids moléculaire (MW ≥300 g.mol^{-1}) et des ions divalents, sans modifier son affinité vis à vis les ions monovalents. Les caractéristiques d'utilisation sont les suivantes (Hong et Elimelech 1997) :

- pression de travail 10 à 40 bars,
- diamètre des pores de la membrane 1 nm,
- débit volumique 50 à 100 l.$h^{-1}m^{-2}$.

Les membranes de nanofiltration sont en fait des membranes classiques greffées par des groupements hydrophiles chargés négativement. Elles ont été testées avec succès en traitement des eaux pour éliminer à la fois les ions divalents (Ca^{+2},Mg^{+2}) et certains pesticides organochlorés. Les sels ionisés multivalents (calcium, magnésium, aluminium, sulfates…) et les composés organiques non ionisés de masse molaire supérieure à environ 250 g/mol sont, par contre, fortement retenus. Ces membranes ont des seuils de coupure entre 300 et 1000 g/mol (Hong et Elimelech 1997).

I.2.4. Sélectivité d'une membrane

La sélectivité d'une membrane pour une substance donnée dépend de sa nature et de sa structure, de l'environnement chimique à proximité de la membrane et des propriétés de la substance à séparer. La sélectivité s'exprime par un taux de rétention *Tr* ou par un facteur de sélectivité (Pontié et al. 2001 ; Altalyan et al. 2016) :

$$T_r = \frac{C_{SA} - C_{SP}}{C_{SA}} = 1 - \frac{C_{SP}}{C_{SA}} \quad \text{(I.2)}$$

Où :

C_{SP}: concentration d'un soluté dans le perméat (kg/m^3) ;

C_{SA}: concentration d'un soluté dans l'alimentation (kg/m^3).

La sélectivité d'une membrane est aussi bien lié au seuil de coupure définit comme la limite de la masse moléculaire à retenir exprimé en molecular weight cutoff (MWCO) (Chabane 2015).

I.2.5. Influence des paramètres opératoires

I.2.5.1. Pression

La pression de fonctionnement (Ptm) doit être supérieure à la pression osmotique. L'étude menée par Sridhar et al. 2002 a montrée que pour une pression appliquée inférieure à 8.9 bar, aucun perméat n'est produit lors du traitement d'un effluent issu d'industrie d'huile végétale. Ainsi, la pression efficace (Ptm-Π) correspond à la pression qui conduit réellement à la production du perméat. Plus cette pression est élevée, plus le flux du perméat est important (Padilla et Tavani 1999 ; Galambos et al. 2004 ; Lin et Yang 2004). De même, le taux de

rétention augmente lorsque la pression augmente (Sridhar et al. 2002 ; Lin et Yang 2004). D'après Lin et al. 2004, la DCO du perméat passe de 6.7 à 3.1 mg.l^{-1} pour des pression de 2.1 et 6.3 bar avec un taux de rétention augmente de 96.9% à 98.6%. Sridhar et al. 2002 ont montré que cette augmentation du taux de rétention est due à l'augmentation de la pression transmembranaire ainsi que des flux du solvant et du flux du soluté. La concentration du soluté dans le perméat diminue. Par conséquent, le taux de rétention augmente.

I.2.5.2. Température

L'augmentation de la température engendre une diminution de la viscosité des solutions, et par conséquent une augmentation de la densité de flux du perméat (Bilstad 1995), d'environ 3% par degré Celsius. Le taux de rétention diminu, la diffusivité de l'eau et des solutés augmente avec la température (Girard et Fukumoto 2000). Pour une pression donnée, le compactage de la membrane est d'autant plus important que la température est élevée. Il est donc important de vérifier que la membrane et le module aient une tenue suffisante à la température de fonctionnement choisie en général, la température maximale recommandée est de l'ordre de 50°C.

I.2.5.3. Taux de récupération "Taux de conversion"

Le taux de récupération "y" d'un système d'OI/NF est le rapport entre le débit de permeat "Q_P" et le débit d'alimentation "Q_A" (Chabane 2015) :

$$y = \frac{Q_P}{Q_A} \qquad \text{(I.3)}$$

L'augmentation du débit d'alimentation engendre une augmentation de la vitesse de circulation le long de la membrane. Plus elle est importante, plus la concentration aux abords de la membrane sont moins importantes aussi que la densité de flux du perméat au cours du temps (Cornelissen et al. 2001). Plus le taux de récupération est élevé, plus le flux diminue, aussi la concentration côté retentât augmente (Sagne 2008).

I.3. Interactions et propriétés de séparation

Dans le cas des molécules organiques, les propriétés de séparation ne sont pas facilement prédictibles. En effet, il a été montré qu'elles résultent de trois interactions qui affectent leur rétention lors d'OI/NF et peuvent être signalées comme suit : encombrements

stériques (effet du tamisage) (Altalyan et al. 2016), répulsion électrostatique (effet de charge) et des interactions hydrophobes /adsorption (Sagne 2008).

Rappelons tout d'abord la structure du polyamide, polymère majoritairement utilisé comme couche active des membranes d'osmose inverse et de nanofiltration (Figure I.9).

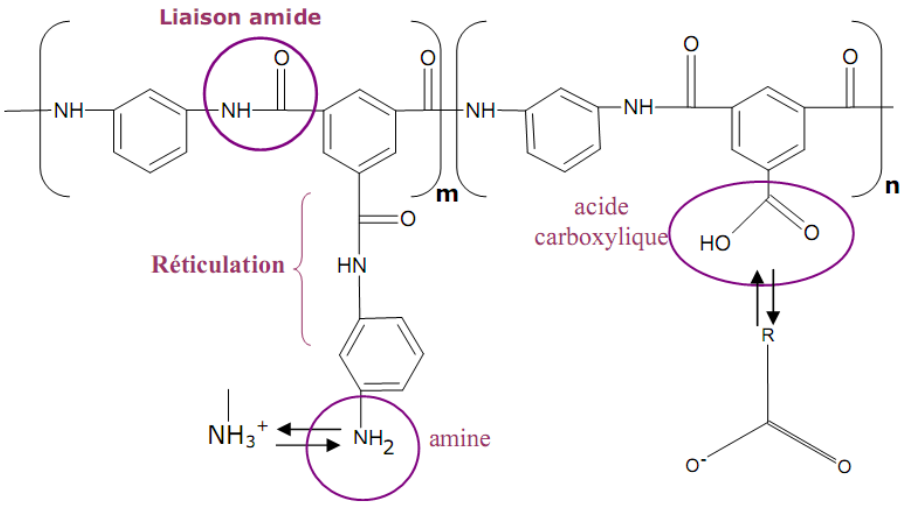


Figure I.9. La molécule de polyamide et ses équilibres acido-basiques (Sagne 2008).

De par cette structure, plusieurs types d'interactions peuvent se produire entre les molécules en solution et la surface de la membrane (Sagne 2008) :

- Les fonctions acides carboxyliques et amines libres (non engagées dans une liaison de réticulation) sont susceptibles de s'ioniser, arborant ainsi respectivement une charge négative et positive et pouvant alors s'engager dans des liaisons électrostatiques fortes.

Elles peuvent également participer à des liaisons hydrogènes.

- Les noyaux aromatiques peuvent développer des interactions π-π avec le nuage électronique d'une autre molécule aromatique, pour former un nuage d'électrons global plus stable.
- Les liaisons amides peuvent développer des interactions de Lifshitz-van der Waals plus faibles, par les doublets d'électrons libres des atomes d'azote et d'oxygène.

I.3.1. Encombrement stériques (effet du tamisage)

La taille moléculaire des molécules organiques et le seuil de coupure (Sc ou MWCO) de membrane ont confirmé l'importance des paramètres dimensionnels pour la rétention des molécules sur les membranes d'OI/NF (Yangali-Quintanilla et al. 2010a ; Dolar et al. 2011). Plusieurs auteurs ont sélectionné les descripteurs de taille moléculaire sachant que la rétention est liée aux mécanismes d'exclusion de taille entre membranes-solutés (Kiso et al. 2001a ; Ozaki

et al. 2002 ; Schafer et al. 2003 ; Kimura et al. 2004 ; Nghiem et al. 2004 ; Yangali-Quintanilla et al. 2009b-2010b).

I.3.1.1. État physique de la membrane

Les propriétés physiques de la membrane comme le nombre de pores, la forme ainsi que la rugosité de la surface jouent un grand rôle dans la rétention des solutés.

La porosité est exprimée par la taille, la distribution de la taille et le nombre de pores effectifs dans la couche active. Le terme "pore" pour les membranes d'OI/NF correspond à un espace vide dans le polymère par lequel le transport du fluide peut avoir lieu sous l'action d'une force motrice. Dans une membrane dense, les pores sont des vides de moins de 0.5 nm de diamètre qui sont formés naturellement dans les polymères solides. Ces pores, ou défauts de densité, sont causés par des irrégularités de l'entremêlement des chaînes de polymères. Ils peuvent aussi être créés par les contractions locales des polymères cristallisables lors de leur cristallisation. Ces pores peuvent être circulaires ou non, fermés ou ouverts et forment un réseau continu d'interconnections. Ils sont susceptibles d'accueillir de petites molécules qui au cours de leur transport à travers ce pore interagissent fortement avec ses parois (Kosutic et al. 2000).

Les indicateurs de porosité de membrane peuvent être le rejet du sel ou le seuil de coupure, attribués par le fabricant en cours de l'expérience.

A. Seuil de coupure ou Seuil de rétention "Sc ou MWCO"

Les membranes d'OI/NF sont caractérisées par leur seuil de coupure. Il s'agit de la masse molaire de la plus petite molécule arrêtée à 90% (ou 95% selon le fournisseur) par la membrane dans des conditions opératoires données. Il est exprimé généralement en Dalton (unité équivalente à des grammes par mole). Le seuil de coupure renseigne l'utilisateur sur la plage d'utilisation de la membrane, mais sa valeur exacte ne doit être considérée qu'à titre indicatif. En effet, les molécules ont des géométries variées et il est possible que des espèces de masse molaire supérieure au seuil de coupure soient retenues par la membrane, de même que des espèces de masse moléculaire légèrement inférieure la traversent. Il est donc important de retenir que le seuil de coupure est défini par rapport à un soluté donné. On peut ainsi caractériser une membrane d'OI/NF, pour une séparation donnée, dans des conditions fixées, par une courbe de variation du taux de rejet d'un soluté en fonction de sa masse moléculaire (Figure I.10) (Lanteri et al. 2009).

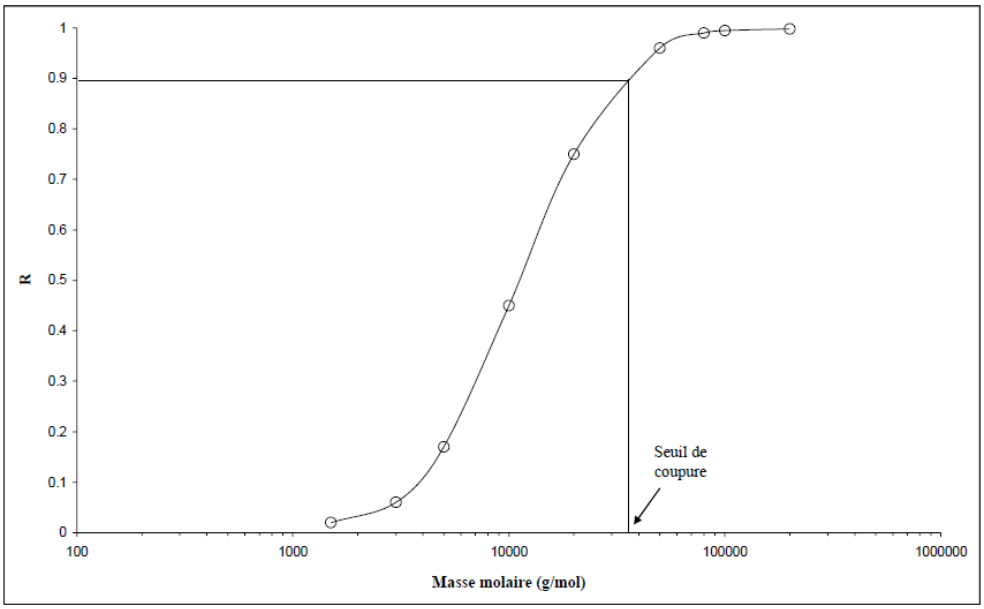


Figure I.10. Illustration du principe de détermination du seuil de coupure d'une membrane, la rétention en fonction de la masse molaire (défini ici pour un taux de rejet de 0.90) (Lanteri et al. 2009).

Plus le **Sc** d'une membrane est faible, plus elle peut retenir de petites molécules ou des colloïdes. Le **Sc** est relié principalement à la taille des pores de la membrane, il est influencé par (Anselme et Jacobs 1996) :

- ✓ la forme de la molécule à filtrer ;
- ✓ sa charge ;
- ✓ le pH ;
- ✓ la pression d'opération ;
- ✓ le flux de perméation ;
- ✓ la charge de la membrane.

B. Rejets des sels

- **Rejet du sel de chlorure de sodium SR(NaCl)**

Le premier paramètre donnant une indication sur la "porosité" de la membrane est son taux de rétention en chlorure de sodium (NaCl), mesuré par le fabricant dans des conditions standards, qui lui sont spécifiques. Il a été bien corrélé au taux de rétention de certains pesticides (Kiso et al. 2000), de polysaccharides et d'alcools (Kiso et al. 2001a ; Bellona et al. 2004). Étude d'Agenson et al. 2003 sur 36 composés organiques semi-volatils et volatils d'intérêt environnemental, de masse molaire (MW) comprise entre 70 et 300 g/mol a montré qu'en

général les taux de rétention les plus forts sont obtenus pour les membranes ayant un taux de rétention en NaCl le plus élevé.

- **Rejet du sel de sulfate de magnésium SR($MgSO_4$)**

Le rejet du sel de sulfate de magnésium incorpore les encombrements stériques et les effets de répulsion électrostatique qui sont liés à la structure de la membrane et les conditions d'exploitation. L'utilisation du seuil de coupure était acceptable à des fins de modélisation, toute fois les membranes d'OI/NF avec une large gamme de MWCO est difficile à estimer la rétention de contaminants. Le rejet du sel de sulfate de magnésium peut être plus approprié (Yangali-Quintanilla et al. 2009b et 2010b).

Kosutic et al. 2000 ont étudié la porosité des membranes commerciales en polyamides et ont rapporté que les différences de performances sont principalement dues par des différences de porosité. Par ailleurs, ils ont indiqué que la variation de la pression du travail a une influence sur cette porosité et donc sur les performances des membranes. Pour certaines membranes, l'augmentation de la pression conduit à un effet de compaction qui fait diminuer la densité de flux du perméat et la taille des pores tout en augmentant leur nombre.

Le pH peut avoir une influence sur la porosité : à pH élevé, la répulsion entre les groupements acides au sein de la membrane peut entrainer une augmentation de la taille des pores (Berg et al. 1997).

I.3.1.2. Taille moléculaire

La taille moléculaire est un paramètre principal qui détermine la rétention de molécules organiques avec les membranes d'OI/NF. Plusieurs paramètres peuvent être utilisés pour représenter la taille moléculaire : le poids, la longueur, la largeur, la profondeur et la largeur équivalente qui sont calculés à partir de la structure moléculaire. Van der Bruggen et al. 1998 ont montré qu'il existe une bonne corrélation entre la rétention et les différents paramètres de taille (Figure I.11).

A. poids moléculaire "MW"

Le poids moléculaire est un paramètre facilement disponible, il fournit seulement une estimation approximative de la taille moléculaire. Il est utilisé pour avoir une première idée sur la taille relative des molécules étudiées (Van der Bruggen et al. 1998). Il ne peut pas être utilisé pour la modélisation, car il ne représente pas la géométrie de la molécule (Van der Bruggen et

al. 1999). Pour les molécules non chargées, peu polaires, la masse molaire peut fournir une première information sur la rétention (Kiso et al. 2001a) avec de bonnes corrélations (Ozaki et Li 2002 ; Sadmani 2014b). Cependant, il ne renseigne pas sur l'encombrement stérique de la molécule. L'évaluation de la taille et de la géométrie de la molécule couplée à la taille de pores peut être un meilleur descripteur que la masse molaire (Bellona et al. 2004). En effet, des pesticides ayant une section et non une masse molaire plus faible sont moins bien retenus (Berg et al. 1997). Le poids moléculaire est considéré comme une variable de descripteur moins significatif pour la modélisation de la rétention (Yangali-Quintanilla et al. 2008).

B. Longueur moléculaire "*Length*"

La longueur moléculaire (*Length*) a été définie comme la distance entre les deux atomes les plus éloignés de la molécule (en tenant compte des forces des Van der Waals). La ligne droite reliant les deux atomes les plus éloignés a été définie comme la L-axe (Santos et al. 2006 ; Kiso et al. 1992 et 2001b).

C. Largeur moléculaire "*Width*"

La molécule est projetée sur le plan perpendiculaire à l'axe L. Le rectangle de surface minimale entourant la projection de la molécule est sélectionné. L'aire du rectangle (*S*) est calculée et la largeur moléculaire (*Width*) est définie comme étant la moitié de la racine carrée de S. La racine carrée de S correspond à la longueur du côté du carré dont la superficie est égale à la surface du rectangle (Santos et al. 2006 ; Kiso et al. 1992 et kiso et al. 2010) :

$$Width = \frac{1}{2}\sqrt{S_{min}} \qquad \text{(I.4)}$$

D. Profondeur moléculaire "*Depth*"

La largeur moléculaire et la profondeur moléculaire (la largeur supérieure à la profondeur) sont mesurées par la projection de la molécule sur le plan perpendiculaire à l'axe de la longueur (Kiso et al. 1992 ; Santos et al. 2006) :

$$Depth = \frac{1}{2}\sqrt{S_{max}} \qquad \text{(I.5)}$$

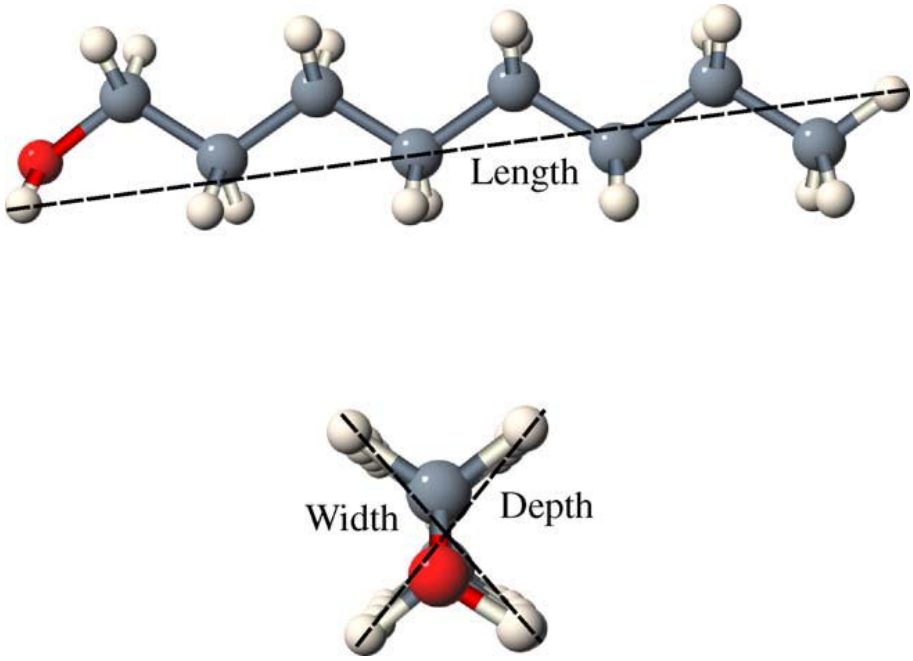


Figure I.11. Calcul des dimensions de l'octanol (*Length*, *Width* et *Depth*) de la structure 3D (Santos et al. 2006).

E. Largeur équivalente moléculaire "*EqWidth*"

La largeur moléculaire équivalente est définie comme étant la moyenne géométrique de la largeur et de la profondeur (Kiso et al. 1992 ; Santos et al. 2006 ; Linares et al. 2011).

On peut calculer la largeur équivalente par la formule suivante (Linares et al. 2011) :

$$EqWidth = \sqrt{Width * Depth} \quad \text{(I.6)}$$

I.3.2. Répulsions électrostatiques (effet de charge)

Les interactions électrostatiques ont souvent été rapportées comme participantes de manière importante au mécanisme de rétention. Elles s'établissent entre deux entités chargées (molécules organiques et membranes) (Van der Bruggen et al. 1998).

I.3.2.1. Charge de la surface

En général les membranes d'OI/NF thin-film composite (TFC) portent une charge négative (Norberg et al. 2007) qui permet de minimiser l'adsorption des composés colmatant chargés négativement et d'augmenter la rétention des sels ionisés dans l'eau. La charge de la membrane est généralement quantifiée par mesure du potentiel zêta (ζ) (Childress et Elimelech

1996 ; Childress et Desmukh 1998 ; Ozaki et Li 2002 ; Ardhaoui et al. 2003 ; Schäfer et al. 2004 ; Bellona et Drewes 2005).

Le pH de l'alimentation a une influence sur la charge de la membrane, il joue sur la protonation des groupes fonctionnels (Childress et Elimelech 1996). A pH élevé, la membrane est chargée négativement du fait de la dissociation du groupement carboxylique et à pH faible elle est chargée positivement du fait de l'ionisation du groupement amines (Ozaki et Li 2002). Le pH pour lequel le potentiel zêta est nul est appelé point isoélectrique ou pH d'ionisation de la membrane (Desmukh et Childress 2001).

I.3.2.2. Charge du soluté

Comme nous venons de voir, les membranes d'osmose inverse et de nanofiltration présentent une charge généralement négative. Il est donc clair qu'un soluté chargé négativement sera mieux retenu, par répulsion électrostatique, qu'un soluté neutre. Le respect de l'électroneutralité implique que les composés chargés positivement le seront également (Shim et al. 2002). Les molécules ionisables sont donc en général mieux retenues sous leur forme ionisée (Berg et al. 1997 ; Ozaki et Li 2002). Cet effet est d'autant plus marqué que la molécule est petite, car sa densité de charge est alors plus élevée.

Pour des molécules de masses molaires proches, celles de même signe que la membrane sont mieux retenues que les neutres, elles-mêmes mieux retenues que celles de charges opposées à celles de la membrane (Van der Bruggen et al. 1999 ; Hu et al. 2003). Pour ces dernières, l'attraction électrostatique favorise leur passage. Pour les plus grosses molécules les effets stériques restent prédominants (Van der Bruggen et al. 1999).

I.3.2.3. Influence du pH

Comme nous l'avons vu précédemment, la charge des membranes dépend du pH de la solution avec laquelle elles sont en contact (Childress et Desmukh 1996 ; Childress et Desmukh 1998 ; Schäfer et al. 2004 ; Bellona et Drewes 2005). Suite à cela, il peut se créer des interactions électrostatiques entre la membrane et les solutés chargés, ionisés ou non, présents en solution. Le positionnement de ceux-ci près de la surface modifie alors la charge apparente de la membrane (Sagne 2008). Par ailleurs, si le pH n'a pas d'influence sur le taux de rétention des molécules non chargées (Kiso et al. 2001b ; Agenson et al. 2003), alors il joue un grand rôle sur celui des molécules ionisables puisqu'il gouverne leur ionisation. En effet, la charge des espèces acido-basiques faibles varie avec le pH.

I.3.3. Interaction hydrophobes/adsorption

L'adsorption est fortement liée à l'hydrophobicité des molécules, à leur capacité à former des liaisons hydrogènes et à leur polarité.

I.3.3.1. Hydrophobicité de la membrane

La mesure de l'angle de contact α que forme l'eau sur une membrane, permet d'avoir une idée de l'hydrophobicité de cette membrane. Cette dernière est supposée gouverner la mouillabilité de la membrane en particulier par l'eau et donc avoir une influence sur ses performances (Ardhaoui et al. 2003). Il existe trois méthodes de mesure de l'angle de contact : la bulle captive, la technique de la lame de Wilhelmy et la goutte déposée (Gumi et al. 2003). Par l'utilisation de la méthode de la goutte déposée sur une surface solide (Figure I.12), deux cas peuvent être présentés (Sagne 2008) :

- ✓ α est supérieur à 90°, la membrane est dite hydrophobe ;
- ✓ α est inférieur à 90°, la membrane est dite hydrophile.

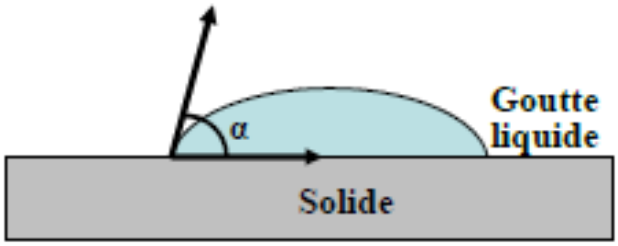


Figure I.12. Mesure de l'angle de contact par la méthode de la goutte déposée (Sagne 2008).

La plupart des membranes d'OI/NF sont hydrophiles possédant des valeurs de l'angle de contact généralement inférieur à 90°.

I.3.3.2. Hydrophobicité des solutés

Le caractère hydrophobe ou hydrophile des solutés va donc jouer un rôle sur leur rétention. Il est évalué par **le logarithme du coefficient de partage octanol-eau** ($Logk_{ow}$), plus souvent exprimé sous forme logarithmique ($logk_{ow}$, $logP$). S'il est supérieur à 2, la molécule est hydrophobe (HB) ; s'il est inférieur, la molécule est hydrophile (HL) (Connell 1989 et 1990 ; Yangali-Quintanilla et al.2009b).

Le logarithme du coefficient de partage n-octanol/eau, *LogP*, a été utilisé comme un indice hydrophobe, et est défini comme suit (Équation I.7) (Taheran et al. 2016) :

$$Logk_{ow} = \log(\frac{C_0}{C_W}) \quad \text{(I.7)}$$

Où C_0 et C_W sont les concentrations des molécules dans les couches n-octanol et l'eau (Kiso et al. 2001a ; Yangali-Quintanilla et al. 2010b ; Taheran et al. 2016).

Pour les solutés ionisables, le caractère hydrophobe est fonction du pourcentage des chargé contre les espèces non chargées, il dépend donc du pH. Par conséquent, un coefficient de partage apparent à un pH donné est calculé pour les solutés ionisables. Le coefficient de distribution octanol-eau dépend du pH, *D*, est le rapport entre les concentrations d'équilibres de toutes les espèces (ionisé et non) d'une molécule dans l'octanol à la même espèce dans la phase aqueuse. Le logarithme du coefficient de distribution peut être calculé à partir du *pKa* et $Logk_{ow}$ du soluté (Équations I.8 et I.9) (Verliefde et al. 2008).

Pour les acides, (Verliefde et al. 2008) **:**

$$\mathrm{LogD_{(pH)}} = \mathrm{logk_{ow}} - \log(1 + 10^{(\mathrm{pH} - \mathrm{pKa})}) \quad \text{(I.8)}$$

Pour les basses, (Verliefde et al. 2008) **:**

$$\mathrm{LogD_{(pH)}} = \mathrm{logk_{ow}} - \log(1 + 10^{(\mathrm{pKa} - \mathrm{pH})}) \quad \text{(I.9)}$$

On parle également des molécules polaires/apolaires. Les molécules hydrophobes sont en général apolaires (Sagne 2008). Le dipôle est dirigé vers la membrane de telle sorte que le côté du dipôle de charge opposée est plus proche de la membrane. Cette orientation n'est pas statique, afin elle doit être considérée comme une tendance statistique des molécules en mouvement rapide d'avoir cette orientation préférentielle. Le dipôle dirigé est donc attiré dans les pores et entre ainsi plus facilement dans la structure de la membrane (Van der Bruggen et al. 1998).

Les molécules ayant un moment dipolaire élevé ont des rétentions inférieures par rapport à des molécules non polaires. Cet effet, est indépendant du signe de la charge de la membrane (positif ou négatif) (Van der Bruggen et al. 1999).

I.4. Applications des procédés d'OI et NF

Les différentes applications envisagées pour les procédés OI et NF sont résumées dans le tableau I.1.

Tableau I.1. Applications des procédés l'OI et NF.

Application	Références
Industries diverses : ▪ Récupération des métaux ▪ Production de l'eau ▪ Traitement de surface ▪ Effluents industriels	(Nielsen 1995 ; Kimura et al. 2009 ; Comerton et al. 2008 ; Metaiche 2014 ; Ravikumar et al. 2014 ; Acero et al. 2016).
Industrie agro-alimentaire : ▪ Filière œnologique ▪ Lactosérum ▪ Recyclage des protéines ▪ Effluents agro-alimentaires ▪ Industrie pharmaceutique et biotechnologique : Sérums, principes actifs, antibiotique.	(Alkhatim 1998 ; Yangali-Quintanilla et al. 2010b et 2009b ; Shahmansouri et al. 2013 ; Metaiche 2014 ; Sadmani et al. 2014a et 2014b).
Eau potable, dessalement, saumure : ▪ Élimination nano-polluants organiques, élimination des particules, adoucissement, déminéralisation.	(Alkhatim 1998 ; Marty1999 ; Abdel-Jawad et al. 2002 ; Abbas et El Bastaki 2005 ; Metaiche 2014 ; Ricci et al. 2015 ; Donghan et al. 2016).

I.5. Avantages et inconvénients des procédés d'OI et NF

Les avantages et les inconvénients des procédés d'OI et NF sont regroupés dans le tableau I.2.

Tableau I.2. Avantages et inconvénients de procédés d'OI et NF, (Aptel et Buckley 1996).

Avantages	Inconvénients
▪ Un taux de conversion élevé (supérieur à 55%). ▪ Une consommation spécifique d'énergie basse (3 à 5 Kwh/m^3). ▪ Un coût d'investissement relativement bas. ▪ Une durée de mise en œuvre réduite. ▪ Opération à température ambiante ce qui évite la dégradation de molécules fragiles dans le domaine agro-alimentaire. ▪ Pas d'intervention de réactifs chimiques comme des agents d'extraction qui sont des sources de pollution.	▪ Sensibilité à la variation de la qualité d'eau brute et surtout aux pollutions. ▪ Nécessité d'un personnel qualifié. ▪ Exploitation relativement difficile surtout au niveau du prétraitement. ▪ Un coût de maintenance élevé dû aux frais de renouvellement des membranes d'une durée de vie ne dépassant pas 7 ans à présent. ▪ Baisse de la perméabilité et modification de la sélectivité en cas de colmatage des membranes.

I.6. Conclusion

La rétention des molécules organiques par les membranes de la nanofiltration et d'osmose inverse dépend d'un grand nombre de facteurs liés à la fois à la membrane (structure physique, charge, hydrophobicité), à la molécule (taille, structure, charge, polarité ou hydrophobicité) et aux conditions expérimentales (pH, pression, taux de récupération et température). Les interactions mises en place sont à la fois d'ordre stériques, électrostatiques et dues au caractère hydrophobe/hydrophile.

Chapitre II

Revue critique sur la modélisation de la rétention des MO lors d'OI/NF par les modèles QSAR

II.1. Introduction

La chémoinformatique est la discipline qui a pour est étudié et résolu des problèmes relatifs à la chimie en appliquant des méthodes informatiques. Cette discipline est devenue une véritable interface entre la chimie et l'informatique (Hamadache 2014). En d'autres termes, la grande quantité de données issues des expériences ne peut être traitée et analysée que par des méthodes informatiques.

Ces dernières années, il y a eu plusieurs tentatives pour faire progresser l'utilisation des relations quantitatives structure-activité (QSAR) comme une approche viable pour développer des modèles pilotés par les données pour décrire la performance des procédés membranaires (Dornier et al. 1995 ; Sahoo et al. 2006 ; Libotean et al. 2009). Peu d'études ont également appliqué les modèles QSAR à la modélisation de la rétention des composés organiques par les membranes d'osmose inverse (OI) et de nanofiltration (NF) (Libotean et al. 2008 ; Yangali-Quintanilla et al. 2008, 2009b et 2010b ; Shahmansouri et al. 2013 ; Sadmani et al. 2014a ; Sadmani et al. 2014b).

Ce chapitre a été consacré à une revue bibliographique a été scindé en trois parties. Dans la première partie, on parlera sur les méthodes d'apprentissage statistiques (la régression linéaire multiple et les réseaux de neurones artificiels). On passera dans la deuxième partie à une revue des travaux de la rétention des molécules organiques lors d'osmose inverse et nanofiltration. La troisième partie a été consacrée à une revue des travaux de la prédiction de la rétention par les modèles QSAR. Une conclusion clôturera ce chapitre et sera une synthèse par laquelle on justifiera le choix du thème de cette thèse.

II.2. Méthodes d'apprentissages statistiques

Au cours de la dernière décennie, les algorithmes d'apprentissage statistique ont suscité beaucoup d'intérêt dans le milieu académique et au sein d'entreprises de diverses industries. Ils ont été implantés avec succès pour l'accomplissement des tâches prédictives reliées à des processus statistiques observés pour lesquels on peut identifier plusieurs variables explicatives (Hamadache 2014). Cette étude se concentre sur une classe particulière de ces algorithmes : la régression linéaire multiple et les réseaux de neurones artificiels.

II.2.1. Régression linéaire multiple (RLM)

La régression linéaire multiple "RLM" est une méthode de modélisation couramment plus utilisée et connue (Sarkhosh et al. 2012), elle est utilisée comme une base pour un certain nombre de méthodes multivariées. RLM est constituée d'une relation quantitative entre un groupe de variables prédictives X_i et une réponse y_i comme représentée dans l'équation (II.1) :

$$y = A_0 + \sum_{i=1}^{n} A_i X_i \qquad \text{(II. 1)}$$

Où y est la sortie de la régression linéaire multiple "RLM", X_i représente les entrées, A_i représente les coefficients et A_0 est la constante (appelée intercepté) et n est le nombre des entrées.

II.2.2. Réseaux de neurones artificiels (RNA)

II.2.2.1. Neurone biologique

L'idée de réseaux de neurones artificiels a été inspirée de la manière par laquelle les neurones biologiques traitent les informations. Le cerveau humain est composé de cellules distinctes appelées neurones formant un ensemble dense d'environ 10 à 100 milliards d'unités interconnectées. Tel qu'il est illustré sur la figure II.1, le neurone est une cellule composée d'un corps cellulaire et d'un noyau. Le corps cellulaire se ramifie pour former les dendrites. C'est par les dendrites que l'information est acheminée de l'extérieur vers le soma (corps du neurone) (Khaouane 2013).

L'information est traitée alors par le corps cellulaire. Si le potentiel d'action dépasse un certain seuil, le corps cellulaire répond par un stimulus. Le signal transmis par le neurone est cheminé ensuite le long de l'axone (unique) pour être transmis aux autres neurones. La transmission entre deux neurones n'est pas directe. En fait, il existe un espace intercellulaire de quelques dizaines d'Angströms entre l'axone du neurone afférent et les dendrites du neurone efférent. La jonction entre deux neurones s'appelle synapse (Tzanakou 2000).

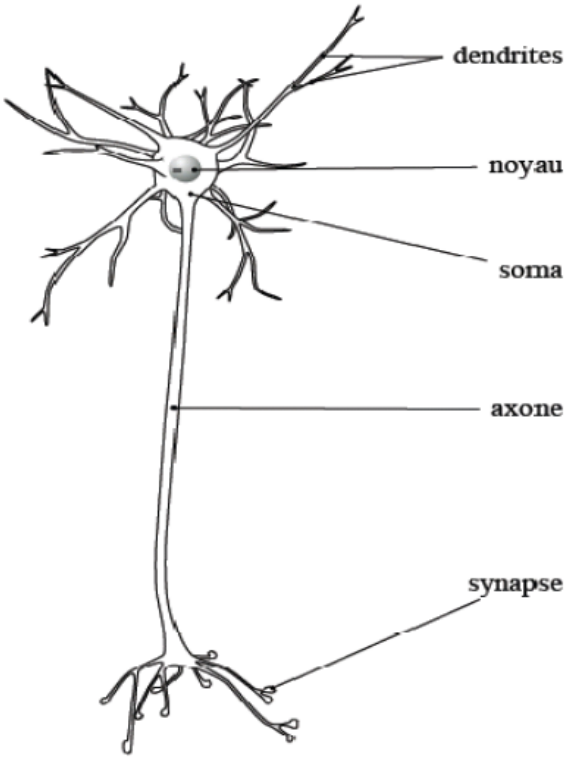


Figure II.1. Neurone biologique (Talbi 2011).

II. 2.2.2. Neurone formel

Le neurone formel est l'élément essentiel d'un réseau de neurones. C'est un opérateur mathématique très simple, dont on peut facilement calculer la valeur numérique (Hamadache 2014). Son fonctionnement est schématisé sur la figure II.2.

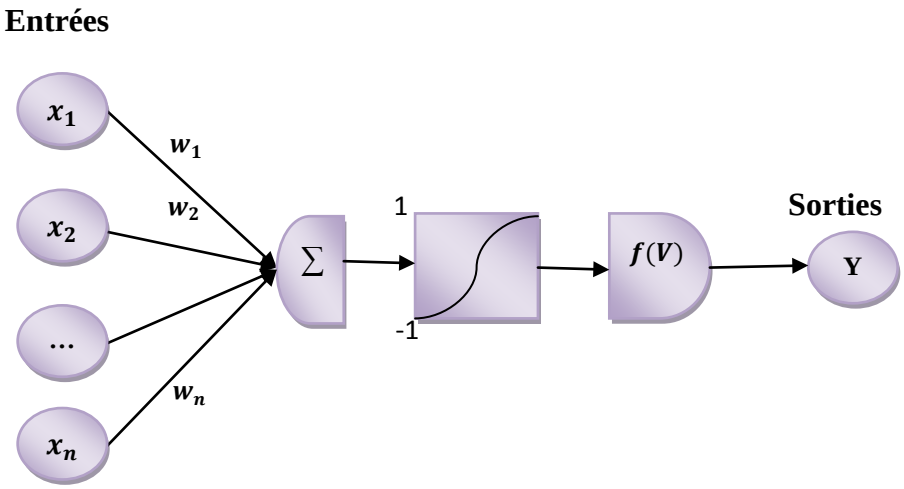


Figure II.2. Neurone formel.

Un neurone formel est une fonction algébrique paramétrée, non linéaire en ses paramètres. Ses entrées peuvent être les sorties d'autres neurones ou des entrées de signaux extérieurs. Sa sortie est une fonction non linéaire *f* d'une combinaison linéaire *v* des entrées (x_i). Le potentiel *v* le plus fréquemment utilisé est la somme pondérée des entrées x_i pondérées par les coefficients (w_i) également appelés poids de connexions (Hamadache 2014) :

$$v = w_0 + \sum_{i=1}^{n} (w_i x_i) \qquad \text{(II. 2)}$$

Précisons que le poids w_0 est affecté à une entrée constante appelée "biais". Une fonction *f*, appelée fonction d'activation, est appliquée à ce potentiel *v*. Le choix d'une fonction d'activation se révèle dans certains cas être un élément constitutif important des réseaux de neurones. Les fonctions mathématiques les plus souvent utilisées dans notre travail sont regroupées dans le tableau II.1. À la différence des neurones biologiques dont l'état est binaire, la plupart des fonctions de transfert sont continues, offrant une infinité de valeurs possibles comprises dans l'intervalle [0, +1] ou [-1, +1] (Khaouane 2013). Un neurone formel, réalise donc l'opération suivante :

$$y = f(v) = f\left(w_0 + \sum_{i=1}^{n} (w_i x_i)\right) \qquad \text{(II. 3)}$$

Tableau II.1. Fonctions de transfert usuelles, $f(x)$ (Si-Moussa 2010).

	Fonction de transfert			
	Linéaire	**Logistique (Sigmoïde logarithmique)**	**Tangente hyperbolique**	**Exponentielle**
Notation	Purelin	Logsig	Tansig	Exponential
Équation de fonction de transfert	$f(x) = x$	$f(x) = \frac{1}{1+e^{-x}}$	$f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$	$f(x) = e^{-x}$

II.2.2.3. Architecture d'un réseau de neurones

Un RNA est formé d'un ensemble de neurones formels interconnectés. Deux grands types d'architectures peuvent être distingués (Khaouane 2013 ; Hamadache 2014) :

- ✓ Les réseaux de neurones non bouclés : Ils réalisent une (ou plusieurs) fonction(s) algébrique(s) des entrées, par composition des fonctions réalisées par chaque neurone. Les réseaux de neurones non bouclés sont des outils statiques utilisés principalement pour effectuer des tâches d'approximation de fonctions non linéaires, de modélisation de processus statiques non linéaires.
- ✓ Les réseaux de neurones bouclés : Ils constituent quant à eux un système dynamique "à temps discret", régi par une (ou plusieurs) équation(s) aux différences non linéaires, résultant de la composition des fonctions réalisées par chacun des neurones et des retards associés à chacune des connexions. Ils sont utilisés pour effectuer des tâches de modélisation et d'adaptation de systèmes dynamiques, de commande de processus ou de filtrage.

Dans notre étude, nous avons adopté le modèle le plus couramment utilisé, le Perceptron Multi-couches (PMC), illustré par la figure II.3 et dont les principes sont exposés dans ce qui suit.

II.2.2.4 Perceptron multicouches

Le PMC est un réseau de neurones non bouclé caractérisé par une structure bien particulière. La figure II.3 illustre la structure de ce type de réseau. Ces neurones sont organisés en couches successives où les informations circulent dans un seul sens, de la couche d'entrée vers la couche de sortie. Les neurones d'une même couche ne sont pas interconnectés. Un neurone ne peut envoyer son résultat qu'à un neurone situé dans une couche postérieure à la sienne. Le réseau est formé de plusieurs couches, qui s'appellent, par convention :

- ✓ couche d'entrée : cette couche représente toujours une couche associée aux entrées du système, composée des cellules d'entrée qui correspondent aux *k* variables d'entrée.
- ✓ couche(s) cachée(s) : chaque couche est composée de p neurones. Ces neurones n'ont aucun lien avec l'extérieur et sont appelés neurones cachés.
- ✓ couche de sortie : elle est constituée par la valeur de sortie de la combinaison linéaire des fonctions de transferts de la couche cachée.

Les neurones sont reliés entre eux par des connexions pondérées. Ce sont les poids de ces connexions qui gouvernent le fonctionnement du réseau et programment une application de l'espace des entrées vers l'espace des sorties, à l'aide de la transformation non linéaire (Khaouane 2013 ; Hamadache 2014).

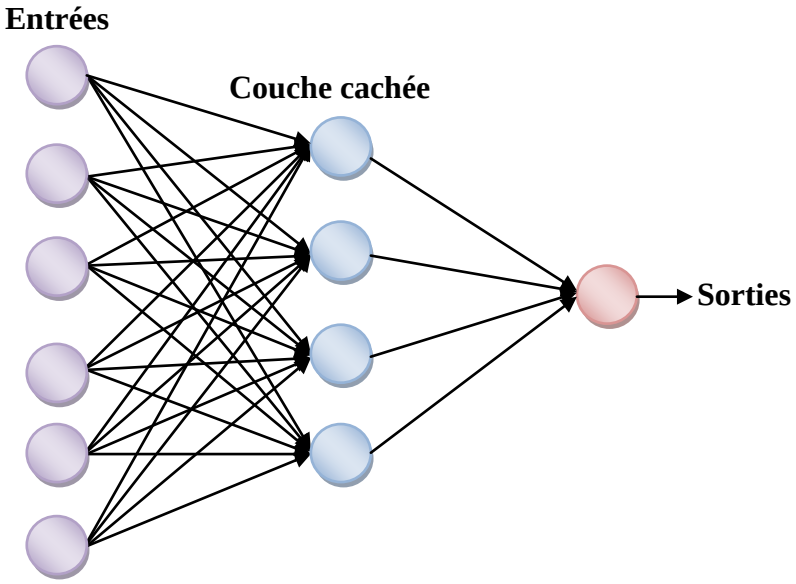
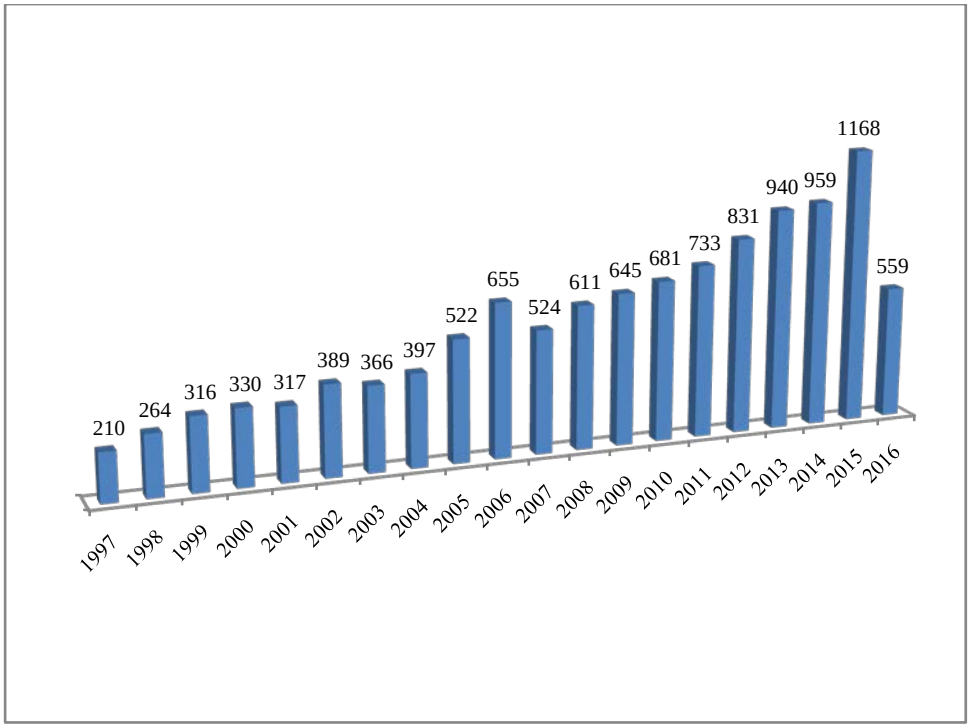


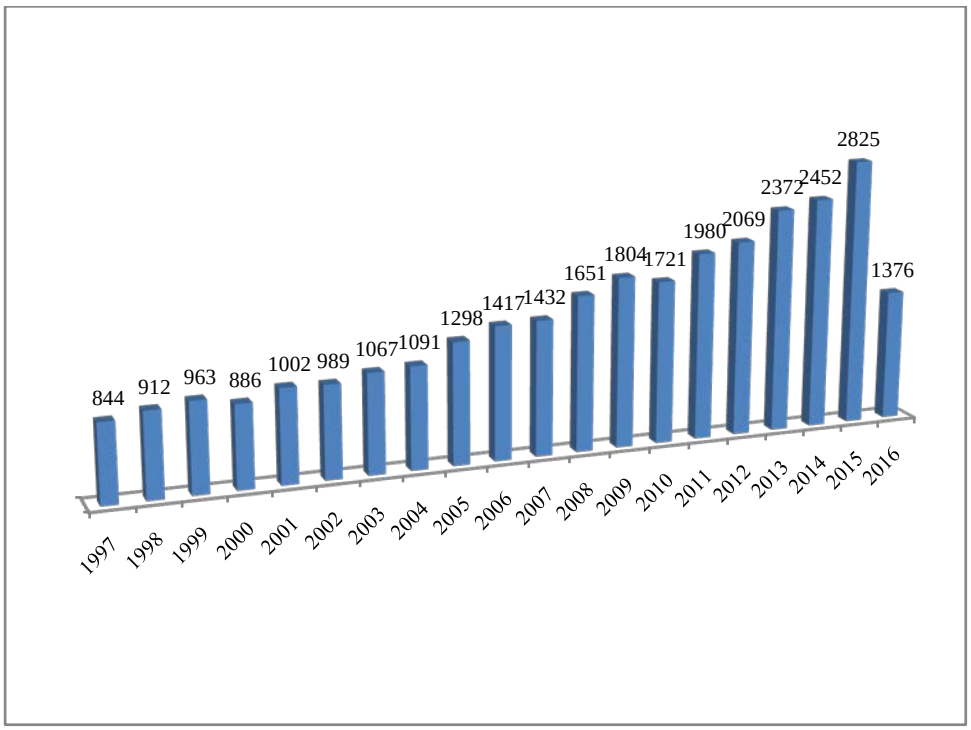
Figure II.3. Structure d'un réseau de neurones non bouclé.

II.3. Revue des travaux de la rétention des MO lors d'OI/NF

De nombreux travaux réalisés au niveau international mettent en évidence la rétention des molécules organiques par d'OI/NF. La base de données consultée est Science Direct. La figure II.4, reporte le nombre de publications de 1997 à 2016 pour les membranes de la nanofiltration et d'osmose inverse.



(a)



(b)

Figure II.4. Nombre de publications traitant la rétention des molécules organiques par (a) : nanofiltration (b) osmose inverse, (ScienceDirect le 05/04/2016).

II.4. Revue des travaux de la prédiction de la rétention par les modèles QSAR

Dans cette partie de la revue bibliographique, on parlera de la littérature scientifique consacrée à l'utilisation des relations quantitatives structure-activité (QSAR) en vue de la prédiction de la rétention des molécules organiques (neutres et ioniques) lors d'osmose inverse et de nanofiltration. Il existe un nombre limité de modèles QSAR, disponibles dans la littérature, pour la prédiction de la rétention. Par cette revue bibliographique, nous aspirons à montrer tout l'intérêt grandissant et l'utilité que revêt cette notion de QSAR. Aussi, à titre d'illustration, nous avons reporté le nombre de publications dédiées à l'étude de la rétention par la méthode QSAR et ce de l'année 2008 à l'année 2014. Sur la base de l'examen de la littérature, nous relaterons ci-après, quelques publications intéressantes qui ont retenu notre attention.

Libotean et al. 2008, ont examiné la rétention des composés organiques lors des membranes d'osmose inverse par un modèle relation quantitative structure-activité (QSAR). Ils ont utilisé une base de données de 144 rétentions de 50 molécules organiques neutres et 4 membranes d'osmose inverse de type polyamide et une membrane d'acétate de cellulose. Pour ce faire, ils ont exploité 45 descripteurs de molécules organiques et 9 propriétés des membranes d'osmose inverse. L'erreur absolue moyenne et l'écart-type de l'erreur absolue trouvés pour toutes les fractions prédites étaient jusqu'à 0.066 (erreur relative moyenne de 70.9%). Ils ont constaté que le moment dipolaire et la taille de la molécule organique sont des descripteurs moléculaires les plus importants pour caractériser la rétention.

Yangali-Quintanilla et al. 2008, ont modélisé la rétention des composés pharmaceutiques PhACs (Ibuprofène, Diclofénac, Acide clofibrique, Naproxène, Primidone et Phénacetine) et des composés organiques (Acide dichloroacétique, Acide trichloroacétique, Chloroforme, Bromoforme, Trichloroethene, Perchloroethene, Carbontetrachloride et Carbontetrabromide) par deux membranes de nanofiltration et six membranes d'osmose inverse. Ils ont exploité une base de données de 64 rétentions de 14 molécules organiques neutres et ioniques. Ils ont pu conclure que la rétention des composés neutres par les membranes d'osmose inverse et de nanofiltration a été bien prédite par le volume molaire, la longueur, la largeur équivalente, le caractère hydrophobe/hydrophile "$logk_{ow}$" et le moment dipolaire. Aussi ont montré que le poids moléculaire est une variable de descripteurs pauvres pour la modélisation de rétention.

Ils ont développé quatre modèles de régression linéaire multiples (RLM) deux modèles corrèlent la rétention avec la polarité des molécules ioniques (Équations II.4 et II.5), un modèle corrèle la rétention avec la taille et la polarité des molécules neutres (Équation II.6) et un modèle relie la rétention avec la taille et l'hydrophobicité des molécules neutres (Équation II.7). Les erreurs standards d'estimations du jeu de test étaient égales à 5.58% pour le premier modèle, 1.45% pour le deuxième modèle, 10.14% pour le troisième modèle et 12.56% pour quatrième le modèle. Ils ont trouvé des coefficients de régressions de 0.628 pour le premier modèle, 0.435 pour deuxième modèle, 0.952 pour le troisième modèle et 0.868 pour le quatrième modèle.

Les équations linéaires de la rétention des molécules organiques par les membranes d'osmose inverse par les modèles de régression linéaire multiple (RLM) sont :

$$Rétention = 101.104 - 8.885Dipole \quad \text{(II. 4)}$$

$$Rétention = 85.184 + 1.559Dipole \quad \text{(II. 5)}$$

$$Rétention = -118.889 + 2.503MV - 29.830Dipole - 86.339Length \quad \text{(II. 6)}$$

$$Rétention = -163.619 - 30.423Logk_{ow} + 386.160EqWidth \quad \text{(II. 7)}$$

Yangali-Quintanilla et al. 2009b, ont développé des modèles QSAR pour la prédiction de la rétention des molécules organiques neutres lors des membranes d'osmose inverse et de nanofiltration en polyamide. Les modèles QSAR combinent des interactions entre les propriétés de soluté, les caractéristiques de la membrane et les conditions de fonctionnement. Ils ont utilisé une base de données de 161 rétentions de 50 composés organiques neutres (produits pharmaceutiques, perturbateurs endocriniens, pesticides, alcools, phénols et solvants) par six membranes de nanofiltration et neuf membranes d'osmose inverse. Ils ont subdivisé leur base de données en trois sous ensemble, 60% pour la base d'apprentissage, 20% pour la validation et 20% pour la base du test. Ils ont trouvé que les descripteurs de taille de soluté (longueur moléculaire et largeur équivalente) et le descripteur d'hydrophobicité représenté par $Logk_{ow}$ étaient les principales propriétés physico-chimiques des composés organiques neutres capables de décrire la rétention. En ce qui concerne les caractéristiques de la membrane, ils ont constaté que la caractéristique de la membrane dont le rejet du sel de magnésium (SR "$MgSO_4$") peut être appliqué pratiquement à prédire la rétention des solutés organiques. Ils ont obtenu de coefficient de régression de R^2 0.97 et une erreur standard de déviation de 5%.

Ils ont proposé un modèle QSAR (Équation II.8), ce modèle corrèle la rétention avec les propriétés des molécules organiques neutres (la taille et l'hydrophobicité) et les propriétés de la membrane (la porosité) :

$$Rétention = 183.920EqWidth + 31.830Length - 0.549Logk_{ow} + 883.294SR(MgSO4) - 945.13 \quad \text{(II. 8)}$$

Yangali-Quintanilla et al. 2010b, ont modélisé la rétention des contaminants émergents (produits pharmaceutiques, les perturbateurs endocriniens) par les membranes de nanofiltration. Ils ont utilisé l'analyse en composantes principales, moindres carrés partiels et régression linéaire multiple pour trouver les modèles QSAR qui combinent les interactions entre les caractéristiques de la membrane, les conditions d'exploitation de filtration et les propriétés de composés pour prédire la rétention. Ils se sont fondés dans leur étude sur une base de données de 106 rétentions de 66 des molécules organiques neutres et ioniques par les membranes de nanofiltration. Ils ont trouvé des coefficients de régressions de 0.75 et 0.84. Ils ont conclu que MWCO était acceptable à des fins de modéliser, mais le rejet de sel de magnésium peut être plus approprié.

Leurs modèles QSAR ont permis de montrer que les variables les plus pertinentes qui influent sur la rétention des solutés organiques étaient *LogD*, rejet de sel de magnésium, largeur équivalente, la profondeur et la longueur. Ils ont développé deux modèles QSAR (Équations II.9 et II.10), ils corrèlent la rétention avec les propriétés molécules organiques neutres et ioniques (la taille et l'hydrophobicité) et propriétés de la membrane (la porosité) :

$$Rétention = 265.150EqWidth - 117.356Depth + 81.662Length - 5.229LogD + 1358.090SR(MgSO_4) - 144.817 \quad \text{(II.9)}$$

$$Rétention = 265.150EqWidth - 117.356Depth + 81.662Length - 5.229logD - 0.272MWCO - 62.565 \quad \text{(II.10)}$$

Shahmansouri et al. 2013, ont développé des modèles QSAR pour prédire la rétention des contaminants organiques neutres (perturbateurs endocriniens, composés pharmaceutiques, pesticides, saccharides et alcools) par membrane de nanofiltration (NF-270). Ils ont sélectionné plusieurs méthodes telles que la régression linéaire multiple (RLM), moindre carré partiel (PLS) et réseau de neurones artificiel (RNA). Ils ont testé une base de données de 255 rétentions de 67 molécules organiques. Ils ont montré que les valeurs de la rétention ont pu être prédites avec des valeurs des coefficients de régressions R^2 allant jusqu'à 0.86 pour le modèle non linéaire (RN) et 0.82 pour le modèle linéaire (RLM). Leurs conclusions indiquent que la rétention des composés organiques neutres était évaluée principalement à l'exclusion de la taille.

Sadmani et al. 2014a, ont étudié l'impact des colloïdes et des cations d'eau naturels sur la rétention de composés pharmaceutiques actifs et des perturbateurs endocriniens par nanofiltration (NE.70). Ils se sont basés dans leurs études sur une base de données expérimentale de 51 rétentions de 10 composés pharmaceutiques actifs et 7 composés de perturbateurs endocriniens. Ils ont développé un modèle de régression linéaire multiple (RLM) (Équation II.11), ce modèle corrèle la rétention avec la taille et l'hydrophobicité des molécules organiques neutres et ioniques. Ils ont rapporté un coefficient de régression $R^2 = 0.795$.

$$\text{Rétention} = -69.754 + 0.419\text{MW} - 9.030\text{logD} + 4.099\text{length} \quad \text{(II.11)}$$

Sadmani et al. 2014b, ont étudié la rétention des composés pharmaceutiques actifs (PhACs) et des perturbateurs endocriniens par nanofiltration. Ils ont sélecté une base de données expérimentale de 51 rétentions de 10 composés pharmaceutiques actifs et 7 composés de perturbateurs endocriniens. Ils ont proposé deux modèles de régressions linéaires multiples (RLM), un modèle corrèle la rétention avec la taille et l'hydrophobicité des molécules organiques neutres et ioniques (Équation II.12) et un modèle relie la rétention avec la taille et la polarité des molécules organiques neutres (Équation II.13). Ils ont constaté des coefficients de régressions $R^2 = 0.777$ pour premier modèle et $R^2 = 0.774$ pour deuxième modèle.

$$Rétention = -56.586 + 0.148MW - 8.391logD + 0.138MV \quad \text{(II.12)}$$

$$Rétention = -44.212 + 0.412MV + 4.345Dipole \quad \text{(II.13)}$$

Pour une signification de problème rencontré lors de la séparation des molécules organiques par les procédés d'OI et NF, des modèles ont été développé pour étudier la performance de ces procédés avec l'utilisation des analyses et des expériences qui ont été réalisées. Parmi ces modèles, nous rencontrons le QSAR (RNA et RLM). Mais il y a un nombre réduit de travaux cités dans la littérature relative aux travaux sur la modélisation par les modèles QSAR dans la rétention des molécules organiques "MO" par les membranes d'OI et NF durant les années 2008-2014. Le tableau II.2 énumère les seuls exemples notables d'études la prédiction de la rétention des molécules organiques lors d'osmose inverse et de nanofiltration par les modèles QSAR publiées au fil des années.

Tableau II.2. Vue d'ensemble des travaux antérieurs sur les modèles de prédiction de la rétention des molécules organiques lors d'osmose inverse et de nanofiltration.

Nombre et nature de MO étudiés	Type de membrane	Taille de base de données	Méthode de Modélisation	Références
17 Molécules organiques neutres et ioniques	NF	51	QSAR-RLM	Sadmani et al. 2014a
17 Molécules organiques neutres et ioniques	NF	51	QSAR-RLM	Sadmani et al. 2014b
67 Molécules organiques neutres	NF	255	QSAR-RN	Shahmansouri et al. 2013
66 Molécules organiques neutres et ioniques	NF	106	QSAR-RLM	Yangali-Quintanilla et al. 2010b
50 Molécules organiques neutres	OI et NF	161	QSAR-RN	Yangali-Quintanilla et al. 2009b
14 Molécules organiques neutres et ioniques	OI et NF	64	QSAR-RLM	Yangali-Quintanilla et al. 2008
50 Molécules organiques neutres	OI	144	QSAR-RN	Libotean et al. 2008

A l'heure actuelle, la modélisation moléculaire permet de calcul de données physico-chimiques beaucoup plus précises, appelées descripteurs moléculaires. Ces derniers peuvent être reliés de manière quantitative à la rétention des molécules par des approches de type QSAR/QSPR (Quantitative Structure Activity Relation/Quantitative Structure Property Relation) (Hansch et al. 1996). Ces méthodes sont souvent appliquées dans les domaines de chromatographie (Mazza et al. 2001 ; Kaliszan et Markuszewski 2003) et en biologie (Kulkarni et al. 2002). Dans le cas du traitement des effluents industriels complexes, pour lesquels de nombreux types d'interactions peuvent avoir une influence simultanée. Peu d'études ont également appliqué QSAR à la modélisation de la rétention des composés organiques par les

membranes d'osmose inverse (OI) et de la nanofiltration (NF). La figure II.5 représente le nombre de publications traitant la modélisation de la rétention des molécules organiques lors d'osmose inverse et de nanofiltration.

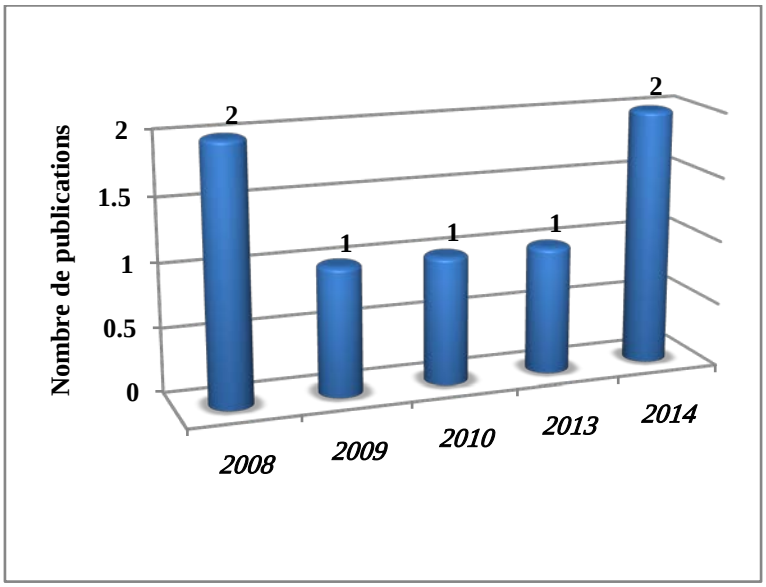


Figure II.5. Nombre de publications traitant la modélisation de la rétention des molécules organiques lors d'osmose inverse et de nanofiltration.

Seules Yangali-Quintanilla et al. 2009b et 2010b qui ont fait une modélisation de la rétention des composés organiques par les membranes d'OI/NF en tenant compte des indicateurs de la porosité (MWCO) et le rejet du sel de sulfate de magnésium "SR($MgSO_4$)". Ils ont découvert qu'il est possible de caractériser une membrane en termes de paramètre stérique/taille SR($MgSO_4$) plutôt que de MWCO.

II.5. Conclusion

La littérature met en évidence que la modélisation peut permettre de valider les hypothèses simplificatrices émises lors de la formulation mathématique des phénomènes de transfert. Peu de travaux ont appliqué les relations quantitatives structure-activité (QSAR) dans la modélisation des processus membranaires, ils peuvent être considérés comme des outils importants pour la prédiction de la rétention des molécules organiques lors des membranes d'osmose inverse et de nanofiltration.

Une analyse détaillée des articles publiés sur les modèles QSAR de la rétention des molécules organiques lors d'osmose inverse et de nanofiltration montre clairement que ces modèles ont été tirés à partir d'ensemble de données limitées de molécules organiques (ne

dépassant pas 70 pour la plupart). Aussi, malgré l'importance des conditions opératoires dans la rétention, la plupart de ces modèles ont une utilité limitée dans la prédiction.

Certains auteurs proposent un modèle linéaire exprimant la rétention des molécules organiques lors d'osmose inverse et de nanofiltration. Par contre, d'autres préconisent un modèle non linéaire obtenu par utilisation des réseaux de neurones artificiels (RNA). Même si les résultats de prédiction quantitative de la rétention sont satisfaisants pour les deux types de modèles, il est mentionné que la capacité de prévision de celui de non linéaire est de loin la meilleure.

En guise de conclusion, cette revue bibliographique nous a servi à fixer l'objectif de ce travail qui consiste en la prédiction de la rétention des molécules organiques lors d'osmose inverse et de nanofiltration en utilisant des réseaux de neurones artificiels et de la régression linéaire. Toutefois, nous avons fait en sorte de combler toutes les lacunes relevées dans les publications antérieures et ce en :

- ✓ choisissant une base la plus large possible ;
- ✓ utilisant une large base de descripteurs moléculaires, caractéristiques de membrane et les conditions opératoires et en déterminer les plus pertinents ;
- ✓ procédant à une optimisation pour le choix de nombre de neurone optimal dans la couche cachée et pour le choix des fonctions de transfert.

Chapitre III
Matériels et méthodes

III.1. Introduction

Les réseaux de neurones artificiels (RN) ont acquis une grande popularité dans la dernière décennie en raison de leur parallélisme élevé, de leur robustesse (Basheer et Hajmeer 2000) de leur capacité inhérente à extraire à partir des données expérimentales, les relations fortement non-linéaires et complexes entre les variables du problème (Si-Moussa 2010) sans aucune connaissance détaillée du système (Bryjak et al. 2004).

Ce chapitre comportera trois axes. Dans le premier axe, on examinera les données collectées, prétraitement et analyse. Le second axe sera consacré aux techniques les plus récentes pour établir des modèles QSAR à savoir la régression linéaire multiple (RLM), les réseaux de neurones artificiels (RNA) et le développement des modèles QSAR. Le troisième axe traitera l'analyse de la sensibilité basée sur la méthode de "poids".

III.2. Collecte des données, prétraitement et analyse

Dans cette étude, nous avons utilisé les données disponibles dans la littérature (Van der Bruggen et al. 1999 ; Ozaki et al. 2002 ; Kimura et al. 2003 et 2004 ; Bellona et al. 2005 ; Xu et al. 2005 ; Yoon et al. 2005 ; Verliefde et al. 2007 ; 2008 ; 2009a et 2009b ; Kim et al. 2007 ; Comerton et al. 2008 et 2009 ; Yangali-Quintanilla et al. 2008 ; 2009a ; 2009b ; 2010a et 2010b ; Huang et al. 2011 ; Plakas et al. 2011 ; Linares et al. 2011 ; Dolar et al. 2012 et 2013 ; Sadmani et al. 2014a et 2014b). La taille de la base de données et le nombre de composés organiques sont montrés dans le tableau III.1.

Tableau III.1. Taille de la base de données et le nombre de composés organiques utilisés dans ce travail.

Base de données (BD)	QSAR	Taille de base de données	Nombre de composés organiques
BD1	QSAR1	956	82
BD2	QSAR2	701	81
BD3	QSAR3	701	81

La liste des 82 composés organiques (neutres et ioniques) constitue les produits pharmaceutiques, les perturbateurs endocriniens, les pesticides, les alcools, les phénols et les solvants et sont répertoriés dans l'annexe A.

Le choix des variables d'entrées et de sortie a été basé sur les interactions entre les propriétés des molécules organiques, les caractéristiques de la membrane et les conditions opératoires de filtration pour la rétention des molécules organiques par les membranes d'OI/NF. Les entrées considérées dans cette étude sont : les descripteurs moléculaires (le poids moléculaire, $LogK_{ow}$, *LogD*, le moment dipolaire, la longueur moléculaire, la largeur moléculaire, la profondeur moléculaire et la largeur équivalente moléculaire), les caractéristiques de la membrane (le poids moléculaire de coupure "MWCO", le rejet de sel de sulfate de magnésium "SR(MgSO4)", la charge de surface de la membrane "potentiel zêta" et l'hydrophobicité de la membrane "angle de contact") et les conditions opératoires (pH, la pression et le taux de récupération).

Les valeurs de $LogK_{ow}$, *LogD* et moment dipolaire ont été obtenues soit à partir de la littérature ou calculées par les logiciels (Tableau III.2). Sur l'annexe B, nous donnons la capture d'écran de deux sites et deux logiciels utilisés.

Tableau III.2. Logiciels utilisés pour calculer $LogK_{ow}$, *LogD* et le moment dipolaire.

	Logiciels
$\boldsymbol{LogK_{ow}}$	Kowwin, epa, http://www.epa.gov/oppt/exposure/pubs/episuite.htm (consulté en janvier 2013)
LogD	Chemspider, http://www.chemspider.com (consulté en janvier 2013)
Moment dipolaire	ChemBio Ultra (ChemBioDraw, ChemBio3D)

Les valeurs de la largeur équivalente moléculaire (*EqWidth*) ont été obtenues à partir de la littérature ou calculées par l'équation I.6 (voir le premier chapitre). Le nombre total de variables (les propriétés des molécules organiques, les caractéristiques membranaires et les conditions d'opératoires) a été réduit en utilisant des matrices de corrélation (voir annexe C). Chaque matrice regroupe l'intensité de la liaison qui existe entre ces paramètres. Cette intensité de liaison a été obtenue par le calcul des coefficients de corrélation de Pearson (R) (Équation III.1) (Khaouane 2013) :

$$R(Z,Y) = \frac{\sum_i (z_i - \bar{z})(y_i - \bar{y})}{\sqrt{\sum_i (z_i - \bar{z})^2}\sqrt{\sum_i (y_i - \bar{y})^2}} \qquad \text{(III. 1)}$$

Où z_i est la i^{ème} valeur du vecteur *Z* (MW, *LogD*, *Length*, *EqWidth*, MWCO, SR ($MgSO_4$), angle de contact, potentiel zêta, pH, pression, taux de récupération et rétention) ; *yi* est la i^{ème} valeur

du vecteur *Y* (MW, *LogD*, *Length*, *EqWidth*, MWCO, SR ($MgSO_4$), angle de contact, potentiel zêta, pH, pression, taux de récupération et rétention) ; $\bar{z}$ et $\bar{y}$ sont respectivement les moyennes arithmétiques des vecteurs *Z* et *Y*. Le degré de coefficient de corrélation de Pearson doit être modéré ($R. < 0.75$) (Khaouane 2013), On élimine l'entrée qui possède la même valeur pour plus de 0.75 des échantillons.

Pour chaque base de données (BD1, BD2 et BD3), un modèle QSAR a été élaboré et un ensemble de variables d'entrées et de sortie a été identifié (Tableau III.3).

Tableau III.3. Entrées et sortie utilisées pour le développement des modèles QSAR (QSAR1, QSAR2 et QSAR3).

	QSAR 1	**QSAR 2**	**QSAR 3**
Entrées	– MW – *Log D* – Moment dipolaire – *Length* – *EqWidth* – MWCO – Potentiel zêta – Angle de contact – pH – Pression – Taux de récupération	– MW – *Log D* – Moment dipolaire – *Length* – *EqWidth* – SR (MgSO4) – Potentiel zêta – Angle de contact – pH – Pression – Taux de récupération	– MW – *Log D* – Moment dipolaire – *Length* – *EqWidth* – MWCO – SR (MgSO4) – Potentiel zêta – Angle de contact – pH – Pression – Taux de récupération
Sortie	Rétention	Rétention	Rétention

L'analyse statistique des données d'entrées et de sortie, en termes du minimum "min", de la moyenne arithmétique (Équation III.2) "$\bar{x}$.", du maximum "max" et de la déviation standard (écart type) "DS" (Équation III.3) (Khaouane 2013), sont présentées dans les tableaux III.4, III.5 et III.6.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \text{(III. 2)}$$

$$DS = \left(\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 \right)^{1/2} \quad \text{(III. 3)}$$

Tableau III.4. Analyse statistique des données d'entrée et de sortie (BD1).

	Min	**Moyenne**	**Max**	**DS**
MW (g/mol)	30.0300	228.7852	359.0000	56.2123
LogD	-3.7000	1.1666	5.9300	1.9107
Moment dipolaire (Debye)	0.0000	2.9373	10.9100	1.9680
***Length* (nm)**	0.4700	1.1803	1.8600	0.2660
***EqWidth* (nm)**	0.2800	0.6706	0.9200	0.1246
MWCO (Dalton)	100.0000	228.5042	400.0000	73.6627
Potentiel zêta (mv)	-87.0000	-26.3868	3.1000	23.3559
Angle de contact (°)	20.1000	49.0953	94.7000	12.9343
pH	2.6800	7.3252	10.0100	1.0963
Pression (Kpa)	240.0000	1150.0833	4100.0000	887.3774
Taux de récupération (%)	0.0300	30.4271	90.0000	30.4319
Rétention (%)	0.0000	74.5900	100.0000	27.5432

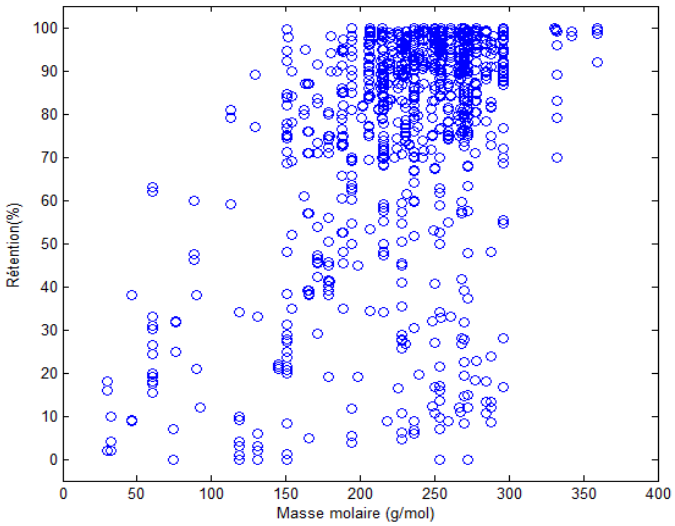
Tableau III.5. Analyse statistique des données d'entrée et de sortie (BD2).

	Min	**Moyenne**	**Max**	**DS**
MW (g/mol)	30.0300	220.9730	342.3000	57.9954
LogD	-3.7000	1.01061	5.9300	1.8676
Moment dipolaire (Debye)	0.0000	2.8131	9.7800	1.7782
***Length* (nm)**	0.4700	1.1971	1.8600	0.2920
***EqWidth* (nm)**	0.3800	0.6916	0.9200	0.1159
SR($MgSO_4$)	0.9600	0.9793	1.0000	0.0112
Potentiel zêta (mv)	-48.0400	-18.5079	1.0300	9.8039
Angle de contact (°)	33.1000	47.7787	70.7000	8.6766
pH	2.6800	7.2800	10.0100	1.1145
Pression (Kpa)	240.0000	1304.0209	2500.0000	938.2896
Taux de récupération (%)	0.5000	22.6818	90.0000	28.7509
Rétention (%)	0.0000	74.9709	100.0000	26.9937

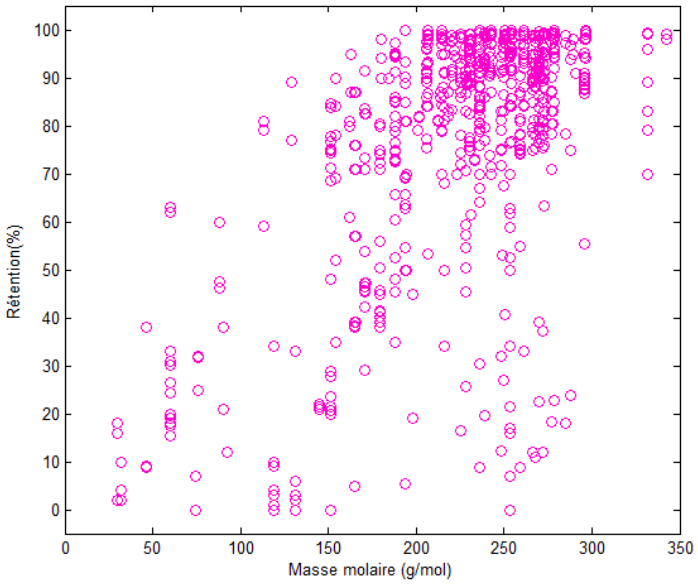
Tableau III.6. Analyse statistique des données d'entrée et de sortie (BD3).

	Min	Moyenne	Max	DS
MW (g/mol)	30.0300	221.1318	342.3000	58.1541
LogD	-3.7000	1.0073	5.9300	1.8701
Moment dipolaire (Debye)	0.0000	2.8387	10.9100	1.8202
***Length* (nm)**	0.4700	1.1967	1.8600	0.2919
***EqWidth* (nm)**	0.3800	0.6908	0.9200	0.1163
MWCO (Dalton)	100.0000	212.4108	300.0000	48.1371
SR($MgSO_4$)	0.9600	0.9792	1.0000	1.1220
Potentiel zêta (mv)	-48.0400	-18.5498	1.0300	9.8060
Angle de contact (°)	28.0000	47.8201	70.7000	8.7304
pH	2.6800	7.2793	10.0100	1.1149
Pression (Kpa)	240.0000	1310.8683	4100.0000	952.7401
Taux de récupération (%)	0.5000	22.3930	90.0000	28.4831
Rétention (%)	0.0000	74.9394	100.0000	26.9982

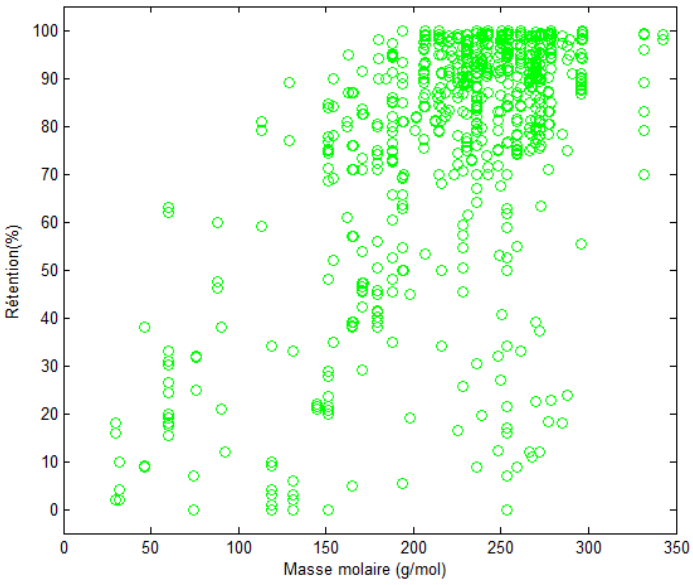
La figure III.1 représente les rétentions expérimentales en fonction de poids moléculaires pour les trois bases de données (les rétentions en fonction de poids moléculaires sont indiquées par des cercles bleus pour BD1 "956 points", des cercles roses pour BD2 "701 points" et des cercles verts pour BD3 "701 points"). Quand le poids moléculaire varié de 150g/mol à 300g/mol, les rétentions expérimentales qui ont des pourcentages supérieures à 60.5 % s'agglomèrent les unes aux autres.



(a)



(b)



(c)

Figure III.1. Rétentions expérimentales en fonction de poids moléculaires : (a) BD1, (b) BD2, (c) BD3.

Pour chaque base de données, les échantillons ont été divisés au hasard en trois sous-ensembles : 80% pour la phase d'apprentissage, 10% pour la phase de validation et 10% pour la phase de test (Figure III.2).

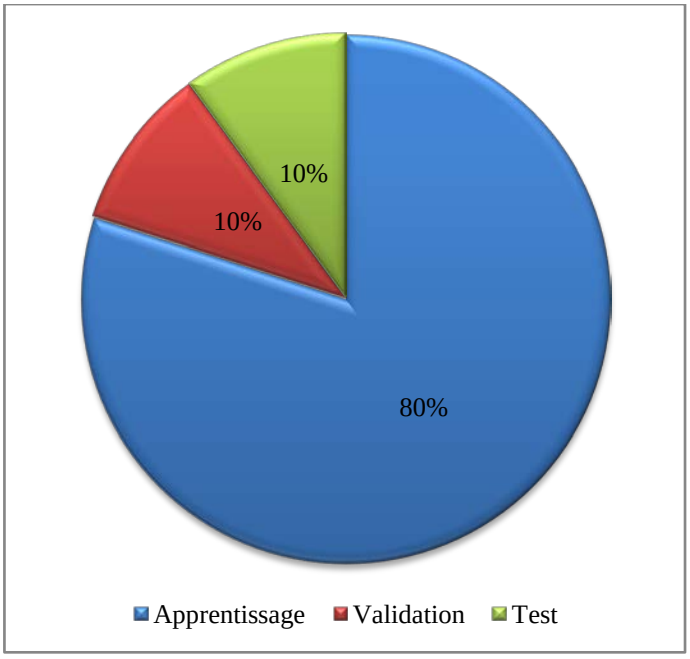


Figure III.2. Représentation des pourcentages de division de la base de données.

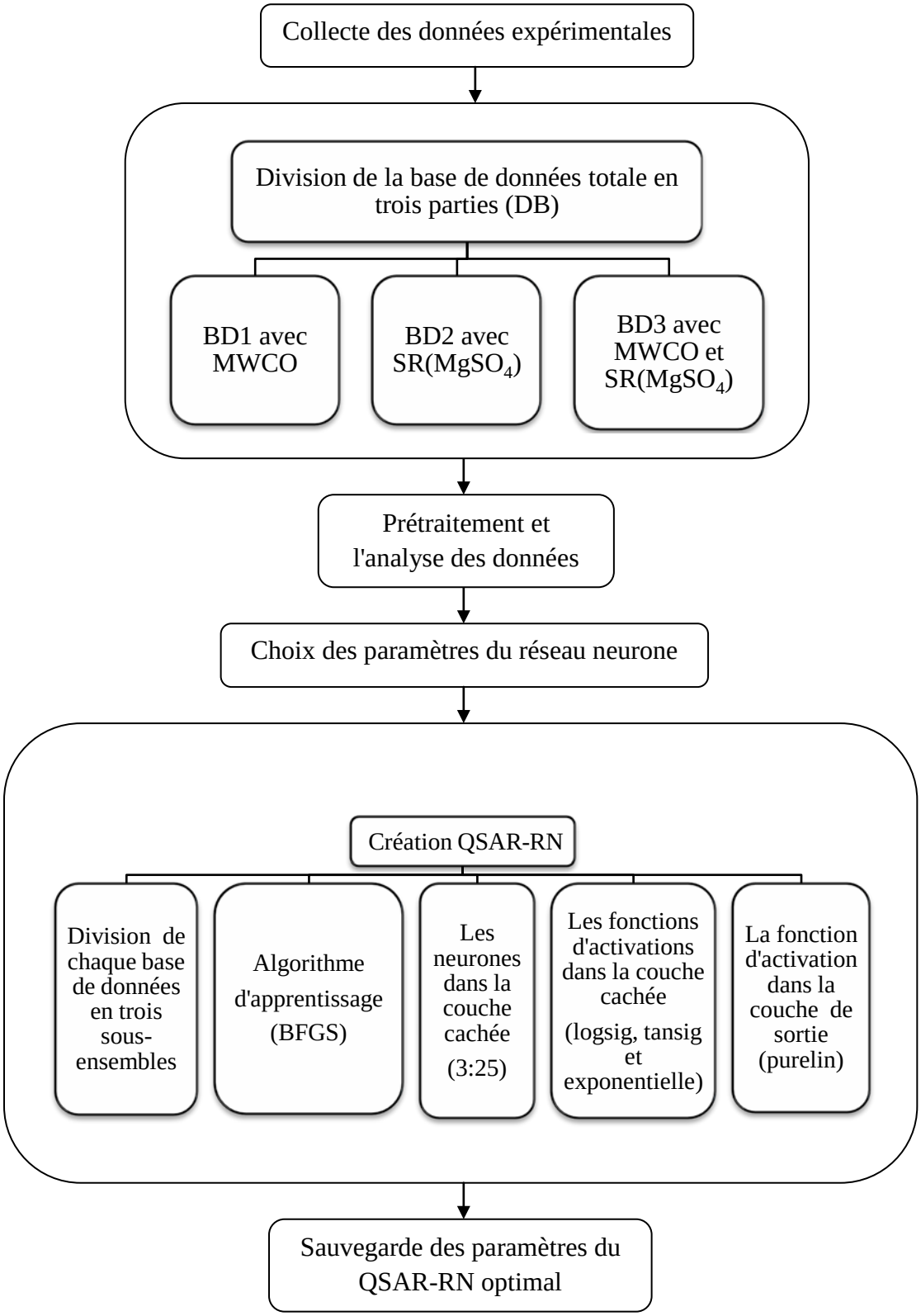
III.3. Développement des modèles QSAR

Cette étape concerne la création d'une relation statistiquement significative de la propriété à modéliser. Beaucoup de méthodes de développement de modèles sont largement utilisées. Les deux méthodes utilisées pour l'établissement des modèles QSAR sont les réseaux de neurones artificiels (RNA), la régression linéaire multiple (RLM).

III.3.1. Réseaux de neurones artificiels

Les réseaux de neurones artificiels (RNA) sont particulièrement bien adaptés pour les modèles QSAR/QSPR en raison de leur capacité à extraire des informations non-linéaires présentes dans la matrice de données (Yangjeh et Jenagharad 2009). Pour trouver un modèle non linéaire entre les propriétés des molécules organiques, les caractéristiques de la membrane et les conditions opératoires de filtration pour la rétention des molécules organiques par les membranes d'OI/NF, l'utilisation de la méthode RNA est nécessaire et est devenue une technique de modélisation importante. Elle est largement utilisée pour les études de QSAR non linéaires et est un outil puissant pour construire des modèles prédictifs.

Dans ce travail, une procédure basée sur le développement et l'optimisation de l'architecture d'un QSAR-RN est décrite. Elle est basée, comme décrite sur l'organigramme ci dessous.


Collecte des données expérimentales
Division de la base de données totale en trois parties (DB)
BD1 avec MWCO
BD2 avec SR($MgSO_4$)
BD3 avec MWCO et SR($MgSO_4$)
Prétraitement et l'analyse des données
Choix des paramètres du réseau neurone
Création QSAR-RN
Division de chaque base de données en trois sous-ensembles
Algorithme d'apprentissage (BFGS)
Les neurones dans la couche cachée (3:25)
Les fonctions d'activations dans la couche cachée (logsig, tansig et exponentielle)
La fonction d'activation dans la couche de sortie (purelin)
Sauvegarde des paramètres du QSAR-RN optimal

L'algorithme d'apprentissage utilisé dans ce travail est le BFGS quasi-Newton (trainbfg). Chaque RN contient trois couches de neurones ou des nœuds : une couche d'entrée avec onze neurones pour RN1 et RN2 et douze neurones pour RN3 dans la couche d'entrée, une couche cachée avec un nombre de neurones actifs optimisés lors de l'apprentissage et une couche de sortie avec une unité générant la valeur estimée de la rétention (R). Le nombre de neurones dans la couche cachée a été varié de 3 à 25 neurones. Les fonctions d'activations « La tangente hyperbolique sigmoïde (tansig), la sigmoïde logarithmique (logsig) et l'exponentielle » ont été utilisées dans la couche cachée. La fonction de transfert linéaire (purelin) a été utilisée dans la couche de sortie.

La procédure d'élaboration du QSAR-RN décrite ci-dessus a été réalisée en utilisant STATISTICA (Annexe B). Ce logiciel est utilisé pour optimiser l'architecture des trois modèles neuronaux (QSAR-RN1, QSAR-RN2 et QSAR-RN3) employés pour la rétention des molécules organiques par les membranes d'OI/NF.

III.3.2. Régression linéaire multiple

La régression linéaire multiple est un outil statistique dont l'objectif est d'établir une relation linéaire entre une propriété d'intérêt (rétention par exemple) et un ensemble de variables (les propriétés des molécules organiques, les caractéristiques de la membrane et les conditions opératoires de filtration) et ce suivant l'équation III.4 :

$$Rétention = a_0 + \sum_{i=1}^{n} a_i d_i \qquad \text{(III. 4)}$$

où d_i représente les entrées, a_i leur coefficient et a_0 est une constante (appelée intercepté).

Les mêmes bases de données utilisées pour le développement de QSAR-RN ont été utilisées pour développer QSAR-RLM par utilisation de la fonction MATLAB ***"regress***" (voir annexe B).

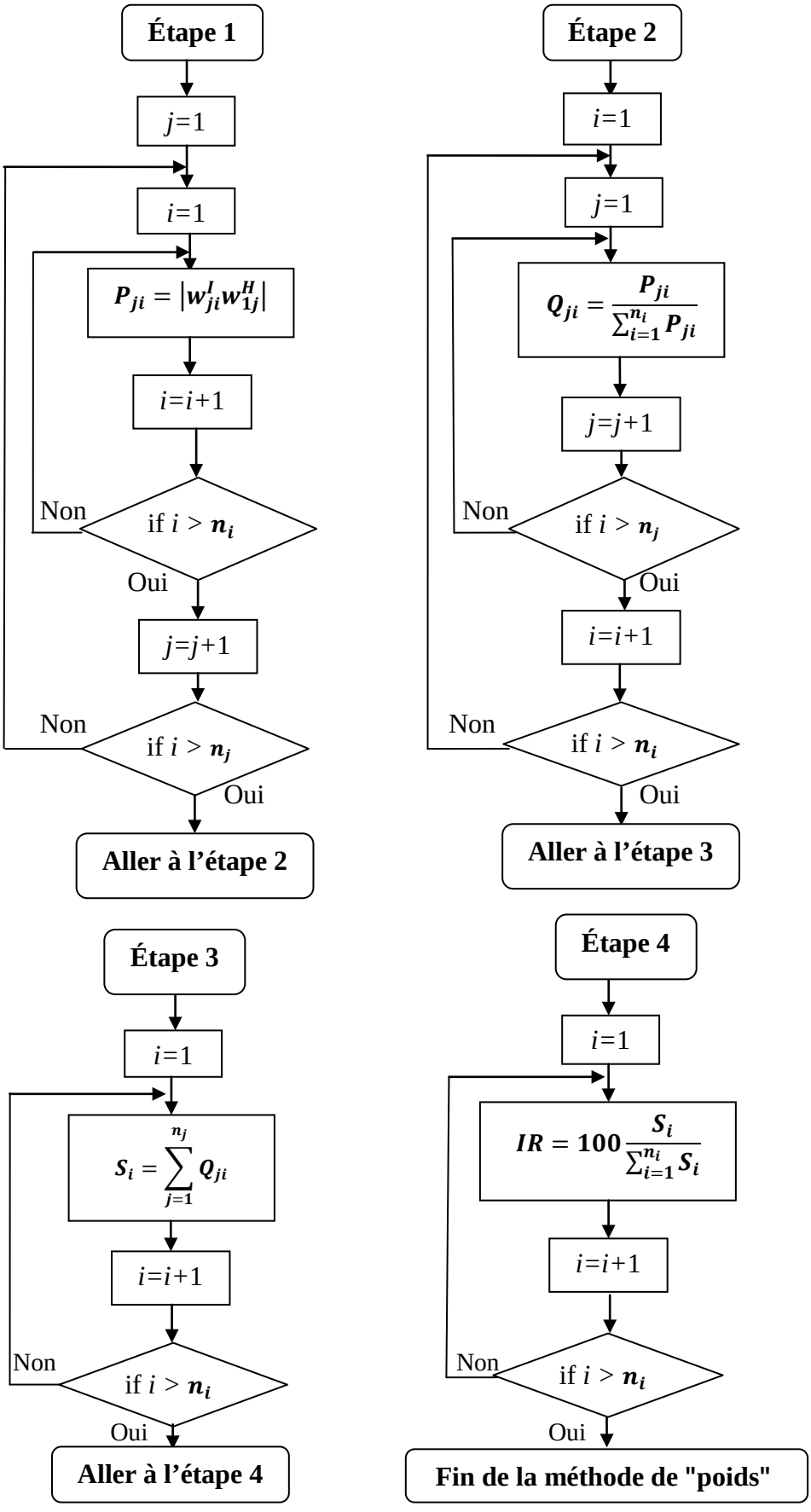
III.4. Analyse de sensibilité par la méthode de "poids"

La contribution des variables d'entrées (MW, *LogD*, *Lenght*, *EqWidth*, MWCO, SR, angle de contact, potentiels zêta, pH, pression et taux de récupération) sur la sortie a été déterminée par une analyse de sensibilité en utilisant la méthode de "poids" pour chaque réseau de neurone (QSAR-RN1, QSAR-RN2 et QSAR-RN3).

Cette méthode, proposée d'abord par Garson (Garson 1991) et reprise par Goh (Goh 1995), dnne une quantification de l'importance relative (*IR*) des différentes entrées sur la sortie du RN (Gevrey et al. 2003). Elle se base sur le partitionnement des poids de connexions en :

- Poids de connexions entrée-cachée (w_{ji}^{I}) ;
- Poids de connexions cachée-sortie (w_{1j}^{H}).

Le processus de calcul, compose de quatre étapes, est schématisé sur l'organigramme ci-dessous (Khaouane 2013).


Étape 1
j=1
i=1
$P_{ji} = |w^I_{ji} w^H_{1j}|$
i=i+1
Non
if i > n_i
Oui
j=j+1
Non
if i > n_j
Oui
Aller à l'étape 2
Étape 2
i=1
j=1
$Q_{ji} = \frac{P_{ji}}{\sum_{i=1}^{n_i} P_{ji}}$
j=j+1
Non
if i > n_j
Oui
i=i+1
Non
if i > n_i
Oui
Aller à l'étape 3
Étape 3
i=1
$S_i = \sum_{j=1}^{n_j} Q_{ji}$
i=i+1
Non
if i > n_i
Oui
Aller à l'étape 4
Étape 4
i=1
$IR = 100 \frac{S_i}{\sum_{i=1}^{n_i} S_i}$
i=i+1
Non
if i > n_i
Oui
Fin de la méthode de "poids"

III.5. Conclusion

Nous avons décrit ici les outils nécessaires à la modélisation de la rétention des molécules organiques lors d'osmose inverse et de nanofiltration par les modèles QSAR (RNA et RLM). Dans cette étape, nous avons présenté les données collectées, prétraitement et analyse. Après nous avons présenté les techniques les plus récentes pour établir des modèles QSAR : la régression linéaire multiple (RLM) et les réseaux de neurones artificiels (RNA) et le développement des modèles QSAR. Nous avons terminé ce chapitre par une analyse de sensibilité basée sur la méthode de "poids".

Le chapitre suivant sera consacré aux résultats et discussion pour mieux comprendre le phénomène étudié par les différentes méthodes.

Chapitre IV
Résultats et discussion

IV.1. Introduction

Les résultats de la modélisation de la rétention des molécules organiques par les membranes d'OI/NF ont fait objet d'une publication dans le "The Korean Journal of Chemical Engineering" (Ammi et al. 2015). Dans ce chapitre, les résultats obtenus sont discutés, il est organisé en quatre parties :

- ✓ Dans la première partie, on présentera les performances des modèles QSAR-RN. Ensuite l'extraction de formules mathématiques à partir de ces modèles. Par la suite, on présentera les résultats de l'analyse de sensibilité par la méthode de "poids". Une comparaison entre la rétention expérimentale et calculée par QSAR-RN3 est ensuite présentée.
- ✓ Une discussion sur les modèles QSAR- RLM et les résultats de l'analyse de sensibilité par la méthode de régression linéaire est présentée dans la deuxième partie.
- ✓ La troisième partie est consacrée à une étude comparative des modèles QSAR-RN avec QSAR- RLM et avec les modèles QSAR de la littérature.
- ✓ La quatrième partie est consacrée à l'étude de la rétention des composés organiques en fonction du pH.

IV.2. Modèles QSAR-RN

IV.2.1. Performances des modèles QSAR-RN

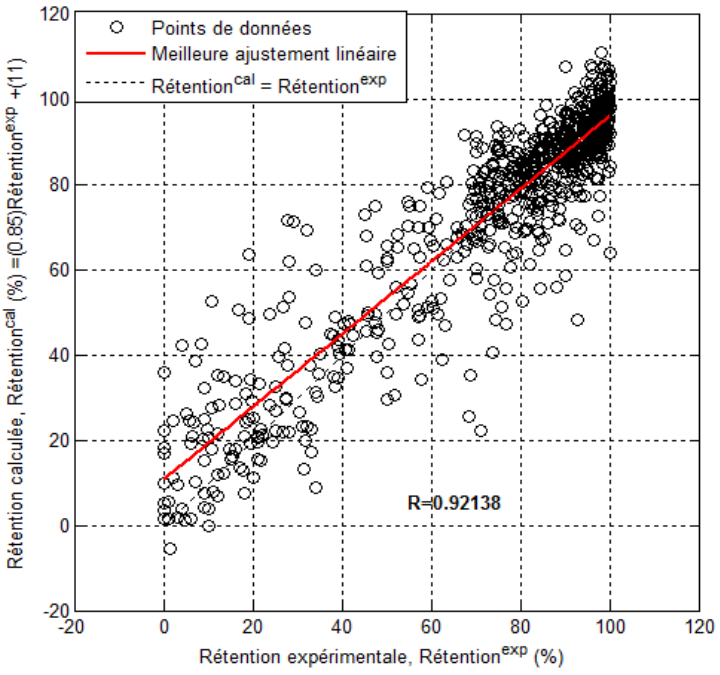
Dans cette étude, pour chaque base de données (BD1, BD2 et BD3) les échantillons, ont été divisés au hasard en trois sous-ensembles : 80% pour la phase d'apprentissage, 10% pour la phase de validation et 10% pour la phase de test. Le tableau IV.1 montre la structure des modèles QSAR-RN optimisés.

Table IV.1. Structure des modèles RN optimisés.

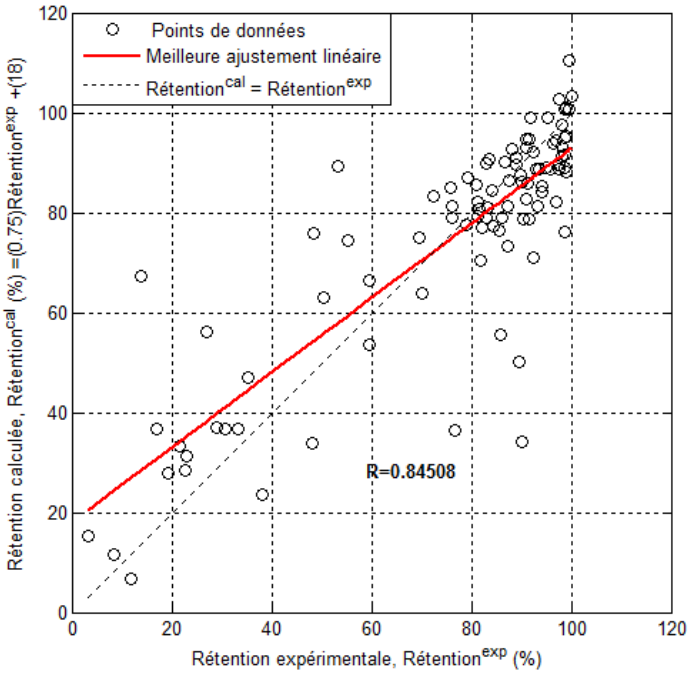
Modèles de QSAR-RN	Algorithme d'apprentissage	Couche entrée	Couche cachée		Couche sortie	
		Nombre de neurone	Nombre de neurone	Fonction d'activation	Nombre de neurone	Fonction d'activation
QSAR-RN1	BFGS quasi-Newton ***(trainbfg)***	11	15	Tansig	1	purelin
QSAR-RN2	BFGS quasi-Newton ***(trainbfg)***	11	24	Tansig	1	purelin
QSAR-RN3	BFGS quasi-Newton ***(trainbfg)***	12	24	Tansig	1	purelin

Selon la discussion précédente, trois modèles de réseaux neurones ont été développés dans le but de prédire la rétention des molécules organiques par les membranes d'OI/NF. L'intrigue et les paramètres de la régression linéaire sont franchement obtenus en utilisant la fonction MATLAB "***postreg***". Le pouvoir de prédiction d'un tel modèle peut être estimé par le coefficient de corrélation *R* calculé selon (Équation III.1), une valeur de *R* supérieure à 0.5 est généralement considérée comme satisfaisante, une valeur supérieure à 0.9 est excellente (Golbraikh et Tropsha 2002).

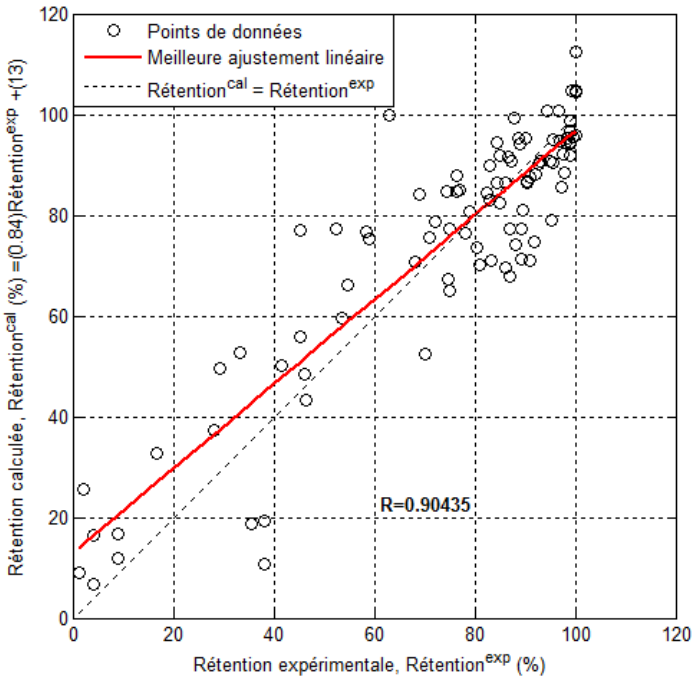
La figure IV.1 montre les courbes de régression linéaire des taux de rétentions calculés par le premier modèle neuronal "QSAR-RN1" optimisé avec les taux de rétentions expérimentales avec des vecteurs de régression approchant de l'idéal ([α, β, R] = [0.8515, 10.9695, 0.9214] "phase d'apprentissage" ; [α, β, R] = [0.7484, 18.2583, 0.8451] "phase de validation" ; [α, β, R] = [0.8384, 13.1733, 0.9043] "phase de test" et [α, β, R] = [0.8407, 11.8426, 0.9128] "phase totale"). Les coefficients des corrélations sont généralement considérés comme excellentes ($0.90 \leq R \leq 1.00$) pour la phase d'apprentissage, la phase de test et la phase totale. Le coefficient de corrélation est couramment considéré comme satisfaisant ($0.50 \leq R < 0.90$) pour la phase de validation, ce qui montre une bonne robustesse du premier modèle neuronal "QSAR-RN1" établi et la possibilité de la prédiction des différents paramètres qui caractérisent la rétention des molécules organiques lors d'OI/NF.



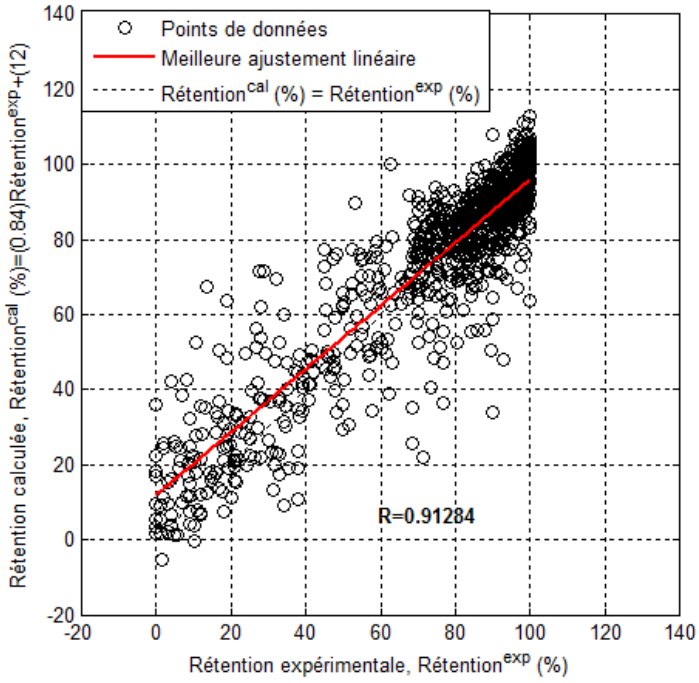
(a)


Points de données
Meilleure ajustement linéaire
Rétention^cal = Rétention^exp
R=0.84508
Rétention calculée, Rétention^cal (%) =(0.75)Rétention^exp +(18)
Rétention expérimentale, Rétention^exp (%)
0
20
40
60
80
100
120

(b)


Points de données
Meilleure ajustement linéaire
Rétention^cal = Rétention^exp
R=0.90435
Rétention calculée, Rétention^cal (%) =(0.84)Rétention^exp +(13)
Rétention expérimentale, Rétention^exp (%)
0
20
40
60
80
100
120

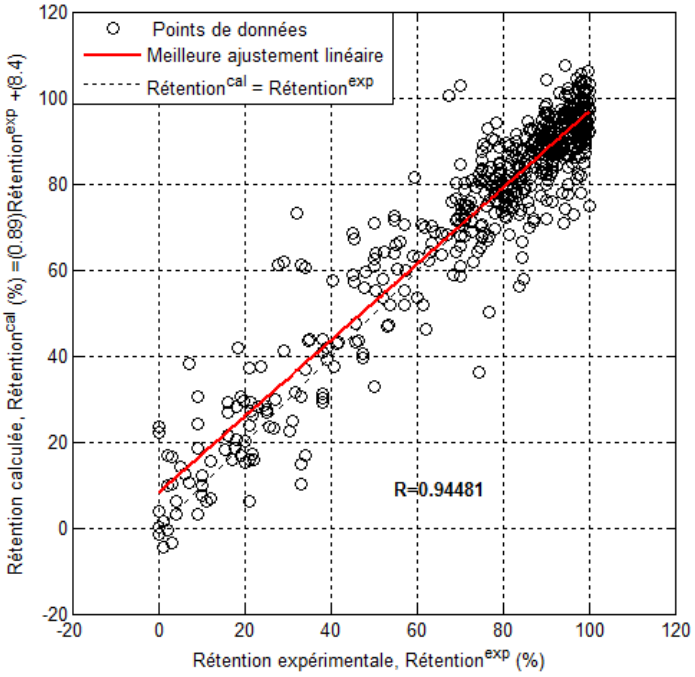
(c)



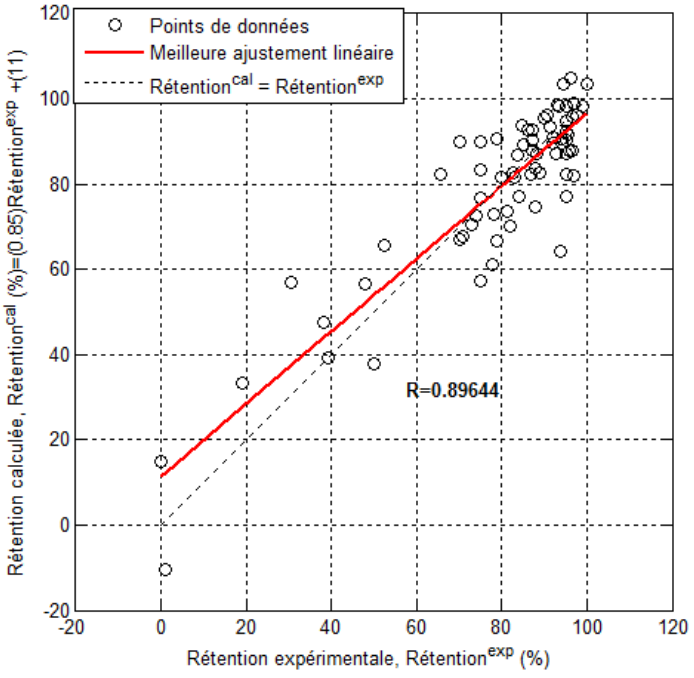
(d)

Figure IV.1. Comparaison entre les valeurs expérimentales et les valeurs calculées pour QSAR-RN1 : (a) la phase d'apprentissage, (b) la phase de validation, (c) la phase de test, (d) la phase totale.

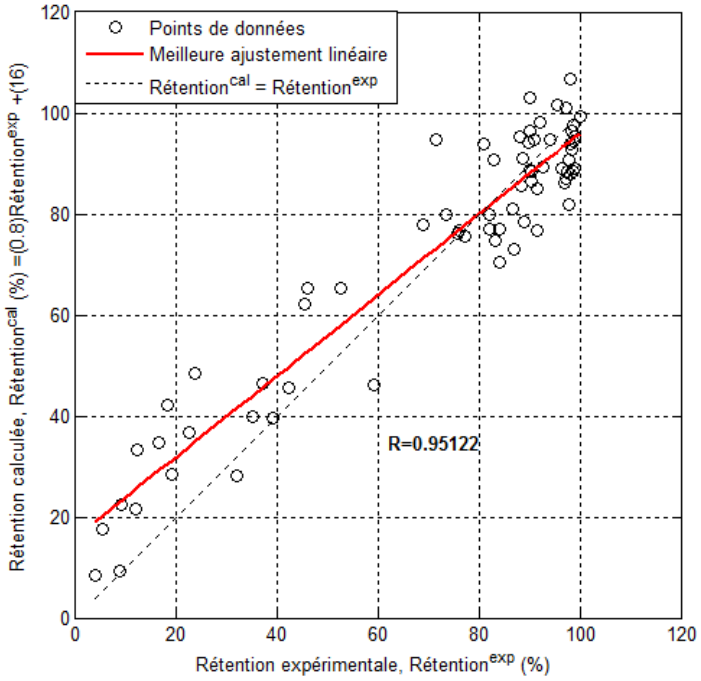
Pour évaluer la performance de deuxième modèle neuronal "QSAR-RN2" obtenu, la représentation graphique des valeurs prédites de la rétention en fonction des valeurs expérimentales pour les ensembles d'apprentissage, de validation, de test et totale obtenues est donnée sur la figure IV.2. Elle montre les courbes de régression linéaire de deuxième modèle neuronal avec des vecteurs de régression approchant de l'idéal ([α, β, R] = [0.8865, 8.3836, 0.9448] "phase d'apprentissage" ; [α, β, R] = [0.8522, 11.4167, 0.8964] "phase de validation" ; [α, β, R] = [0.8024, 15.9668, 0.9512] "phase de test" et [α, β, R] = [0.8731, 9.5688, 0.9419] "phase totale"). Les coefficients des corrélations sont généralement considérés comme excellentes ($0.90 \leq R \leq 1.00$) pour la phase d'apprentissage, la phase de test et la phase totale. Le coefficient de corrélation est couramment considéré comme satisfaisant ($0.50 \leq R < 0.90$) pour la phase de validation, ce qu'implique une bonne robustesse du deuxième modèle neuronal "QSAR-RN2" établi et la possibilité de la prédiction des différents paramètres qui caractérisent la rétention des molécules organiques lors d'OI/NF.


Points de données
Meilleure ajustement linéaire
Rétention^cal = Rétention^exp
R=0.94481
Rétention calculée, Rétention^cal (%) =(0.89)Rétention^exp +(8.4)
Rétention expérimentale, Rétention^exp (%)

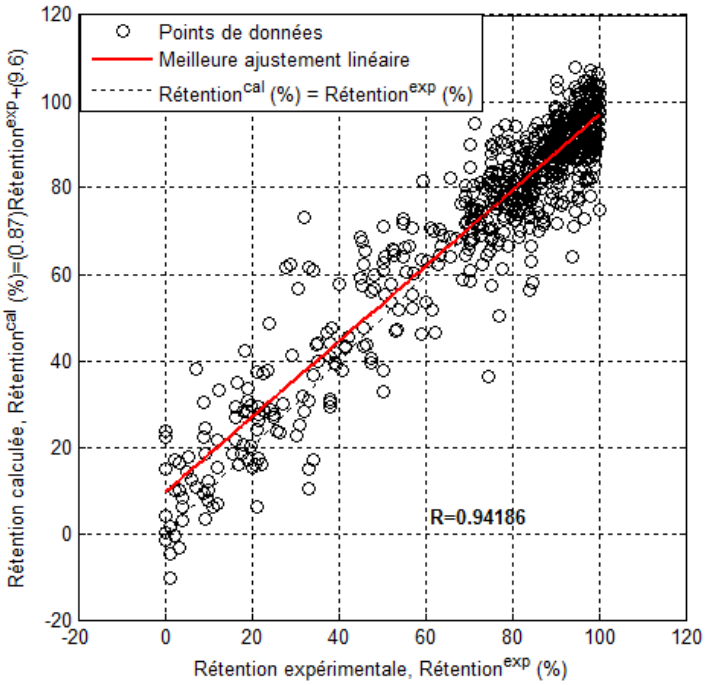
(a)


Points de données
Meilleure ajustement linéaire
Rétention^cal = Rétention^exp
R=0.89644
Rétention calculée, Rétention^cal (%)=(0.85)Rétention^exp +(11)
Rétention expérimentale, Rétention^exp (%)

(b)



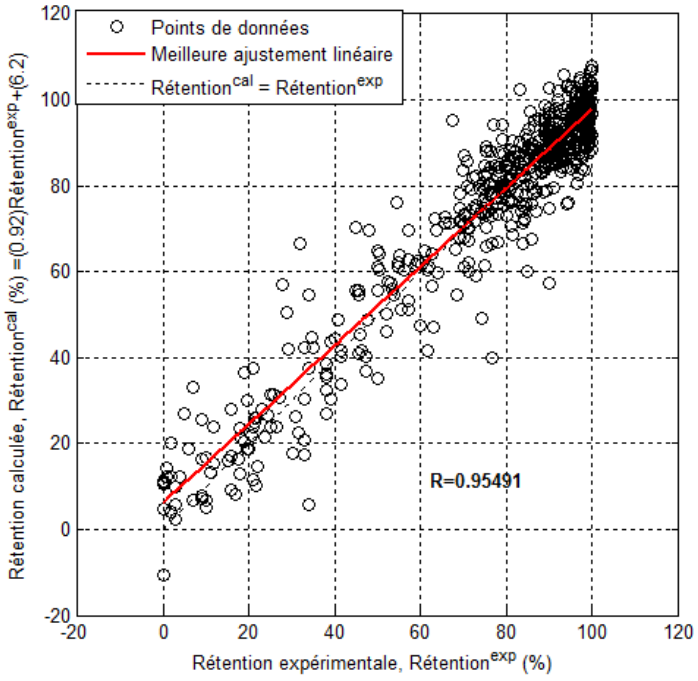
(c)



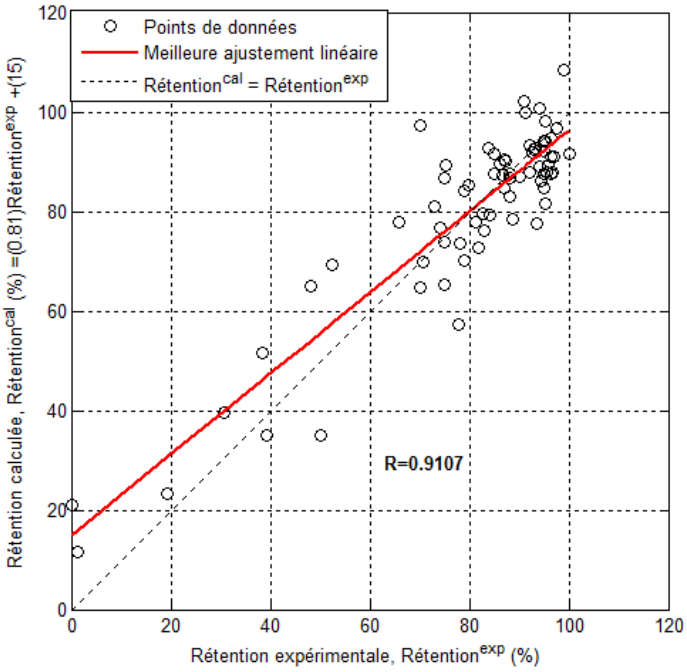
(d)

Figure IV.2. Comparaison entre les valeurs expérimentales et les valeurs calculées pour QSAR-RN2 : (a) la phase d'apprentissage, (b) la phase de validation, (c) la phase de test, (d) la phase totale.

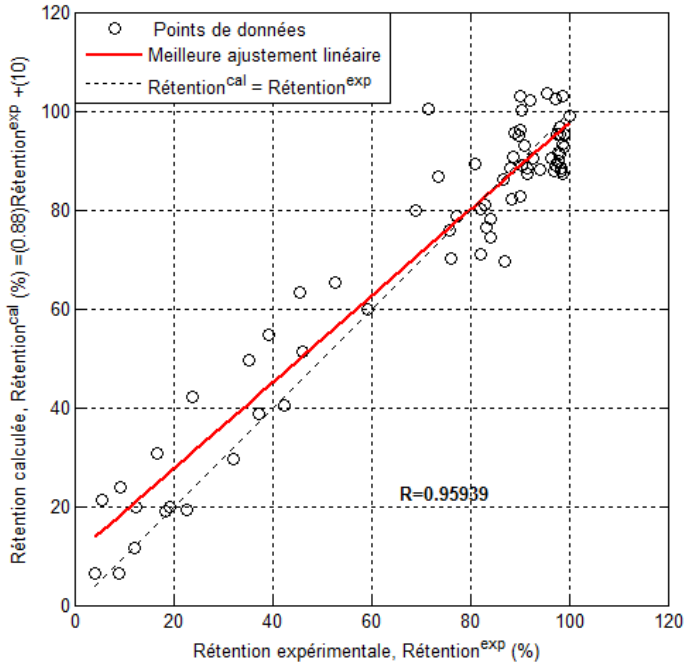
La Figure IV.3 visualise la comparaison entre les valeurs expérimentales et les valeurs calculées par le troisième modèle neuronal "QSAR-RN3". Elle montre les courbes de régression linéaire du troisième modèle neuronal avec des vecteurs de régression approchant de l'idéal ([α, β, R] = [0.8141, 15.0969, 0.9549] "phase d'apprentissage" ; [α, β, R] = [0.8754, 10.3062, 0.9107] "phase de validation" ; [α, β, R] = [0.9050, 7.2802, 0.9594] "phase de test" et [α, β, R] = [0.9050, 7.2802, 0.9527] "phase totale"). Les coefficients des corrélations sont généralement considérés comme excellentes ($0.90 \leq R \leq 1.00$) pour la phase d'apprentissage, la phase de validation, la phase de test et la phase totale, ce qui montre une bonne robustesse du troisième modèle neuronal "QSAR-RN3" établi et la possibilité de la prédiction des différents paramètres qui caractérisent la rétention des molécules organiques lors d'OI/NF.



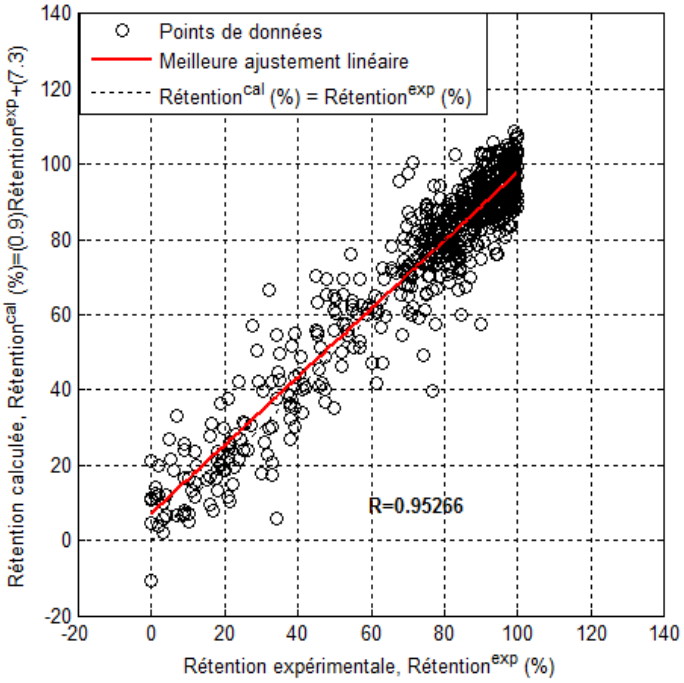
(a)


Points de données
Meilleure ajustement linéaire
Rétention^{cal} = Rétention^{exp}
Rétention calculée, Rétention^{cal} (%) =(0.81)Rétention^{exp} +(15)
Rétention expérimentale, Rétention^{exp} (%)
R=0.9107
0
20
40
60
80
100
120

(b)


Points de données
Meilleure ajustement linéaire
Rétention^{cal} = Rétention^{exp}
Rétention calculée, Rétention^{cal} (%) =(0.88)Rétention^{exp} +(10)
Rétention expérimentale, Rétention^{exp} (%)
R=0.95939
0
20
40
60
80
100
120

(c)



(d)

Figure IV.3. Comparaison entre les valeurs expérimentales et les valeurs calculées pour QSAR-RN3 : (a) la phase d'apprentissage, (b) la phase de validation, (c) la phase de test, (d) la phase totale.

La performance de chaque modèle a été évaluée en termes de coefficient de corrélation (R) et la racine de l'erreur quadratique moyenne (*REQM*). *REQM* a été utilisée pour déterminer l'erreur de modélisation entre les valeurs expérimentales et calculées. La racine de l'erreur quadratique moyenne est définie comme suit (Khaouane 2013 ; Hemmati-Sarapardeh et al. 2014 ; Faisal et al. 2015 ; Kolluri et al. 2015) :

$$REQM = \sqrt{\frac{\sum_{i=1}^{n}(Y_{i,exp} - Y_{i,cal})^2}{n}} \qquad (IV.1)$$

Où n est le nombre total de points de données, $Y_{i,exp}$ est la valeur expérimentale et $Y_{i,cal}$ représente la valeur calculée à partir du modèle QSAR.

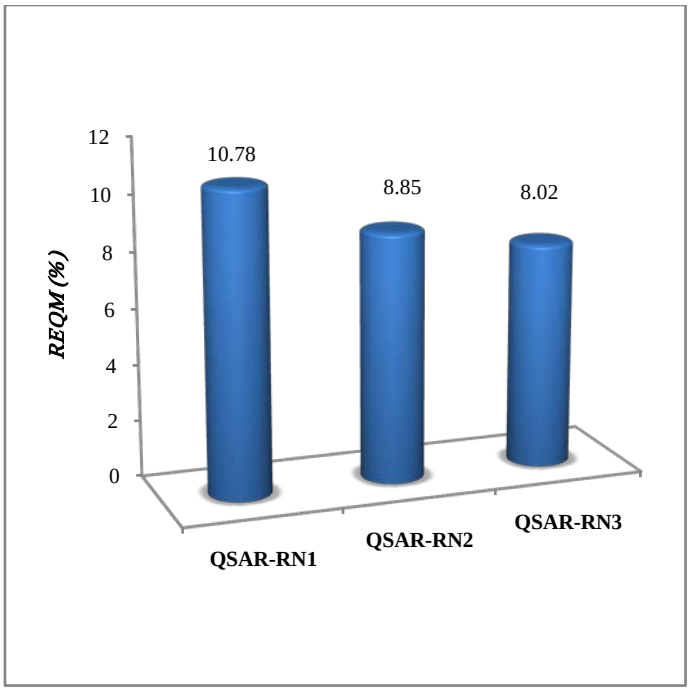
Les vecteurs de régression linéaire pour les trois modèles neuronaux (QSAR-RN1, QSAR-RN2 et QSAR-RN3) sont présentés dans le tableau IV.2. Ce-ci montre que les paramètres de la régression les moins favorables sont ceux obtenus dans QSAR-RN1. Il est prouvé que l'approche neuronale proposée donne des résultats satisfaisants avec des valeurs de

vecteur de régression approchant à l'idéal [équation linéaire : $y^{cal} = \alpha y^{exp} + \beta$, ***α*=1** (pente), ***β*=0** (ordonnée à l'origine), ***R*=1** (coefficient de corrélation)] dans l'ajustement des profils des QSAR-RN1, QSAR-RN2 et QSAR-RN3.

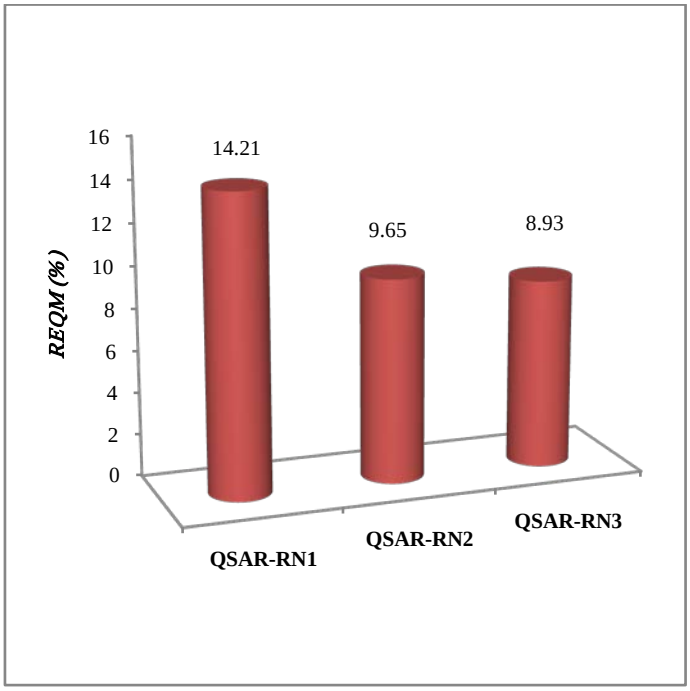
Table IV.2. Vecteurs de régression linéaire issue de l'équation linéaire : [équation linéaire : $y^{cal} = \alpha y^{exp} + \beta$, ***α***(pente), ***β***(ordonnée à l'origine), ***R***(coefficient de corrélation)] et la racine de l'erreur quadratique moyenne (***REQM***).

		α	***β***	***R***	***REQM***
QSAR-RN1	**Phase d'apprentissage**	0.8515	10.9695	0.9214	10.7808
	Phase de validation	0.7484	18.2583	0.8451	14.2107
	Phase de test	0.8384	13.1733	0.9043	11.5299
	Totale	0.8407	11.8426	0.9128	11.2430
QSAR-RN2	**Phase d'apprentissage**	0.8865	8.3836	0.9448	8.8558
	Phase de validation	0.8522	11.4167	0.8964	9.6501
	Phase de test	0.8024	15.9668	0.9512	10.1455
	Totale	0.8731	9.5688	0.9419	9.0742
QSAR-RN3	**Phase d'apprentissage**	0.8141	15.0969	0.9549	8.0246
	Phase de validation	0.8754	10.3062	0.9107	8.9305
	Phase de test	0.9050	7.2802	0.9594	8.8490
	Totale	0.9169	6.2129	0.9527	8.2047

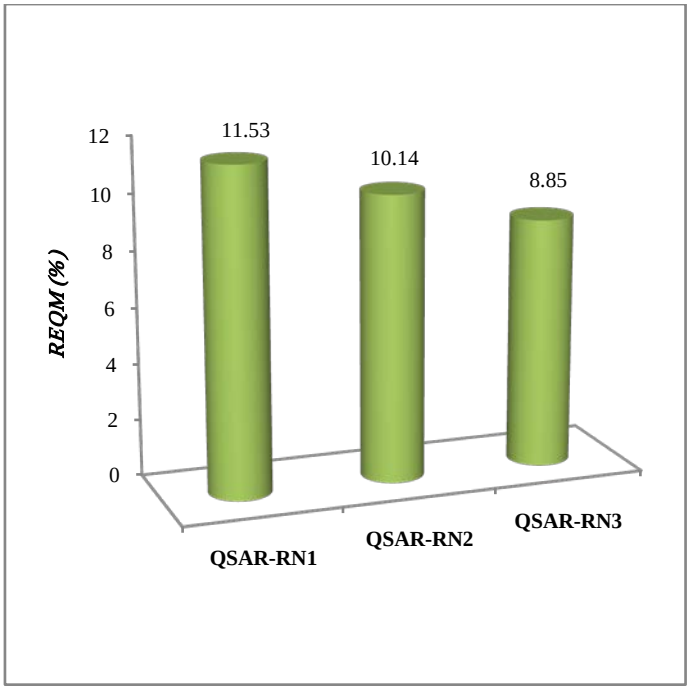
La comparaison entre les trois modèles QSAR-RN pour chaque phase (apprentissage, validation, test et totale), en termes de la *REQM* est présentée sur l'histogramme de la figure IV.4, elle montre que le troisième modèle neuronal "QSAR-RN3" a donné des erreurs plus faibles que d'autres modèles (QSAR-RN1 et QSAR-RN2). Cette comparaison indique que le QSAR-RN3 possède une capacité de modélisation élevée par rapport aux modèles (QSAR-RN1 et QSAR-RN2).


12
10
8
6
4
2
0
REQM (%)
10.78
8.85
8.02
QSAR-RN1
QSAR-RN2
QSAR-RN3

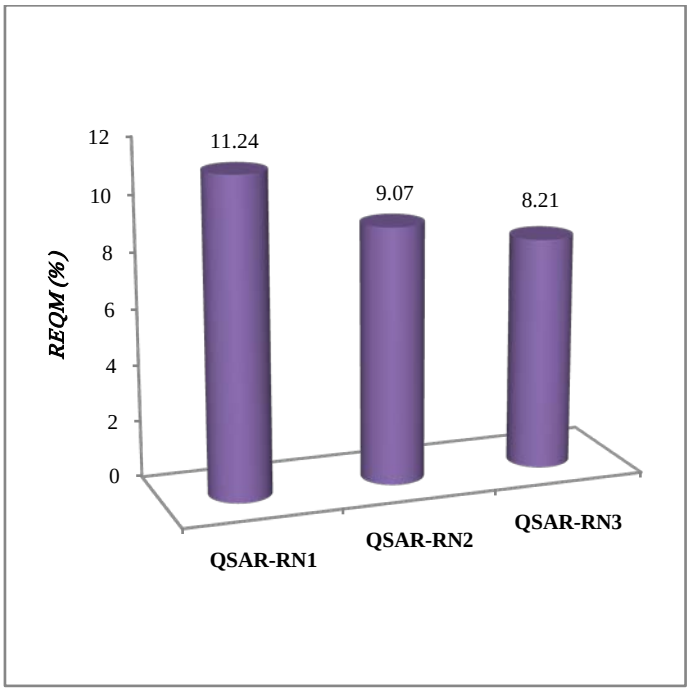
(a)


16
14
12
10
8
6
4
2
0
REQM (%)
14.21
9.65
8.93
QSAR-RN1
QSAR-RN2
QSAR-RN3

(b)



(c)



(d)

Figure IV.4. Performances des modèles neuronaux (QSAR-RN*1*, QSAR-RN*2* et QSAR-RN*3*) pour : (a) la phase d'apprentissage, (b) la phase de validation, (c) la phase de test, (d) la phase totale.

IV.2.2. Expressions mathématiques des modèles neuronaux optimisés

À partir du QSAR-RN optimisé, représenté sur la figure IV.5, on peut exprimer l'assimilation de la rétention des molécules organiques (neutres et ioniques) par les membranes d'OI/NF par un modèle mathématique y incorporant toutes les entrées *Ei* (MW, *LogD*, moment dipolaire, *Length*, *EqWidth*, *MWCO*, SR($MgSO_4$), angle de contact, potentiel zêta, pH, pression et taux de récupération) comme suit :

Les sorties Z_j de la couche cachée :

$$Z_j = f_H\left[\sum_{i=1}^{11} w_{ji}^I E_i + b_j^H\right] = \frac{exp\left(\sum_{i=1}^{11} w_{ji}^I E_i + b_j^H\right) - exp\left(-\sum_{i=1}^{11} w_{ji}^I E_i + b_j^H\right)}{exp\left(\sum_{i=1}^{11} w_{ji}^I E_i + b_j^H\right) + exp\left(-\sum_{i=1}^{11} w_{ji}^I E_i + b_j^H\right)} \quad (IV.2)$$

j=1, 2, …,15

La sortie "rétention" :

$$Rétention\ (RN1) = f_0\left[\sum_{j=1}^{15} w_{1j}^H Z_j + b_1^o\right] = \sum_{j=1}^{15} w_{1j}^H Z_j + b_1^o \quad (IV.3)$$

La combinaison des équations (VI.2) et (VI.3) mène à la formule mathématique suivante qui décrit l'assimilation de rétention en prenant en compte toutes les entrées *Ei* (MW, *LogD*, moment dipolaire, *Length*, *EqWidth*, *MWCO*, angle de contact, potentiel zêta, *pH*, pression et taux de récupération) :

$$Rétention\ (RN1) = \sum_{j=1}^{15} w_{1j}^H \frac{exp\left(\sum_{i=1}^{11} w_{ji}^I E_i + b_j^H\right) - exp\left(-\sum_{i=1}^{11} w_{ji}^I E_i + b_j^H\right)}{exp\left(\sum_{i=1}^{11} w_{ji}^I E_i + b_j^H\right) + exp\left(-\sum_{i=1}^{11} w_{ji}^I E_i + b_j^H\right)} + b_1^o \quad (IV.4)$$

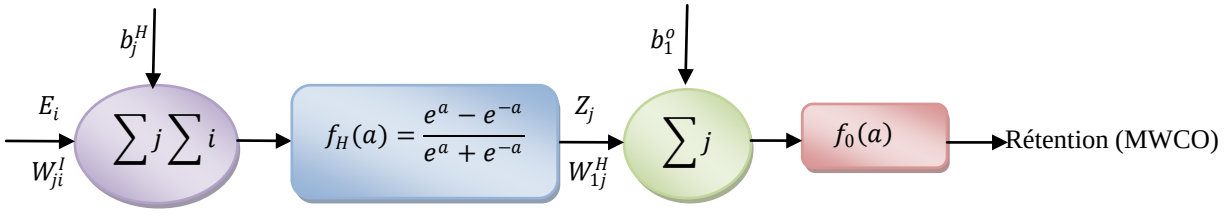


Figure IV.5. Représentation schématique du QSAR-RN1 optimisés.

De même, la rétention (SR "$MgSO_4$") et la rétention (MWCO et SR "$MgSO_4$") peuvent être exprimées par des équations mathématiques extraites des QSAR-RN2 et QSAR-RN3 optimisées comme suit :

$$Rétention\ (RN2) = \sum_{j=1}^{24} w_{1j}^H \frac{exp\left(\sum_{i=1}^{11} w_{ji}^I E_i + b_j^H\right) - exp\left(-\sum_{i=1}^{11} w_{ji}^I E_i + b_j^H\right)}{exp\left(\sum_{i=1}^{11} w_{ji}^I E_i + b_j^H\right) + exp\left(-\sum_{i=1}^{11} w_{ji}^I E_i + b_j^H\right)} + b_1^o \quad (IV.5)$$

$$Rétention\ (RN3) = \sum_{j=1}^{24} w_{1j}^{H} \frac{exp\left(\sum_{i=1}^{12} w_{ji}^{I} E_i + b_j^{H}\right) - exp\left(-\sum_{i=1}^{12} w_{ji}^{I} E_i + b_j^{H}\right)}{exp\left(\sum_{i=1}^{12} w_{ji}^{I} E_i + b_j^{H}\right) + exp\left(-\sum_{i=1}^{12} w_{ji}^{I} E_i + b_j^{H}\right)} + b_1^{o} \qquad \text{(IV. 6)}$$

Ou w_{ji}^{I} représente le poids synaptique de la connexion du neurone *i* de la couche d'entrée au neurone *j* de la couche cachée (15 lignes x 11 colonnes pour RN1, 24 lignes x 11 colonnes pour RN2 et RN3); w_{1j}^{H} représente le poids synaptique de la connexion du neurone *j* de la couche cachée au neurone de la couche de sortie (15 lignes x 1 colonnes pour RN1, 24 lignes x 1 colonnes pour RN2 et RN3); b_j^{H} représente les biais du neurone *j* de la couche cachée (15 lignes pour RN1, 24 lignes pour RN2 et RN3) et b_1^{o} représente les biais du neurone *j* de la couche de sortie (1 ligne pour RN1, RN2 et RN3). Les matrices des poids et biais, des trois QSAR-RN optimisés, sont donnés dans les tableaux IV.3, IV.4 et IV.5.

Les équations mathématiques, dérivant des modèles neuronaux optimisés, contiennent juste le degré exigé de complexité et incluent dans leurs formules les caractéristiques pertinentes de la rétention des molécules organiques (neutres et ioniques) par les membranes d'OI/NF à savoir les propriétés des molécules organiques, les caractéristiques de la membrane et les conditions opératoires de filtration. Ils peuvent donc facilement être appliqués dans des boucles de contrôle du processus de la rétention des molécules organiques par les membranes d'OI/NF.

Tableau IV.3. Poids et biais du QSAR-RN1 optimal.

Connexions Couche d'entrée-Couche cachée												**Connexions Couche cachée-Couche de sortie**	
Poids											**Biais**	**Poids**	**Biais**
$w_{j,1}^{I}$	$w_{j,2}^{I}$	$w_{j,3}^{I}$	$w_{j,4}^{I}$	$w_{j,5}^{I}$	$w_{j,6}^{I}$	$w_{j,7}^{I}$	$w_{j,8}^{I}$	$w_{j,9}^{I}$	$w_{j,10}^{I}$	$w_{j,11}^{I}$	b_{j}^{H}	$w_{I,j}^{H}$	b_{1}^{0}
0.38674	-0.54036	-0.67138	0.65097	-0.88769	1.43593	-1.87786	-2.36854	-0.12673	1.21642	0.46707	-0.63310	-0.14232	0.45101
-0.18181	-0.38531	-1.60174	1.49593	-0.41662	-0.12234	3.33611	-0.96803	-1.21867	1.77047	-0.27249	-0.44357	-0.50611	
-0.78903	-1.46645	-0.32800	0.45214	-0.03833	-0.18646	0.52582	0.04644	-1.37065	-0.92767	0.06487	0.64998	0.20779	
1.63317	0.59088	-0.18504	-1.86023	-0.78768	0.64211	0.42551	-0.39858	-1.21832	-1.60476	-1.46163	0.21620	-1.05660	
1.78689	-1.42073	0.45775	-0.00996	-0.86958	-5.46996	-1.93987	1.27965	3.26427	-0.46953	-0.60615	-4.42260	0.41303	
-0.15742	0.96063	-0.41441	-1.37880	0.17450	-1.77087	0.82407	4.98539	3.39085	1.78831	1.20171	1.81029	0.97004	
-0.80446	2.55599	-1.16163	1.10102	-1.38130	0.54555	-0.51344	-5.19086	-1.59735	0.94328	-0.55916	0.27757	0.33437	
-1.98251	0.12342	-0.70445	3.52640	3.78150	-0.20327	-2.16937	-0.34972	-0.16697	0.58902	-0.31291	-2.09229	1.02699	
1.41396	1.51530	-1.48366	-1.11441	-0.05982	1.24289	-0.87840	6.08366	-0.62559	-1.06082	-0.14849	2.08409	-0.41800	
-2.83309	-0.15947	0.36298	2.67229	-0.99764	-1.10609	-0.08320	-0.14505	0.60368	0.91117	-0.23021	0.75323	0.10275	
-0.76408	-1.32107	-0.60460	0.07260	-2.06274	1.14848	0.08498	-1.04055	-1.49689	-0.80727	-1.25600	0.37273	0.29835	
-0.59986	-0.79565	0.04594	-0.89737	-0.78066	0.00412	-0.46397	-2.20712	-0.58105	-0.55956	0.08697	-0.11492	0.41096	
1.70796	-0.35643	-0.88384	-1.04446	-1.59278	0.75338	1.94786	0.29430	0.45165	-2.96784	3.41431	1.54408	-0.43092	
-0.85029	0.79890	0.26068	-0.98262	1.46500	1.79328	0.49939	-0.36951	-0.80502	0.18942	3.26083	1.23654	-0.66704	
0.13073	-0.17722	1.42984	-1.37246	-1.05626	-3.44771	-0.91333	-1.91665	-1.79437	0.20217	-4.01924	2.52992	1.29211	

Tableau IV.4. Poids et biais du QSAR-RN2 optimal.

Connexions Couche d'entrée-Couche cachée												**Connexions Couche cachée-Couche de sortie**	
Poids											**Biais**	**Poids**	**Biais**
$w_{j,1}^{I}$	$w_{j,2}^{I}$	$w_{j,3}^{I}$	$w_{j,4}^{I}$	$w_{j,5}^{I}$	$w_{j,6}^{I}$	$w_{j,7}^{I}$	$w_{j,8}^{I}$	$w_{j,9}^{I}$	$w_{j,10}^{I}$	$w_{j,11}^{I}$	b_j^H	$w_{I,j}^{H}$	b_1^0
0.1914	-0.2949	-0.3262	0.0931	1.1888	-5.3343	1.8525	-1.1684	-0.1461	0.2923	1.7750	0.2803	-0.4630	
0.4354	-0.8307	-0.3299	-0.3959	-0.2795	-1.3399	-0.1345	1.4734	0.5868	0.5177	-4.1003	-0.2135	-0.1657	
0.3119	-1.9181	-0.0223	0.4790	0.3637	0.4257	-1.4574	0.2292	-0.0277	-0.6077	-0.5646	-0.8106	-1.6519	
-0.0804	0.2924	-0.0553	-0.3412	-0.5750	0.4157	-1.5647	0.8862	-0.5941	-0.0037	0.3626	-2.0673	-0.4173	
-1.0319	-2.1136	-0.0632	-0.1330	-1.3480	-1.5664	-1.0750	0.5078	-1.2361	-0.2380	-0.3333	-0.2025	-0.2466	
0.5523	-1.9090	-0.0395	-0.0080	-1.1457	-1.5057	0.2684	-1.0856	-0.1705	-1.5016	-0.1106	-0.2718	0.2628	
0.7353	-1.4637	-0.2270	-0.0427	-0.5695	-2.7241	0.7121	-1.2804	0.4764	-1.9158	1.0803	0.0284	-0.2065	
0.1165	0.9292	-0.8589	-1.4166	1.2869	2.3802	-0.1957	-0.1372	-0.6848	-1.9799	1.2562	-1.2506	0.8877	
-0.0492	-0.1452	0.2112	-1.5338	1.5966	-0.8021	0.0579	-0.9593	-0.2269	0.0877	0.5200	0.5102	0.6790	
-0.1481	1.5834	-0.4918	0.1283	-1.1907	-1.2296	0.7342	1.4447	-0.8272	1.5418	-0.1320	1.4819	-0.2487	
-0.2257	-1.0656	0.9907	0.8352	1.7177	-0.0254	1.2358	-1.3427	-1.9933	-0.3446	0.5314	3.0098	-0.2930	
-0.8221	0.5182	1.3404	0.2210	0.9525	-0.4486	-1.1853	0.6689	1.3360	0.0453	-0.6988	-0.0225	0.6369	-0.3034
-0.1934	0.8910	0.0586	-0.1152	-3.7988	-1.9427	0.5098	4.1967	-0.5808	0.7106	-0.4957	0.9152	-0.3483	
-0.9766	1.3115	0.0093	0.3995	1.2022	-0.9268	-0.3699	-1.2328	0.5023	1.2669	-0.0493	1.1409	-0.6367	
-0.7533	0.4738	0.4747	-0.6338	2.6451	-2.2424	-3.2366	1.1148	0.0110	-0.4272	-0.2877	0.4751	1.4030	
1.1713	0.1005	0.2573	0.9681	0.6013	-0.2318	0.7429	-0.6421	-0.8660	0.6994	-0.7969	1.7598	-0.0702	
-0.0535	-1.7095	-0.7147	-0.0264	4.4593	1.7996	-0.5816	0.2743	0.0266	0.4486	-1.4696	0.5898	-0.8265	
0.2335	1.7818	-0.0752	-0.2636	-4.0319	-0.9455	0.0792	-1.7976	0.1913	0.0279	1.0068	-0.9854	0.6741	
-0.1478	-1.0775	0.1370	-0.7670	-10.2145	0.3983	-0.9575	-0.2677	-0.1174	-0.6576	1.5657	-0.2829	-1.2712	
1.0697	-0.9582	0.3662	0.0810	-1.3994	0.2127	-2.9493	-1.0286	-0.0980	-0.6365	0.8390	0.1544	0.3400	
0.9293	0.0349	0.0346	0.1090	1.5279	-0.0046	-2.1193	1.0653	-1.0642	0.6024	-1.0243	0.0780	0.2008	
1.2169	-0.4850	0.2226	-0.3578	-0.6974	1.0833	-0.7201	1.6076	-0.1941	1.7511	-1.5536	-0.3010	2.0855	
0.0559	0.1372	-0.0615	0.5879	-0.2658	0.0496	0.8220	-0.2692	0.0143	0.4619	2.5167	0.0690	0.3759	
0.7003	-0.1589	0.0873	-0.7853	1.1794	-1.2963	-1.1564	-0.9431	0.1620	4.2840	-2.9147	2.7972	-0.4097	

Tableau IV.5. Poids et biais du QSAR-RN3 optimal.

Connexions Couche d'entrée-Couche cachée													**Connexions Couche cachée-Couche de sortie**	
Poids												**Biais**	**Poids**	**Biais**
$w_{j,1}^{I}$	$w_{j,2}^{I}$	$w_{j,3}^{I}$	$w_{j,4}^{I}$	$w_{j,5}^{I}$	$w_{j,6}^{I}$	$w_{j,7}^{I}$	$w_{j,8}^{I}$	$w_{j,9}^{I}$	$w_{j,10}^{I}$	$w_{j,11}^{I}$	$w_{j,12}^{I}$	b_j^H	$w_{i,j}^{H}$	b_1^0
1.2854	-1.8110	1.2033	1.1669	-0.3234	1.9348	2.5426	0.5099	2.8689	-1.1388	2.2969	-0.0183	-1.1782	-0.2899	
-0.3309	-0.7865	-1.0769	0.2567	-0.0557	1.8980	1.9743	-0.3933	-3.1707	-0.4258	-2.9562	-2.0019	-0.4935	0.1931	
0.0817	-0.1693	-0.0045	0.0805	-4.1337	0.2386	5.8230	-0.1324	-1.7576	-0.1384	0.8646	1.4040	-0.3201	-1.4965	
-0.2184	1.1707	0.8502	0.3444	-3.8757	-0.4751	-6.5488	-0.0926	0.6131	-0.0629	1.9330	0.5307	-0.0828	-0.4724	
-0.0853	-0.7371	0.4701	0.2248	1.4110	-0.6634	-12.6810	0.4219	0.9646	0.6610	1.3993	1.3820	-0.9371	-0.3642	
0.9866	-0.6041	-1.9117	1.6166	-1.2836	-0.2203	1.6268	-3.4555	0.4848	0.7080	-0.8495	-0.5396	-0.8815	-0.1877	
0.4969	-1.5230	-0.8926	1.0377	-0.4202	-0.3916	-0.9289	-2.6582	0.3566	-1.5108	0.2305	0.6667	0.1667	-0.1119	
-0.1122	0.4251	0.7406	-0.0707	2.7956	-0.2546	2.4472	-2.0889	-0.6988	0.3732	-0.2269	0.4556	0.3649	-0.7290	
0.3188	-0.1519	0.2663	0.1387	-3.1392	-0.4739	-0.3035	-1.3303	0.4125	0.1381	0.0449	0.4545	-1.0927	3.1918	
0.4304	-0.1559	1.7808	0.0902	-1.2629	-0.7545	-0.5835	-8.0826	-0.5062	-0.8314	-0.7221	0.3124	0.7223	-0.4561	
2.1216	-1.2364	-2.1482	0.6386	1.4247	-0.1205	2.0403	1.3523	-3.1430	-0.6296	-1.2258	-0.6047	-1.3203	-0.3105	
-0.8330	0.7422	-0.0819	0.7639	1.3270	2.9852	-6.3765	6.0555	-0.4947	-0.5698	-0.2961	0.7934	-0.9866	-1.8561	0.2705
0.8322	-0.5758	-0.0191	-0.7493	-1.2527	-0.1475	1.2025	0.5461	1.1164	0.1066	-1.9200	-1.9920	2.4606	-0.2271	
1.3019	0.4550	0.0367	-0.3489	-0.3700	1.3712	-2.8960	-0.5934	-0.2706	0.2422	0.5159	0.4583	-2.4519	0.2955	
1.3666	2.6276	-0.5638	0.8175	0.7125	-0.3854	-0.3685	0.7192	-0.2643	0.5034	-1.0021	1.0191	2.6712	-2.1050	
-1.7871	0.9994	-1.0980	0.2191	-0.3304	-1.3698	-0.2886	0.6866	0.7810	0.0522	-0.9065	0.1995	-0.4000	0.6856	
-0.6726	-0.9846	-3.4156	-0.3107	-0.2121	-2.9606	-1.2198	0.0404	0.4967	0.2929	0.8530	-0.7365	0.9833	-0.3523	
1.1317	1.7329	-1.3929	-0.8621	0.4680	-0.1736	0.2379	-0.1239	-0.5962	0.6281	-1.9435	0.9050	-0.9428	-0.0653	
1.1850	0.2924	3.1065	-0.6943	0.3818	-0.6338	1.0731	-0.6548	-0.9006	0.7725	0.2944	-1.2990	-0.4773	-0.6945	
-2.3795	-0.1095	0.4269	0.4506	0.3267	-1.4457	0.5670	0.7484	-0.9394	-0.5595	-1.0606	1.5089	0.4730	-0.8208	
0.9162	-0.3753	-0.9653	-1.4830	-0.1642	-0.8066	-0.2858	-0.0574	0.5286	0.4085	0.6903	-1.0554	0.0661	0.4988	
0.4680	-0.2759	-0.8620	-0.2258	0.0822	-0.4156	2.2270	-0.7575	0.0346	-0.3555	-2.0641	-0.5969	-0.4830	0.1006	
0.0585	1.4000	-1.0255	-0.5415	-0.2315	0.2758	0.8020	-1.8112	0.9104	-0.3071	-1.1167	-0.5397	0.7356	-0.3301	
0.3967	0.1803	3.5247	-0.5610	1.0046	-2.7121	-0.0113	1.1586	-0.1120	-0.5255	-1.8988	-2.9348	-1.6846	1.3947	

L'architecture des réseaux de neurones optimaux obtenus est donnée sur la figure IV.6, ci-après.

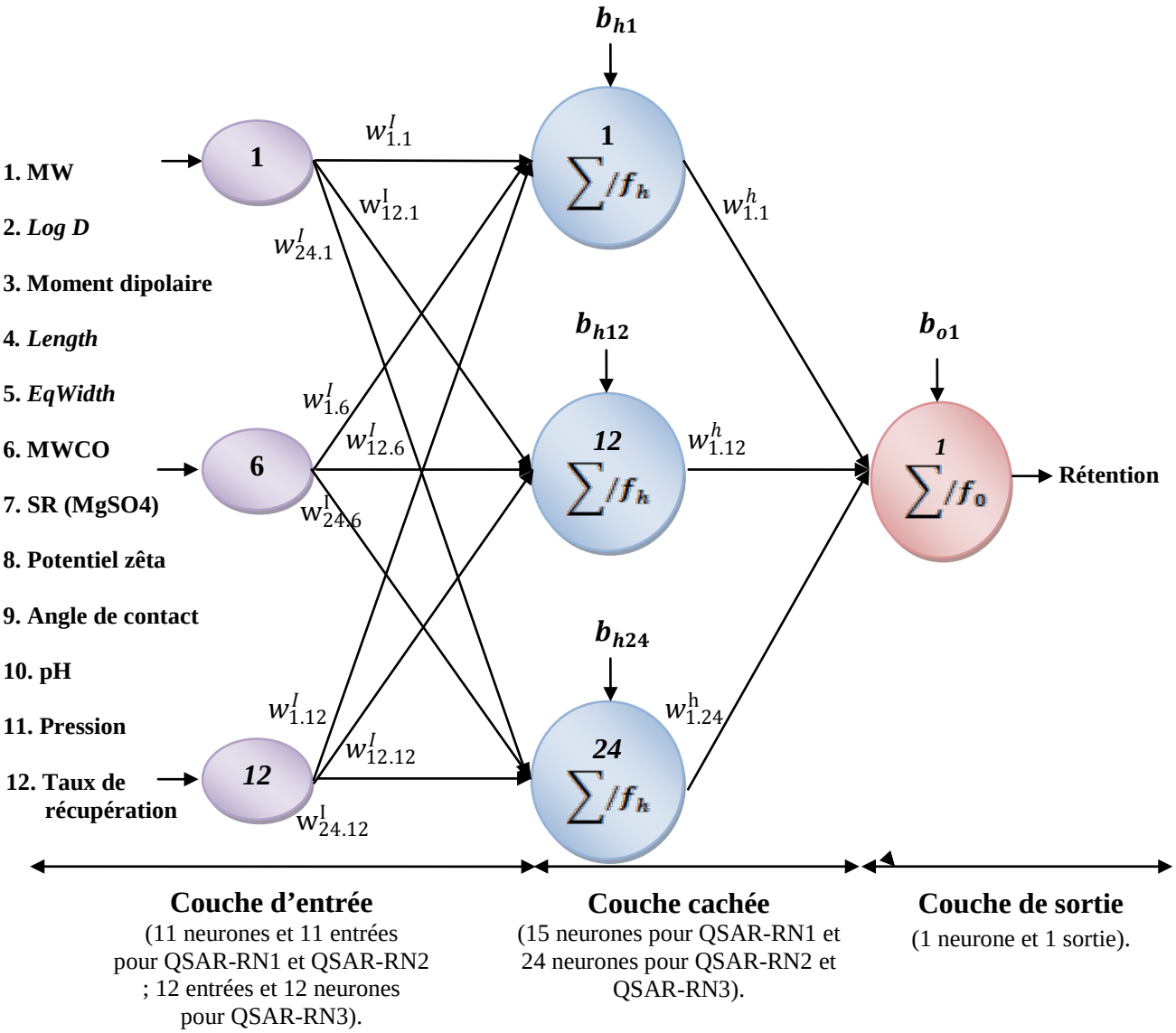
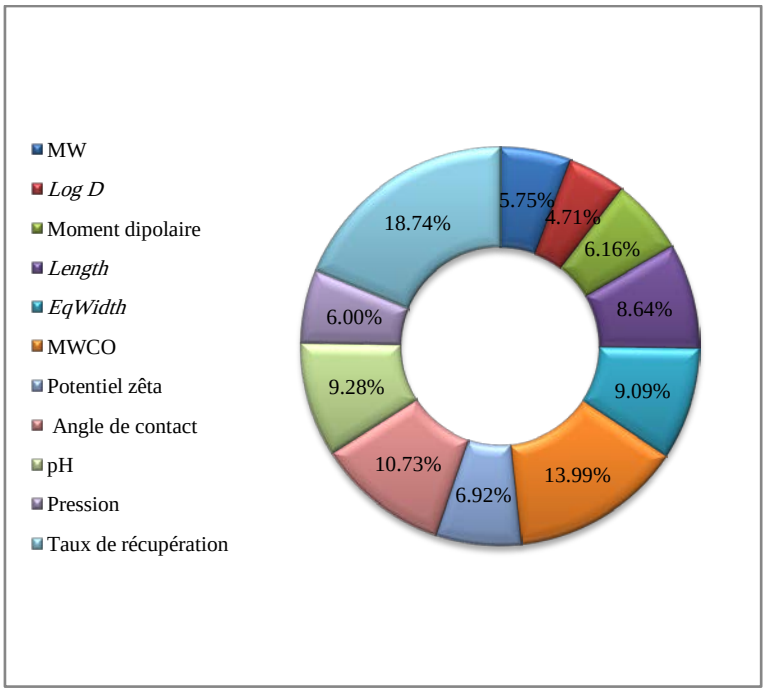


Figure IV.6. Architecture finale du réseau de neurone optimal de type feedforward pour la prédiction de la rétention des molécules organiques (neutres et ioniques) par les membranes d'OI/NF.

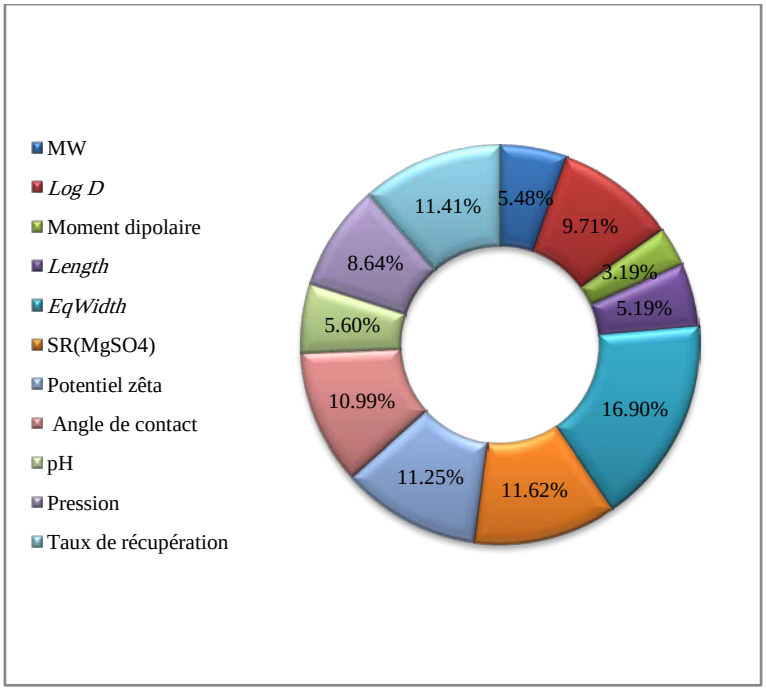
IV.2.3. Analyse de sensibilité par la méthode de "poids"

Pour voir la contribution ainsi que le profil de variation de chaque variable d'entrée (MW, *LogD*, *Lenght*, *EqWidth*, MWCO, SR, angle de contact, potentiels zêta, pH, pression et taux de récupération) sur la sortie de chaque réseau (QSAR-RN1, QSAR-RN2 et QSAR-RN3), On a procédé à une analyse de sensibilité par méthode de "poids".

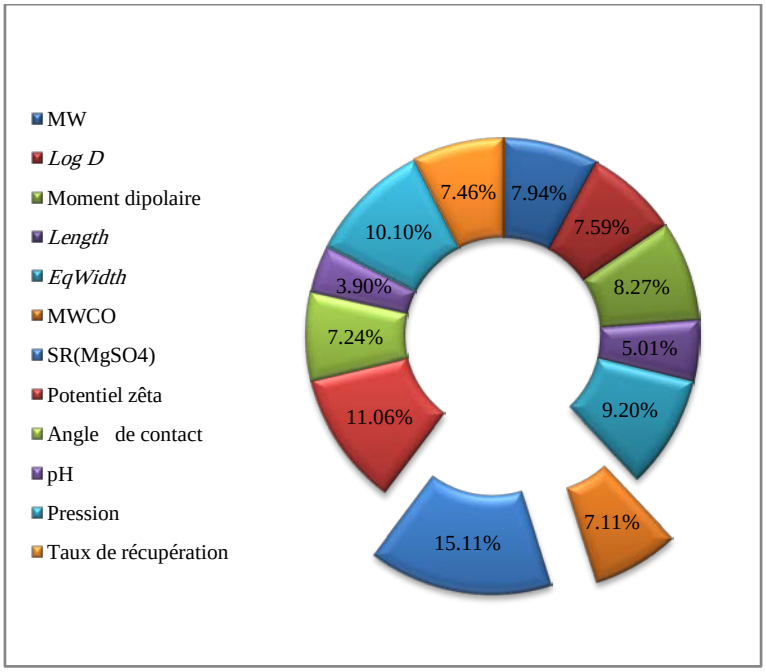
Nous avons appliqué cette méthode aux QSAR-RN1, QSAR-RN2 et QSAR-RN3 en utilisant leurs poids respectifs illustrés dans les tableaux IV.3, IV.4 et IV.5. Les résultats de contributions sont montrés sur la figure IV.7. Les variables les plus pertinentes qui peuvent influer sur la rétention de composés organiques (neutres et ioniques) par les membranes d'OI/NF sont : **(i)** taux de récupération, MWCO, angle de contact, pH, *EqWidth*, *Length*, potentiel zêta, moment dipolaire, pression, MW et *LogD* pour RN1; **(ii)** *EqWidth*, SR ($MgSO_4$), taux de récupération, potentiel zêta, angle de contact, *LogD*, pression, pH, MW, *Length*, et moment dipolaire pour RN2 et **(iii)** SR ($MgSO_4$), potentiel zêta, pression, *EqWidth*, moment dipolaire, MW, *LogD*, taux de récupération, angle de contact, MWCO, *Length* et pH pour RN3. Toutes les entrées exercent les contributions les plus importantes et ceux pour chaque QSAR-RN.



(a)



(b)



(c)

Figure IV.7. Présentation des pourcentages de l'importance relative (*IR*) de chaque variable d'entrée sur la rétention : (a) QSAR-RN1, (b) QSAR-RN2, (c) QSAR-RN3.

D'après la figure IV.7 (c), il est évident que le paramètre stérique/taille MWCO est moins efficace que le paramètre stérique/taille SR($MgSO_4$) pour prédire la rétention des molécules organiques lors d'OI/NF (*IR* (MWCO) = 7.11% et *IR* "SR ($MgSO_4$)" = 15.11%).

Ainsi, le rejet de sel de magnésium "SR ($MgSO_4$)" est un bon indicateur de porosité que le seuil de coupure "MWCO".

Les résultats de l'analyse de sensibilité par la méthode "poids" montrent qu'il est plus simple de caractériser une membrane en termes de SR($MgSO_4$) plutôt que MWCO. Ils sont en accord avec les conclusions de Yangali-Quintanilla et al. 2009b et 2010b.

IV.2.4. Comparaison entre la rétention expérimentale et calculée par QSAR-RN3

La figure IV.8 visualise une comparaison entre les données expérimentales et les résultats de prédictions du troisième modèle neuronal en fonction des masses molaires des molécules organiques. Elle montre une concordance acceptable entre les données expérimentales de la littérature de taux de la rétention des molécules organiques (indiquées par des cercles vides) et les résultats de prédictions du modèle neuronal (indiqués par des marqueurs pleins en bleu), ce qui suggère de bonnes capacités prédictives du modèle neuronal conçu "QSAR-RN3" et cela pour les poids moléculaires situés dans l'intervalle de poids moléculaire ($30.03 \leq MW \leq 342.30$).

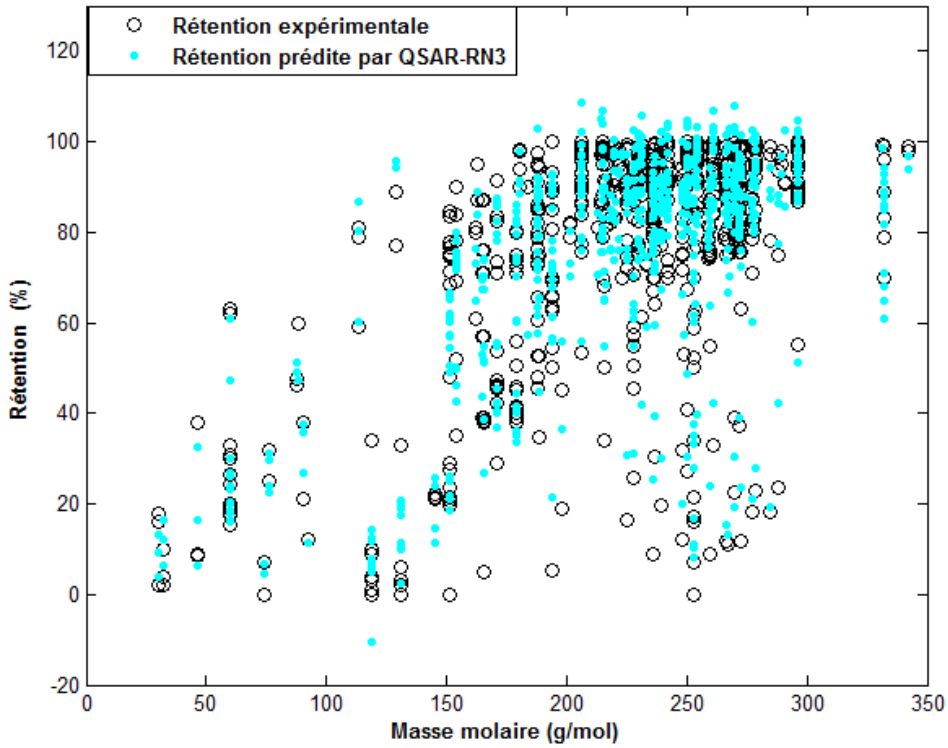


Figure IV.8. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de la masse molaire des MO.

La comparaison entre les données expérimentales et les résultats de prédictions du troisième modèle neuronal en fonction du logarithme des coefficients de distributions "LogD" des molécules organiques est illustrée dans la figure IV.9. Elle suggère une concordance acceptable entre les données expérimentales de la littérature de taux de la rétention des molécules organiques (indiquées par des cercles vides) et les résultats de prédictions du modèle neuronal (indiqués par des marqueurs pleins en vert), ce qui suggère de bonnes capacités prédictives du modèle neuronal conçu "QSAR-RN3" et cela pour les logarithmes des coefficients de distributions situés dans l'intervalle de logarithme du coefficient de distribution ($-3.70 \leq \text{LogD} \leq 5.93$).

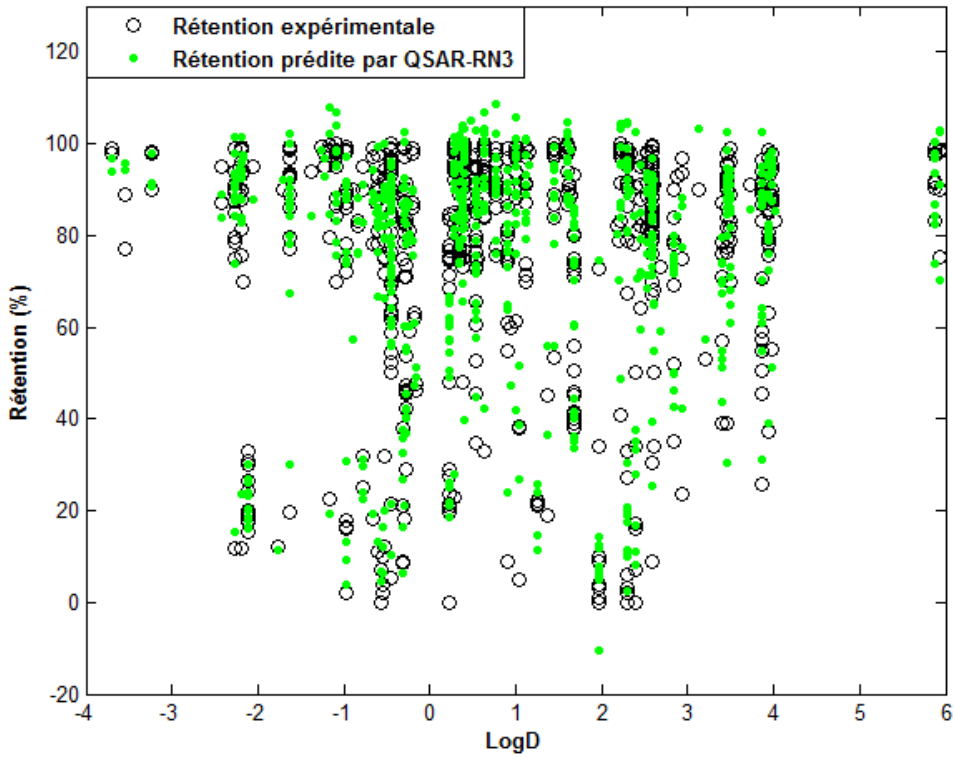


Figure IV.9. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de *LogD* des MO.

La figure IV.10 illustre une comparaison entre les données expérimentales et les résultats de prédictions du troisième modèle neuronal en fonction des moments dipolaires des molécules organiques. Elle montre une concordance acceptable entre les données expérimentales de la littérature de taux de la rétention des molécules organiques (indiquées par des cercles vides) et les résultats de prédictions du modèle neuronal (indiqués par des marqueurs pleins en mauve), ce qui suggère de bonnes capacités prédictives du modèle

neuronal conçu "QSAR-RN3" et cela pour les moments dipolaires situés dans l'intervalle de moment dipolaire (0.00 ≤ moment dipolaire ≤ 10.91).

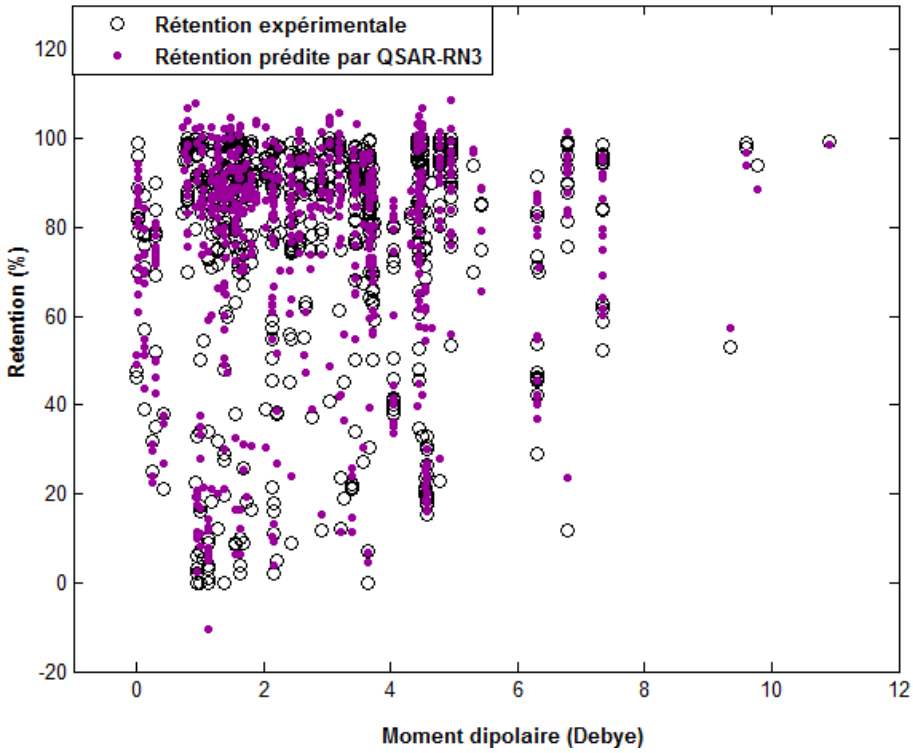


Figure IV.10. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de moment dipolaire des MO.

La comparaison entre les données expérimentales et les résultats de prédictions du troisième modèle neuronal en fonction des longueurs moléculaires des molécules organiques est présentée dans la figure IV.11. Elle montre une concordance acceptable entre les données expérimentales de la littérature de taux de la rétention des molécules organiques (indiquées par des cercles vides) et les résultats de prédictions du modèle neuronal (indiqués par des marqueurs pleins en orange), ce qui suggère de bonnes capacités prédictives du modèle neuronal conçu "QSAR-RN3" et cela pour les longueurs moléculaires situées dans l'intervalle de la longueur moléculaire (0.47 ≤ Length ≤ 1.86).

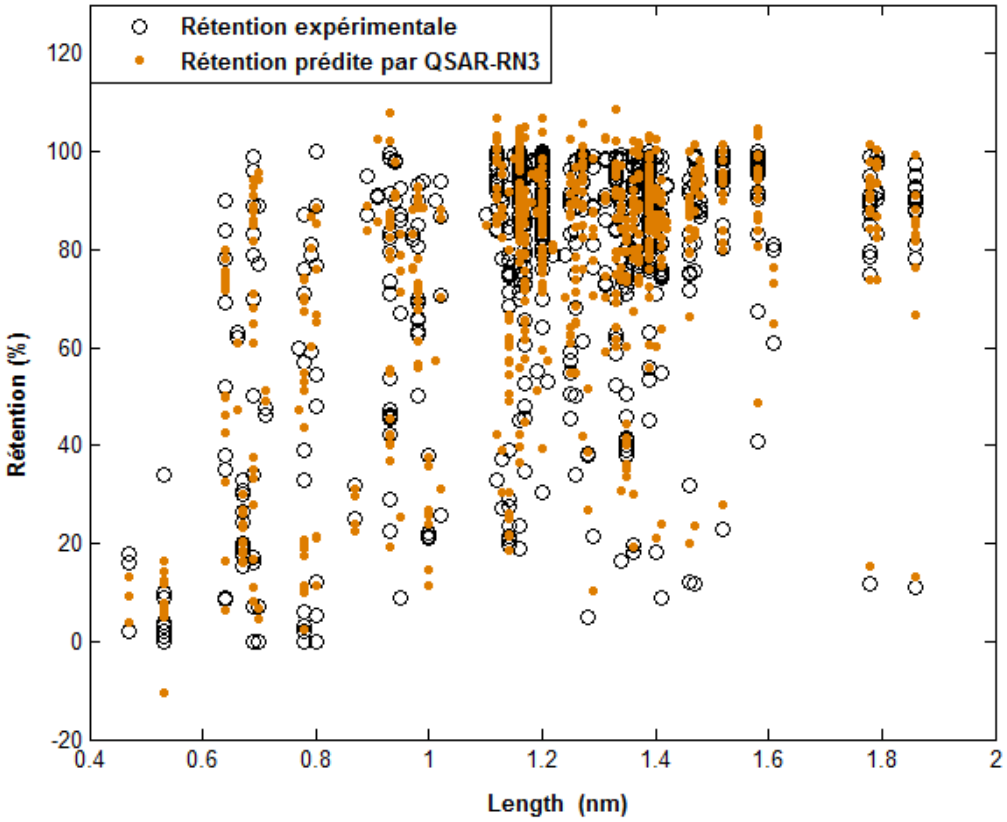


Figure IV.11. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de la longueur des MO.

La figure IV.12 montre une comparaison entre les données expérimentales et les résultats de prédictions du troisième modèle neuronal en fonction des largeurs équivalentes moléculaires des molécules organiques. Elle montre une concordance acceptable entre les données expérimentales de la littérature de taux de la rétention des molécules organiques (indiquées par des cercles vides) et les résultats de prédictions du modèle neuronal (indiqués par des marqueurs pleins en rose), ce qui suggère de bonnes capacités prédictives du modèle neuronal conçu "QSAR-RN3" et cela pour les largeurs équivalentes moléculaires situées dans l'intervalle de la largeur équivalente moléculaire ($0.38 \leq \text{EqWidth} \leq 0.92$).

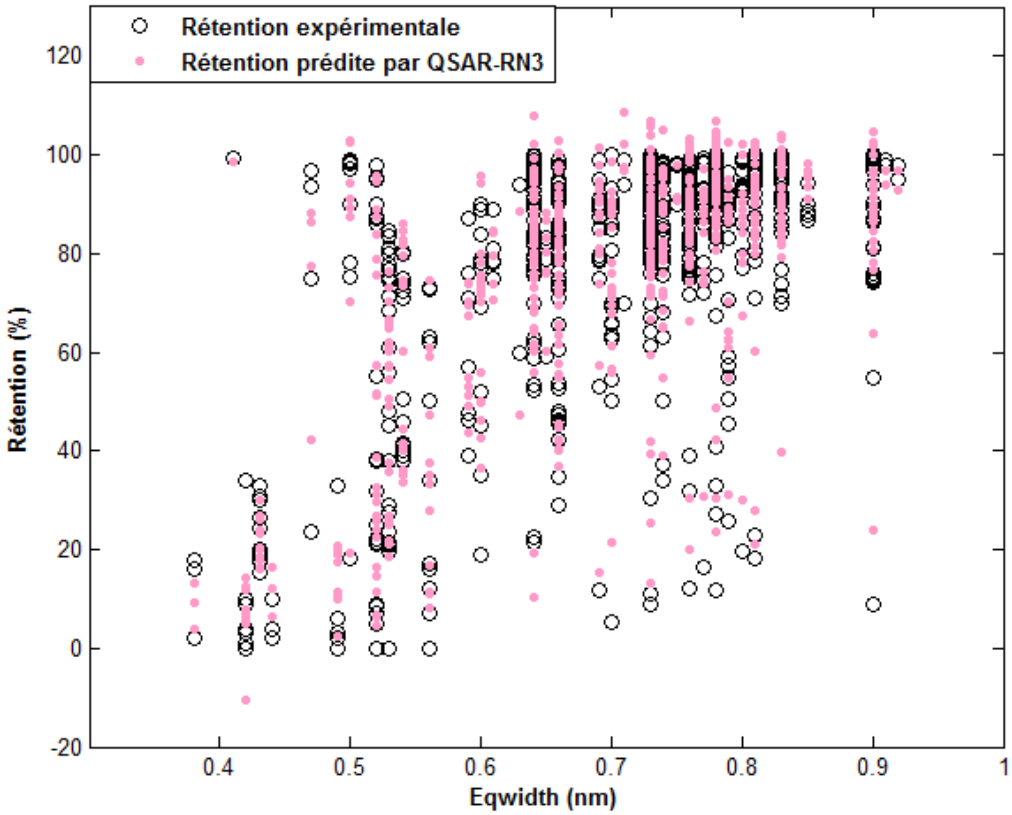


Figure IV.12. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de la largeur équivalente des MO.

La comparaison entre les données expérimentales et les résultats de prédictions du troisième modèle neuronal en fonction des seuils de coupure des membranes est présentée dans la figure IV.13. Elle montre une excellente concordance entre les données expérimentales de la littérature de taux de la rétention des molécules organiques (indiquées par des cercles vides) et les résultats de prédictions du modèle neuronal (indiqués par des marqueurs pleins en rouge), ce qui suggère de bonnes capacités prédictives du modèle neuronal conçu "QSAR-RN3" et cela pour les seuils de coupure situés dans l'intervalle du seuil de coupure ($100 \leq MWCO \leq 300$).

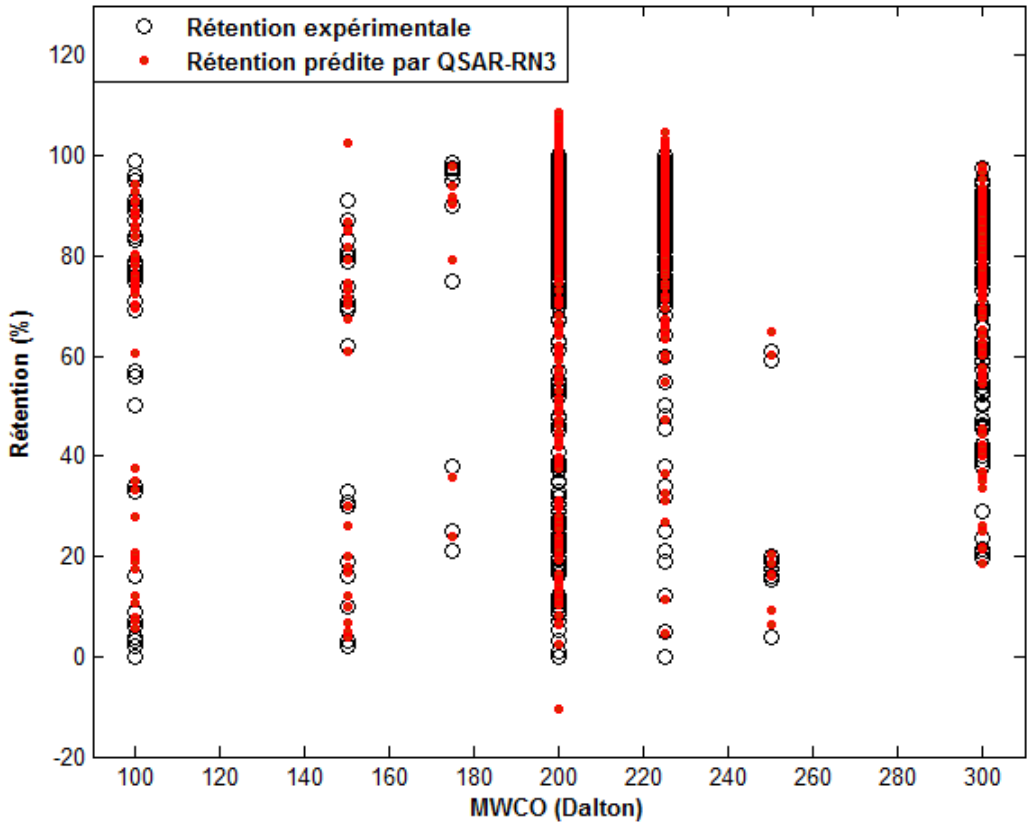


Figure IV.13. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de MWCO des MO.

La figure IV.14 montre une comparaison entre les données expérimentales et les résultats de prédictions du troisième modèle neuronal en fonction des rejets de sels de magnésium "SR($MgSO_4$)" des membranes. Elle montre une excellente concordance entre les données expérimentales de la littérature de taux de la rétention des molécules organiques (indiquées par des cercles vides) et les résultats de prédictions du modèle neuronal (indiqués par des marqueurs pleins en bleu foncé), ce qui suggère de bonnes capacités prédictives du modèle neuronal conçu "QSAR-RN3" et cela pour les rejets de sels de magnésium situés dans l'intervalle du rejet de sel de magnésium ($0.96 \leq SR(MgSO_4) \leq 1.00$).

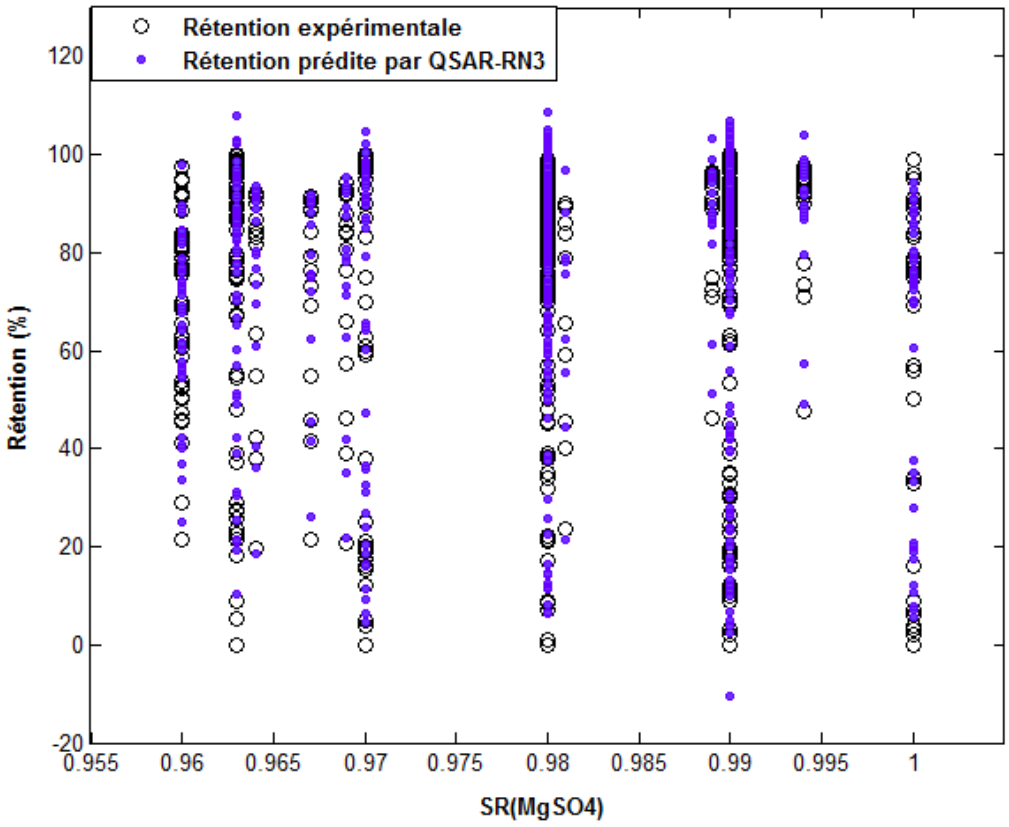


Figure IV.14. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de SR($MgSO_4$) des MO.

La comparaison entre les données expérimentales et les résultats de prédictions du troisième modèle neuronal en fonction des potentiels zêtas des membranes est illustrée dans la figure IV.15. Elle montre une excellente concordance entre les données expérimentales de la littérature de taux de la rétention des molécules organiques (indiquées par des cercles vides) et les résultats de prédictions du modèle neuronal (indiqués par des marqueurs pleins en bleu clair), ce qui suggère de bonnes capacités prédictives du modèle neuronal conçu "QSAR-RN3" et cela pour les potentiels zêtas situés dans l'intervalle du potentiel zêta (-0.48 ≤ potentiel zêta ≤ 1.03).

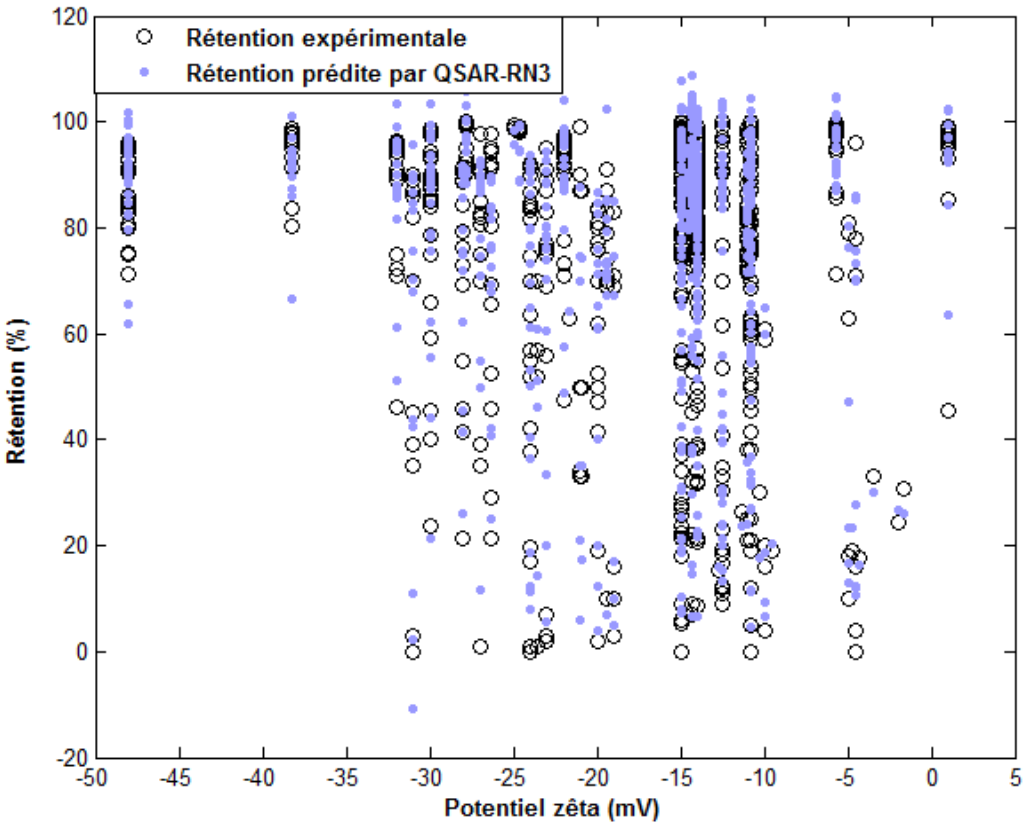


Figure IV.15. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de potentiel zêta des MO.

La figure IV.16 montre une comparaison entre les données expérimentales et les résultats de prédictions du troisième modèle neuronal en fonction des angles de contacts des membranes. Elle montre une concordance acceptable entre les données expérimentales de la littérature de taux de la rétention des molécules organiques (indiquées par des cercles vides) et les résultats de prédictions du modèle neuronal (indiqués par des marqueurs pleins en vert), ce qui suggère de bonnes capacités prédictives du modèle neuronal conçu "QSAR-RN3" et cela pour les angles de contacts situés dans l'intervalle d'angle de contact (28.00 ≤ angle de contact ≤ 70.70).

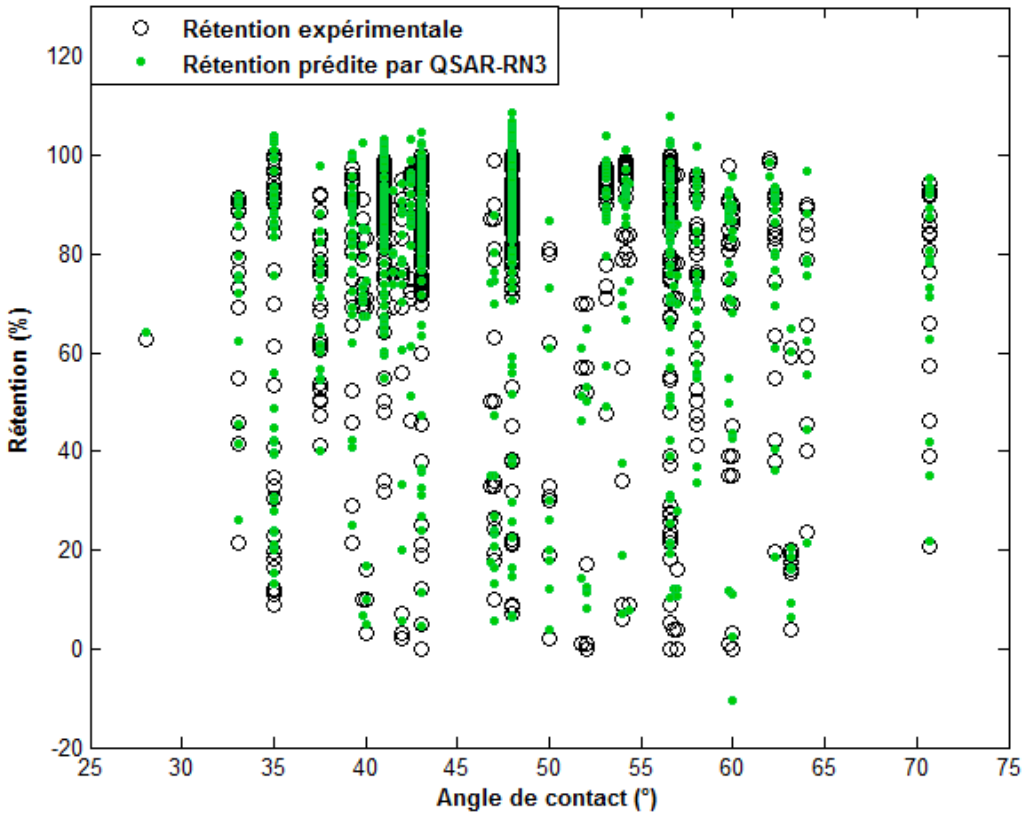


Figure IV.16. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction d'angle de contact des MO.

La comparaison entre les données expérimentales et les résultats de prédictions du troisième modèle neuronal en fonction des pH est présentée dans la figure IV.17. Elle montre une excellente concordance entre les données expérimentales de la littérature de taux de la rétention des molécules organiques (indiquées par des cercles vides) et les résultats de prédictions du modèle neuronal (indiqués par des marqueurs pleins en mauve), ce qui suggère de bonnes capacités prédictives du modèle neuronal conçu "QSAR-RN3" et cela pour les pH situés dans l'intervalle du pH ($2.68 \leq pH \leq 10.01$).

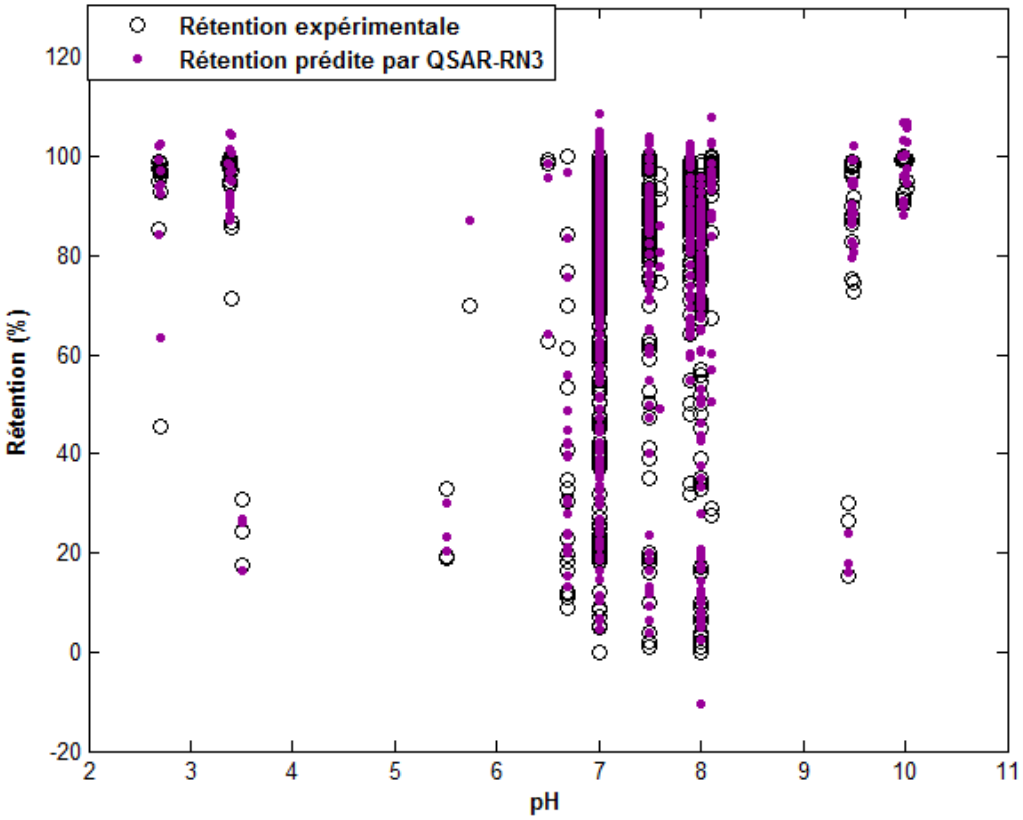


Figure IV.17. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de pH des MO.

La figure IV.18 montre une comparaison entre les données expérimentales et les résultats de prédictions du troisième modèle neuronal en fonction des pressions. Elle montre une excellente concordance entre les données expérimentales de la littérature de taux de la rétention des molécules organiques (indiquées par des cercles vides) et les résultats de prédictions du modèle neuronal (indiqués par des marqueurs pleins en rose), ce qui suggère de bonnes capacités prédictives du modèle neuronal conçu "QSAR-RN3" et cela pour les pressions situées dans l'intervalle du pression ($240 \leq$ pression ≤ 4100).

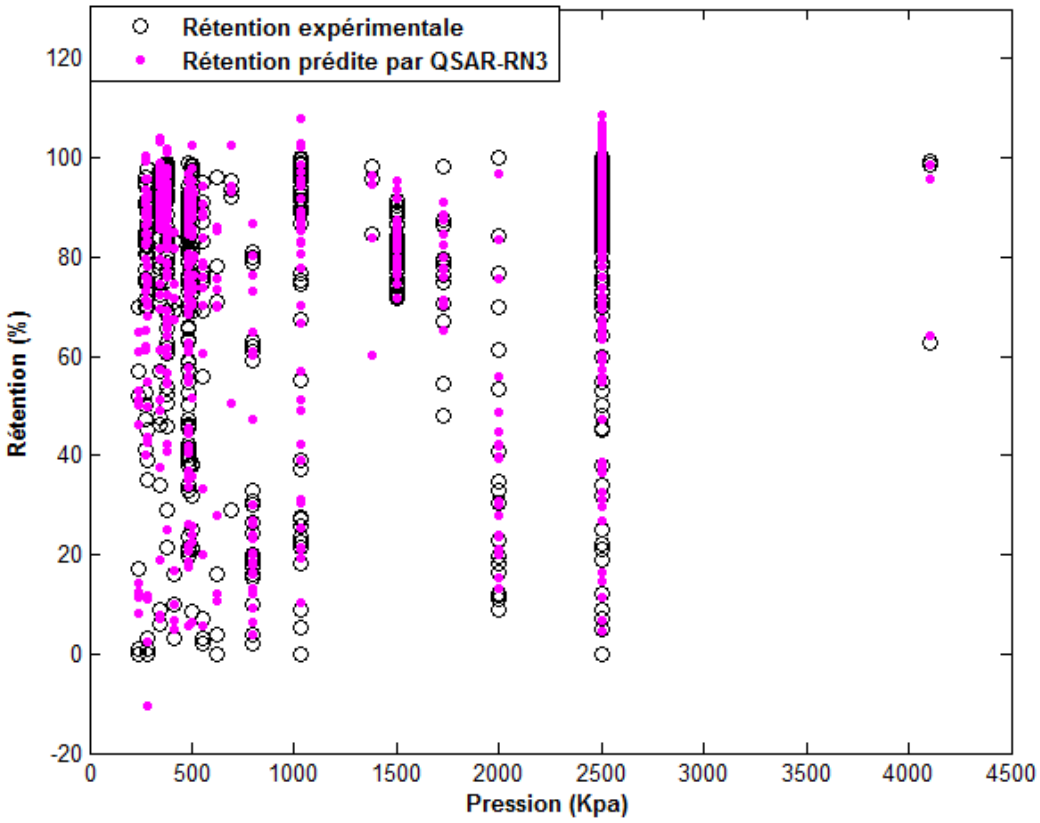


Figure IV.18. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de pression des MO.

La comparaison entre les données expérimentales et les résultats de prédictions du troisième modèle neuronal en fonction des taux de récupérations est présentée dans la figure IV.19. Elle montre une excellente concordance entre les données expérimentales de la littérature de taux de la rétention des molécules organiques (indiquées par des cercles vides) et les résultats de prédictions du modèle neuronal (indiqués par des marqueurs pleins en orange), ce qui suggère de bonnes capacités prédictives du modèle neuronal conçu "QSAR-RN3" et cela pour les taux de récupérations situés dans l'intervalle du taux de récupération ($0.50 \leq$ taux de récupération ≤ 90.00).

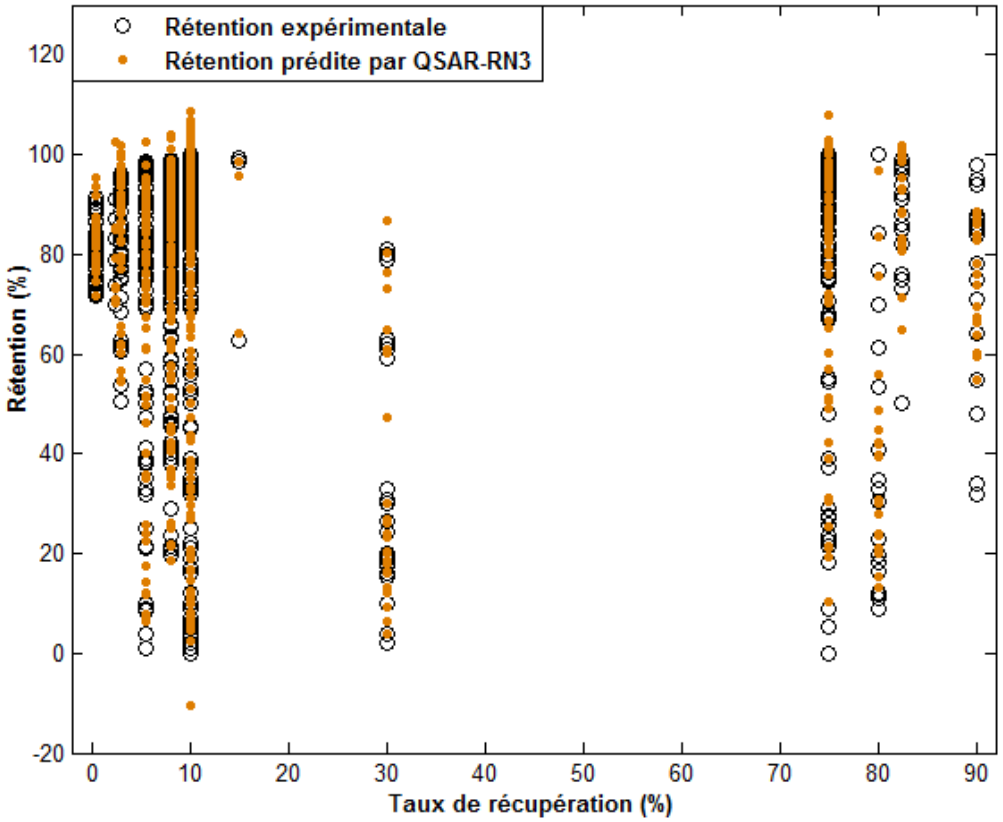


Figure IV.19. Comparaison entre la rétention calculée par QSAR-RN3 et la rétention expérimentale en fonction de taux de récupération des MO.

IV.3. Modèles QSAR- RLM

Les modèles obtenus pour la prédiction de la rétention des molécules organiques (neutres et ioniques) par les membranes d'OI/NF sont des modèles linéaires (équations IV.7, IV.8 et IV.9) :

QSAR-RLM1

$$\text{QSAR} - \text{RLM1} = -40.4736 + 0.1697MW - 1.2353LogD + 1.4790moment\ dipolaire + 22.8340Length + 51.6208EqWidth - 0.1210MWCO - 0.2496potential\ zeta + 0.5339angle\ de\ contact + 1.2114pH + 0.0020pression - 0.1488taux\ de\ récupération \quad \text{(IV.7)}$$

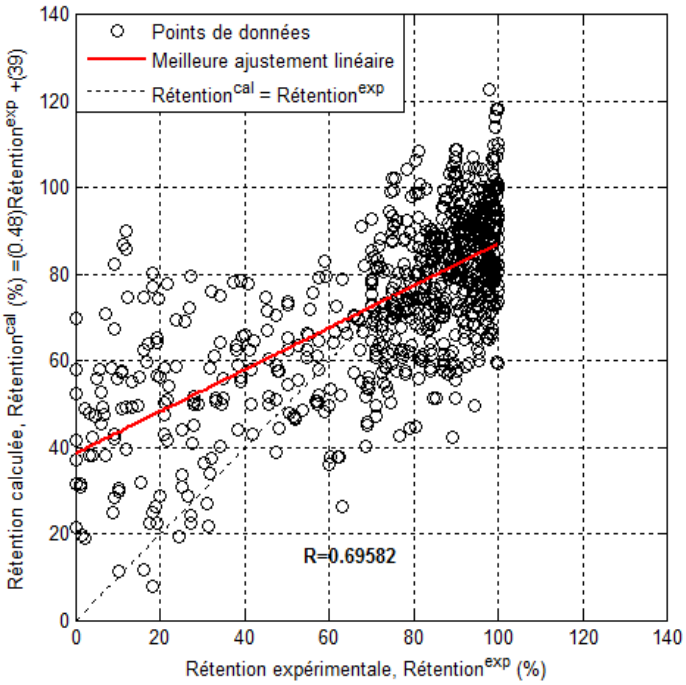
QSAR-RLM2

$$\text{QSAR} - \text{RLM2} = -155.3461 + 0.1330MW + 0.1327LogD + 0.6828moment\ dipolaire + 15.7248Length + 71.9421EqWidth + 105.9362SR - 0.3252potential\ zeta + 0.2592angle\ de\ contact + 1.0009pH + 0.0029pression - 0.1315taux\ de\ récupération \quad \text{(IV.8)}$$

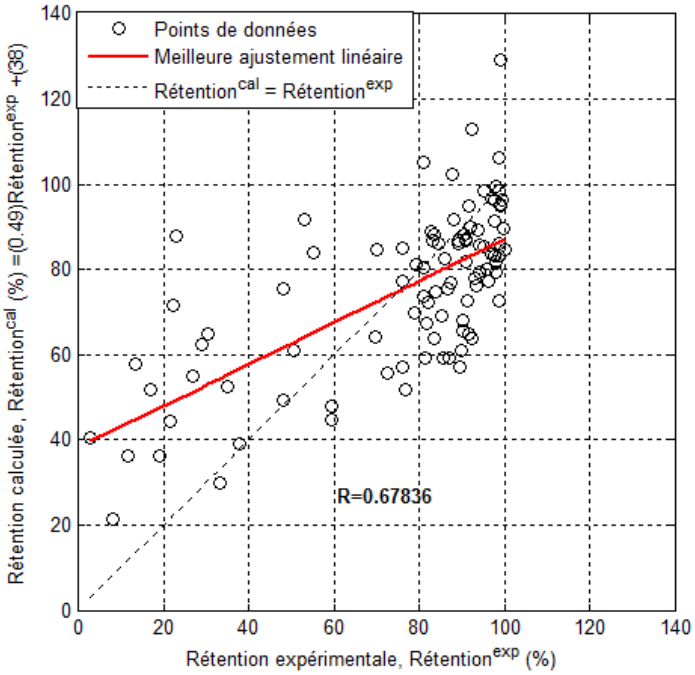
QSAR-RLM3

$$\text{QSAR} - \text{RLM3} = 289.3667 + 0.1030MW + 0.0051LogD + 1.1699moment\ dipolaire + 21.2466Length + 90.4811EqWidth - 0.1512MWCO - 327.7187SR - 0.3417potential\ zeta + 0.2828angle\ de\ contact + 0.7164pH + 0.0025pression - 0.1666taux\ de\ récupération \quad \text{(IV.9)}$$

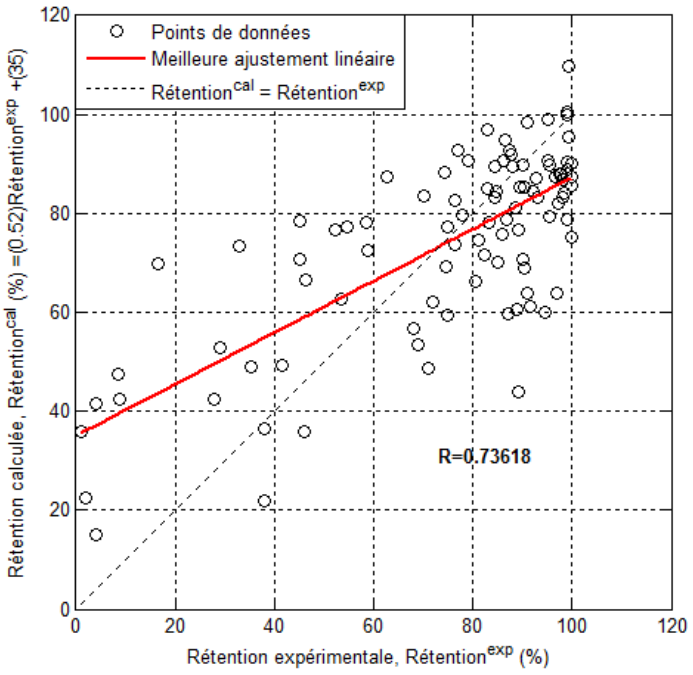
La figure IV.20 visualise la comparaison entre les valeurs expérimentales et les valeurs calculées par le premier modèle QSAR-RLM. Elle montre les courbes de régression linéaire de modèle QSAR-RLM1. Les vecteurs de régression linéaire, ([α, β, R] = [0.4833, 38.6502, 0.6958] "phase d'apprentissage" ; [α, β, R] = [0.4895, 38.1303, 0.6784] "phase de validation" ; [α, β, R] = [0.5215, 35.0610, 0.7362] "phase de test" et [α, β, R] = [0.4874, 38.2604, 0.6980] "phase totale"), indiquent une capacité prédictive modérée de modèle QSAR-RLM1. Les coefficients des corrélations sont généralement considérés comme satisfaisants ($0.50 \leq R < 0.90$) pour la phase d'apprentissage, la phase de validation, la phase de test et la phase totale de modèle QSAR-RLM1, ce qui implique une robustesse acceptable de modèle QSAR-RLM1 établi et la possibilité de la prédiction des différents paramètres qui caractérisent la rétention des molécules organiques lors d'OI/NF.



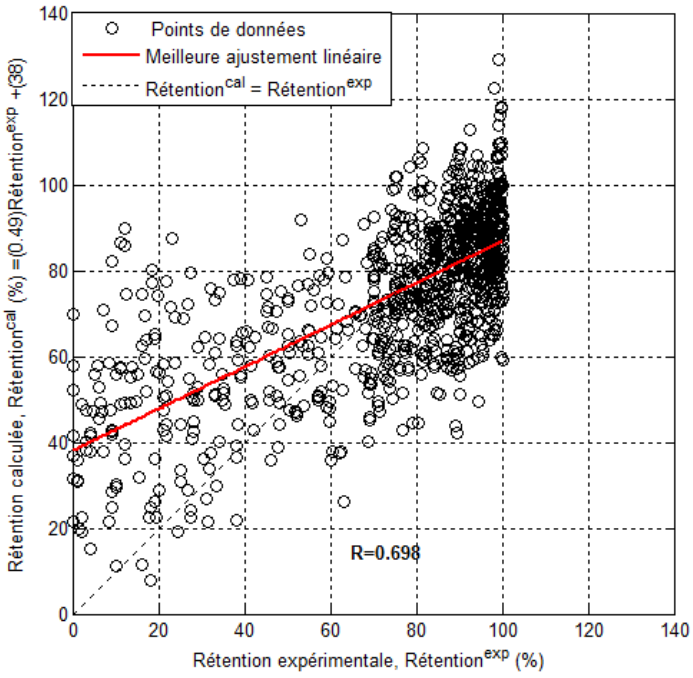
(a)


Points de données
Meilleure ajustement linéaire
Rétention^{cal} = Rétention^{exp}
R=0.67836
Rétention calculée, Rétention^{cal} (%) =(0.49)Rétention^{exp} +(38)
Rétention expérimentale, Rétention^{exp} (%)
0
20
40
60
80
100
120
140

(b)


Points de données
Meilleure ajustement linéaire
Rétention^{cal} = Rétention^{exp}
R=0.73618
Rétention calculée, Rétention^{cal} (%) =(0.52)Rétention^{exp} +(35)
Rétention expérimentale, Rétention^{exp} (%)
0
20
40
60
80
100
120

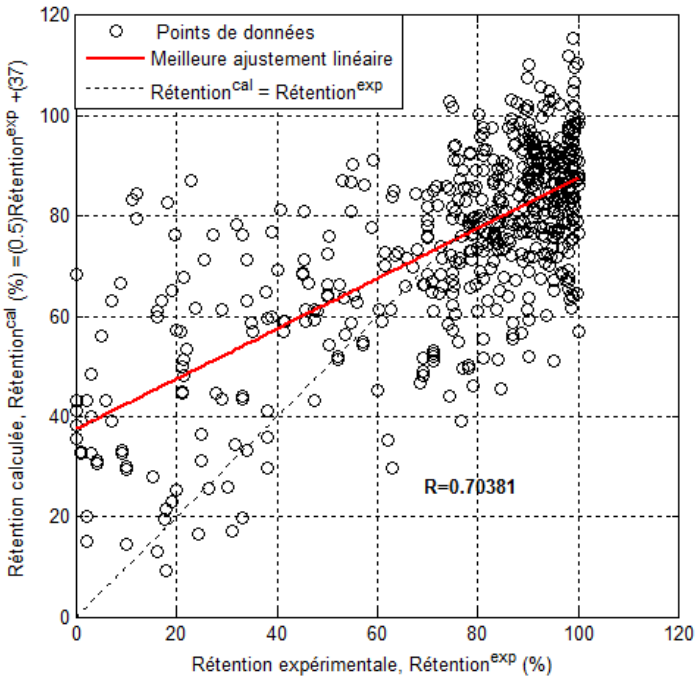
(c)



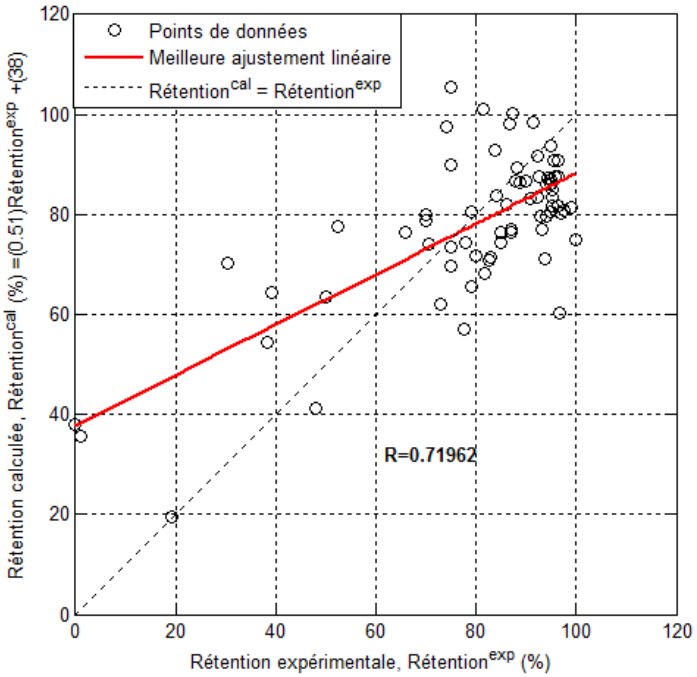
(d)

Figure IV.20. Modèle QSAR-RLM1 : les valeurs calculées vs. Les valeurs expérimentales pour les ensembles de données : (a) phase d'apprentissage, (b) phase de validation, (c) phase de test, (d) totale.

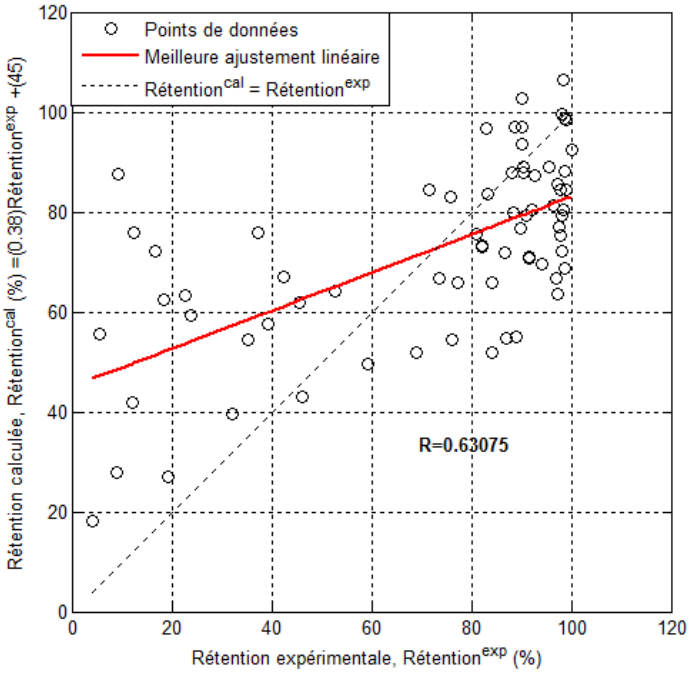
La figure IV.21 montre les courbes de régression linéaire des taux de rétentions calculés par le modèle de QSAR-RLM2 avec les taux de rétentions expérimentales avec des vecteurs de régression approchant d'approuvable ([α, β, R] = [0.5004, 37.4867, 0.7038] "phase d'apprentissage" ; [α, β, R] = [0.5059, 37.6789, 0.7196] "phase de validation" ; [α, β, R] = [45.2731, 0.3799, 0.6308] "phase de test" et [α, β, R] = [0.4860, 38.5344, 0.6971] "phase totale"). Les coefficients des corrélations sont généralement considérés comme satisfaisants ($0.50 \leq R < 0.90$) pour la phase d'apprentissage, la phase de validation, la phase de test et la phase totale, ces résultats montrent que le modèle QSAR-RLM2 possède une capacité de modélisation acceptable pour caractériser la rétention des molécules organiques lors d'OI/NF.


Points de données
Meilleure ajustement linéaire
Rétention^cal = Rétention^exp
R=0.70381
Rétention calculée, Rétention^cal (%) =(0.5)Rétention^exp +(37)
Rétention expérimentale, Rétention^exp (%)
0
20
40
60
80
100
120

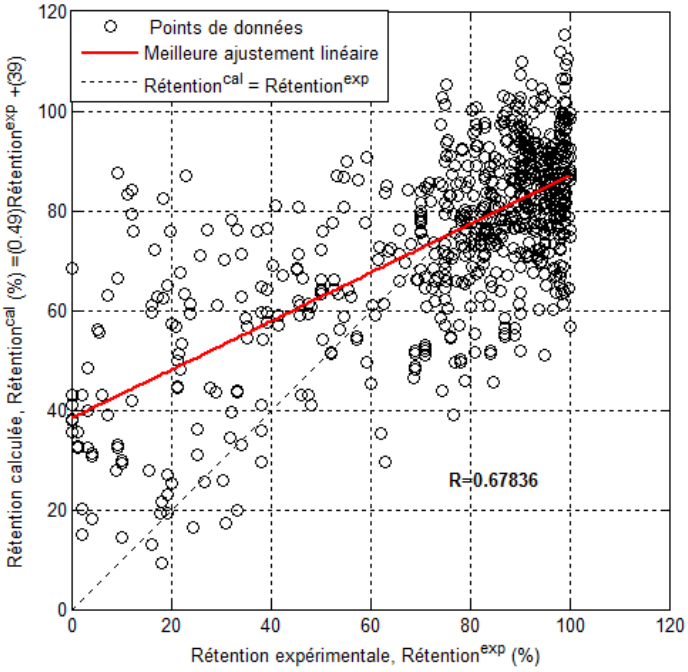
(a)


Points de données
Meilleure ajustement linéaire
Rétention^cal = Rétention^exp
R=0.71962
Rétention calculée, Rétention^cal (%) =(0.51)Rétention^exp +(38)
Rétention expérimentale, Rétention^exp (%)
0
20
40
60
80
100
120

(b)



(c)

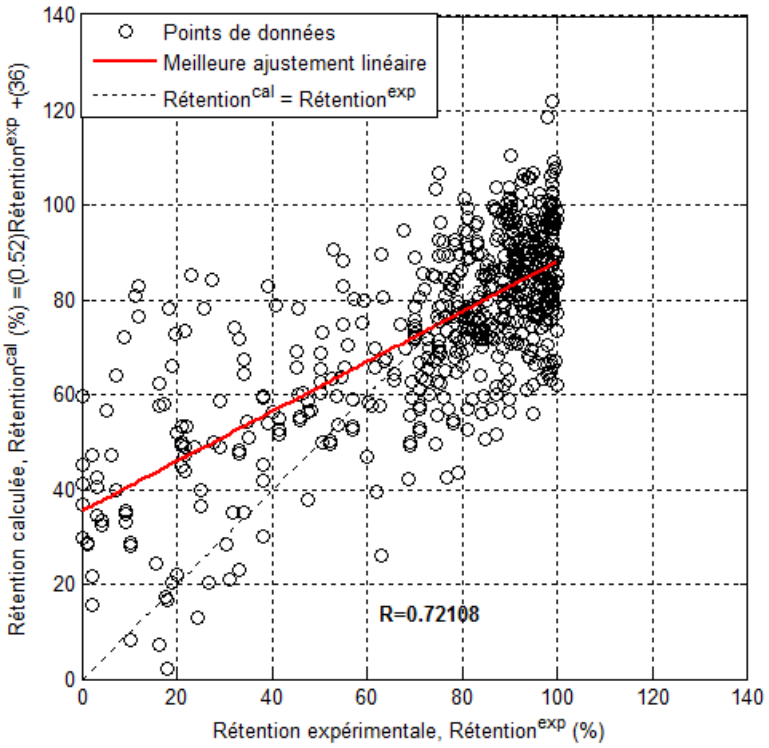


(d)

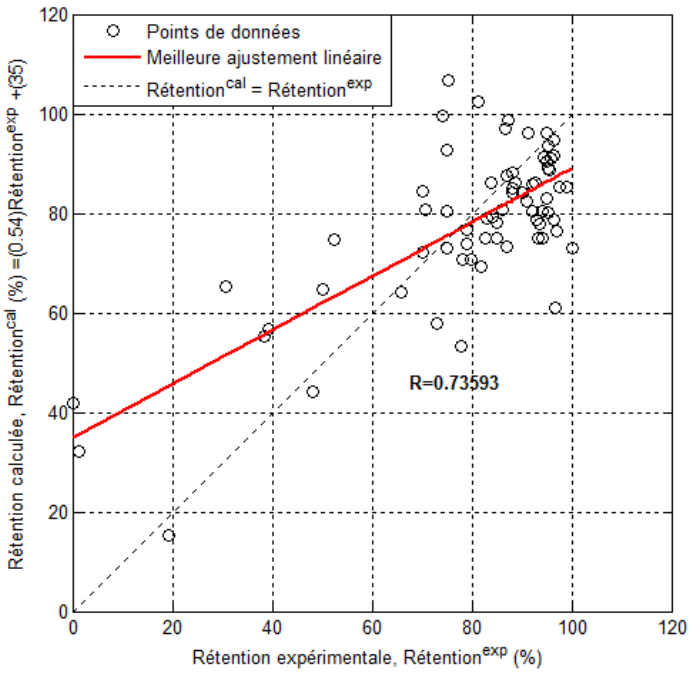
Figure IV.21. Modèle QSAR-RLM2 : les valeurs calculées vs. Les valeurs expérimentales pour les ensembles de données : (a) phase d'apprentissage, (b) phase de validation, (c) phase de test, (d) totale.

Pour évaluer la performance de modèle QSAR-RLM3 obtenu, la représentation graphique des valeurs prédites de la rétention en fonction des valeurs expérimentales pour les

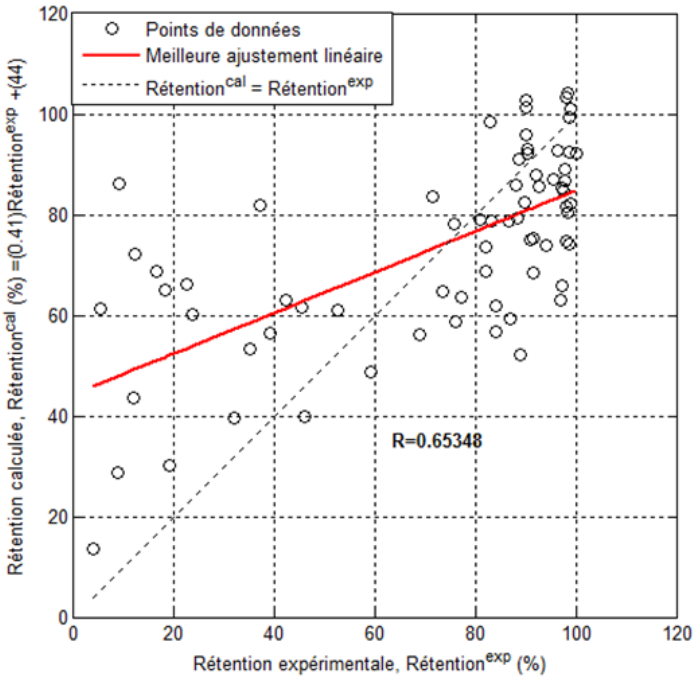
ensembles d'apprentissage, de validation, de test et totale obtenues est donnée sur la figure IV.22. Elle montre les courbes de régression linéaire de modèle QSAR-RLM3 avec des vecteurs de régression approchant d'acceptable ([α, β, R] = [0.5250, 35.5444, 0.7211] "phase d'apprentissage" ; [α, β, R] = [0.5417, 35.0678, 0.7359] "phase de validation" ; [α, β, R] = [0.4050, 44.4571, 0.6535] "phase de test" et [α, β, R] = [0.5111, 36.6951, 0.7147] "phase totale"). Les coefficients des corrélations sont généralement considérés comme satisfaisants ($0.50 \leq R < 0.90$) pour la phase d'apprentissage, la phase de validation, la phase de test et la phase totale, ce qui implique une acceptable robustesse de modèle QSAR-RLM3 établi et la possibilité de la prédiction des différents paramètres qui caractérisent la rétention des molécules organiques lors d'OI/NF.



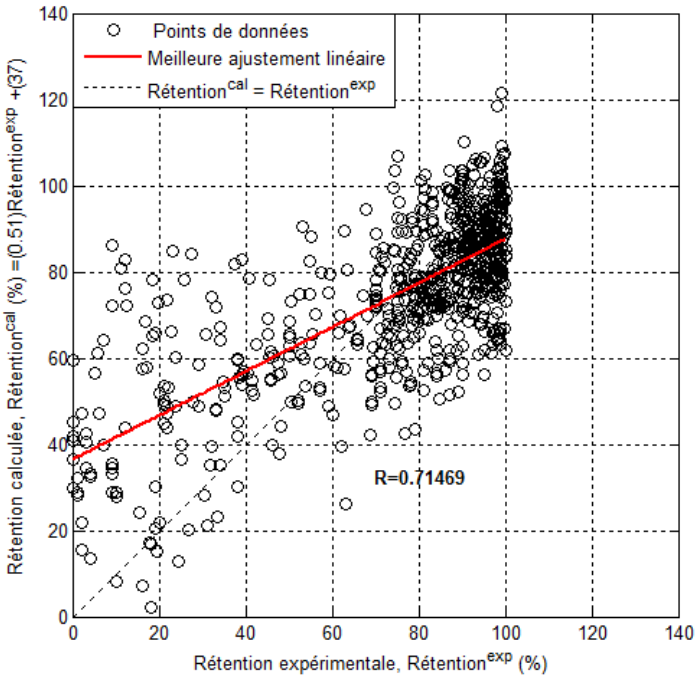
(a)


Points de données
Meilleure ajustement linéaire
Rétention^{cal} = Rétention^{exp}
R=0.73593
Rétention calculée, Rétention^{cal} (%) =(0.54)Rétention^{exp} +(35)
Rétention expérimentale, Rétention^{exp} (%)
0
20
40
60
80
100
120

(b)


Points de données
Meilleure ajustement linéaire
Rétention^{cal} = Rétention^{exp}
R=0.65348
Rétention calculée, Rétention^{cal} (%) =(0.41)Rétention^{exp} +(44)
Rétention expérimentale, Rétention^{exp} (%)
0
20
40
60
80
100
120

(c)



(d)

Figure IV.22. Modèle QSAR-RLM3 : les valeurs calculées vs. Les valeurs expérimentales pour les ensembles de données : (a) phase d'apprentissage, (b) phase de validation, (c) phase de test, (d) totale.

Le tableau IV.6 montre les vecteurs de régression linéaire et les performances pour chaque modèle de QSAR-RLM. Ce tableau indique que les paramètres de la régression les plus favorables sont ceux obtenus par QSAR-RLM3. Il est clair que les résultats obtenus dans cette étude indiquent que les modèles de QSAR-RLM proposés ont une capacité prédictive modérée pour les processus de la rétention des molécules organiques (neutres et ioniques) par les membranes d'OI/NF.

Tableau IV.6. Vecteurs de régression linéaire issue de l'équation linéaire : [équation linéaire : $y^{cal} = \alpha y^{exp} + \beta$, ***α***(pente), ***β***(ordonnée à l'origine), ***R***(coefficient de corrélation)], l'erreur absolue moyenne ***(EAM)***, la racine de l'erreur quadratique moyenne (***REQM***) et l'erreur standard de prédiction ***(ESP)***.

		α	***β***	***R***	***EAM***	***REQM***	***ESP***
QSAR-RLM 1	**Phase d'apprentissage**	0.4833	38.6502	0.6958	14.8272	19.9219	26.7756
	Phase de validation	0.4895	38.1303	0.6784	15.1439	19.4816	25.5854
	Phase de test	0.5215	35.0610	0.7362	14.4548	18.1956	24.4095
	Totale	0.4874	38.2604	0.6980	14.8217	19.7134	26.4290
QSAR-RLM 2	**Phase d'apprentissage**	0.5004	37.4867	0.7038	14.3946	19.1989	25.7103
	Phase de validation	0.5059	37.6789	0.7196	12.0425	15.1105	18.8986
	Phase de test	45.2731	0.3799	0.6308	18.1881	23.7644	32.9829
	Totale	0.4860	38.5344	0.6783	14.5385	19.3437	25.8124
QSAR-RLM 3	**Phase d'apprentissage**	0.5250	35.5444	0.7211	14.0435	18.7243	25.0747
	Phase de validation	0.5417	35.0678	0.7359	11.3255	14.7056	18.3922
	Phase de test	0.4050	44.4571	0.6535	17.3757	23.2357	32.2492
	Totale	0.5111	36.6951	0.7147	14.1048	18.8702	25.1807

L'estimation des divers paramètres inconnus des modèles (Tableaux IV.7, IV.8 et IV.9) révèle que seuls trois coefficients des modèles polynomiaux sont jugés significatifs (P-value < 0.05) : L'effet linéaire de MW, *LogD,* moment dipolaire, *Length, Eqwidth,* MWCO, potentiel zêta, angle de contact, pression et taux de récupération pour QSAR-RLM1 ; MW, *Length, Eqwidth,* potentiel zêta et taux de récupération pour QSAR-RLM2 et MW, moment dipolaire, *Length, Eqwidth,* MWCO, potentiel zêta, angle de contact, pression et taux de récupération pour QSAR-RLM3.

Tableau IV.7. Coefficients du modèle estimés par analyse de régression linéaire multiple (QSAR-RLM1).

	Coeff. SC	**Std. Err**	**P-value**
Constante/Intercepté	-40.4736	7.4613	$7.38\ 10^{-8}$ [a]
MW	0.1698	0.0147	$7.89\ 10^{-29}$ [a]
LogD	-1.2353	0.3982	0.0020 [a]
Moment dipolaire	1.4790	0.3860	$1.35\ 10^{-4}$ [a]
Length	22.8341	3.0821	$2.83\ 10^{-13}$ [a]
Eqwidth	51.6208	6.4357	$3.09\ 10^{-15}$ [a]
MWCO	-0.1210	0.0101	$1.29\ 10^{-30}$ [a]
Potentiel zêta	-0.2496	0.0362	$1.02\ 10^{-11}$ [a]
Angle de contact	0.5339	0.0542	$7.52\ 10^{-22}$ [a]
pH	1.2114	0.6379	0.0579
Pression	0.0020	0.0009	0.0380 [a]
Taux de récupération	-0.1488	0.0255	$7.10\ 10^{-9}$ [a]

[a]Termes significatifs (P-value < 0.05)

Tableau IV.8. Coefficients du modèle estimés par analyse de régression linéaire multiple (QSAR-RLM2).

	Coeff. SC	**Std. Err**	**P-value**
Constante/Intercepté	-155.3461	74.8329	0.0383 [a]
MW	0.1331	0.0198	$3.62\ 10^{-11}$ [a]
LogD	0.1327	0.4814	0.7829
Moment dipolaire	0.6828	0.4604	0.1385
Length	15.7248	3.4179	$5.01\ 10^{-6}$ [a]
Eqwidth	71.9422	9.4101	$7.02\ 10^{-14}$ [a]
SR($MgSO_4$)	105.9362	73.9451	0.1524
Potentiel zêta	-0.3252	0.0928	$4.85\ 10^{-4}$ [a]
Angle de contact	0.2592	0.0937	0.0058
pH	1.0009	0.7360	0.1743
Pression	0.0029	0.0011	0.0094
Taux de récupération	-0.1315	0.0297	$1.07\ 10^{-5}$ [a]

[a]Termes significatifs (P-value < 0.05)

Tableau IV.9. Coefficients du modèle estimés par analyse de régression linéaire multiple (QSAR-RLM3).

	Coeff. SC	**Std. Err**	**P-value**
Constante/Intercepté	289.3667	104.8497	0.0059
MW	0.1030	0.0200	$3.27\ 10^{-7}$ [a]
LogD	0.0051	0.4704	0.9913
Moment dipolaire	1.1700	0.4570	0.0107 [a]
Length	21.2466	3.4649	$1.46\ 10^{-9}$ [a]
Eqwidth	90.4811	9.7067	$1.53\ 10^{-19}$ [a]
MWCO	-0.1512	0.0256	$5.29\ 10^{-9}$ [a]
SR($MgSO_4$)	-327.7187	102.9066	0.0015
Potentiel zêta	-0.3417	0.0906	$1.76\ 10^{-4}$ [a]
Angle de contact	0.2828	0.0915	0.0021 [a]
pH	0.7164	0.7201	0.3202
Pression	0.0025	0.0011	0.0231 [a]
Taux de récupération	-0.1666	0.0296	$2.56\ 10^{-8}$ [a]

[a]Termes significatifs (P-value < 0.05)

L'analyse de la régression peut être utilisée comme une technique d'analyse de sensibilité. Les techniques de régression permettent d'évaluer la sensibilité des entrées de modèles, en tenant compte de l'impact simultané des autres paramètres du modèle sur le résultat (Cullen et Frey 1999). L'analyse de la régression comme une méthode d'analyse de sensibilité est appliquée dans divers domaines tels que la science vétérinaire (Taylor et al. 1992), sciences sociales (Barrett et al. 1986), sciences de l'alimentation et la sécurité alimentaire. L'importance des entrées a été indiquée par les erreurs standards dans les modèles de régressions linéaires multiples. L'importance (Std. Err) de chaque d'entrées sur la rétention pour les trois modèles est montrée dans les tableaux IV.7, IV.8 et IV.9 ci dessus.

On constate que les entrées "*LogD*, moment dipolaire, *Length*, *EqWidth*, SR($MgSO_4$) et pH" ont des importances très éloignées de zéro et ceux pour chaque QSAR-RLM. Cette analyse de sensibilité a montré à posteriori l'importance des six paramètres choisis pour la modélisation de la rétention des molécules organiques lors d'OI/NF. D'après le tableau IV.9, on remarque

que le seuil de coupure "MWCO" est moins efficace que le rejet de sel de magnésium "SR ($MgSO_4$)" pour prédire la rétention des molécules organiques lors d'OI/NF (l'importance de (MWCO) = 0.02 et l'importance de (SR "$MgSO_4$") = 102.91). Aussi, il est plus simple de caractériser une membrane en termes de SR ($MgSO_4$) plutôt que MWCO. Ces résultats sont en accord avec les résultats de l'analyse de sensibilité par la méthode "poids" et les conclusions de Yangali-Quintanilla et al.2009b et 2010b.

IV.4. Comparaison entre les modèles QSAR

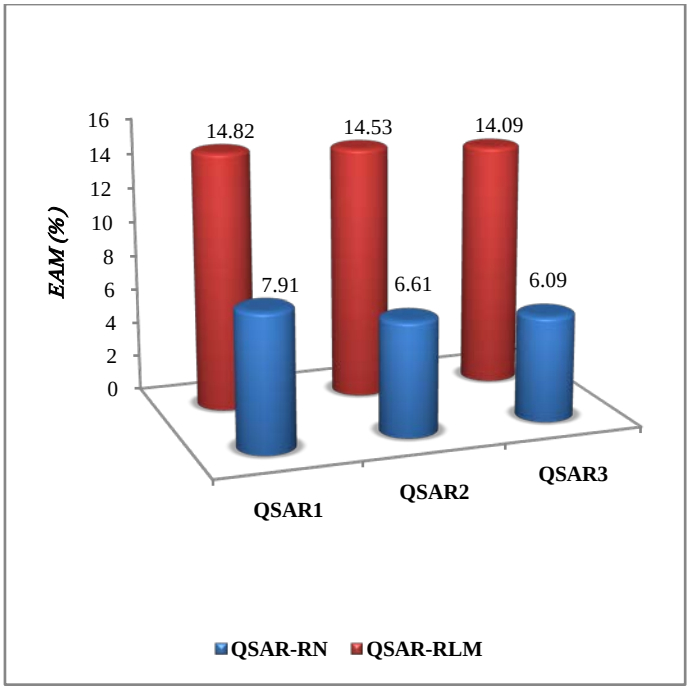
IV.4.1. Comparaison entre les modèles QSAR-RN et QSAR-RLM

La figure IV.23 montre une comparaison entre les modèles QSAR-RN et QSAR-RLM en termes de l'erreur absolue moyenne *(EAM*), la racine de l'erreur quadratique moyenne (*REQM*) et l'erreur de standard de prédiction (*ESP*), pour la totalité de l'ensemble de données. La comparaison des résultats entre les modèles de QSAR-RN et les modèles QSAR-RLM a montré la supériorité des modèles QSAR-RN. *EAM*, *REQM* et *ESP* sont définies comme suit (Khaouane 2013 ; Hemmati-Sarapardeh et al. 2014 ; Shokir et al. 2014 ; Adib et al. 2015 ; Faisal et al. 2015) :

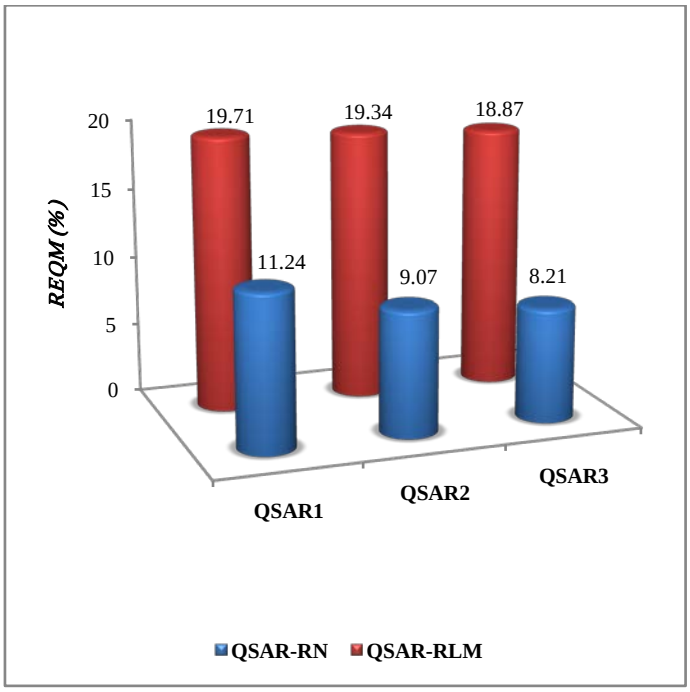
$$EAM = \frac{1}{n}\sum_{i=1}^{n}\left|(y_{i,exp} - y_{i,cal})\right| \qquad \text{(IV. 10)}$$

$$ESP(\%) = \frac{REQM}{Y_e} * 100 \qquad \text{(IV. 11)}$$

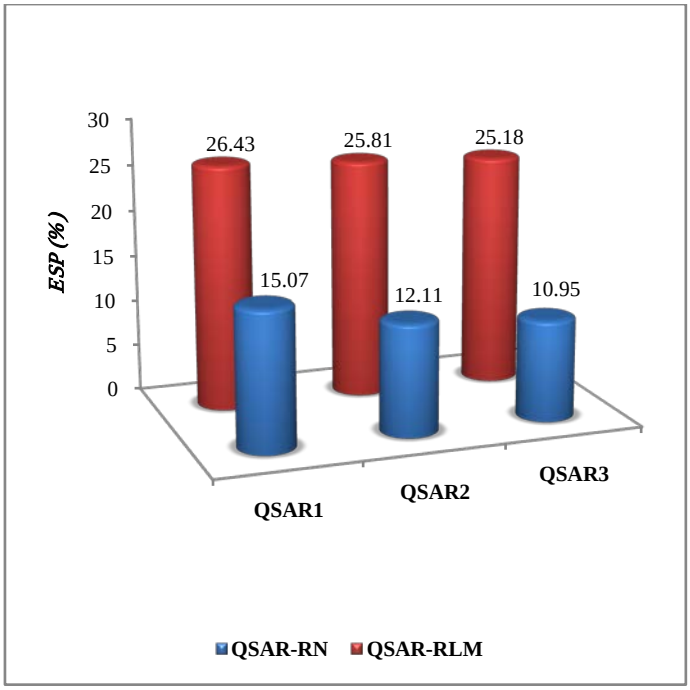
Où n est le nombre total de données, $y_{i,exp}$ est la valeur expérimentale, $y_{i,cal}$ représente la valeur calculée à partir du modèle QSAR et Y_e est la valeur moyenne des données expérimentales.


EAM (%)
16
14
12
10
8
6
4
2
0
14.82
14.53
14.09
7.91
6.61
6.09
QSAR1
QSAR2
QSAR3
QSAR-RN
QSAR-RLM

(a)


REQM (%)
20
15
10
5
0
19.71
19.34
18.87
11.24
9.07
8.21
QSAR1
QSAR2
QSAR3
QSAR-RN
QSAR-RLM

(b)



(c)

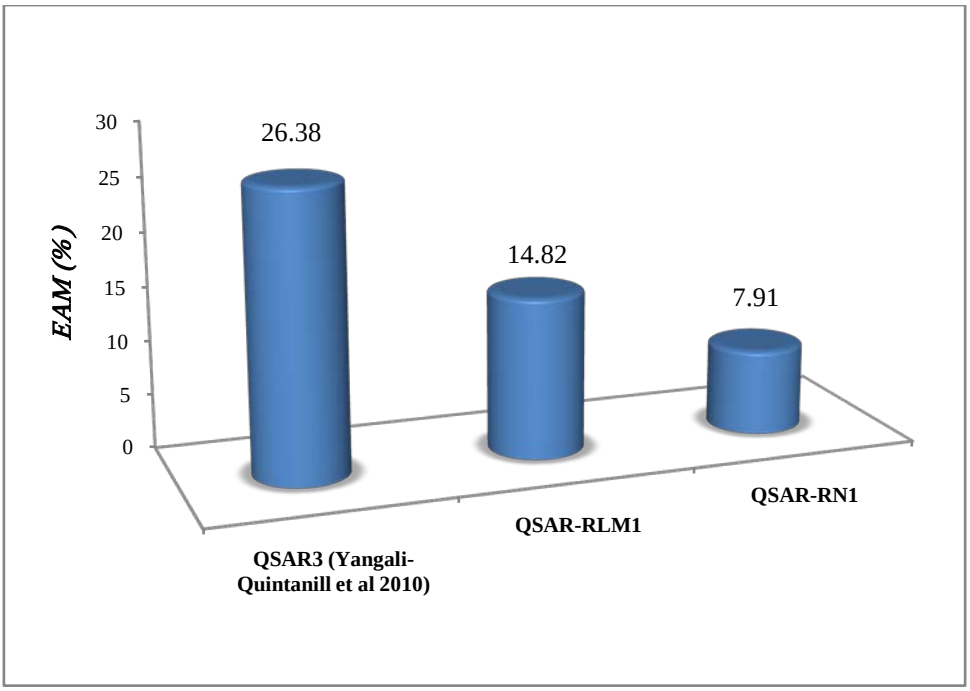
Figure IV.23. Comparaison des résultats de la base totale entre le modèle de QSAR-RN et le modèle QSAR-RLM : (a) *EAM*, (b) *REQM*, (c) *ESP*.

IV.4.2. Comparaison entre nos modèles QSAR et les modèles QSAR de la littérature

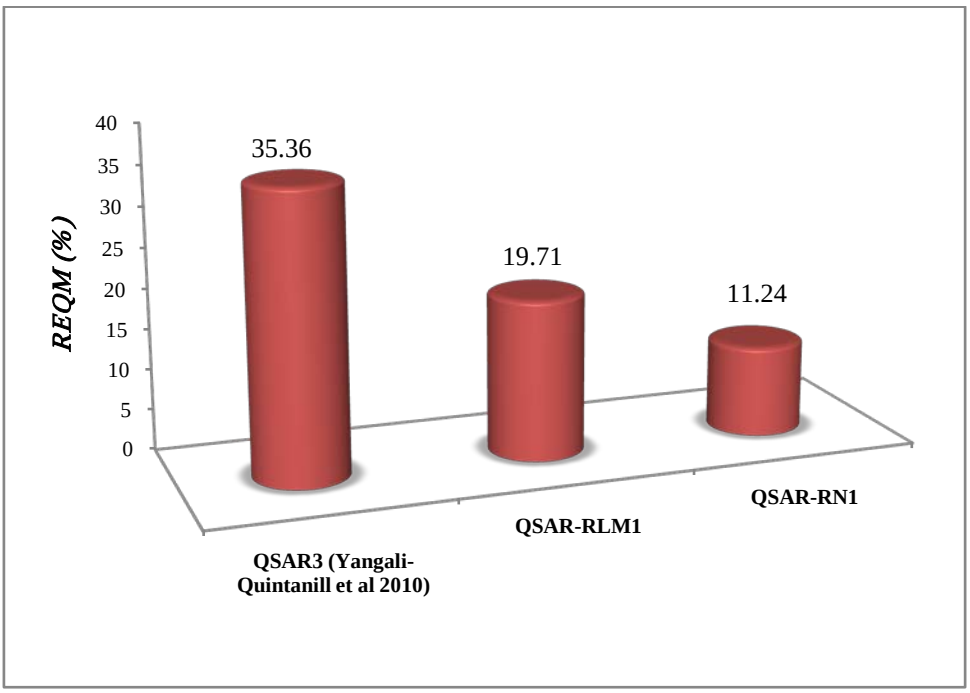
Comme mentionné dans le chapitre II, un nombre limité de modèles QSAR sur la prédiction de la rétention des molécules organiques (neutres et ioniques) par les membranes d'OI/NF ont déjà été mis au point. Certains d'entre eux étaient basés sur le poids moléculaire de coupure "MWCO" ou le rejet de sel de sulfate de magnésium "SR($MgSO_4$)" comme des indicateurs de porosités.

Afin d'établir les modèles QSAR-RN et QSAR-RLM développés comme une alternative plausible aux modèles QSAR de la littérature, une comparaison entre les trois approches a été faite en termes de l'erreur absolue moyenne (*EAM*), la racine de l'erreur quadratique moyenne (*REQM*) et l'erreur standard de prédiction (*ESP*). La figure IV.24 représente une comparaison entre nos modèles (QSAR-RN1 et QSAR-RLM1) et le modèle QSAR développé par Yangali-Quintanilla et al. 2010b "le troisième modèle QSAR est évoqué par l'équation II.10 qui est déjà décrite dans le premier chapitre". Les modèles QSAR-RN et QSAR-RLM développés dans ce travail ont donné des erreurs plus faibles que le modèle QSAR de Yangali-Quintanilla et al.

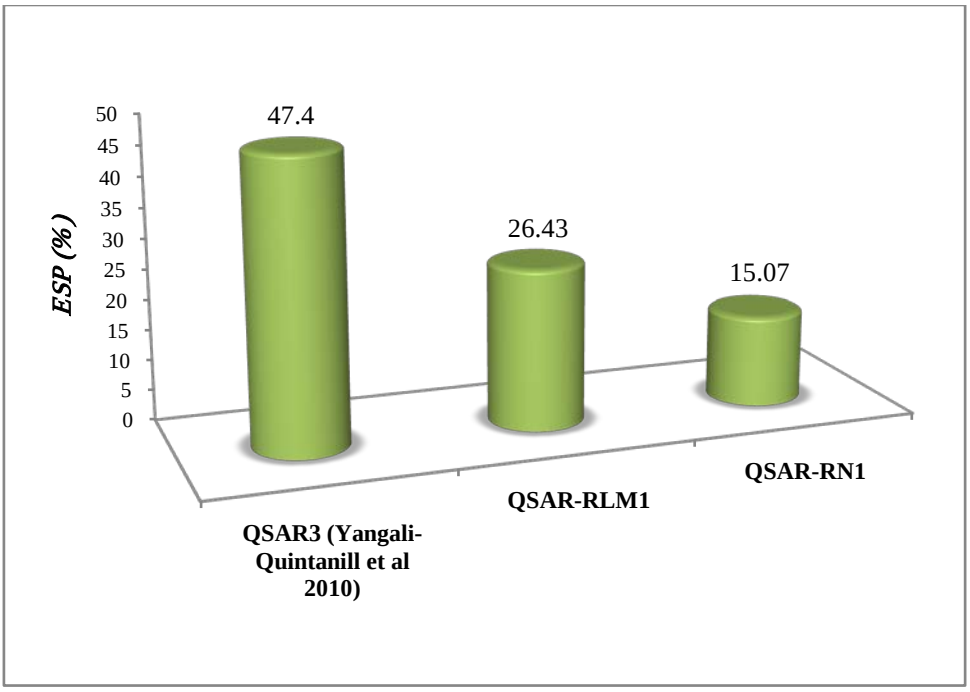
2010b, ce qui montre que nos modèles possèdent une capacité de modélisation élevée par rapport au modèle de Yangali-Quintanilla et al. 2010b.



(a)



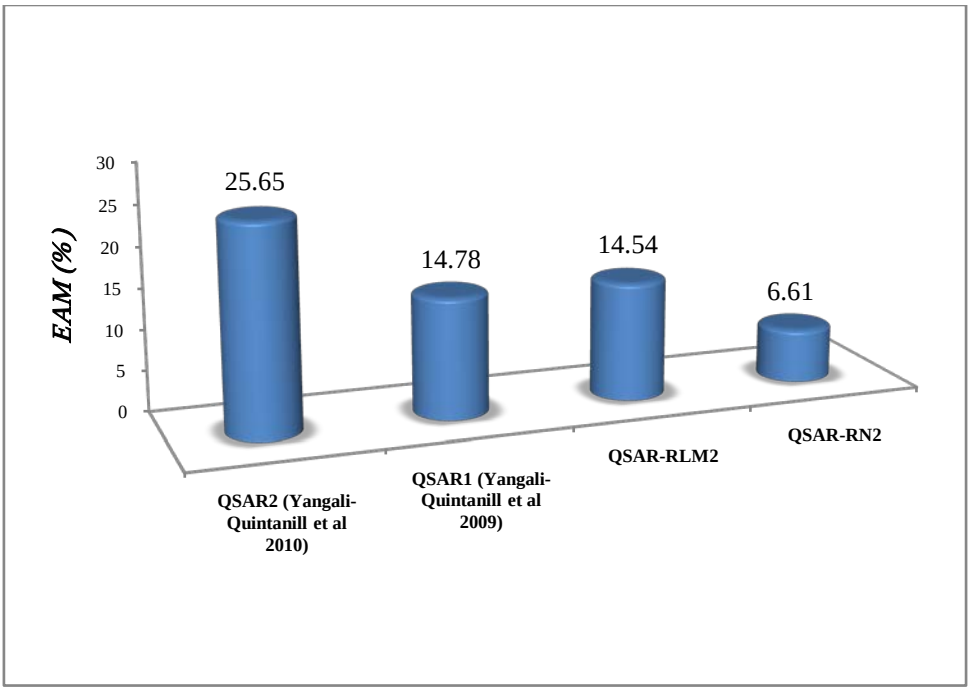
(b)



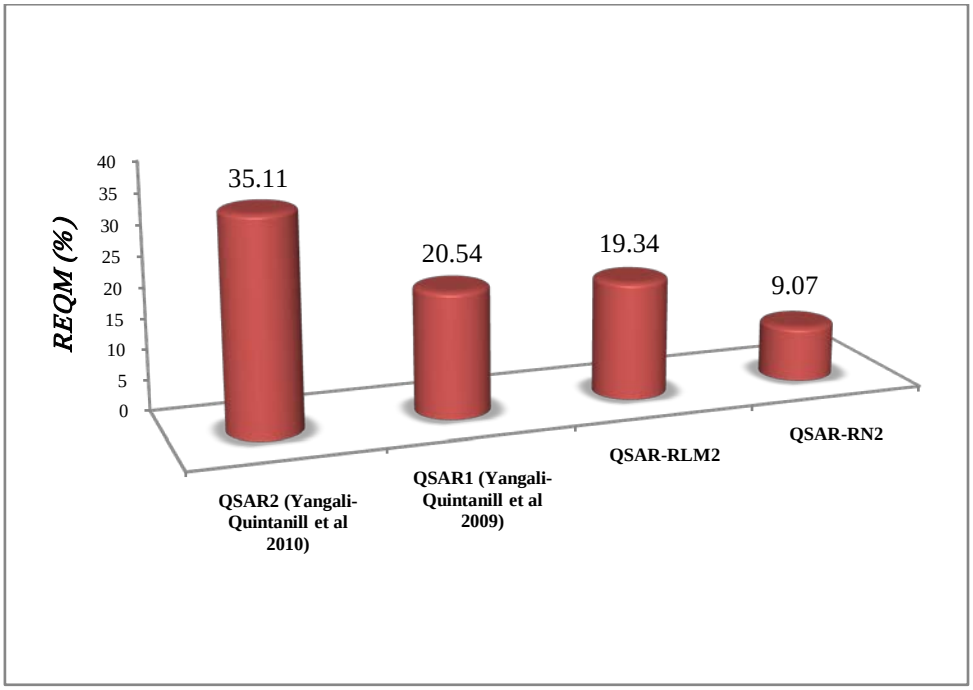
(c)

Figure IV.24. Comparaison entre nos modèles QSAR et les modèles QSAR de la littérature "MWCO est indicateur de porosité" : (a) *EAM*, (b) *REQM*, (c) *ESP*.

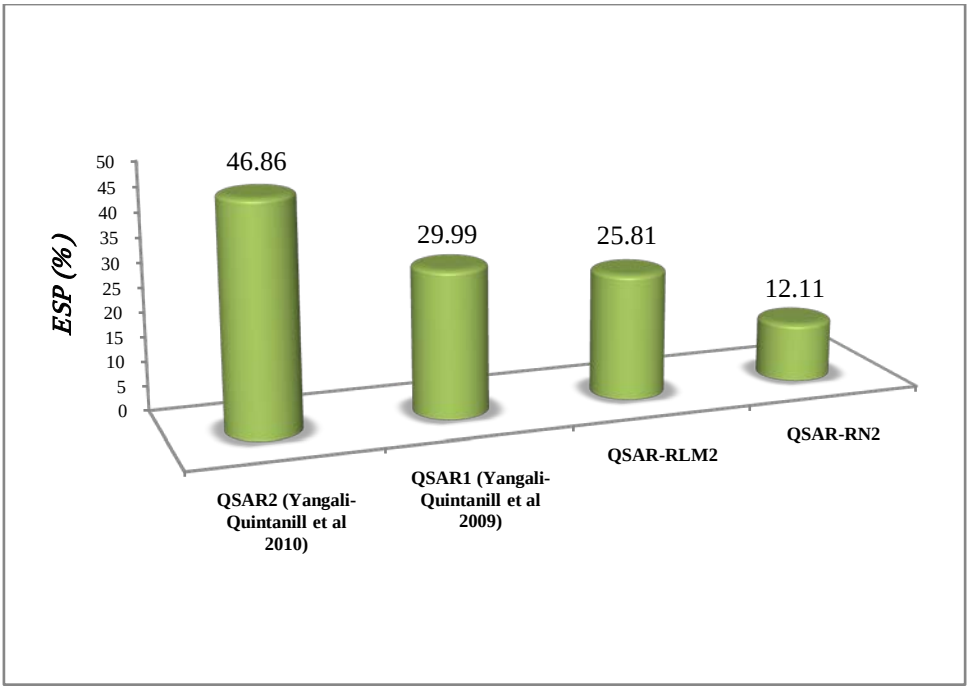
La figure IV.25 représente la comparaison entre nos modèles (QSAR-RN2 et QSAR-RLM2) et les modèles QSAR développés par Yangali-Quintanilla et al. 2009b et 2010b "le premier modèle QSAR est indiqué par l'équation II.8 et le deuxième modèle QSAR est représenté par l'équation II.9 qui sont déjà décrits dans le premier chapitre". La performance de chaque modèle a été évaluée en termes de l'erreur absolue moyenne (*EAM*), la racine de l'erreur quadratique moyenne (*REQM*) et l'erreur standard de prédiction (*ESP*). Les modèles QSAR-RN et QSAR-RLM développés dans ce travail ont donné des erreurs plus faibles que les modèles QSAR de Yangali-Quintanilla et al. 2009b et 2010b, ces résultats montrent la supériorité de nos modèles QSAR.


EAM (%)
30
25
20
15
10
5
0
25.65
14.78
14.54
6.61
QSAR2 (Yangali-Quintanill et al 2010)
QSAR1 (Yangali-Quintanill et al 2009)
QSAR-RLM2
QSAR-RN2

(a)


REQM (%)
40
35
30
25
20
15
10
5
0
35.11
20.54
19.34
9.07
QSAR2 (Yangali-Quintanill et al 2010)
QSAR1 (Yangali-Quintanill et al 2009)
QSAR-RLM2
QSAR-RN2

(b)



(c)

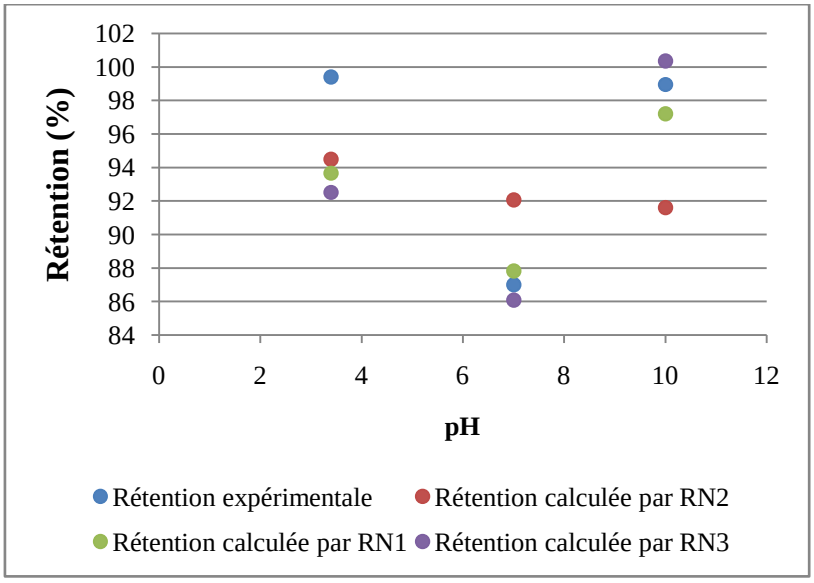
Figure IV.25. Comparaison entre nos modèles QSAR et les modèles QSAR de la littérature "SR($MgSO_4$) est un indicateur de porosité" : (a) *EAM*, (b) *REQM*, (c) *ESP*.

IV.5. Rétention des composés organiques en fonction du pH

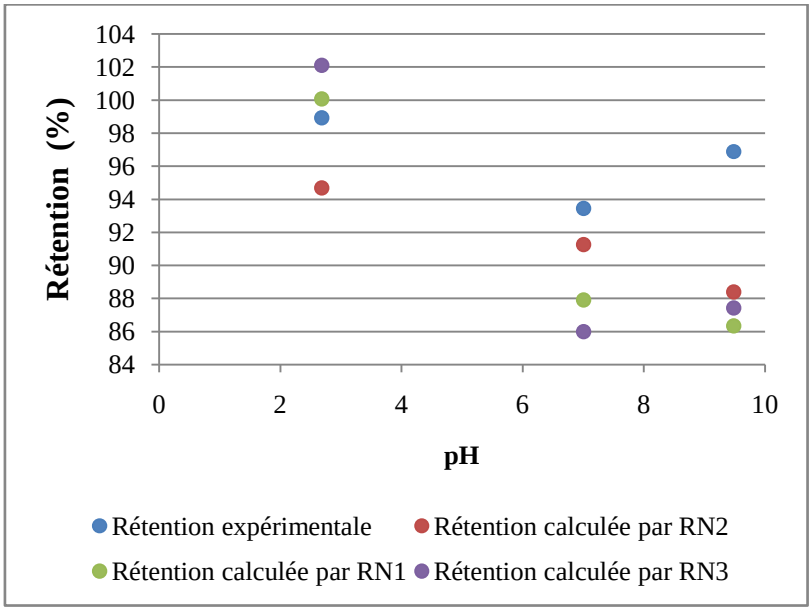
Les valeurs expérimentales et calculées pour rétention des composés organiques sélectionnés (chargés positivement, neutres, mais également chargés négativement) à différents pH sont illustrées dans les figures IV.26, IV.27 et IV.28 à la fois pour Trisep TS80 et Desal HL membranes, et pour chacun des modèles (RN1, RN2 et RN3).

Les valeurs de la rétention calculées sont en accord avec les valeurs de la rétention expérimentales pour les composés organiques chargés positivement (Clenbutérol et le Salbutamol) (Figure IV.26), ces résultats sont interprétés que s'il y a une bonne validation pour les modèles RN. À pH 10, les valeurs de la rétention calculées sont en accords modérés des valeurs de la rétention expérimentales pour Salbutamol, parce qu'il y a une validation acceptable pour les modèles RN. Les valeurs de la rétention pour tous les composés organiques (chargés positivement) à pH neutre (7) sont faibles pour les membranes Trisep TS80 et Desal HL, car à pH 7, tous les composés organiques sont également chargés positivement, mais les deux membranes deviennent chargées négativement, la répulsion électrostatique se change en attraction électrostatique (Verliefde et al. 2008). À pH 10, la membrane est chargée

négativement de plus à pH 7, mais la plupart des composés organiques (chargés positivement) sont neutres au lieu de charge positive à ce pH. Par conséquent, pas des interactions électrostatiques entre la membrane et les composés organiques ne devraient être produites à un pH de 10, mais la rétention est plus élevée qu'à pH 7 (Verliefde et al. 2008).



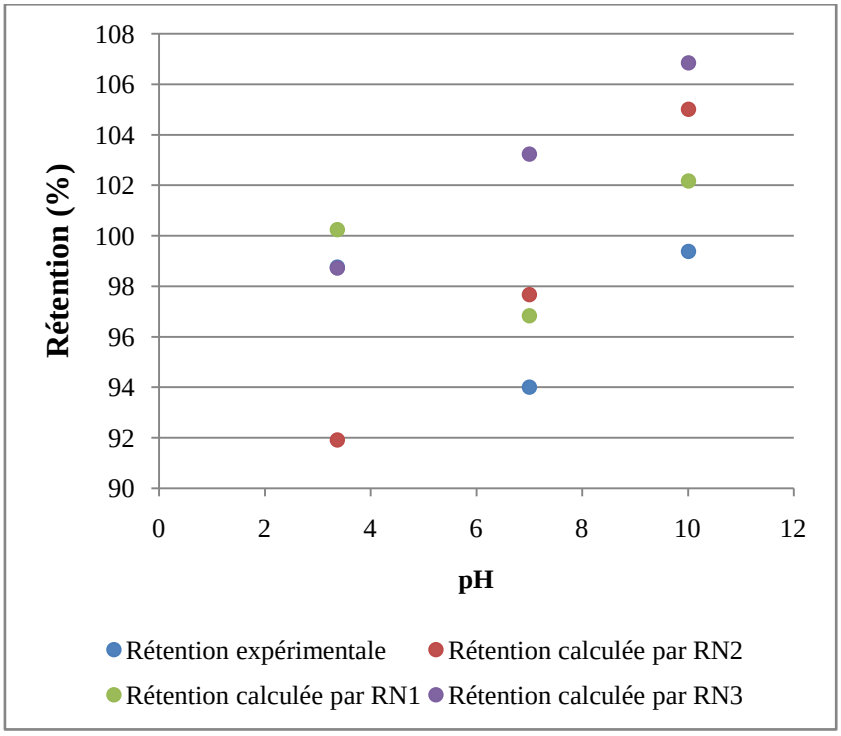
(a)



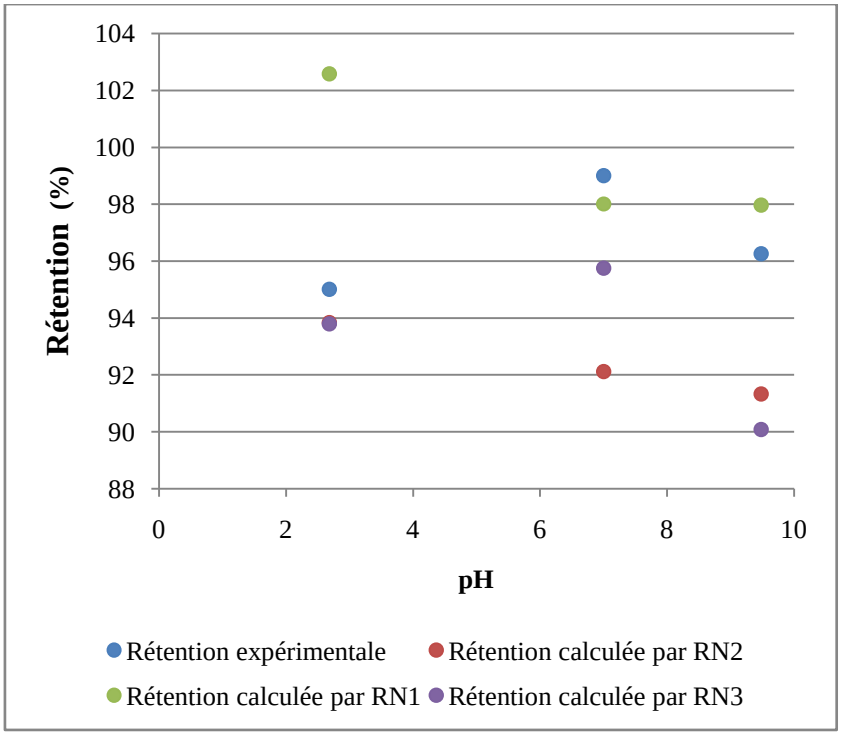
(b)

Figure IV.26. Comparaison entre la rétention expérimentale et calculée des composés organiques chargés positivement en fonction du pH de l'eau d'alimentation avec Trisep TS80 (Clenbuterol, (a)) et Desal HL (Salbutamol, (b)) pour chaque modèle (RN1, RN2 et RN3).

Les valeurs de la rétention calculées sont en accord avec les valeurs de la rétention expérimentales pour les composés organiques neutres (Cyclophosphamide et Pentoxifylline) (Figure IV.27), les valeurs de la rétention pour toutes les molécules organiques sont relativement élevées (≥90%) pour les deux membranes. Ces résultats sont interprétés que s'il y a une grande validation des modèles RN. Pour les composés organiques neutres, le pH n'a presque aucun effet sur la rétention, toutes les valeurs de la rétention calculées et expérimentales sont relativement constantes dans tout l'intervalle de pH testé (Verliefde et al. 2008).



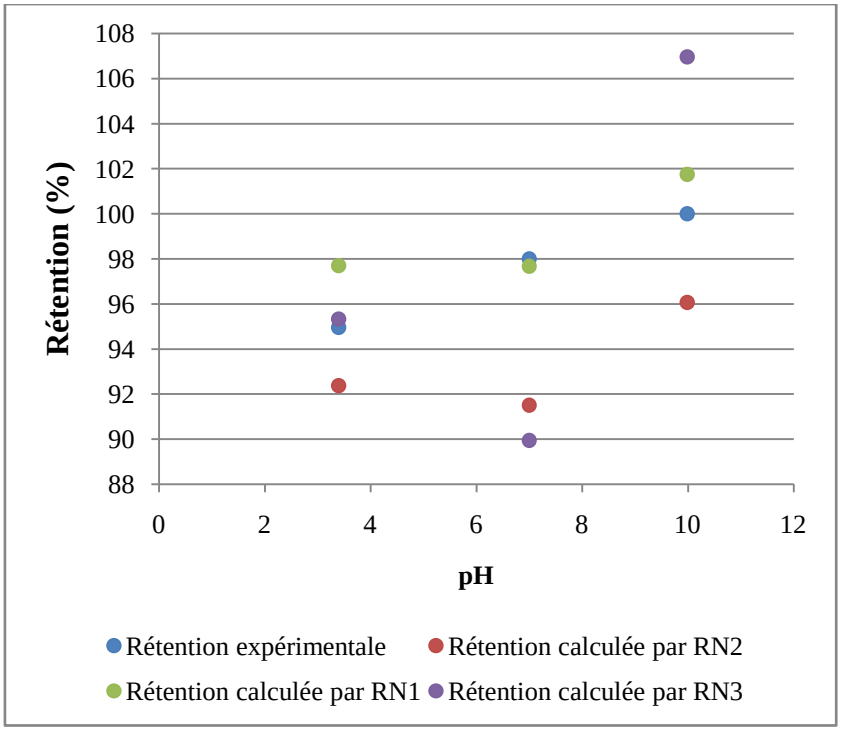
(a)



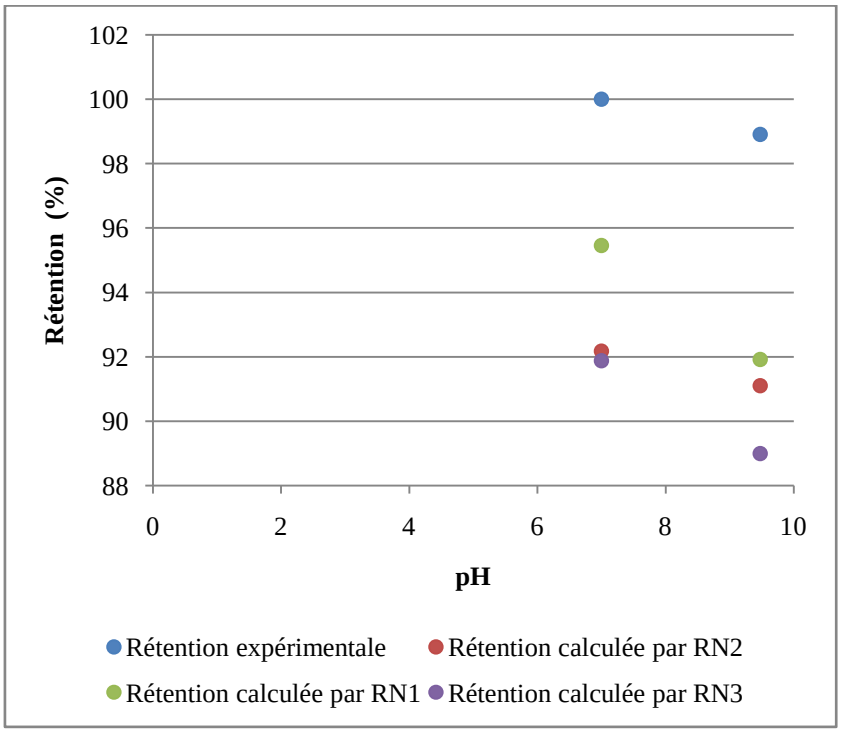
(b)

Figure IV.27. Comparaison entre la rétention expérimentale et calculée des composés organiques neutres en fonction du pH de l'eau d'alimentation avec Trisep TS80 (Cyclophosphamide, (a)) et Desal HL (Pentoxifylline, (b)) pour chaque modèle (RN1, RN2 et RN3).

Les valeurs de la rétention calculées sont en accord avec les valeurs de la rétention expérimentales pour les composés organiques chargés négativement (Acide clofibrique et Fenoprofen) (Figure IV.28), ces résultats sont interprétés que s'il existe une excellente validation pour les modèles RN. Les valeurs de la rétention pour tous les composés organiques (chargés négativement) à pH neutre et haut (7 et 10) sont élevées, en raison de la répulsion électrostatique entre les composés organiques chargés négativement et la surface de la membrane chargée négativement (Verliefde et al. 2008). À pH inférieur (pH<7), les membranes deviennent neutres (ou même chargées positivement), et le dissocier, les composés organiques chargés négativement (acides) retrouver leur proton et devenir neutres. Par conséquent, aucune répulsion électrostatique n'existe plus et la rétention de composés organiques est seulement déterminée par l'encombrement stérique et les interactions hydrophobes entre les composés organiques et la membrane (Verliefde et al. 2008).



(a)



(b)

Figure IV.28. Comparaison entre la rétention expérimentale et calculée des composés organiques chargés négativement en fonction du pH de l'eau d'alimentation avec Trisep TS80 (Acide clofibrique, (a)) et Desal HL (Fenoprofen, (b)) pour chaque modèle (RN1, RN2 et RN3).

Il peut être observé pour les deux membranes que les valeurs de la rétention de composés organiques chargés positivement sont inférieures aux valeurs de la rétention de composés organiques neutres, qui sont à leur tour plus faibles que les valeurs de la rétention de composés organiques chargés négativement. Cela a été expliqué par les interactions électrostatiques entre les solutés et la surface de la membrane : pour les solutés à charge négative, la répulsion de charge existe entre les solutés et la surface de la membrane chargée négativement. Pour les solutés neutres, aucune interaction de charge avec la surface de la membrane n'existe plus, mais pour les solutés chargés positivement, les charges approuvent l'attraction de charge (Verliefde et al. 2008).

Les résultats des modèles RN sont en harmonie avec la réponse de Verliefde et al. 2008, les modèles RN peuvent être utilisés pour prédire les rétentions de composés organiques neutres et ioniques à une large gamme de pH.

IV.6. Conclusion

Ce présent chapitre a donné une raison de plus pour l'utilisation généralisée des modèles QSAR-RN et QSAR-RLM pour la modélisation de la rétention des molécules organiques (neutres et ioniques) par les membranes d'osmose inverse et de nanofiltration. Le but de la présente étude était de développer trois modèles QSAR-RN de type feed-forward et trois modèles QSAR-RLM capables de résumer les interactions entre les caractéristiques de la membrane, les conditions d'opératoire de filtration et les propriétés physico-chimiques des composés organiques (neutres et ioniques).

Les modèles QSAR-RN développés dans ce travail ont donné des erreurs plus faibles que les modèles QSAR-RLM et les modèles QSAR de la littérature. Les erreurs absolues moyennes pour l'ensemble de données totales étaient 7.9138 pour QSAR-RN1, 6.6125 pour QSAR-RN2 et 6.0995 pour QSAR-RN3, les racines des erreurs quadratiques moyennes étaient 11.2430 pour QSAR-RN1, 9.0742 pour QSAR-RN2 et 8.2047 pour QSAR-RN3 et les erreurs standards de prédictions étaient 15.0730 pour QSAR-RN1, 12.1087 pour QSAR-RN2 et 10.9485 pour QSAR-RN3. Les erreurs absolues moyennes pour l'ensemble de données totales étaient 14.8217 pour QSAR-RLM1, 14.5385 pour QSAR-RLM2 et 14.1048 pour QSAR-RLM3, les racines des erreurs quadratiques moyennes étaient 19.7134 pour QSAR-RLM1, 19.3437 pour QSAR- RLM2 et 18.8702 pour QSAR-RLM3 et les erreurs standards de prédictions étaient 26.4290 pour QSAR-RLM1, 25.8124 pour QSAR-RN2 et 25.1807 pour

QSAR-RLM3. Les erreurs absolues moyennes pour les modèles QSAR de la littérature étaient 14.7812 pour QSAR1, 25.6549 pour QSAR2 et 26.3838 pour QSAR3, les racines des erreurs quadratiques moyennes étaient 20.8394 pour QSAR1, 35.1141 pour QSAR2 et 35.3575 pour QSAR3 et les erreurs standards de prédictions étaient 29.9998 pour QSAR1, 46.8567 pour QSAR2 et 47.4025 pour QSAR3.

L'analyse de sensibilité par la méthode de "poids" a révélé que les variables les plus importantes qui influent sur la prédiction de la rétention de composés organiques (neutres et ioniques) sont : MW, *LogD*, *Length*, *Eqwidth*, MWCO, SR ($MgSO_4$), angle de contact, potentiel zêta, pH, pression et taux de récupération. D'autre part, l'analyse de sensibilité par la méthode de "régression linéaire" appliquée aux trois modèles QSAR-RLM a démontré l'importance des six paramètres (*LogD*, moment dipolaire, *Length*, *EqWidth*, SR($MgSO_4$) et pH) choisis pour modéliser la prédiction de la rétention de composés organiques (neutres et ioniques).

La rétention de sel de sulfate de magnésium peut être un paramètre forfaitaire possible pour la modélisation de la rétention des molécules organiques (neutres et ioniques) par les membranes d'OI/NF.

Conclusion générale et perspectives

Les réseaux de neurones possèdent une propriété remarquable dont l'apport est manifesté pour la modélisation des processus non linéaires : ce sont des approximateurs universels parcimonieux. Les modèles des réseaux neurones artificiels (RNA) ont finalement été appliqués pour la prévision de séparation des substances organiques par les membranes d'osmose inverse et de nanofiltration qui ont été abordés par d'autres chercheurs d'une façon réductible dont la taille de la base de données et les paramètres calculés ont donné des résultats encourageants.

L'objectif de cette thèse a porté sur l'utilisation des réseaux de neurones et des régressions linéaires multiples pour la modélisation du processus de la rétention des composés organiques (neutres et ioniques) par les membranes d'osmose inverse et de nanofiltration. Pour cela six modèles QSAR "trois modèles neuronaux de type feed-forward et trois régressions linéaires multiples" sont construits, il est capable de résumer les interactions entre les caractéristiques de la membrane, les conditions opératoires de filtration et les propriétés des composés organiques (neutres et ioniques). Les trois modèles neuronaux de type feedforward (QSAR-RN) caractérisés par une structure similaire (onze neurones pour QSAR-RN1 et QSAR-RN2 et douze neurones pour QSAR-RN3 dans la couche d'entrée, une couche cachée et un neurone dans la couche de sortie) sont construits dans le but de prédire la rétention des composés organiques (neutres et ioniques). Une bonne concordance, entre les valeurs expérimentales et les valeurs prédites, pour chaque modèle neuronal a été obtenue (les racines des erreurs quadratiques moyennes pour la base de données totale étaient 11.2430% pour QSAR-RN1, 9.0742% pour QSAR-RN2 et 8.2047% pour QSAR-RN3). Les valeurs prédites par QSAR-RLM pour la rétention étaient en accord modérées avec les valeurs expérimentales (les racines des erreurs quadratiques moyennes pour la base de données totale étaient 19.7134 pour QSAR-RLM1, 19.3437 pour QSAR- RLM2 et 18.8702 pour QSAR-RLM3).

En outre, la comparaison des résultats entre les modèles de QSAR-RN et les modèles QSAR-RLM a montré la supériorité des modèles QSAR-RN et une autre comparaison entre les modèles QSAR développés dans ce travail et les modèles QSAR de Yangali-Quintanilla et al montre que nos modèles QSAR ont donné beaucoup mieux (les racines des erreurs quadratiques moyennes étaient 20.8394 pour QSAR1, 35.1141 pour QSAR2 et 35.3575 pour QSAR3). D'autre part, l'analyse de sensibilité a révélé que les variables les plus importantes qui influent sur la rétention de composés organiques (neutres et ioniques) sont : MW, *LogD*, *Length*, *EqWidth*, MWCO, SR ($MgSO_4$), angle de contact, potentiel zêta, pH, pression et taux de

récupération. Aussi, l'analyse de sensibilité par régression linéaire appliquée aux trois modèles QSAR-RLM a démontré l'importance des six paramètres (*LogD*, moment dipolaire, *Length*, *EqWidth*, SR($MgSO_4$) et pH) choisis pour modéliser la prédiction de la rétention de composés organiques (neutres et ioniques). En plus le rejet de sel de sulfate de magnésium peut être un paramètre forfaitaire possible pour la modélisation de la rétention des molécules organiques (neutres et ioniques) par d'OI/NF.

À la lumière des résultats obtenus, on peut tirer les conclusions suivantes :

- Les réseaux de neurones sont des outils statistiques, qui permettent d'ajuster des fonctions non linéaires très générales à des ensembles de points comme toute méthode statistique ;
- Les réseaux de neurones fonctionnent comme une boîte noire et les paramètres du réseau (poids et biais) n'ont pas de signification physique. L'application des réseaux neuronaux ne développe pas directement de la théorie de description du système, mais rend faisable la modélisation et la simulation des procédés d'osmose inverse et de nanofiltration ;
- La modélisation par réseau de neurones est une méthode qui est basée seulement sur des données expérimentales et/ou semi-expérimentales (déduites des modèles analytiques, empiriques, numériques, …), donc la qualité et la quantité de données décident de l'exactitude et la robustesse du modèle ;
- Les réseaux de neurones peuvent être très robustes et performants même en présence de peu de données expérimentales à condition que ces dernières soient statistiquement bien réparties dans le domaine d'entrée du modèle neuronal, ce qui est le cas avec le plan composite centré ;
- Les formules mathématiques extraites des modèles neuronaux peuvent trouver des applications intéressantes dans la prédiction, l'optimisation et même le contrôle des procédés d'osmose inverse et de nanofiltration.

De ce travail de thèse ont émergé de nombreuses pistes de recherche qui feront l'objet de futurs travaux.

- ✓ Les perspectives concernent plus particulièrement la modélisation de la rétention des molécules organiques par les membranes d'OI/NF, mais en utilisant d'autres indicateurs

de porosité, on propose de faire une étude avec le rejet de sel SR(NaCl) et le rejet de sel SR($CaCl_2$).

- ✓ Les perspectives concernent plus particulièrement la modélisation d'autres processus (microfiltration, ultrafiltration) pour des fins de prédiction.
- ✓ Nous pouvons aussi étendre l'idée de ce travail à d'autres techniques récentes de modélisation et d'optimisation telles que les approches de modélisation types "boîte grise" et d'autres méta-heuristiques d'optimisation (simulated annealing, bee colony optimization, ...).

Références bibliographiques

Abbas A. et Al-Bastaki N. (2005)."Modeling of an RO water desalination unit using neural networks". Chemical Engineering Journal. **114** (1) : 139–143.

Abdel-Jawad M., Al-Shammari S. et Al-Sulaimi J. (2002). "Non-conventional treatment of treated municipal wastewater for reverse osmosis". Desalination. **142** (1) : 11-18.

Acero J. L, Benitez F. J., Francisco J. R. et Fernando T. (2016). "Micropollutants removal from retentates generated in ultrafiltration and nanofiltration treatments of municipal secondary effluents by means of coagulation, oxidation, and adsorption processes". Chemical Engineering Journal. **289** : 48–58.

Adib H., Hassanajili S., Sheikhi-Kouhsar M., Salahi A. et Mohammadi T. (2015). "Experimental and computational investigation of polyacrylonitrile ultrafiltration membrane for industrial oily waste water treatment". Korean J Chem Eng. **32** (1) :159-167.

Agenson K. O., Oh J. I. et Urase T. (2003). "Retention of a wide variety of organic pollutants by different nanofiltration/reverse osmosis membranes : controlling parameters of process". Journal of Membrane Science. **225** (1-2) : 91-103.

Alkhatim H. S. (1998)."Treatment of whey effluents from dairy industries by nanofiltration membranes". Desalination. **119** (1-3) : 177-183.

Altalyana N., Jonesa B., Bradda J., Nghiemb L. D. et Alyazichia Y. M. (2006)."Removal of volatile organic compounds (VOCs) from groundwater byreverse osmosis and nanofiltration". Journal of Water Process Engineering. **9** : 9-21.

Ammi Y., Khaouane L. et Hanini S. (2015)."Prediction of the rejection of organic compounds (neutral and ionic) by nanofiltration and reverse osmosis membranes using neural networks". Korean Journal of Chemical Engineering. **32** (11) : 2300-2310.

Annane K. (2015). "Extraction et transport des ions de cadium (II) par membranes d'affinité dans des dispositifs microfluidiques". Thèse de doctorat. Université de Mouloud Mammeri. Tizi-Ouzou. Algérie.

Anselme C. et Jacobs E. P. (1996). "Ultrafiltration, In : Water Treatment Membrane Process". Chapitre 10, McGraw-Hill : 88.

Antoine B. (2004). "Etude de la nanofiltration pour son intégration dans les procédés dans de production d'acides organiques". Thèse de doctorat. Université Paul Sabatier. Toulouse. Françe.

Aptel Ph. et Buckley C. A. (1996). "Categories of membrane operations, In : Water Treatment Membrane Process ". Chapitre 2, McGraw-Hill : 24.

Ardhaoui M., Cherichi H., Ognier S. et Ghoul M. (2003)."Caractérisation des propriétés de surface des membranes de nanofiltration: développement d'un outil d'aide à la décision". Récents Progrès en Génie des Procédés 90 (SFGP 2003, St Nazaire) : 32-39.

Baker R. W. (2004). "Memebrane Technology and application". 2nd ed, John Wiley and Sons Chichester : 287-295.

Barrett, G. V., Alexander R. A., Anesgart M. N. et Doverspike D. (1986). "Frequently Encountered Problems in the Application of Regression-Analysis to the Investigation of Sex-Discrimination in Salaries". Public Personnel Management. **15** (2) : 143-157.

Basheer I. A. et Hajmeer M. (2000)."Artificial neural networks : Fundamentals, computing, design, and application". Journal of Microbiological Methods. **43** (1) : 3-31.

Bellona C. et Drewes J. E. (2005). "The role of membrane surface charge and solute physico-chemical properties in the rejection of organic acids by NF membranes". Journal of Membrane Science. **249** (1) : 227-234.

Bellona C., Drewes J. E., Xu P. et Amy G. (2004). "Factors affecting the rejection of organic solutes during NF/RO treatment - a literature review". Water Research. **38** (12) : 2795-2809.

Berg P., Hagmeyer G. et Gimbel R. (1997). "Removal of pesticides and other micropollutants by nanofiltration". Desalination. **113** (2-3**)** : 205-208.

Bilstad T. (1995). "Nitrogen separation from domestic wastewater by reverse osmosis". Journal of Membrane Science. **102** : 93-102.

Bouamar M. et Ladjal M. (**2007)**. "Système multicapteur utilisant les réseaux de neurones artificiels pour la surveillance des eaux potables". 4th International Conference : Sciences of Electronic, Technologies of Information and Telecommunications March. Tunisia : 25-29.

Bouranene, S. (2008). "Etude de la rétention de soluté neutres et ioniques par des membranes de nanofiltration". Thèse de doctorat. Université de Franche-Comté. École Doctorale Louis Pasteur. Besançon. France.

Bryjak J., Ciesielski K. et Zbicinski I. (2004). "Modelling of glucoamylase thermal inactivation in the presence of starch by artificial neural network". Journal of Biotechnology. **114** (1-2) : 177-185.

Chabane M. (2015)."Étude de la résistance au colmatage et de la rétention des composes toxiques par membranes". Thèse de doctorat. Université d'Aboubekr Belkaid de Tlemcen. Tlemcen. Algérie.

Chatkaew T. (2009). "Procédés hybrides à membranes pour le prétraitement d'eau de mer avant dessalement par osmose inverse". Thèse de doctorat. Université de Toulouse. Toulouse. France.

Chikhi M. (2008). "Etude de la compléxation des cations métalliques en vue de leur séparation par un procédé membranaire". Thèse de doctorat. Université Mentouri. Constantine. Algérie.

Childress A. E. et Desmukh S. S. (1998). "Effects of humic substances and anionic surfactants on the surface charge and performance of reverse osmosis membranes". Desalination. **118** (1-3) : 167-174.

Childress A. E. et Elimelech M. (1996). "Effect of solution chemistry on the surface charge of polymeric reverse osmosis and nanofiltration membranes". Journal of Membrane Science. **119** (2) : 253-268.

Cliff M. A., Dever M. C., Hall J. W. et Girard B. (1995). "Development and Evaluation of Multiple Regression Models for Prediction of Sweet Cherry Liking". Food Research International. **28** (6) : 583-589.

Comerton A. M., Andrews R. C. et Bagley D. M. (2009). "The influence of natural organic matter and cations on the rejection of endocrine disrupting and pharmaceutically active compounds by nanofiltration". Water Research. **43** : 613 – 622.

Comerton A. M., Andrews R. C., Bagley D. M. et Hao C. (2008). "The rejection of endocrine disrupting and pharmaceutically active compounds by NF and RO membranes as a function of compound and water matrix properties". Journal of Membrane Science. **313** : 323–335.

Conell, D. W. **(1989)**. "Bioaccumulation of Xenobiotic Compounds". CRC Press : 75–110.

Connell D. W. (1990). "Bioaccumulation of Xenobiotic Compounds". CRC Press. Boca Raton, FL : 75-110.

Cornelissen E. R., Sijbers P., Van den Berkmortel H., Koning J., De Wit A., De Nil F. et Van Impe J. F. (2001). "Reuse of leachate waste-water using MEMBIOR technology and reverse osmosis". Membrane Technology. **136** : 6-9.

Cullen A.C., et Frey H.C. (1999). "Probabilistic Techniques in Exposure Assessment". Plenum Press : New York.

Desmukh S. S. et Childress A. E. (2001). "Zeta potential of commercial RO membranes : influence of source water type and chemistry". Desalination. **140** (1) : 87-95.

Dolar D., Ana V., Danijela A. et Kresimir., K. (2011). "Effect of water matrices on removal of veterinary pharmaceuticals by nanofiltration and reverse osmosis membranes". Journal of Environmental Sciences. **23** (8) : 1299–1307.

Dolar D., Ignjatic Zokic T., Kosutic K., Asperger D. et Mutavdzic Pavlovic D. (2012). "RO/NF membrane treatment of veterinary pharmaceutical wastewater: comparison of results obtained on a laboratory and a pilot scale Environmental science and pollution". Research International. **19** (4) :1033-1042.

Dolar D., Zokić T. I., Košutić K., Periša M. et Babić S. (2013)."Photolysis of enrofloxacin and removal of its photodegradation products from water by reverse osmosis and nanofiltration membranes". Separation and Purification Technology. **115** (13) : 1-8.

Donghan G., Yuhong D. et Ruiliang Y. (2016). "Performance analysis of an organic Rankine cycle for a reverse osmosis desalination system using zeotropic mixtures". Desalination. **381** : 38–46.

Dornier M., Decloux M., Trystram G. et Lebert A. (1995)."Dynamic modeling of cross flow microfiltration using neural networks". Journal of Membrane Science. **98** (3) : 263–273.

Dreyfus G., Martinez J. M., Samuelides M., Gordon M. B., Badran F., Thiria S. et Hérault L. (2002). "Réseaux de neurones : Méthodologie et applications". Éditions Eyrolles, Paris : 408.

Ecosse D., (2001). "Techniques alternatives en vue de subvenir a la penurie d'eau dans le monde. Mem. D.E.S.S. (Qualite et Gestion de l'Eau)", Fac. Sciences, Amiens : 62.

Faisal A., Cho H., Kim J., Seong N. et Yeo Y. (2015). "A real-time model based on least squares support vector machines and output bias update for the prediction of NO x emission from coal-fired power plant". Korean Journal of Chemical Engineering. **32** (6) : 1029-1036.

Fell C. J. D. **(1995)**. "Reverse Osmosis in : Membrane Separations Technology. Principles and applications". (eds . R. D. Noble and S.A. Stern). New York, Elsevier : 113-142.

Galambos I., Molina J. M., Jaray P., Vatai G. et Bekassy-Molnar E. (2004). "High organic content industrial wastewater treatment by membrane filtration". Desalination. **162** : 117-120.

Garson G. D. (1991). "Interpreting neural-network connection weights". Artificial Intelligence Expert. **6** (4) : 46-51.

Gevrey M., Dimopoulos I. et Lek S. (2003). "Review and comparison of methods to study the contribution of variables in artificial neural network models". Ecological Modelling. **160** (3) : 249-264.

Girard B., Fukumoto L. R. (2000). "Membrane processing of fruit juices and beverages : a review". Critical Reviews in Biotechnology. **20** (2) : 109-175.

Goh A. T. C. (1995). "Back-propagation neural networks for modeling complex systems". Artificial Intelligence in Engineering. **9** (3) : 143-151.

Golbraikh A., Tropsha A. (2002). " Beware of q2! " J. Mol. Graph. Model. **2** : 269-276.

Gumi T., Valiente M., Khulbe K. C., Palet C. et Matsuura T. (2003). "Characterization of activated composite membranes by solute transport, contact angle measurement, AFM and ESR". Journal of Membrane Science **212** (1-2) : 123-134.

Hamadache M. (2014). "Modélisation en vue de la prédiction ou de la corrélation de l'activité toxicologique à partir de la structure moléculaire". Thèse de doctorat. Université de Médéa. Médéa. Algérie.

Hamlaoui M. L., Kherrat R., Marrakchi M., Jaffrezic-Renault N. et Walcarius A. (2002). "Development of an ammonium ISFET sensor with a polymeric membrane including zeolite". Materials Science and Engineering **21** (1-2) : 25-28.

Hansch C., Leo A. et Hoekman D., (1996). "Exploring QSAR - Hydrophobic, Electronic, and Steric Constants". American Chemical Society. **39** (5) :1189–1190.

Hemmati-Sarapardeh A., Majidi S. M. J., Mahmoudi B., Ramazani S. A. A. et Mohammadi A (2014). "Experimental measurement and modeling of saturated reservoir oil viscosity". Korean Journal of Chemical Engineering. **31** (7) : 1253-1264.

Hu J. Y., Ong S. L., Shan J. H., Kang J. B. et Ng W. J. (2003). "Treatability of organic fractions derived from secondary effluent by reverse osmosis". Water Research. **37** (19) : 4801-4809.

Hong S. et Elimelech M. (1997). "Chemical and physical aspects of natural organic matter (NOM) fouling of nanofiltration membranes". Journal of Membrane Science **132** (2) : 159-181.

Huang H., Cho H., Schwab K. et Jacangelo J. G. (2011). "Effects of feedwater pretreatment on the removal of organic microconstituents by a low fouling reverse osmosis membrane". Desalination. **281**: 446–454.

Kaliszan R. et Markuszewski M. J. (2003). "Studies on correlation between structure of solutes and their retention". Chemia Analityczna. **48** (3) : 373-395.

Khaouane L. (2013). "Etude et Modélisation de la Biosynthèse des Antibiotiques à partir de Différentes Souches Productrices – Cas de Pleuromutiline". Thèse de doctorat. Université de Médéa. Médéa. Algérie.

Kim T-U., Drewes JE., Scott Summers R., Amy GL. (2007). "Solute transport model for trace organic neutral and charged compounds through nanofiltration and reverse osmosis membrane". Water Research. **41** (17) : 3977-3988.

Kimura K., Amy G., Drewes J., Heberer T., Kim T. et Watanabe Y. (2003). "Rejection of organic micropollutants (disinfection byproducts, endocrine disrupting compounds, and pharmaceutically active compounds) by NF/RO membranes". Journal of Membrane Science. **227** (1–2) : 113–121.

Kimura K., Iwase T., Kita S. et Watanabe Y., (2009). "Influence of residual organic macromolecules produced in biological wastewater treatment processes on removal of pharmaceuticals by NF/RO membranes". Water Research. **43** (15) : 3751–3758.

Kimura K., Toshima S., Amy G. et Watanabe Y. (2004). "Rejection of neutral endocrine disrupting compounds (EDCs) and pharmaceutical active compounds (PhACs) by RO membranes". Journal of Membrane Science. **245** (1-2) : 71-87.

Kiso Y., Kltao T., Jinnob K. et Miyae M. (1992)."The effects of molecular width on permeation of organic solute through cellulose acetate reverse osmosis membranes". Journal of Membrane Science. **74** : 95-103.

Kiso Y., Kon T., Kitao T. et Nishimura K. (2001a). "Rejection properties of alkyl phthalates with nanofiltration membranes". Journal of Membrane Science. **182** (1-2) : 205-214.

Kiso Y., Muroshige K., Oguchi T., Yamada T., Hhirose M., Ohara T. et Shintani T. (2010). "Effect of molecular shape on rejection of uncharged organic compounds by nanofiltration membranes and on calculated pore radii". Journal of Membrane Science. **358** (1-2) : 101–113.

Kiso Y., Nishimura Y., Kitao T. et Nishimura K. (2000). "Rejection properties of non-phenylic presticides with nanofiltration membranes". Journal of Membrane Science. **171** (2) : 229-237.

Kiso Y., Sugiura Y., Kitao T. et Nishimura K. (2001b). "Effects of hydrophobicity and molecular size on rejection of aromatic pesticides with nanofiltration membranes". Journal of Membrane Science. **192** : 1–10.

Kolluri S., Esfahani I., Garikiparthy P. et Yoo C. (2015). "Evaluation of multivariate statistical analyses for monitoring and prediction of processes in an seawater reverse osmosis desalination plant". Korean Journal of Chemical Engineering. **32** (8) : 1486-1497.

Kosutic K., Kastelan-Kunst L. et Kunst B. (2000). "Porosity of some commercial reverse osmosis and nanofiltration polyamide thin-film composite membranes". Journal of Membrane Science. **168** (1-2) : 101-108.

Kulkarni A., Han Y. et Hopfinger A. J. (2002). "Predicting caco-2 cell permeation coefficients of organic molecules using membrane-interaction QSAR analysis". Journal of Chemical Information and Computer Sciences. **42** (2) : 331-342.

Lanteri Y. (2009). "Transport a travers des membranes de nanofiltration : caractérisation des propriétés électronique". Thèse de doctorat. Université de Franche-COMTE. Besançon. France.

Libotean D., Giralt J., Giralt F., Rallo R., Wolfe T. et Cohen Y. (2009). " Neural network approach for modeling the performance of reverse osmosis membrane desalting". Journal of Membrane Science. **326** (2) : 408-419.

Libotean D., Giralt J., Rallo R., Cohen Y., Giralta F., Ridgway H. F., Rodriguez G. et Phipps D., (2008). "Organic compounds passage through RO". Journal of Membrane Science. **313** (1-2) : 23–43.

Lin S. H. et Yang C. R. (2004). "Chemical and physical pretreatment of chemical mechanical polishing wastewater from semiconductor fabrication". Journal of Hazardous Materials. **108** (1-2) : 103-109.

Linares R.V., Yangali-Quintanilla V., Li Z. et Amy G. (2011). "Rejection of micropollutants by clean and fouled forward osmosis membrane". Water Research. **45** (20) : 6737 -6744.

Marty P. (1999). "Traitement des effluents par filtration membranaire". Industries Alimentaires et Agricoles. **116** (10) : 50-55.

Mazza C. B., Sukumar N., Breneman C. M. et Cramer S. M., (2001). "Prediction of protein retention in ion-exchange systems using molecular descriptors obtained from crystal structure". Analytical Chemistry. **73** (22) : 5457-5461.

Metaiche M. (2006). "Optimisation des systèmes de dessalement par osmose inverse : conception, paramètres de fonctionnement et simulation numérique". Thèse de doctorat. École nationale polytechnique. Alger. Algérie.

Metaiche M. (2014). "Technologie membranaire". Faculté des sciences et des sciences appliquées de l'université de Bouira, Algérie. Chapitre 1: 4-6.

Nielsen C. E., (1995). "Membrane filtration for strong effluents". Water Management International, Water and Wastewater Treatment.77-80.

Nghiem L. D., Schäfer A. I. et Elimelech M., (2004). "Removal of natural hormones by nanofiltration membranes : measurement, modelling, and mechanisms". Environmental Science & Technology . 38 (6) : 1888–1896.

Norberg D., Hong S., Taylor J. et Zhao Y. (2007). "Surface characterization and performance evaluation of commercial fouling resistant low-pressure RO membranes". Desalination. **202** (1-3) : 45-52.

Ozaki H. et Li H. (2002). "Rejection of organic compounds by ultra-low pressure reverse osmosis membrane". Water Research. **36** (1) : 123-130.

Padilla A. P. et Tavani E. L. (1999). "Treatment of an industrial effluent by reverse osmosis". Desalination. **126** (1) : 219-226.

Petrinic I., Korenak J., Povodnik D. et Hélix-Nielsen C. (2015). "A feasibility study of ultrafiltration/reverse osmosis (UF/RO)-based wastewater treatment and reuse in the metal finishing industry". Journal of Cleaner Production. **101** : 292-300.

Plakas K. V. et Karabelas A. J. (2011). "A systematic study on triazine retention by fouled with humic substances NF/ULPRO membranes". Separation and Purification Technology. **80** (2) : 246–26.

Pontié M. et coll. (2001). "Les techniques séparatives à membranes, Théorie, applications et perspectives". Paris : Union Internationale pour les applications de l'électricité : 212.

Ravikumar Y. V. L., Kalyani S., Satyanarayana S. V. et Sridhar, S. (2014). "Processing of pharmaceutical effluent condensate by nanofiltration and reverse osmosis membrane techniques". Journal of the Taiwan Institute of Chemical Engineers. **45** (1) : 50-56.

Ricci B. C., Ferreira C. D., Aguiar, A. O. et Amaral M. C. S. (2015). " Integration of nanofiltration and reverse osmosis for metal separation and sulfuric acid recovery from gold mining effluent". Separation and Purification Technology. **154** : 11-21.

Roorda J. H., (2004). "Filtration characteristics in dead end ultrafiltration of wwtp-effluent". Thèse de doctorat. Université Delft. Delft. Netherlands.

Sadmani A. H. M. A., Andrews R. C. et Bagley M. D. (2014a). "Impact of natural water colloids and cations on the rejection of pharmaceutically active and endocrine disrupting compounds by nanofiltaration". Journal of Membrane Science. **450** : 272-281.

Sadmani A. H. M. A., Andrews R.C. et Bagley M. D. (2014b)."Nanofiltration of pharmaceutically active and endocrine disrupting compounds as a function of compound interactions with DOM fractions and cations in natural water". Separation and Purification Technology. **122** : 462–471.

Sahoo GB. et Ray C. (2006). "Predicting flux decline in crossflow membranes using artificial neural networks and genetic algorithms". Journal of Membrane Science. **283** (1-2) : 147-157.

Sagne C. (2008). "Etude des mécanismes de transfert de molécules organiques en osmose inverse". Thèse de doctorat. École nationale supérieure de chimie de Montpellier. Montpellier. France.

Santos J. L. C., de Beukelaar P., Vankelecom I. F. J., Velizarov S. et Crespo J. G. (2006). "Effect of solute geometry and orientation on the rejection of uncharged compounds by nanofiltration". Separation and Purification Technology. **50** (1) : 122–131.

Sarkhosh, M., Jahan, B., et Ghasemi, M. A. (2012). "A quantitative structure- property relationship of gas chromatographic/mass spectrometric retention data of 85 volatile organic compounds as air pollutant materials by multivariate methods". Chem. Cent. J. 6(Suppl 2) : S4.

Schäfer A. I., Nghiem L. D. et Waite T. D. (2003). "Removal of the natural hormone estrone from aqueous solutions using nanofiltration and reverse osmosis". Environmental Science & Technology. **37** (1) : 182-188.

Schäfer A. I., Pihlajamäki A., Fane A. G., Waite T. D. et Nyström M. (2004). "Natural organic matter removal by nanofiltration: effects of solution chemistry on retention of low molar mass acids versus bulk organic matter". Journal of Membrane Science. **242** (1-2) : 73-85.

Shahmansouri A. et Christopher B. (2013). "Application of quantitative structure property relationships (QSPRs) to predict the rejection of organic solutes by nanofiltration". Separation and Purification Technology. **118** : 27–638.

Shim Y., Lee H. J., Lee S., Moon S. H. et Cho J. (2002). "Effects of natural organic matter and ion species on membrane surface charge". Environnemental Science and Technology. **36** (17) : 3864- 3871.

Shokir E-M., Al-Homadhi E., Al-Mahdy O. et El-Midany A. H. (2014). "Development of artificial neural network models for supercritical fluid solvency in presence of co-solvents". Korean Journal of Chemical Engineering. **31** (8) :1496-1504.

Si-Moussa C. (2010). "Modélisation des équilibres de phases a pressions élevées par les réseaux de neurones artificiels". Thèse de doctorat. École nationale supérieure polytechnique. Alger. Algérie.

Sridhar S., Kale, A., Khan A. A. (2002). "Reverse osmosis of edible vegetable oil industry effluent". Journal of Membrane Science. **205** (1-2) : 83-90.

Talbi M. L. (2011). "Analyse et traitement du signal électrocardiographique (ECG)." Thèse de Doctorat, Université Mentouri de Constantine. Constantine, Algérie.

Taylor B. C., Stott J. L., Thurmond M. A. et Picanso J. P. (1992)."Alteration in Lymphocyte Subpopulations in Bovine Leukosis Virus-Infected Cattle". Veterinary Immunology and Immunopathology. **31** (1-2) : 35-47.

Taheran M., Brar S. K., Verma M., SurampalliR. Y., Zhang T. C. et Valero J. R (2016). "Membrane processes for removal of pharmaceutically active compounds (PhACs) from water and wastewaters". Science of The Total Environment. **547** : 60-77.

Triki Z. (2014). "Études, Analyses et Optimisation de la Consommation Énergétique des Unités de Dessalement pour les Sites Isolés". Thèse de doctorat, Université Constantine 1. Constantine. Algérie.

Tzanakou E. M. (2000). "Supervised and unsupervised pattern recognition : Feature extraction and computational intelligence". Taylor & Francis Group, London, UK.

Van der Bruggen B., Schaep J., Maes W., Wilms D. et Vandecasteele C. (1998). "Nanofiltration as a treatment method for the removal of pesticides from ground waters". Desalination. **117**: 139-147.

Van der Bruggen B., Schaep J., Wilms D. et Vandecasteele C. (1999). "Influence of molecular size, polarity and charge on the retention of organic molecules by nanofiltration". Journal of Membrane Science. **156** (1) : 29-41.

Verliefde A. R. D., Cornelissen E. R., Heijman S. G. J., Petrinic I., Luxbacher T., Amy G. L., Van der Bruggen B. et van Dijk J. C. (2009a) "Influence of membrane fouling by (pretreated) surface water on rejection of pharmaceutically active compounds (PhACs) by nanofiltration membranes". Journal of Membrane Science. **330** (1-2) : 90–103.

Verliefde A. R. D., Cornelissen E. R., Heijman S. G. J., Verberk J. Q. J. C., Amy G. L., Van der Bruggen B. et van Dijk J.C.(2009b). "Construction and validation of a full-scale model for rejection of organic micropollutants by NF membranes".Journal of Membrane Science. **339** (1-2) : 10-20.

Verliefde A. R. D., Cornelissen E. R., Heijman S. G. J., Verberk J .Q. J. C., Amy G. L., Van der Bruggen B. et van Dijk J. C.(2008). "The role of electrostatic interactions on the rejection of organic solutes in aqueous solutions with nanofiltration". Journal of Membrane Science. **322** (1) : 52–66.

Verliefde A. R. D., Heijman S. G. J., Cornelissen E. R., Amy G., Van der Bruggen B. et van Dijk J. C. (2007). "Influence of electrostatic interactions on the rejection with NF and assessment of the removal efficiency during NF/GAC treatment of pharmaceutically active compounds in surface water". Water Research. **41** (15) : 3227 – 3240.

Xu P., Drewes J. E., Bellona C., Amy G., Kim T. U., Adam M. et Heberer T. (2005). "Rejection of Emerging Organic Micropollutants in Nanofiltration–Reverse Osmosis Membrane Applications". Water Environment Research. **77** (1) : 40-8.

Yangali Quintanilla V. (2010a). "Rejection of emerging organic contaminants by nanofiltration and reverse osmosis membranes effects of fouling, modelling and water reuse". Thèse de doctorat. Université de Delft de technologie. Delft. Netherlands.

Yangali-Quintanilla V., Kim T. U., Kennedy M. et Amy G. (2008). "Modeling of RO/NF membrane rejection of PhACs and organic compounds : a statistical analysis". Drinking Water Engineering and Science. **1** (1) : 7–15.

Yangali-Quintanilla V., Sadmani A., McConville M., Kennedy M. et Amy G. (2010b). "A QSAR model for predicting rejection of emerging contaminants (pharmaceuticals, endocrine disruptors) by nanofiltration membranes". Water Research. **44** (2) : 373 – 384.

Yangali-Quintanilla V., Sadmani A., McConville M., Kennedy M. et Amy G. (2009a). "Rejection of pharmaceutically active compounds and endocrine disrupting compounds by clean and fouled nanofiltration membranes". Water Research. **43** (9) : 2349 – 2362.

Yangali-Quintanilla V., Verliefde A., Kime T. U., Sadmani A., Kennedy M. et Amy G. (2009b). "Artificial neural network models based on QSAR for predicting rejection of neutral organic compounds by polyamide nanofiltration and reverse osmosis membranes". Journal of Memebrane Science. **342** (1-2) : 251-262.

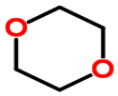
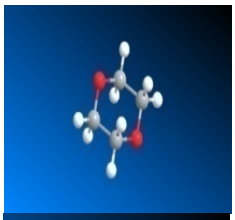
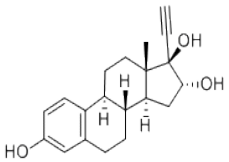
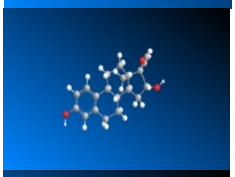
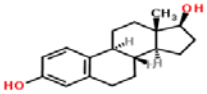
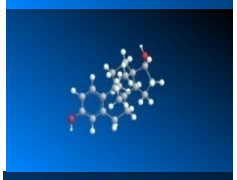
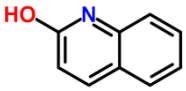
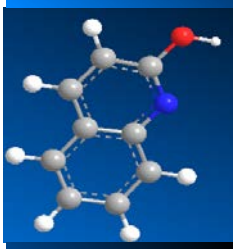
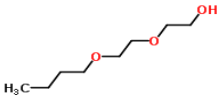
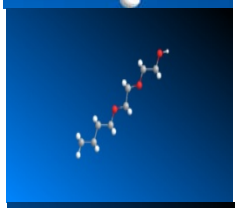
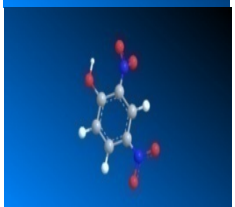
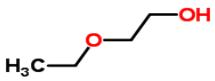
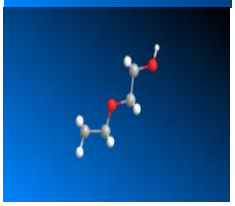
Yangjeh A. H. et Jenagharad M. D. (2009). "Application of a genetic algorithm and an artificial neural network for global prediction of the toxicity of phenols to Tetrahymena pyriformis". Monatsh Chem. **140** : 1279.

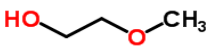
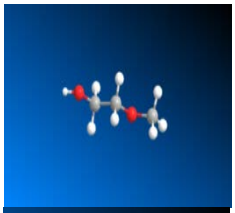
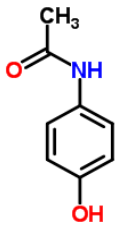
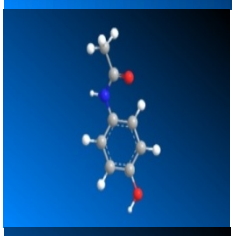
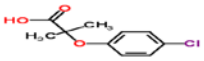
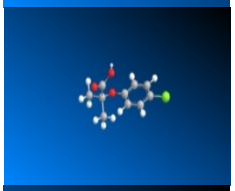
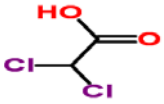
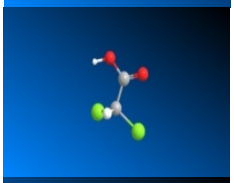
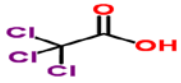
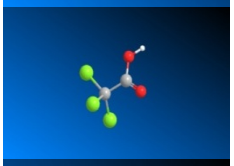
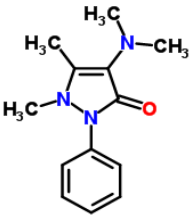
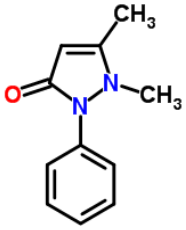
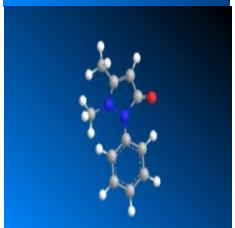
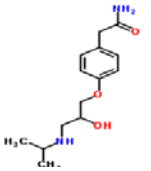
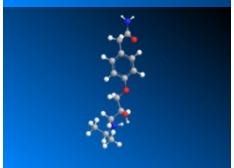
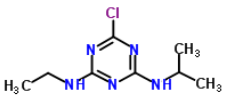
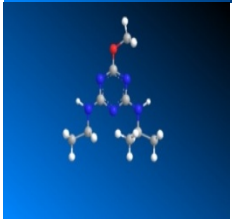
Yoon Y. et Lueptow R. M. (2005). "Removal of organic contaminants by RO and NF membranes". Journal of Memebrane Science. **261** (1-2) : 76–86.

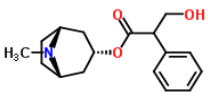
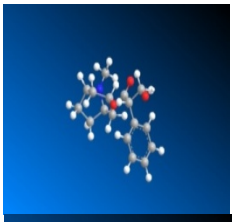
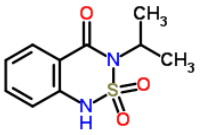
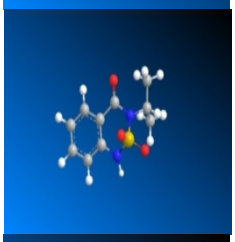
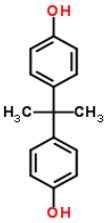
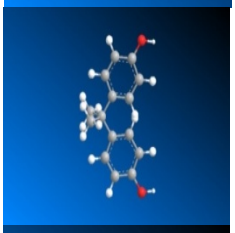
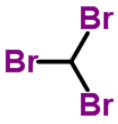
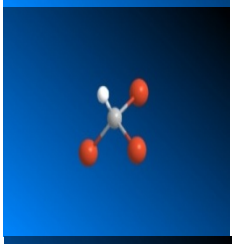
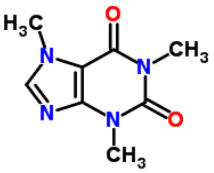
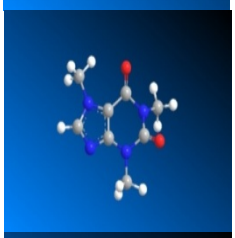
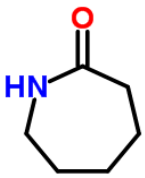
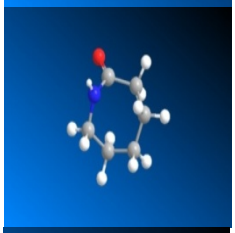
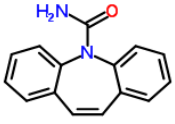
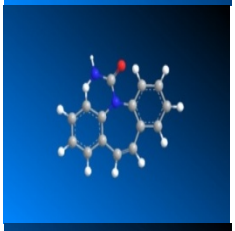
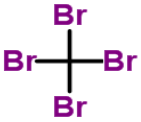
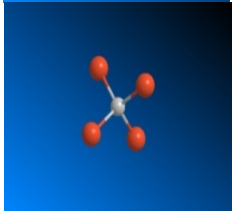
Annexe A

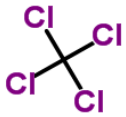
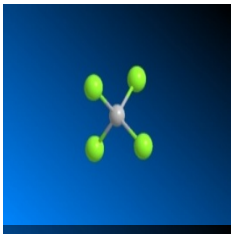
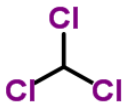
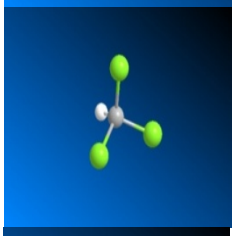
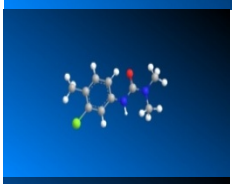
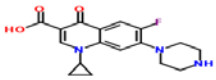
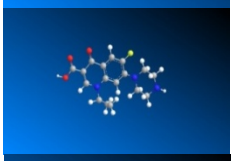
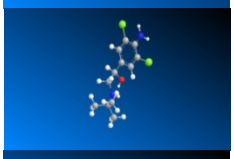
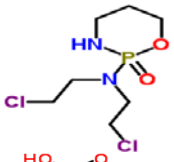
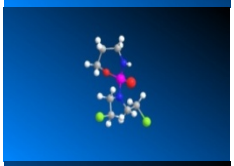
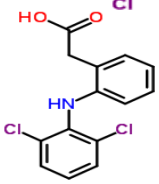
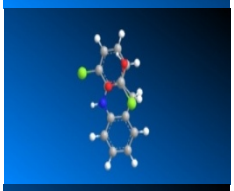
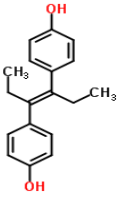
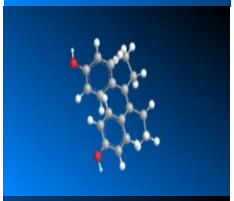
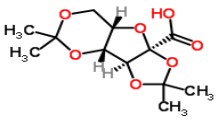
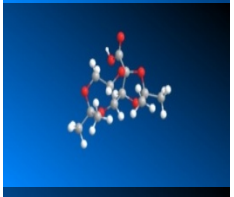
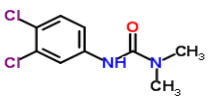
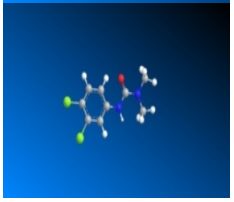
Annexe A.

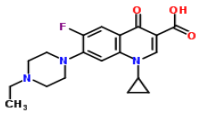
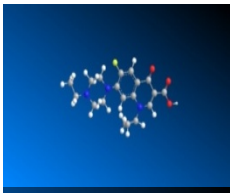
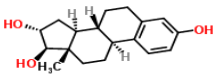
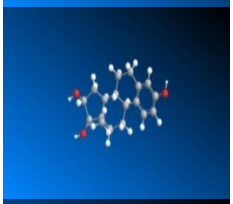
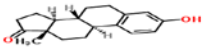
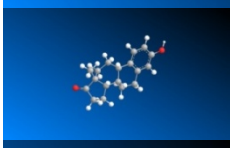
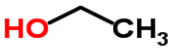
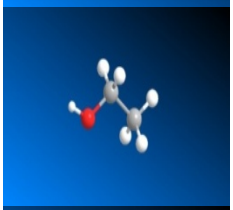
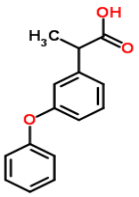
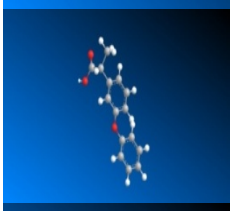
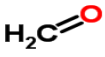
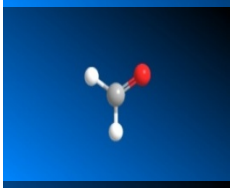
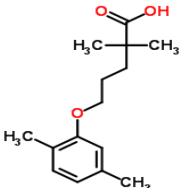
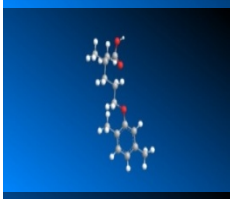
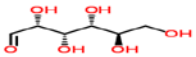
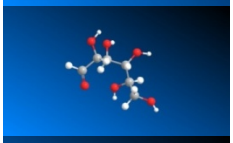
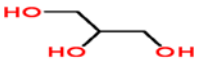
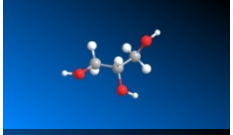
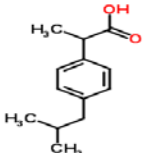
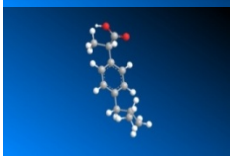
Tableau 1. Liste de 82 molécules organiques utilisées pour développer les modèles QSAR (RN et RLM).

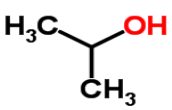
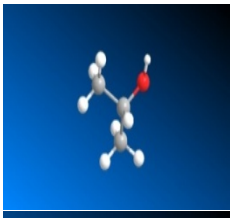
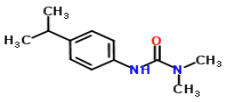
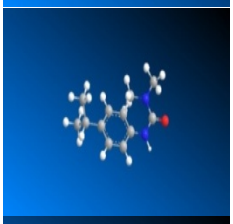
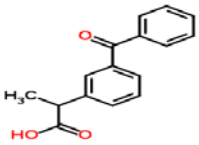
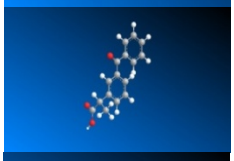
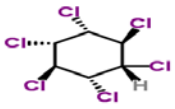
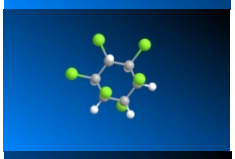
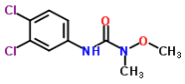
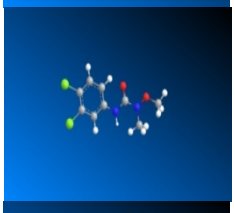
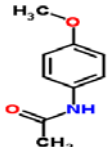
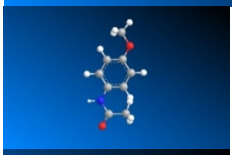
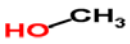
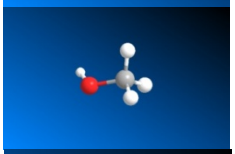
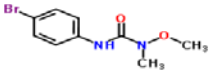
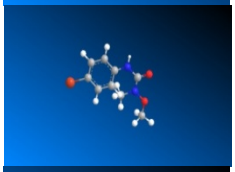
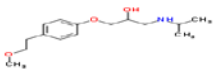
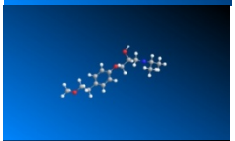
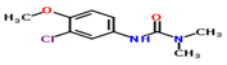
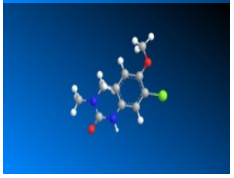
#	Nom	Formule brute	MW (g/mol)	Numéro CAS #	Structure chimique 2D	Structure chimique 3D
01	**1,4-dioxane**	$C_4H_8O_2$	88.03	71-55-6		
02	**17a-ethynylestradiol**	$C_{20}H_{24}O_2$	296.4	4717-40-2		
03	**17β-Estradiol**	$C_{18}H_{24}O_2$	272.39	50-28-2		
04	**2-(1H)-Quinoline**	C_9H_7NO	145.16	59-31-4		
05	**2-(2-Butoxyethoxy)ethanol**	$C_8H_{18}O_3$	162.23	112-34-5		
06	**2,4-dinitrophénol**	$C_6H_4N_2O_5$	184.11	1326-82-5		
07	**2-ethoxyethanol**	$C_4H_{10}O_2$	90.12	110-80-5		

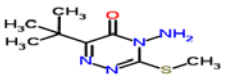
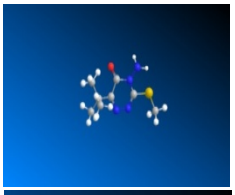
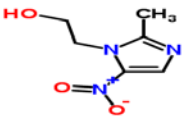
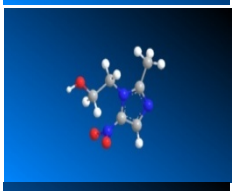
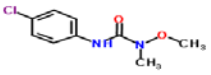
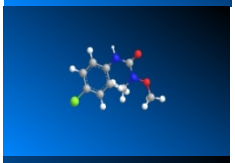
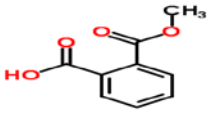
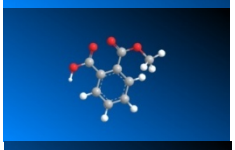
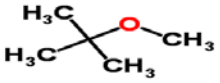
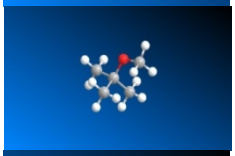
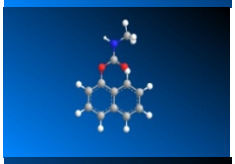
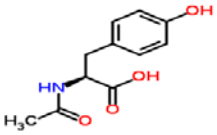
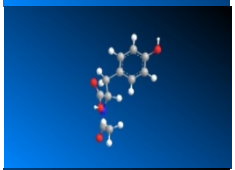
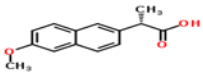
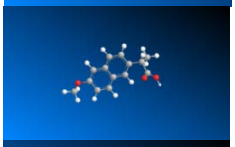
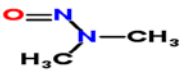
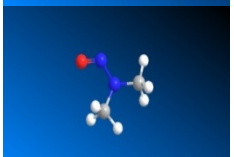
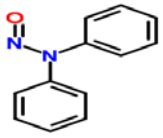
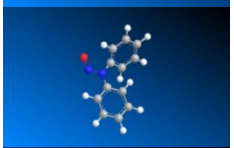
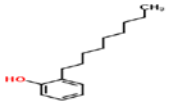
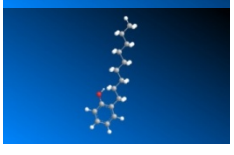
08	**2-methoxyethanol**	$C_3H_8O_2$	76.11	109-86-4		
09	**Acetaminophène**	$C_8H_9NO_2$	151.17	103-90-2		
10	**Acide clofibrique**	$C_{10}H_{11}ClO_3$	214.65	882-09-7		
11	**Acide dichloroacétique**	$C_2H_2Cl_2O2$	128.94	79-43-6		
12	**Acide trichloroacétique**	$C_2HCl_3O_2$	163.39	76-03-9		
13	**Aminopyrine**	$C_{13}H_{17}N_3O$	231.30	58-15-1		
14	**Antipyrine**	$C_{11}H_{12}N_2O$	188.23	60-80-0		
15	**Aténolol**	$C_{14}H_{22}N_2O_3$	266.34	29122-68-7		
16	**Atrazine**	$C_8H_{14}ClN_5$	215.69	1912-24-9		

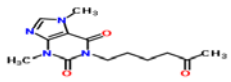
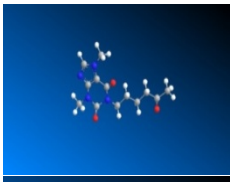
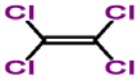
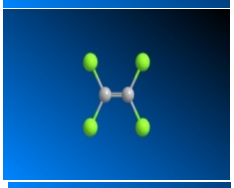
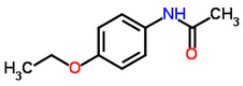
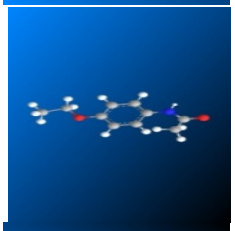
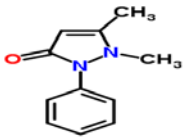
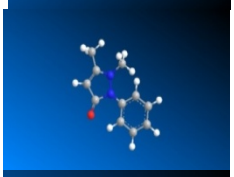
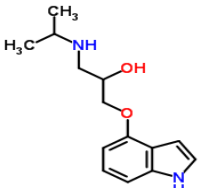
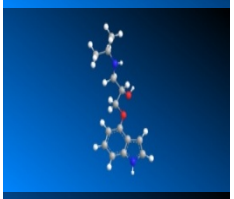
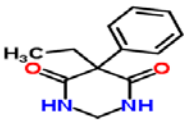
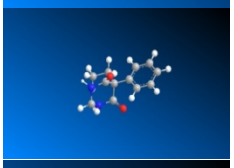
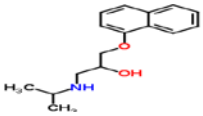
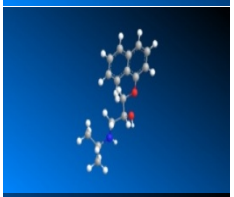
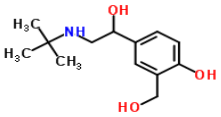
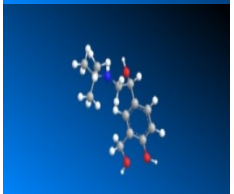
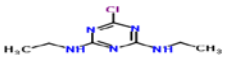
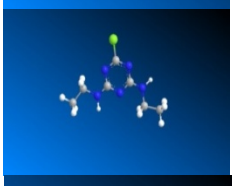
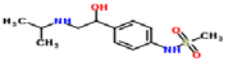
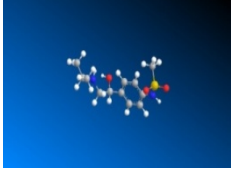
17	**Atropine**	$C_{17}H_{23}NO_3$	289.38	51-55-8		
18	**Bentazon**	$C_{10}H_{12}N_2O_3S$	240.28	25057-89-0		
19	**Bisphénol A**	$C_{15}H_{16}O_2$	228.29	80-05-7		
20	**Bromoforme**	$CHBr_3$	252.73	75-25-2		
21	**Caffeine**	$C_8H_{10}N_4O_2$	194.19	58-08-2		
22	**Caprolactame**	$C_6H_{11}NO$	113.16	105-60-2		
23	**Carbamazépine**	$C_{15}H_{12}N_2O$	236.28	298-46-4		
24	**Carbontetrabromide**	CBr_4	331.63	558-13-4		

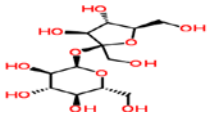
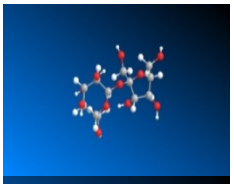
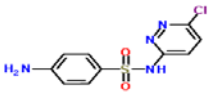
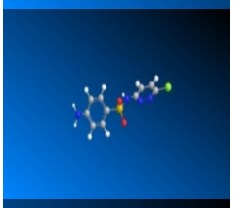
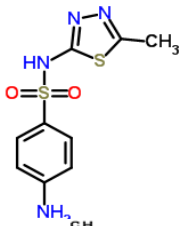
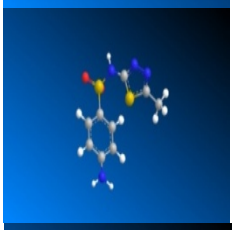
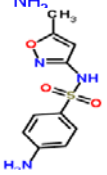
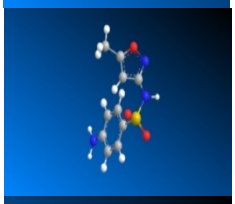
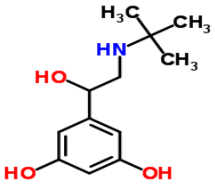
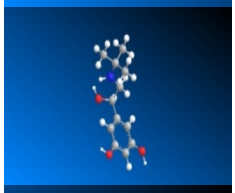
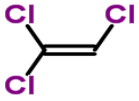
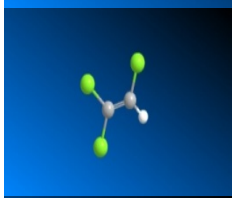
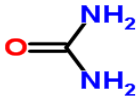
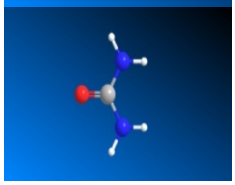
25	**Carbontetrachloride**	CCl_4	153.82	56-23-5		
26	**Chloroforme**	$CHCl_3$	119.38	67-66-3		
27	**Chlorotoluron**	$C_{10}H_{13}ClN_2O$	212.68	15545-48-9		
28	**Ciprofloxacine**	$C_{17}H_{18}FN_3O_3$	331.35	85721-33-1		
29	**Clenbutérol**	$C_{12}H_{18}Cl_2N_2O$	277.20	37148-27-9		
30	**Cyclophosphamide**	$C_7H_{15}Cl_2N_2O_2P$	261.09	50-18-0		
31	**Diclofénac**	$C_{14}H_{11}Cl_2NO2$	296.16	15307-86-5		
32	**Diethylstilbesterol**	$C_{18}H_{20}O_2$	268.35	56-53-1		
33	**Dikegulac**	$C_{12}H_{18}O_7$	274.27	18467-77-1		
34	**Diuron**	$C_9H_{10}Cl_2N_2O$	233.10	330-54-1		

35	**Enrofloxacin**	$C_{19}H_{22}FN_3O_3$	359.40	93106-60-6		
36	**Estriol**	$C_{18}H_{24}O_3$	288.39	50-27-1		
37	**Estrone**	$C_{18}H_{22}O_2$	270.37	53-16-7		
38	**Éthanol**	C_2H_6O	46.07	64-17-5		
39	**Fenoprofen**	$C_{15}H_{14}O_3$	242.28	31879-05-7		
40	**Formaldéhyde**	CH_2O	30.03	50-00-0		
41	**Gemfibrozil**	$C_{15}H_{22}O_3$	250.34	25812-30-0		
42	**Glucose**	$C_6H_{12}O_6$	180.16	50-99-7		
43	**Glycérol**	$C_3H_8O_3$	92.10	56-81-5		
44	**Ibuprofène**	$C_{13}H_{18}O_2$	206.29	15687-27-1		

45	**Isopropanol**	C_3H_8O	60.10	67-63-0		
46	**Isoproturon**	$C_{12}H_{18}N_2O$	206.29	34123-59-6		
47	**Ketoprofène**	$C_{16}H_{14}O_3$	254.29	22071-15-4		
48	**Lindane**	$C_6H_6Cl_6$	290.83	58-89-9		
49	**Linuron**	$C_9H_{10}Cl_2N_2O_2$	249.10	330-55-2		
50	**Methacetin**	$C_9H_{11}NO_2$	165.19	51-66-1		
51	**Méthanol**	CH_4O	32.04	67-56-1		
52	**Métobromuron**	$C_9H_{11}BrN_2O_2$	259.10	3060-89-7		
53	**Métoprolol**	$C_{15}H_{25}NO_3$	267.37	37350-58-6		
54	**Metoxuron**	$C_{10}H_{13}ClN_2O_2$	228.68	19937-59-8		

55	**Metribuzin**	C_8H14N_4OS	214.29	21087-64-9		
56	**Métronidazole**	$C_6H_9N_3O_3$	171.16	443-48-1		
57	**Monolinuron**	$C_9H_{11}ClN_2O_2$	214.65	1746-81-2		
58	**Monomethylphtalate**	$C_9H_8O_4$	180.16	4376-18-5		
59	**MTBE (Methyl tert-butyl ether)**	$C_5H_{12}O$	88.15	1634-04-4		
60	**NAC (standar) Carbaryl (1-NAPHTHYL-N-METHYLCARBAMATE)**	$C_{12}H_{11}NO_2$	201.23	63-25-2		
61	**N-acetyl-L-tyrosine**	$C_{11}H_{13}NO_4$	223.23	537-55-3		
62	**Naproxène**	$C_{14}H_{14}O_3$	230.27	22204-53-1		
63	**NDMA (N-NITROSODIMETHYLAMINE)**	$C_2H_6N_2O$	74.08	62-75-9		
64	**NDPA (DIPHENYLNITROSAMINE)**	$C_{12}H_{10}N_2O$	198.23	86-30-6		
65	**Nonylphénol**	$C_{15}H_{24}O$	220.36	25154-52-3		

66	**Pentoxifylline**	$C_{13}H_{18}N_4O_3$	278.31	006493-05-6		
67	**Perchloroethene**	C_2Cl_4	165.83	127-18-4		
68	**Phenacétine**	$C_{10}H_{13}NO_2$	179.22	62-44-2		
69	**Phenazon**	$C_{11}H_{12}N_2O$	188.23	60-80-0		
70	**Pindolol**	$C_{14}H_{20}N_2O_2$	248.33	13523-86-9		
71	**Primidone**	$C_{12}H_{14}N_2O_2$	218.26	125-33-7		
72	**Propanolol**	$C_{16}H_{21}NO_2$	259.34	525-66-6		
73	**Salbutamol**	$C_{13}H_{21}NO_3$	239.32	18559-94-9		
74	**Simazine**	$C_7H_{12}ClN_5$	201.66	122-34-9		
75	**Sotalol**	$C_{12}H_{20}N_2O_3S$	272.37	3930-20-9		

76	**Sucrose**	$C_{12}H_{22}O_{11}$	342.30	57-50-1		
77	**Sulfachloropyridazine**	$C_{10}H_9ClN_4O_2S$	284.73	80-32-0		
78	**Sulfamethizole**	$C_9H_{10}N_4O_2S_2$	270.33	144-82-1		
79	**Sulfaméthoxazole**	$C_{10}H_{11}N_3O_3S$	253.28	723-46-6		
80	**Terbutaline**	$C_{12}H_{19}NO_3$	225.29	23031-25-6		
81	**Trichloroethene**	C_2HCl_3	131.39	79-01-6		
82	**Ureé**	CH_4N_2O	60.06	57-13-6		

Annexe B

Annexe B.

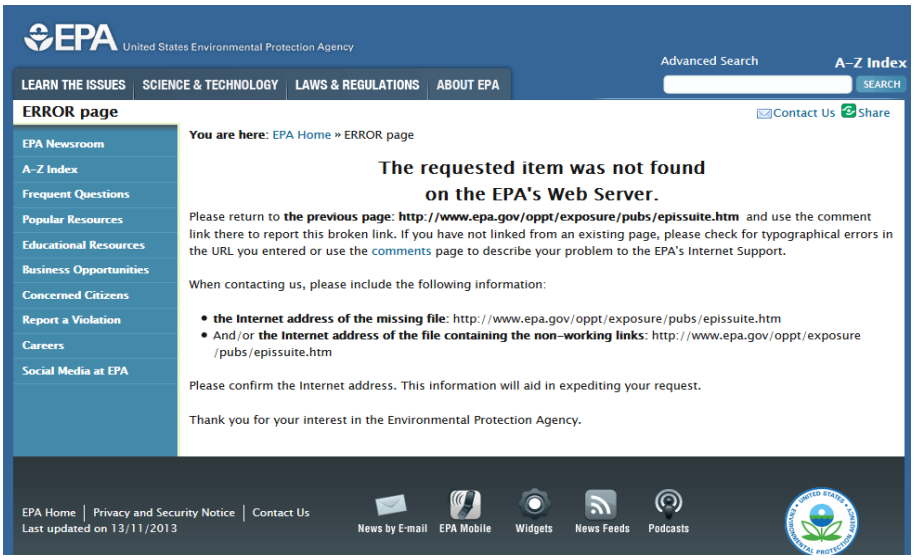


Figure 1. Capture d'écran du site intitulé Kowwin, epa.

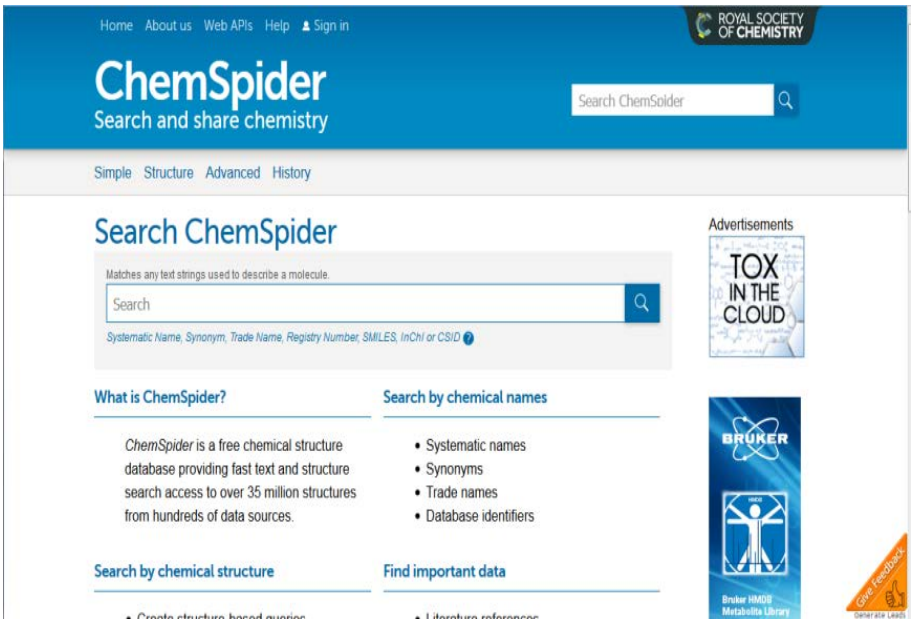
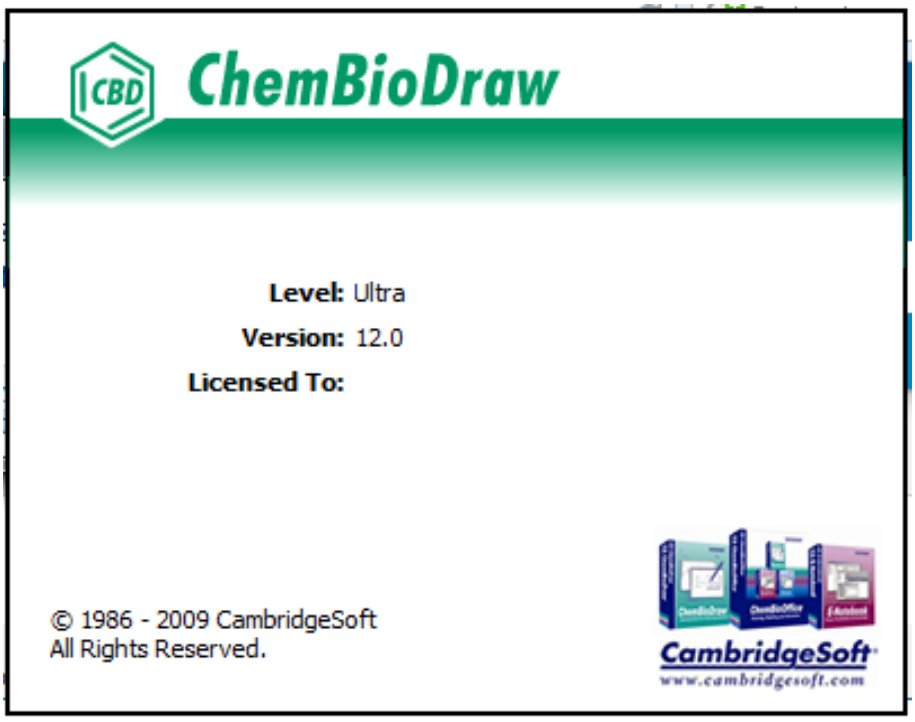
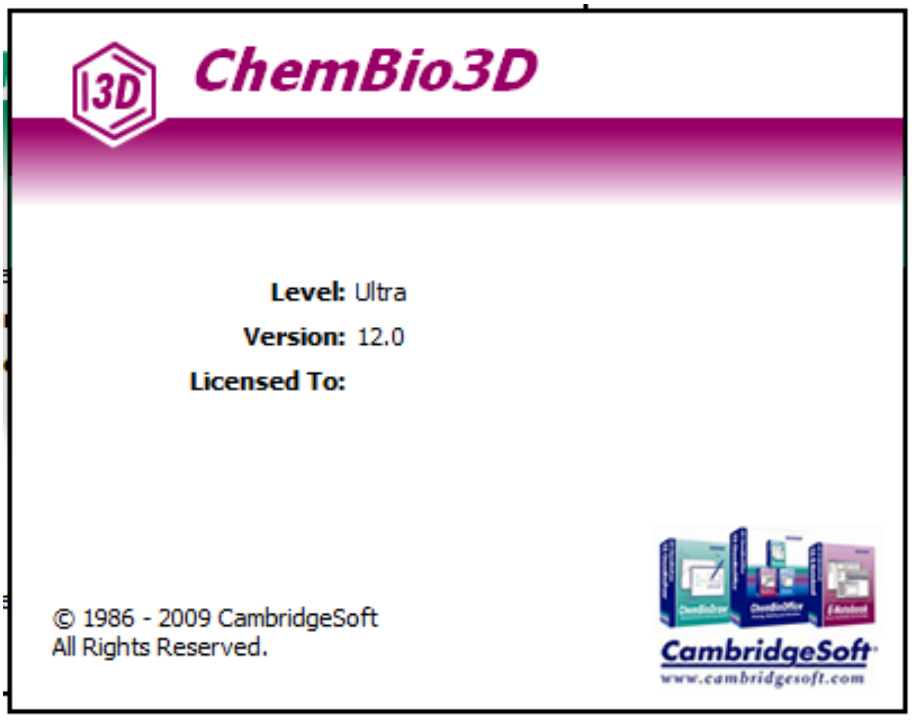


Figure 2. Capture d'écran du site intitulé Chemspider.



(a)



(b)

Figure 3. Capture d'écran des logiciels : (a) ChemBioDraw et (b) ChemBio3D.



Figure 4. Capture d'écran de logiciel STATISTICA.

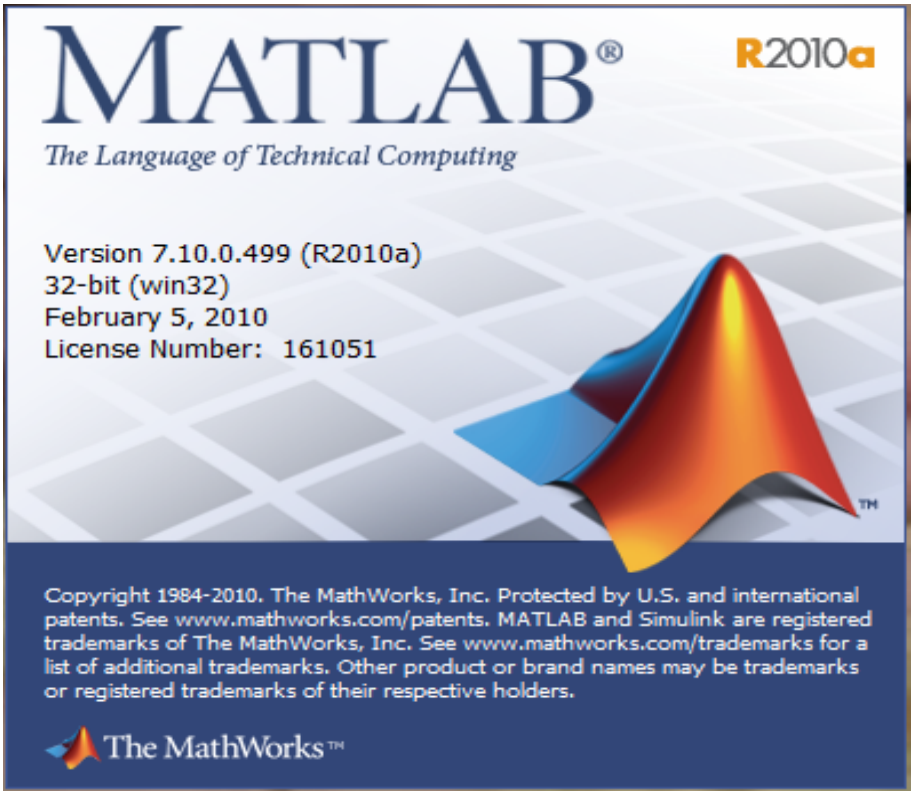


Figure 5. Capture d'écran de logiciel MATLAB 2010.

Annexe C

Annexe C.

Tableau 1. Matrice des coefficients de corrélation pour la première base de données (BD1)

Nom	MW (g/mol)	Log Kow	LogD	Dipole (Debye)	Length (nm)	Width (nm)	Depth (nm)	Eqwidth (nm)	MWCO (dalton)	Potentiel zêta (mv)	Angle de contact (°)	pH	Pression (KPa)	Taux de récupération (%)	Rétention (%)
MW(g/mol)	1.0000	0.3644	0.1980	0.1582	0.4924	0.2510	0.3147	0.3551	0.1433	-0.1273	-0.0398	-0.0086	0.1576	0.1677	0.4903
Log Kow	0.3644	1.0000	0.7333	-0.3099	0.2345	0.1340	0.2440	0.2637	0.0660	-0.1096	0.0646	0.1012	-0.1292	0.0828	0.2089
LogD	0.1980	0.7333	1.0000	-0.3085	0.0171	0.0235	-0.0334	0.0326	0.0183	-0.2058	0.0766	0.1220	-0.3181	0.1030	-0.0126
Dipole	0.1582	-0.3099	-0.3085	1.0000	0.1816	0.1094	-0.2101	-0.1252	0.0543	0.1608	-0.0322	-0.2454	0.0719	-0.1810	0.1606
Length	0.4924	0.2345	0.0171	0.1816	1.0000	0.4165	0.3551	0.4138	0.0827	0.1197	-0.1485	-0.1445	0.3241	-0.0285	0.4343
Width	0.2510	0.1340	0.0235	0.1094	0.4165	1.0000	0.4793	0.6929	-0.1737	0.3042	-0.1690	-0.1360	0.2816	-0.1874	0.3663
Depth	0.3147	0.2440	-0.0334	-0.2101	0.3551	0.4793	1.0000	0.9110	-0.0483	0.1338	-0.2201	-0.0578	0.3835	-0.0657	0.3462
Eqwidth	0.3551	0.2637	0.0326	-0.1252	0.4138	0.6929	0.9110	1.0000	-0.0678	0.1347	-0.2405	-0.0958	0.3640	-0.0542	0.3921
MWCO	0.1433	0.0660	0.0183	0.0543	0.0827	-0.1737	-0.0483	-0.0678	1.0000	-0.4170	0.0533	0.0731	-0.2158	0.2612	-0.2188
Potentiel zêta	-0.1273	-0.1096	-0.2058	0.1608	0.1197	0.3042	0.1338	0.1347	-0.4170	1.0000	0.0792	-0.2735	0.2666	-0.4349	0.0667
Angle de contact	-0.0398	0.0646	0.0766	-0.0322	-0.1485	-0.1690	-0.2201	-0.2405	0.0533	0.0792	1.0000	0.1170	-0.2564	0.0364	0.0875
pH	-0.0086	0.1012	0.1220	-0.2454	-0.1445	-0.1360	-0.0578	-0.0958	0.0731	-0.2735	0.1170	1.0000	-0.0653	0.2565	-0.0281
Pression	0.1576	-0.1292	-0.3181	0.0719	0.3241	0.2816	0.3835	0.3640	-0.2158	0.2666	-0.2564	-0.0653	1.0000	0.0886	0.2404
Taux de récupération	0.1677	0.0828	0.1030	-0.1810	-0.0285	-0.1874	-0.0657	-0.0542	0.2612	-0.4349	0.0364	0.2565	0.0886	1.0000	-0.1186
Rétention	0.4903	0.2089	-0.0126	0.1606	0.4343	0.3663	0.3462	0.3921	-0.2188	0.0667	0.0875	-0.0281	0.2404	-0.1186	1.0000

Tableau 2. Matrice des coefficients de corrélation pour la deuxième base de données (BD2)

Nom	MW (g/mol)	Log Kow	LogD	Dipole (Debye)	Length (nm)	Width (nm)	Depth (nm)	Eqwidth (nm)	SR (MgSO4)	Potentiel zêta (mv)	Angle de contact (°)	pH	Pression (KPa)	Taux de récupération (%)	Rétention (%)
MW(g/mol)	1.0000	0.4286	0.2112	0.0231	0.5670	0.5186	0.6016	0.6790	-0.0885	-0.0077	-0.1023	0.0058	0.2566	0.1345	0.5926
Log Kow	0.4286	1.0000	0.7244	-0.3589	0.2507	0.1734	0.2982	0.3136	0.0051	-0.1369	0.0311	0.0944	-0.0929	0.0240	0.3035
LogD	0.2112	0.7244	1.0000	-0.3237	-0.0188	-0.0179	-0.0560	-0.0021	-0.0883	-0.1908	0.1770	0.1086	-0.3585	-0.0174	0.0550
Dipole	0.0231	-0.3589	-0.3237	1.0000	0.1890	0.2833	-0.0919	0.0394	-0.1789	-0.0836	-0.0542	-0.1574	0.0968	-0.0675	0.1042
Length	0.5670	0.2507	-0.0188	0.1890	1.0000	0.4705	0.4365	0.5193	-0.2034	-0.0152	-0.2005	-0.1263	0.3331	0.0744	0.4964
Width	0.5186	0.1734	-0.0179	0.2833	0.4705	1.0000	0.5584	0.7564	-0.0199	-0.0207	-0.2488	-0.0986	0.2898	-0.0463	0.5090
Depth	0.6016	0.2982	-0.0560	-0.0919	0.4365	0.5584	1.0000	0.9060	0.0013	0.0916	-0.2482	-0.0683	0.3951	0.0423	0.5255
Eqwidth	0.6790	0.3136	-0.0021	0.0394	0.5193	0.7564	0.9060	1.0000	-0.0510	0.0462	-0.2401	-0.0871	0.3822	0.0503	0.5952
SR (MgSO4)	-0.0885	0.0051	-0.0883	-0.1789	-0.2034	-0.0199	0.0013	-0.0510	1.0000	-0.0616	-0.1742	0.0328	0.0754	-0.2125	-0.0096
Potentiel zêta	-0.0077	-0.1369	-0.1908	-0.0836	-0.0152	-0.0207	0.0916	0.0462	-0.0616	1.0000	-0.0872	-0.2662	0.4186	0.2542	-0.1280
Angle de contact	-0.1023	0.0311	0.1770	-0.0542	-0.2005	-0.2488	-0.2482	-0.2401	-0.1742	-0.0872	1.0000	0.0201	-0.3606	-0.0251	-0.0846
pH	0.0058	0.0944	0.1086	-0.1574	-0.1263	-0.0986	-0.0683	-0.0871	0.0328	-0.2662	0.0201	1.0000	-0.0507	0.1734	-0.0064
Pression	0.2566	-0.0929	-0.3585	0.0968	0.3331	0.2898	0.3951	0.3822	0.0754	0.4186	-0.3606	-0.0507	1.0000	0.2993	0.2314
Taux de récupération	0.1345	0.0240	-0.0174	-0.0675	0.0744	-0.0463	0.0423	0.0503	-0.2125	0.2542	-0.0251	0.1734	0.2993	1.0000	-0.0792
Rétention	0.5926	0.3035	0.0550	0.1042	0.4964	0.5090	0.5255	0.5952	-0.0096	-0.1280	-0.0846	-0.0064	0.2314	-0.0792	1.0000

Tableau 3. Matrice des coefficients de corrélation pour la troisième base de données (BD3)

Nom	MW (g/mol)	Log Kow	LogD	Dipole (Debye)	Length (nm)	Width (nm)	Depth (nm)	Eqwidth (nm)	MWCO (dalton)	SR (MgSO4)	Potentiel zêta (mv)	Angle de contact (°)	pH	Pression (KPa)	Taux de récupération (%)	Rétention (%)
MW(g/mol)	1.0000	0.4286	0.2112	0.0231	0.5670	0.5186	0.6016	0.6790	0.1062	-0.0885	-0.0077	-0.1023	0.0058	0.2566	0.1345	0.5926
Log Kow	0.4286	1.0000	0.7244	-0.3589	0.2507	0.1734	0.2982	0.3136	-0.0081	0.0051	-0.1369	0.0311	0.0944	-0.0929	0.0240	0.3035
LogD	0.2112	0.7244	1.0000	-0.3237	-0.0188	-0.0179	-0.0560	-0.0021	-0.0423	-0.0883	-0.1908	0.1770	0.1086	-0.3585	-0.0174	0.0550
Dipole	0.0231	-0.3589	-0.3237	1.0000	0.1890	0.2833	-0.0919	0.0394	0.3111	-0.1789	-0.0836	-0.0542	-0.1574	0.0968	-0.0675	0.1042
Length	0.5670	0.2507	-0.0188	0.1890	1.0000	0.4705	0.4365	0.5193	0.3705	-0.2034	-0.0152	-0.2005	-0.1263	0.3331	0.0744	0.4964
Width	0.5186	0.1734	-0.0179	0.2833	0.4705	1.0000	0.5584	0.7564	0.2522	-0.0199	-0.0207	-0.2488	-0.0986	0.2898	-0.0463	0.5090
Depth	0.6016	0.2982	-0.0560	-0.0919	0.4365	0.5584	1.0000	0.9060	0.1712	0.0013	0.0916	-0.2482	-0.0683	0.3951	0.0423	0.5255
Eqwidth	0.6790	0.3136	-0.0021	0.0394	0.5193	0.7564	0.9060	1.0000	0.2554	-0.0510	0.0462	-0.2401	-0.0871	0.3822	0.0503	0.5952
MWCO	0.1062	-0.0081	-0.0423	0.3111	0.3705	0.2522	0.1712	0.2554	1.0000	-0.7061	-0.0214	0.0656	-0.1579	-0.0142	-0.0339	0.0680
SR (MgSO4)	-0.0885	0.0051	-0.0883	-0.1789	-0.2034	-0.0199	0.0013	-0.0510	-0.7061	1.0000	-0.0616	-0.1742	0.0328	0.0754	-0.2125	-0.0096
Potentiel zêta	-0.0077	-0.1369	-0.1908	-0.0836	-0.0152	-0.0207	0.0916	0.0462	-0.0214	-0.0616	1.0000	-0.0872	-0.2662	0.4186	0.2542	-0.1280
Angle de contact	-0.1023	0.0311	0.1770	-0.0542	-0.2005	-0.2488	-0.2482	-0.2401	0.0656	-0.1742	-0.0872	1.0000	0.0201	-0.3606	-0.0251	-0.0846
pH	0.0058	0.0944	0.1086	-0.1574	-0.1263	-0.0986	-0.0683	-0.0871	-0.1579	0.0328	-0.2662	0.0201	1.0000	-0.0507	0.1734	-0.0064
Pression	0.2566	-0.0929	-0.3585	0.0968	0.3331	0.2898	0.3951	0.3822	-0.0142	0.0754	0.4186	-0.3606	-0.0507	1.0000	0.2993	0.2314
Taux de récupération	0.1345	0.0240	-0.0174	-0.0675	0.0744	-0.0463	0.0423	0.0503	-0.0339	-0.2125	0.2542	-0.0251	0.1734	0.2993	1.0000	-0.0792
Rétention	0.5926	0.3035	0.0550	0.1042	0.4964	0.5090	0.5255	0.5952	0.0680	-0.0096	-0.1280	-0.0846	-0.0064	0.2314	-0.0792	1.0000