REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE YAHIA FARES DE MEDEA



Faculté de **T**echnologie

Département de **G**énie des **P**rocédés et **E**nvironnement

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

Filière : Génie des Procédés
Spécialité : Génie des Procédés et environnement

Modélisation de la rétention des pesticides par nanofiltration et osmose

Présenté par :

- M^{elle} HENTABLI Soumia
- M^{elle} MOULAI MOSTEFA Malika

Proposé et Dirigé par :

- Dr. AMMI Yamina

Co-dirigé par :

- Dr. CHABA MOUNA Siham

Promotion 2023-2024


بسم الله الرحمن الرحيم

Dédicaces

Au-delà des personnes, des lieux et des temps, je dédie ce travail avec ma profonde conviction, à tous ceux qui ont toujours cru et sans doute à la science.

À ceux qui savent concrétiser leurs idées et savent réussir.

À ceux qui ne portent que du bien pour les autres.

À ceux qui ne vivent que dans et par la vérité.

D'un cœur plein d'affection et de tendresse je dédie ce modeste travail à :

- *Mes parents qui me représentent l'exemple de sacrifice et qui ont fait de moi ce que je suis devenu et qui m'ont toujours encouragé et aidé à surmonter les difficultés en me donnant le plus précieux soutient tout le long de mes études.*
- *Mon Mari Ahmed, qui a toujours été à mes côtés et m'a apporté un soutien moral et matériel et m'a encouragé à chaque étape de ce travail*
- *Ma belle famille*
- *Mes sœurs*
- *A toutes mes enseignants à tous les niveaux.*
- *À ma chère binôme malika*

Je vous dis merci

« soumia »

Dédicaces

A ma très chère mère

Patiente, sereine et infatigable ; Toi qui as consacré toute ta vie à moi, ce travail est le tien. L'amour, la tendresse, la bonne éducation dont j'ai bénéficié de toi, ton soutien inestimable, morale, financier, tes bénédictions, ont fait ce que je suis aujourd'hui. Femme au cœur d'or, les mots me manquent pour exprimer ma profonde gratitude et reconnaissance. Puisse Allah, le tout puissant, le très miséricordieux te garde encore longtemps à nos côtés.

A mon très cher père

Trouve dans ce travail qui est aussi le tien, le témoignage de mon amour et de mon infinie reconnaissance. Tu as toujours souhaité pour tes enfants les meilleures conditions de vie. Qu'Allah te garde encore longtemps auprès de tes chers enfants afin que tu puisses goûter aux fruits de ton dur labeur !

Aux personnes dont j'ai bien aimé la présence en ce jour, A ma sœur ***HADJIRA*** *et mes frères* ***DJAMEL, MOHAMED CHERIF*** *et* ***LOKMANE*** *. A la femme de mon frère* ***SOUMIA.***

A mon mari ***MOHAMED AMINE*** *pour sa grande patience, je le remercie chaleureusement surtout pour son soutien moral ininterrompu et ses nombreux conseils tout le long de mon mémoire.*

Sans oublier mon amie dans ce travail ***SOUMIA*** *pour sa patience et son soutien*

A tous ceux que j'aime et ceux qui m'aiment.

À tous ceux qui ont traversé ma vie et qui ont toujours été mes modèles.

Merci pour votre amour et vos encouragements.

« Malika»

Remerciements et gratitude

*Le plus grand merci revient à **ALLAH** qui nous a guidé dans le bon sens durant notre vie qui nous aide à réaliser ce travail.*

*Nous sommes également profondément reconnaissants envers notre superviseur, le Dr. **AMMI Yamina**, Maître de conférences classe A à l'UYFM, qu'il trouve ainsi l'expression de notre profonde reconnaissance pour ses compétences, ses encouragements et ses précieux conseils durant le déroulement de ce travail qui ont éclairé notre chemin vers la réussite de la concrétisation de notre mémoire. Nous remercions également notre CO-Promotrice de thème, le Dr. CHABA MOUNA Siham, pour son appui et sa contribution à la réussite travail.*

Nos vifs remerciements vont également aux membres du jury pour l'intérêt qu'ils ont porté à notre recherche en acceptant d'examiner notre travail et de l'enrichir par leurs propositions.

En fin nous remercions nos parents et tous ceux qui ont participé de près ou de loin à l'achèvement de ce travail.

ملخص

نمذجة إزالة المبيدات الحشرية عن طريق الترشيح النانوي والتناضح

الهدف من هذه الدراسة هو تطوير نموذج عصبي مثالي يعتمد على العلاقات الكمية بين خصائص البنية ($QSPR\text{-}RNA_{optimal}$) لنمذجة الاحتفاظ بالمبيدات الحشرية عن طريق الترشيح النانوي وأغشية التناضح.

للتنبؤ ببقايا المبيدات الحشرية، سيتم بناء نموذج عصبي متقدم ($QSPR\text{-}RNA_{optimal}$) يتميز ببنية (14 خلية عصبية في طبقة الإدخال، 19 خلية عصبية في الطبقة المخفية وخلية عصبية واحدة في طبقة الإخراج). تم استخدام مجموعة من 574 نقطة بيانات لاختبار النموذج الأمثل. تم حجز 80% من البيانات للتدريب و20% لاختبار النموذج الأمثل.

أظهرت مقارنة قيم الاستبقاء المحسوبة خلال مرحلة الاختبار مع تلك المقاسة وجود علاقة جيدة ($R = 0.9627$).

كشف تحليل الحساسية باستخدام طريقة "الوزن" أن المتغيرات الأكثر صلة التي يمكن أن تؤثر على الاحتفاظ بالمبيدات عن طريق الترشيح النانوي وأغشية التناضح لها مساهمة كبيرة (IR>5%)

الكلمات المفتاحية: نمذجة, **إزالة, المبيدات الحشرية, طريق الترشيح النانوي, التناضح.**

Abstract

Modeling the retention of pesticides by nanofiltration and osmosis

The goal assigned to this work is to develop an optimal neural model based on the quantitative structure-property relationship ($QSPR\text{-}AAN_{optimal}$) to model the retention of pesticides by nanofiltration and osmosis membranes.

To predict pesticide residues, a feedforward neural model ($QSPR\text{-}ANN_{optimal}$) characterized by a structure (14 neurons in the input layer, 19 neurons in the hidden layer, and 1 neuron in the output layer) will be built. A set of 574 data points was used to test the optimal model. 80% of the data was reserved for training and 20% for testing the optimal model.

The comparison of the retention values calculated during the test phase with those measured showed a good correlation (R = 0,9627).

The sensitivity analysis by the "weight" method revealed that the most relevant variables that can influence the retention of pesticides by osmosis and nanofiltration membranes exert significant contributions (IR>5%).

Keywords: Modeling, Retention, Pesticides, Nanofiltration, Reverse.

Résumé

Modélisation de la rétention des pesticides par nanofiltration et osmose

Le but assigné à cette étude est de développer un modèle neuronal optimal basé sur des relations quantitatives structure-propriété ($QSPR\text{-}RNA_{optimal}$) pour modéliser la rétention des pesticides par les membranes de nanofiltration et osmose.

Pour prédire les résidus de pesticides, un modèle neuronal du type feedforward ($QSPR\text{-}RNA_{optimal}$) caractérisé par une structure (14 neurones dans la couche d'entrée, 19 neurones dans la couche cachée et 1 neurone dans la couche de sortie) sera construit. Un ensemble de 574 points de données a été utilisé pour tester le modèle optimal. 80 % des données ont été réservées à l'apprentissage et 20 % au test du modèle optimal.

La comparaison des valeurs de rétention calculées lors de la phase de test avec celles mesurées a montré une bonne corrélation (R = 0,9627).

L'analyse de sensibilité par la méthode "poids" a révélé que les variables les plus pertinentes pouvant influencer la rétention des pesticides par les membranes d'osmose et nano filtration ont une contribution significative (IR>5 %).

Mots-clés : Modélisation, Rétention, Pesticides, Nano filtration, Osmose

Abréviations iii
Nomenclature et symboles iv
Liste des tableaux vi
Liste des figures vii

Introduction générale 1

Chapitre I. Rétention des pesticides par les membranes

I.1. Introduction 04
I.2. Généralités 04
I.2.1. Définition de la membrane 04
I.2.2. Classification des membranes 05
I.2.2.1. Classification d'après la nature chimique 05
I.2.2.2. Classification d'après la porosité 05
1.2.2.3. Classification d'après la morphologie 05
1.2.2.4. Classification d'après la géométrie 06
1.2.2.5. Classification d'après le mécanisme de séparation 11
I.3. Osmose et Nano filtration 11
I.3.1. Définitions et principes de fonctionnement 11
I.3.2. Applications des procédés d'osmose inverse et nano filtration 13
I.3.3. Avantages et inconvénients des procédés d'osmose et nano filtration 14
I.4. Performances membranaires 15
I.4.1. Sélectivité 15
I.4.2. Perméabilité 15
I.5. Influence des paramètres opératoires 16
I.5.1. Température 16
I.5.2. Pression 16
I.5.3. Taux de conversion 16
I.6. Interactions membrane-soluté 16
I.6.1. Encombrement stériques (effet du tamisage) 17
I.6.1.1. État physique de la membrane 18
I.6.1.2. Taille moléculaire 20
I.6.2. Répulsions électrostatiques (effet de charge) 22
I.6.2.1. Charge du soluté 22
I.6.2.2. Charge de la surface 23
I.6.2.3. Influence du pH 23
I.6.3. Interaction hydrophobes/adsorption 23
I.6.3.1. Hydrophobicité de la membrane 23
I.6.3.2. Hydrophobicité des solutés 24
I.8. Revue des travaux de la rétention des pesticides lors d'osmose et nano filtration 25
I.9. Conclusion 26

Chapitre II. OSPR-Réseau de Neurones Artificiels

II.1. Introduction 27
II.2. Méthodes d'apprentissages statistiques 28
II.3. Relations quantitatives structure-Activités (QSAR) 28
II.3.1. Outils de QSAR 29
II.3.2. Données biologiques 29
II.3.3. Descripteurs moléculaires 29
II.3.4. Formulation mathématique des QSAR 31
II.4. Réseaux de neurones artificiels (RNA) 32
II.4.1. Histoire des réseaux de neurones artificiels 32
II.4.2. Neurone biologique 33
II.4.3. Neurone formel 34
II.4.4. Architecture d'un réseau de neurones 37
II.4.5. Perceptron multicouches 38
II.4.6. Types d'apprentissage des réseaux de neurones 39
II.4.7. Réseaux de neurones les plus utilisés 40
II.4.8 Conception d'un modèle neuronal 41
II.4.9. Détermination des entrées/sorties du réseau de neurones 41
II.4.10. Construction de la base de données et choix des échantillons 42
II.4.11. Définition de la structure du réseau 43
II.4.12. Apprentissage 44
II.4.13. Evaluation des performances du réseau de neurones 46
II.5. Conclusion 47

Chapitre III. Modélisation de la rétention des pesticides lors d'osmose et nanofiltration par QSPR-RNA

III.1. Introduction 48
III. 2. Collecte des données, prétraitement et analyse 48
III.3. Développement des modèles QSPR-RNA 54
III.4. Performance des modèles neuronaux 56
III.5. Expressions mathématiques des modèles neuronaux optimisés 62
III.6. Analyse de sensibilité par la méthode de "poids" 65
III.7. Diagnostic des points aberrants et domaine d'applicabilité du modèle neuronal optimisé 67
III.8. Conclusion 70
Conclusion générale et perspectives 71
Références bibliographiques
Annexe A
Annexe B

Symbole	Signification
b_0	Biais de la couche de sortie
b_h	Biais de la couche cachée
BFGS	Broyden-Fletcher-Goldfarb-Shanno
D	Descripteur moléculaire
HB	Hydrophobe
HL	Hydrophile
IUPAC	International Union of Pure and Applied Chemistry
NF	nanofiltration
MF	Microfiltration
MWCO	Molecular weight cutt off (Seuil de poids moléculaire)
OD	Osmose Directe
OI	Osmose Inverse
QSAR	Relations quantitatives structure-activités
QSPR	Relations quantitatives structure-propriétés
QSTR	Relations quantitatives structure-toxicités
RFR	Réseaux à fonction de bases radiales
RNA	Réseaux de neurones artificiels
TFC	Thin-film composite
Tr	Taux de rejet
UF	Ultrafiltration
$W_{j,i}^{I}$	Poids de connexions entrée-cachée
$W_{1,j}^{H}$	Poids de connexions cachée-sortie

Lecture	Désignation	Unité
C	Concentration de la solution	(mol.m^{-3})
DCO	Demande Chimique en Oxygène	(mg.L^{-1})
Depth	Profondeur moléculaire	(nm)
EAM	Erreur Absolue Moyenne	(%)
Ei	$I^{éme}$ entrée du modèle	(-)
EQM	Erreur Quadratique Moyenne	(-)
EqWidth	Largeur équivalente moléculaire	(nm)
ESP	Erreur Standard de Prédiction	(%)
Length	Longueur moléculaire	(nm)
log 1/C	Logarithme inverse de l'activité	(-)
Q_A	Débit volumique d'Alimentation	(m^3/h)
Q_P	Débit volumique de Perméat	(m^3/h)
REQM	Racine de l'Erreur Quadratique Moyenne	(%)
T	Température	(k)
W	Poids synaptique de la connexion du neurone	(-)
Width	Largeur moléculaire	(nm)
Y	Taux de récupération	(-)
Zj	Sortie de la couche cachée	(-)
$logk_{ow}$	Logarithme du coefficient de partage octanol-eau	(-)
π	Pression osmotique de la solution	(Pa)

Lettres Grecques

Lecture	Désignation	Unité
A	Pente de la droite du meilleur ajustement linéaire	(-)
B	Ordonnée à l'origine de la droite du meilleur ajustement linéaire	(-)

Exposants

Cal	Calculée
Exp	Expérimentale
H	Cachée (Hidden)
I	Entrée (Input)
O	Sortie (Output)

Indices

Lecture	**Désignation**
Max	Maximum
Min	Minimum
$\bar{x}$	Moyenne

Chapitre I. Membranes de filtration

Tableau I.1.	Classification IUPAC des pores suivants leur taille.	05
Tableau I.2.	Avantages et inconvénients des différents modules membranaires	10
Tableau I.3.	Avantages et inconvénients de procédés d'OI et NF.	14

Chapitre II. QSPR-Réseaux de neurones artificiels

Tableau II.1.	Fonctions de transfert usuelles, *f(n)*.	35
Tableau II.2.	Résumé des caractéristiques des algorithmes d'apprentissage de rétro propagation implantés dans MATLAB®	45

Chapitre III. Modélisation de la rétention des pesticides lors d'osmose et nano filtration par QSPR-RNA

Tableau III.1.	Taille de la base de données et le nombre de pesticides utilisés dans ce travail.	48
Tableau III.2.	Logiciels utilisés pour calculer LogD, le moment dipolaire, Smin, Smax et Length.	49
Tableau III.3.	Entrées et sortie utilisées pour le développement des modèles QSPR-RNA	50
Tableau III.4.	Analyse statistique des données d'entrée et de sortie (BD1).	51
Tableau III.5.	Analyse statistique des données d'entrée et de sortie (BD2).	51
Tableau III.6.	Analyse statistique des données d'entrée et de sortie (BD3).	52
Tableau III.7.	Comparaison entres les modèles neuronaux (QSPR-RNA).	57
Tableau III.8.	Structure de modèle QSPR-RNA$_{optimal}$ optimisés.	58
Tableau III.9.	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 l'erreur quadratique moyenne ***(REQM***), l'erreur absolue moyenne ***(EAM)*** et l'erreur de standard de prédiction ***(ESP)***.	61
Tableau III.10.	Poids et biais du QSPR-RNA $_{optimal}$.	64

Annexe A

Tableau 1	Liste de 26 pesticides utilisés pour développer les modèles (QSPR-RNA).	

Chapitre I. Membranes de filtration

Figure I.1.	Principe de séparation par membrane.	04
Figure I.2.	Membrane symétrique poreuse.	06
Figure I.3.	Représentation Schématique d'une membrane asymétrique.	06
Figure I.4.	Module plan. Schéma d'un montage de type filtre-presse.	07
Figure I.5.	Structure interne d'une membrane à spirale.	08
Figure I.6.	Module tubulaire.	09
Figure I.7.	Module à fibres creuses.	10
Figure I.8.	Principe de l'osmose inverse.	12
Figure I.9.	Molécule de polyamide et ses équilibres acido-basiques.	17
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)	19
Figure I.11.	Calcul des dimensions de l'octanol (Length, Width et Depth) de la structure 3D.	22
Figure I.12.	Mesure de l'angle de contact par la méthode de la goutte déposée.	24
Figure I.13.	Nombre de publications de 2000 à 2024 pour les membranes de la nano filtration, osmose inverse et osmose direct.	25

Chapitre II. QSPR-Réseaux de neurones artificiels

Figure II.1.	Représentation d'un neurone biologique.	34
Figure II.2.	Neurone formel.	34
Figure II.3.	Mise en correspondance neurone biologique / neurone artificiel.	36
Figure II.4.	Réseaux de neurones les plus utilisés.	37
Figure II.5.	Structure d'un Réseaux de neurones non bouclés.	38

Chapitre III. Modélisation de la rétention des pesticides lors d'osmose et nano filtration par QSPR-RNA

Figure III.1.	Rétentions expérimentales des pesticides en fonction de masse molaire : (a) BD1, (b) BD2, (c) BD3.	54
Figure III.2.	Organigramme de développement et l'optimisation de l'architecture d'un QSPR-RNA.	55
Figure III.3.	Comparaison entre les valeurs expérimentales et les valeurs calculées pour QSPR-RNA$_{optimal}$: (a) la phase d'apprentissage, (b) la phase de test, (c) la phase totale.	60
Figure III.4.	Performances des modèles neuronaux (QSAR-RN*1*, QSAR-RN*2* et QSAR-RN*3*) pour : (a) la phase d'apprentissage, (c) la phase de test, (d) la phase totale.	62
Figure III.5.	Représentation schématique du modèle neuronal "QSPR-RNA$_{optimal}$" Optimisé.	63

Figure III.6. Architecture finale du réseau de neurone optimal de type feedforward pour la prédiction de la rétention des pesticides par les membranes d'osmose et nano filtration. 65

Figure III.7. Présentation des pourcentages de l'importance relative (*IR*) de chaque variable d'entrée sur la rétention des pesticides lors d'osmose et nano filtration. 67

Figure III.8. Diagnostic des données potentiellement suspectes et domaine d'applicabilité du modèle neuronal optimal : (a) pour la phase test et (b) pour la phase totale. 69

Annexe B

Figure 1. Capture d'écran du site intitulé Chemspider.

Figure 2. Capture d'écran des logiciels : (a) ChemBioDraw2D et (b) ChemBio3D.

Figure 3. Capture d'écran de logiciel Hyperchem

Figure 4. Capture d'écran de logiciel STATISTICA.

Figure 5. Capture d'écran de logiciel MATLAB 2021.

Introduction générale

Un pesticide est défini comme "une substance ou une association de substances destinée à repousser, détruire ou combattre les ravageurs, y compris les vecteurs de maladies humaines ou animales, ou toutes autres espèces végétales ou animales nocives ou gênantes au cours de la production, la transformation ou l'entreposage des produits agricoles". Considérés comme "tueurs de fléaux" par certains étymologistes, ce sont des molécules dont les propriétés toxiques permettent de lutter contre des organismes nuisibles à l'homme et/ou à son environnement (Hennia 2019). L'utilisation généralisée des pesticides en agriculture a été créditée de l'augmentation spectaculaire des rendements des cultures en réduisant les menaces naturelles et est devenue une partie intégrante des pratiques agricoles. L'utilisation sanitaire des pesticides a été, elle aussi, couronnée de succès à la suite de l'éradication de certaines maladies. Par ailleurs, les pesticides ayant des propriétés antiseptiques et désinfectantes ont eu leur part d'utilisation en médecine. La gamme des pesticides disponibles s'est alors rapidement diversifiée et leur usage a connu un très fort développement au cours des décennies suivantes. Petit à petit, du fait de leur facilité d'emploi et de leur efficacité, ils se sont imposés dans les usages domestiques (anti-moustiques par exemple), dans le traitement de la voierie publique, des plantes ornementales des voies ferrés, des routes, des espaces verts et des jardins. Durant la seconde moitié du vingtième siècle, l'utilisation des pesticides a doublé tous les dix ans. De plus, la résistance acquise par certaines espèces vis-à-vis des pesticides a contraint les utilisateurs de ces produits à augmenter la quantité à utiliser afin que l'efficacité d'action soit optimale d'une part et à la diversification des molécules actives (recherche de nouveaux pesticides plus actifs) d'autre part (Hamdache 2014).

Le problème de la contamination des eaux par les pesticides et autres micropolluants organiques inquiète de plus en plus les producteurs d'eau potable. La diversité de ces produits et le niveau des concentrations atteint dans les eaux brutes ces dernières années exigent des méthodes d'élimination efficaces qui garantissent en bout de chaîne de traitement des concentrations résiduelles inférieures aux limites maximales fixées par les instances européennes : 0.1 ug/L pour chaque substance prise individuellement et 0.5 ug/L pour les totaux pesticides plus métabolites. Pour éliminer les pesticides de l'eau, plusieurs techniques sont utilisées, seules ou en association, avec plus ou moins de succès. Le développement de nouvelles membranes a ouvert des perspectives d'application dans le domaine du traitement de l'eau potable par osmose inverse basse pression et par nanofiltration. En plus de ses capacités d'adoucissement de l'eau, cette dernière peut aussi bien éliminer la matière organique dissoute, que la couleur et certains pesticides. Mais à cause d'une consommation élevée d'énergie, cette technique a vu son essor freiné (Boussahel et al. 2022).

Les méthodes de séparation membranaire sont un procédé de séparation de fluides utilisant comme agent séparant une membrane synthétique qui est une couche mince de matière. L'épaisseur d'une membrane peut varier de 100 nm à un peu plus de 1 cm. Elle permet l'arrêt ou le passage sélectif de certaines substances dissoutes ou non dans un mélange, entre les deux milieux qu'elle sépare. La partie du mélange retenue par la membrane est appelée rétentat (ou concentrat) alors que celle qui traverse cette dernière est appelée perméat. La séparation se fait sous l'action d'une force motrice de transfert selon un mécanisme de séparation défini. Les caractéristiques des membranes sont déterminées par deux paramètres : la perméabilité et la sélectivité. On distingue quatre techniques répertoriées sous les noms de microfiltration (MF), ultrafiltration (UF), nanofiltration (NF), osmose inverse (OI) et osmose direct (OD). 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 répandus et de nombreuses études indiquent qu'il permet de traiter différents types d'effluents en retenant les pesticides. D'autre part, le rôle complexe et encore mal cerné des nombreuses interactions qui peuvent s'établir entre les solutés, l'eau et la membrane. De nombreux modèles de transferts existent 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.

La modélisation est une autre approche permettant de diminuer le nombre d'expériences à effectuer pour déterminer les paramètres optimaux pour la rétention des pesticides lors de nanofiltration et d'osmose. Si elle représente bien le procédé, les influences des paramètres optimaux pourront être évaluées sans aucune expérience et n'auront qu'à être validés sur le site. 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. 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. Bien que les réseaux de neurones artificiels (RNA) 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 comme moyen de calcul des phénomènes de transports par d'osmose inverse et de nanofiltration lors de la rétention des pesticides (Ammi 2016).

L'objectif de ce travail est l'application des QSPR-RNA de type feedforwardet pour modéliser la rétention des pesticides par les membranes d'osmose et nanofiltration. Ainsi, nous avons organisé ce mémoire en trois chapitres :

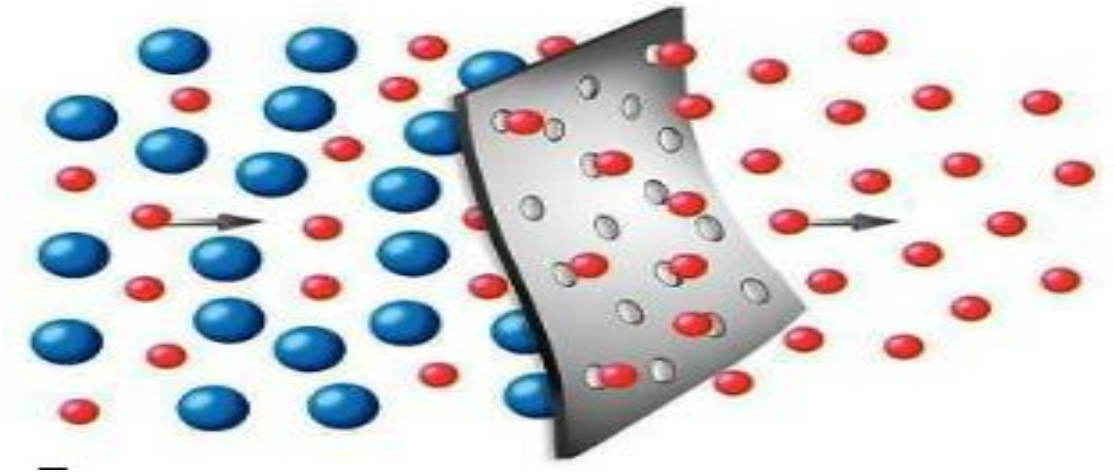
- ✓ Le premier chapitre est consacré à la description de la technologie membranaire, on donnera dans un premier temps un aperçu sur les généralités sur les membranes, ensuite les interactions et les propriétés de séparation, puis les applications et les avantages et une revue des travaux de la rétention des pesticides lors d'osmose et nanofiltration.
- ✓ Le deuxième chapitre est dédié à la présentation des modèles de réseaux de neurones artificiels basés sur la relation quantitative structure-activité "QSPR-RNA".
- ✓ Nous présentons dans le troisième chapitre la modélisation de la rétention des pesticides lors d'osmose et nanofiltration. Il comptera six parties : la première partie présente la collecte des données, prétraitement et analyse, la deuxième partie sera consacrée au développement des modèles neuronaux, nous discuterons les performances des modèles neuronaux dans la troisième partie, les expressions mathématiques de modèle neuronal optimal seront abordées dans la quatrième partie, nous traiterons dans la cinquième partie l'analyse de sensibilité par la méthode de poids et dans la dernière partie nous discuterons le diagnostic des points aberrants et domaine d'applicabilité du modèle neuronal optimise.

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



Chapitre I

Rétention des pesticides par les membranes



I.1 Introduction

Les membranes de filtration peuvent être en osmose inverse "OI", en nano filtration "NF", en osmose direct "OD". Semi-perméables, de structure symétrique, asymétrique ou composite, elles permettent la rétention de solutés ou de particules contenus dans un solvant. Les nombreux procédés d'élaboration sont directement adaptés à la fabrication de membranes à partir d'une classe de matériaux (organique ou minérale). Ces membranes sont ensuite assemblées sous forme de modules, plan, spiralé, tubulaire, ou en fibre creuses, qui sont étroitement reliés à la géométrie des membranes, donc au matériau (Jean-Christophe et al. 2007).

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 de 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 (Nachad 2009) :

1. La dimension et la forme;
2. La nature chimique;
3. L'état physique;
4. La charge électrique, etc (Nachad 2009).

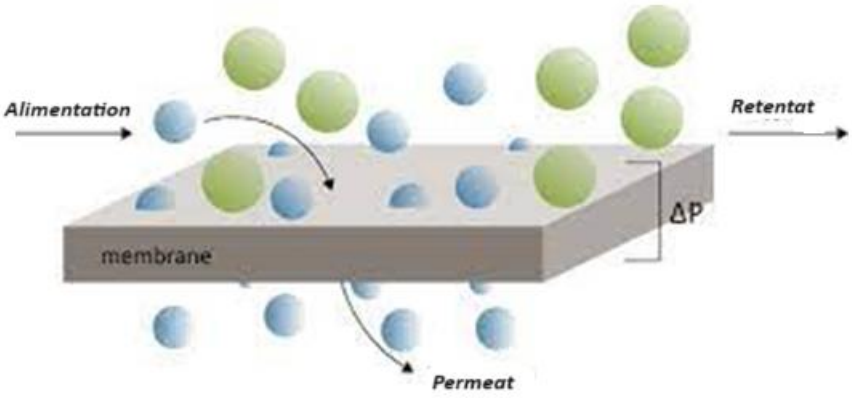


Figure I.1. Principe de séparation par membrane (https://openmembranedatabase.org/membrane-science/).

I.2.2 Classification des membranes

I.2.2.1 Classification d'après la nature chimique

a. Membrane organique : Elles sont fabriquées à partir de polymères organiques (acétate de cellulose, polysulfones, polyamides, etc.). Les qualités de ces matériaux leur confèrent une grande adaptabilité aux différentes applications. Environ 90% des membranes d'ultrafiltration et de microfiltration sont constituées de membranes organiques (Remise 2006).

b. Membrane minérale : Elles sont à la base, de silice, de carbone, d'oxyde de graphite et de métaux frittés. Ces matériaux possèdent une grande résistance physicochimique, mécanique et thermique. Leur durée de vie excède de trois fois celle des membranes organiques mais leur coût est sensiblement plus élevé (Barbar 2006).

I.2.2.2 Classification d'après la porosité

- **Membrane poreuse** : ce type de membrane se compose de pores dont la dimension est supérieure aux dimensions moléculaires. L'IUPAC (International Union of Pure and Applied Chemistry) distingue trois catégories de pores selon leurs tailles (Tableau I.1).

Tableau I.1. Classification IUPAC des pores suivants leur taille (Hamani 2016).

Dénomination	Micropore	Mésopore	Macropore
Taille de pore	<2 nm	2-5 nm	>50nm

- **Membrane dense**

Ce terme signifie une absence de porosité de dimension microscopique, ces membranes ne présentent aucune porosité. Elles consistent en un film dense à travers lequel le perméat est transporté par diffusions sous l'effet d'une grandeur de potentiel électrique, de concentration ou d'une force de pression. Si la solubilité diffère, les membranes denses peuvent séparer des composés de taille voisine (Hamani 2016).

1.2.2.3 Classification d'après la morphologie

Les membranes peuvent également être classées en fonction de leur structure, on distingue alors (Bessiere 2005):

- ✓ **Des membranes isotropes**, dites symétriques, dont les propriétés structurelles sont constantes sur toute l'épaisseur, qu'elles soient denses ou poreuses.

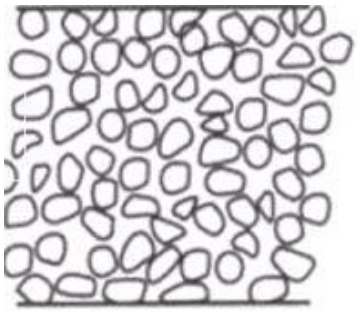


Figure I.2. Membrane symétrique poreuse (https://www.researchgate.net/figure/14-Representation-schematique-dune-membrane-symetrique-a-et-dune-membrane_fig12_319914244).

- ✓ **Des membranes anisotropes**, qualifiées d'asymétriques, dont la structure varie de la surface de la membrane vers l'intérieur (Figure I. 3). Parmi ces dernières, on compte les membranes préparées à partir d'un seul matériau par inversion de phase et les membranes composites qui sont constituées d'un empilement de couches de différents polymères.
- ✓ **Membranes composites**, ce sont des membranes asymétriques constituées d'un assemblage de deux ou plusieurs couches de matériaux et de taille de pores différente La membrane composite comporte deux types de couches : les couches actives (couche filtrant) et le support assurant de bonnes propriétés mécanique (Haddad 2004 ; Xu et al 2017).

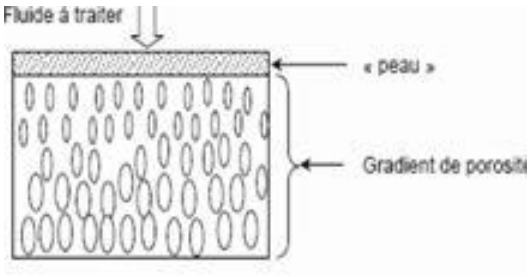


Figure I.3. Représentation Schématique d'une membrane asymétrique (Bessiere 2005).

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 (Tamas 2004).

1.2.2.4 Classification d'après la géométrie

La géométrie des membranes est souvent assimilée à la géométrie des modules. Ainsi, il est courant de parler de membrane plane, spiralée, tubulaire ou fibre creuse. En fait, il n'existe que deux géométries : les membranes planes et les membranes cylindriques (Desclaux, S et al. 2007).

Les membranes en elles-mêmes ne constituent qu'une partie du système. Elles sont agencées en unités, appelées modules, dont la mise en œuvre doit répondre à un certain nombre de critères

d'ordre mécanique (étanchéité et résistance à la pression), hydrodynamique (minimiser les pertes de charge et les volumes morts) ou encore économique (compacité élevée, investissement minimal et facilité d'entretien) (Bessiere 2007). Actuellement quatre types de modules sont commercialisés :

- **Modules plans**

Cette configuration dérivée des filtres pressés est constituée d'un empilement de membranes sous forme de feuilles et de plaque supports. Le fluide à traiter circule entre les membranes de deux plaques adjacentes. Ces plaques assurent le soutien mécanique de la membrane ainsi que le drainage du perméat (Zhou et al 2019).

Ce type de module est très utilisé dans la réalisation des essais de filtration (Bouazizi et al 2016).

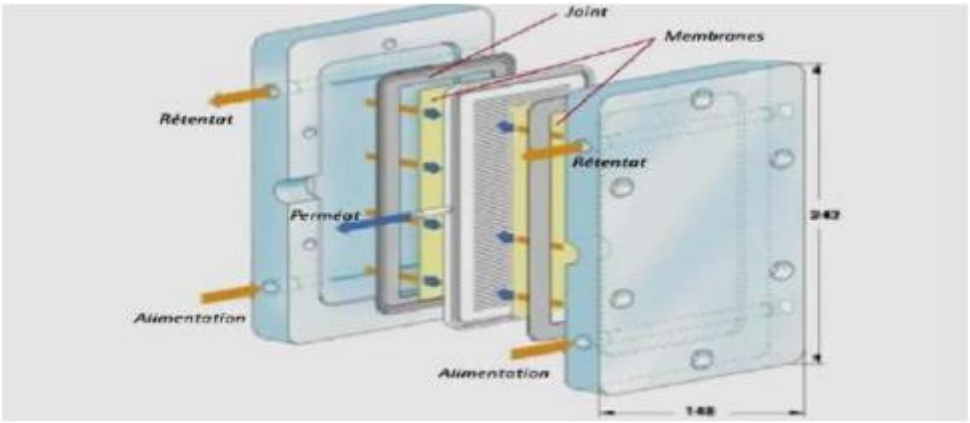


Figure I.4. Module plan. Schéma d'un montage de type filtre-presse (Ramdani et al 2017).

Plusieurs travaux ont été réalisés de ce sens, B. Achiou et al ont étudié l'élaboration de membranes plates de microfiltration en céramique à base de pouzzolane naturelle. Les résultats de la filtration montrent que ces membranes peuvent être utilisées pour le traitement des effluents textiles après avoir éliminé environ 99% de la turbidité. Ces membranes pourraient également être utilisées comme membrane de microfiltration ou comme support pour développer une membrane d'ultrafiltration. Elles peuvent être appliquées aussi pour éliminer totalement les colorants solubles des solutions aqueuses (Achiou et al 2016).

De même, Souad Saja et al, ont élaboré une membrane plate en céramique à base de perlite naturelle pour la microfiltration. Ils ont montré que la porosité de la membrane de microfiltration est de 52.11% avec une taille de pore moyenne de 1.70 µm et ils ont obtenu une perméabilité à l'eau qui atteint une valeur de 1433.46 L/h m^2bar. De plus, la membrane a une résistance mécanique de 21.68 mPa et une bonne résistance chimique en milieu acide et basique. Les résultats expérimentaux montrent que la rétention de la turbidité était supérieure à 97 et 96% respectivement pour les effluents

agro-alimentaires et de tannerie. Ces résultats de la filtration sont très prometteurs et pourraient permettre d'utiliser la membrane préparée comme procédé alternatif dans le traitement des eaux usées industrielles (Saja et al 2017).

- **Modules spirales**

Dans ce type de module, les feuilles de membranes planes sont enroulées autour d'un tube poreux destiné à collecter le filtrat. On obtient ainsi un cylindre multicouche où le perméat s'écoule selon un chemin spiralé vers le tube poreux tandis que l'alimentation circule axialement dans les canaux (Ramdani et al 2017). Les deux feuilles de membranes sont séparées entre elles par des espaceurs, eux aussi sont une sorte de filet. Ces espaceurs ont deux fonctions, former un espace pour la circulation tangentielle de l'eau du côté alimentation et promouvoir des turbulences dans cet écoulement.

Ce type de configuration est surtout utilisé pour des membranes à faible débit (NF et OI) dont il est nécessaire d'augmenter la surface filtrante sans trop augmenter le volume de l'installation (Mona A. Abdel-Fatah.2018).

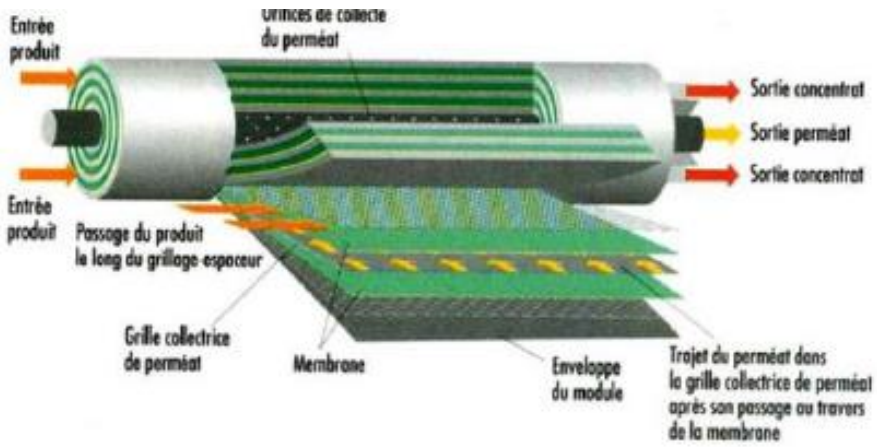


Figure I.5. Structure interne d'une membrane à spirale (Lanteri 2009).

Ce type de module a été largement développé à l'échelle industriel surtout dans les stations de dessalement pour la potabilisation de l'eau de mer, A titre d'exemple, il a été utilisé afin d'étudier la rétention des halogénures de sodium (F, Cl-, Br- et I-) et pour le traitement de l'eau saharienne saumâtre de la région d'El Oued (Kouinine) en Algérie ont utilisés une membrane de nanofiltration organique spiralée composite de surface filtrante 2.6 m2 et un seuil de coupure MWCO = 300 Da. Cette membrane a permis d'aboutir à des performances de dessalement et de défluoruration intéressantes. Donc ils ont confirmé que cette technique est tout à fait recommandable pour aider à réduire la salinité des eaux saumâtres sahariennes hyper fluorées (Lanteri 2009).

- **Modules tubulaires**

Il est constitué d'une ou plusieurs membranes inorganiques de forme tubulaire ayant des diamètres internes entre 7 et 25 mm. L'eau à traiter circule, sous pression, à l'intérieur du tube et le perméat est collecté à l'extérieur du tube (Diop 2010) .Les tubes sont rassemblés en petit nombre (quelques dizaines) dans un module et ils y sont disposés parallèlement les uns aux autres avec un espacement régulier.

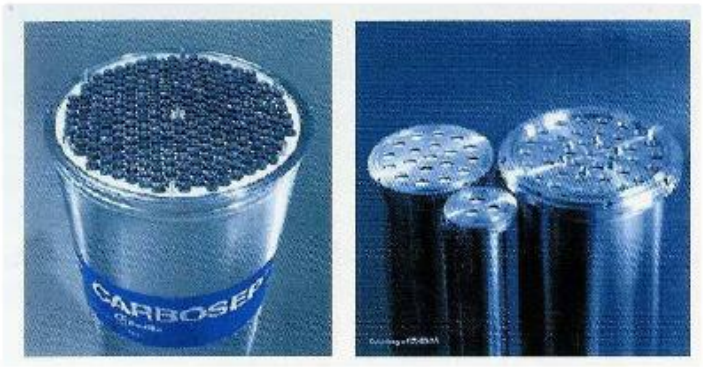


Figure I.6. Module tubulaire (Ramdani al 2017).

Ces modules sont destinés pour les membranes minérales et spécifiquement pour les membranes céramiques. Vinoth Kumar et son équipe ont fabriqué une membrane céramique tubulaire à faible coût en utilisant des mélanges d'argile. Ils ont montré que la membrane préparée offre une meilleure porosité (53%), une forte résistance mécanique et une excellente résistance à la corrosion. De plus, la membrane élaborée est très performante dans le traitement des eaux usées huileuses, et peut être utilisée comme support pour la fabrication des membranes d'ultrafiltration (UF) et de nanofiltration (NF) (Padaki et al 2012).

- **Modules fibres creuses**

Ils rassemblent plusieurs milliers de membranes sous forme de fibres de très petit diamètre (< 2 mm) en un faisceau. Ce type de module est caractérisé surtout par son grand rapport surface de filtration/ volume d'installation, son faible volume mort et sa tenue mécanique (Bouazizi et al 2016). L'eau à traiter circule généralement à l'extérieur de la fibre et le perméat est collecté à l'intérieur des tubes.

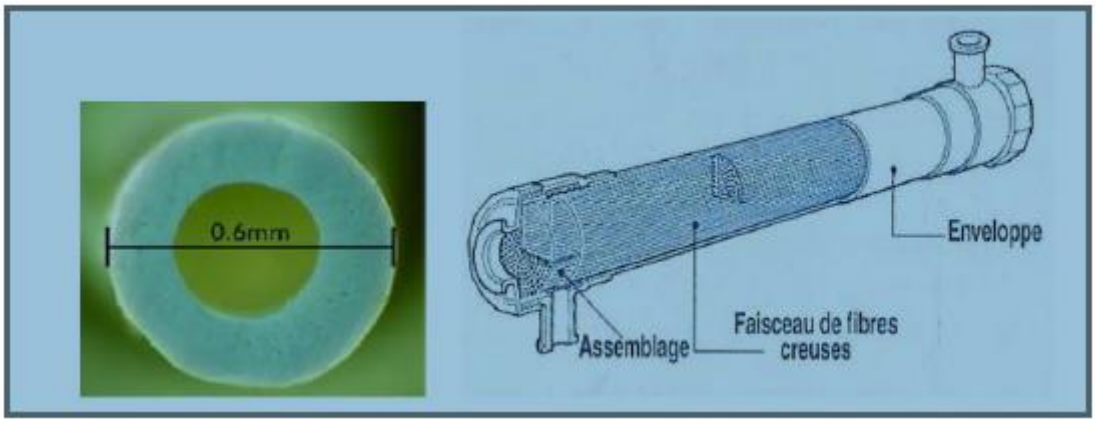


Figure I.7. Module à fibres creuses (Kratbi 2023).

Ces membranes permettent le traitement différentes type d'eau et en utilisant divers types de polymères. Récemment, Liu et al ont fabriqué une membrane de fibre creuse en poly (fluorure de vinylidène) pour séparer les colorants et les sels existants dans les eaux usées textiles. La membrane présentait un rejet élevé de 99,99% pour l'élimination des colorants (rouge Congo) et un rejet faible de moins de 12,0% pour les sels. Elle a également montré une excellente stabilité au cours d'un fonctionnement prolongé avec un mélange de colorants et de sels. Ces résultats ont montré des performances de séparation efficaces et stables pour la séparation des sels et des colorants, ce qui serait très important pour la réutilisation des eaux usées textiles (Yoo et al 2003). Les principaux avantages et inconvénients des modules sont regroupés dans le tableau 1.2 ci-dessous.

Tableau I.2. Avantages et inconvénients des différents modules membranaires (Ramdani et al 2017 ; Diop 2010).

Géométrie	Avantages	Inconvénients
Plane	-Possibilité de récupérer le perméat. -Système souple et modulable. -Changement facile des membranes.	-Nettoyage assez difficile. -Investissement relativementélevé. -Système peu compact (100 à 400 m^2/m^3).
Spirale	-Compacité élevée (300 à 1000 m^2/m^3). -Faible coût d'investissement.	-Prétraitement nécessaire. -Nettoyage très difficile. -Sensible au colmatage.
Tubulaire	-Très peu d'exigences deprétraitement. -Facilité de nettoyage. -Technologie simple.	-Consommation énergétiqueélevée. -Compacité limitée (10 et 300m^2/m^3). -Coût élevé.
Fibre creuse	-Compacité élevée (15000 m^2/m^3). -Faible consommation énergétique. -Faible coût d'investissement.	-Prétraitement nécessaire. -Sensibilité de colmatage dû au faible diamètre des fibres. -Fragiles (Canaux fins).

1.2.2.5 Classification d'après le mécanisme de séparation

Les mécanismes de séparation par membranes sont complexes et encore partiellement compris. Les processus qui jouent un rôle dans la séparation membranaire sont le tamisage, la friction sur les parois des pores des membranes, la diffusion dans le matériau membranaire ou dans les pores des membranes, les forces de surface répulsives ou attractives, notamment la répulsion électrostatique. La classification la plus répandue des membranes d'après le mécanisme de séparation est la suivante (Tamas, A.P 2004):

- **Membranes poreuses "(MF, UF, NF)**, où les effets de tamisage et de friction jouent des rôles important mais où les forces de surface peuvent aussi jouer un rôle important comme en "NF": les pores peuvent être subdivisés en plusieurs catégories:
  - Macrospores-plus grands que 50 nm (MF, UF);
  - Méso pores-2 à 50 nm (UF);
  - Micropores-plus petits que 2 nm (NF).
- **Membranes non poreuses** (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.
- **Membranes échangeuses des ions**. Ces types de membranes 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 faut remarquer que toute la communauté scientifique ne s'entend pas sur cette classification. Par exemple, certains auteurs considèrent que les membranes d'OI sont des membranes poreuses. Il faut aussi noter que la plupart des membranes polymériques sont chargées même si leur charge est plus faible que celles des membranes échangeuses d'ions.

I.3 Osmose et Nanofiltration

I.3.1 Définitions et principes de fonctionnement

L'**osmose inverse (OI)** 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 mises en contact par une membrane sélective sous l'action du gradient de concentration (Figure I.8). Lorsqu'une pression est

appliquée sur le compartiment le plus concentré, le flux de solvant diminue jusqu'à s'annuler 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. La pression efficace correspond donc à la pression de part et d'autre de la membrane (pression transmembranaire, Ptm) diminuée de la différence de pression osmotique (ΔΠ) de part et d'autre de la membrane.

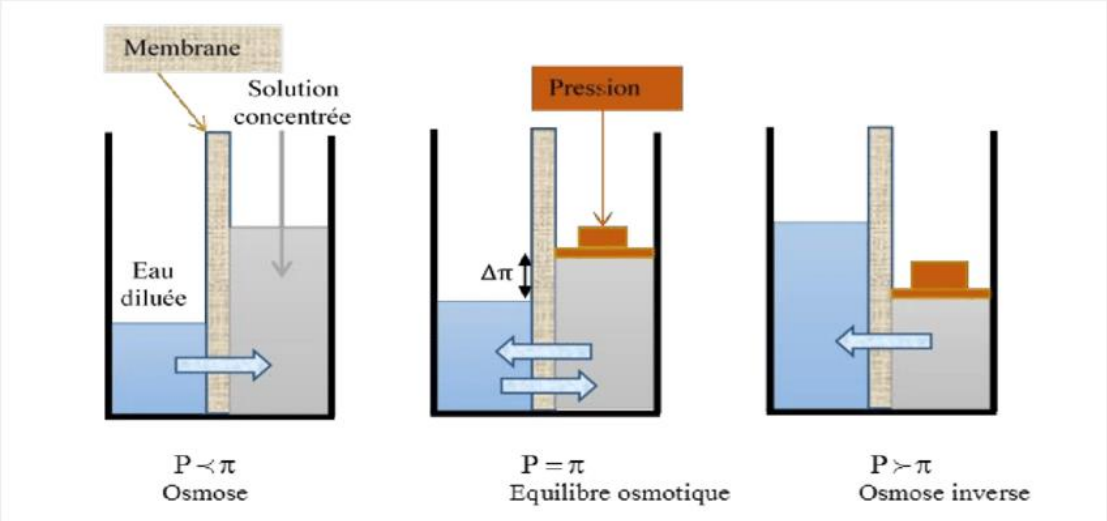


Figure I.8. Principe de l'osmose inverse (Kratbi 2023).

En cas d'une solution diluée, qui est considérée comme une solution idéale thermodynamiquement, la pression osmotique est calculée à partir de l'équation de Van't Hoff (Camille 2008).

$$\pi = CRT \qquad (I.1)$$

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) (ammi 2016)

La nano filtration (NF) 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 :

- Pression de travail 10 à 40 bars,
- Diamètre des pores de la membrane 1 nm,

- Débit volumique 50 à 100 l.h−1 mm^{-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 ($\boldsymbol{Ca^{+2}}$, $\boldsymbol{Mg^{+2}}$) 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 (Ammi 2016).

L'**osmose directe (OD)** est un procédé membranaire qui permet la concentration de liquides grâce à l'utilisation d'une membrane sélective et d'une solution de soutirage riche en sel, en sucre ou autre soluté. Le rôle de la solution de soutirage consiste à attirer l'eau du liquide que l'on cherche à concentrer grâce au phénomène d'osmose. Ce type de filtration ne requiert pas l'utilisation de pressions élevées et peut même être utilisé à température ambiante, ce qui est utile pour préserver les molécules fragiles. L'osmose directe permet d'atteindre un niveau de concentration de 60-65% de solides totaux.

I.3.2 Applications des procédés d'osmose inverse et nanofiltration

Les différentes applications envisagées pour les procédés d'osmose inverse et nanofiltration sont résumées comme suites (Ammi 2016) :

- Industrie agro-alimentaire

- Filière œnologique ;
- Lactosérum ;
- Recyclage des protéines.

- Eau potable, dessalement, saumure

- Élimination nano-polluants organiques ;
- Élimination des particules ;
- Adoucissement.

- Industries diverses

- Récupération des métaux ;
- Production de l'eau ;
- Traitement de surface.

L'osmose directe est un procédé relativement nouveau comparé à l'osmose inverse qui a 60 ans

d'histoire et une expérience de milliard de mètres cubes traités. Néamoins, elle est utilisée pour le traitement des eaux usées industrielles, le traitement des liquides alimentaire au niveau de l'industrie agro-alimentaires, la concentration de la lixiviation de décharge, la purification de l'eau en situations d'urgence, et également dans le contrôle des drogues dans l'organisme (Cath et al 2006).

I.3.3 Avantages et inconvénients des procédés d'osmose et nano filtration

Les avantages et les inconvénients des procédés d'OI et NF sont regroupés dans le tableau I.3.

Tableau I.3. Avantages et inconvénients de procédés d'OI et NF (Ammi 2016).

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 agroalimentaire. – 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.

Les principaux avantages de l'utilisation du procédé d'osmose directe sont qu'il fonctionne à des pressions hydrauliques faibles ou nulles, qu'il a un rejet élevé d'un large éventail de contaminants et qu'il peut avoir une propension à l'encrassement de la membrane plus faible que les procédés à membrane entraînée par pression. Étant donné que la seule pression impliquée dans le processus d'OD est due à la résistance à l'écoulement dans le module de membrane (quelques barres), l'équipement utilisé est très simple et le support de membrane est moins problématique. (Holloway et al. 2005). En outre, pour l'industrie alimentaire et pharmaceutique, l'OD a l'avantage de concentrer le flux d'alimentation sans nécessiter de pressions ou de températures élevées pouvant nuire à la solution d'alimentation. Pour les applications médicales, ce procédé peut aider à la libération lente et précise de médicaments qui ont une faible biodisponibilité en raison de leur solubilité ou perméabilité limitées (Bhatt et al. 2004). L'inconvénient majeur de l'osmose directe réside dans le fait la solution

de soutirage, qui présente une pression osmotique, élevée peut engendrer un flux d'eau important, en plus de nécessiter une température relativement élevée pour la séparation de l'eau et la solution de soutirage (Achilli et al. 2010).

I.4 Performances membranaires

Comme pour tous les matériaux, le choix d'une membrane par rapport à d'autres est basé sur plusieurs critères. Cette partie présente un récapitulatif des critères qui influencent ce choix.

I.4.1 Sélectivité

La sélectivité (T_R) d'une membrane est, en général, définie par le taux de rejet (appelé aussi taux de rétention) de l'espèce (sel, macromolécule, particule) que la membrane est censée retenir, Elle s'exprime par le taux de rétention ou facteur de sélectivité TR (Pontié et al 2001).

$$T_R = 1 - \frac{Co}{Cp} \qquad \text{(I.2)}$$

Avec Co et Cp correspondent à la concentration de l'espèce à retenir dans la solution et la concentration de la même espèce dans le perméat, respectivement. Lorsque T_R est nul, le soluté n'est pas retenu et lorsque T_R = 100, le soluté est entièrement retenu.

I.4.2 Perméabilité

La perméabilité (A) d'une membrane fraction de liquide qui traverse la membrane est une caractéristique intrinsèque de la membrane qui dépend de sa structure. De façon pratique, la perméabilité peut être définie comme étant le rapport entre le flux de perméat (J_P) et la différence de pression transmembranaire effective (ΔPm) :

$$A = \frac{Jp}{\Delta Pm} \qquad \text{(I.3)}$$

La densité de flux de perméat (appelée aussi de perméation) est un débit de perméation unitaire, c'est-à-dire, le rapport entre le débit volumétrique de perméation (Q_P) et la surface effective de la membrane (S).

$$Jp = \frac{Qp}{S} \qquad \text{(I.4)}$$

I.5 Influence des paramètres opératoires

I.5.1 Température

L'influence de la température est due au changement de viscosité de l'alimentation qui augmente la densité du flux de perméat (Bilstad 1995), Maurel 1998) a montré que cette densité augmente de 3% pour chaque élévation d'un degré Celsius. De plus, la diffusivité à la membrane augmente, ce qui entraîne un taux de rétention plus faible (Girard et Fukumoto 2000).

Le compactage de la membrane à une pression donnée s'augmente avec la température, en raison des changements de sa structure (Maurel 1998). Par conséquent, une bonne membrane est celle qui résiste le mieux à ce phénomène.

I.5.2 Pression

Etant donné que la réaction de synthèse d'ammoniac entraine une réduction en volume (décroissance dans le nombre de molécules), le pourcentage d'équilibre de l'ammoniac augmentera avec la pression. En même temps, la réaction est accélérée par une augmentation de pression. Donc la conversion s'améliorera avec une pression plus élevée.

I.5.3 Taux de conversion

Le taux de conversion présente une relation indirecte avec le flux de perméat. Une augmentation du taux entraîne une diminution de la densité du flux (Kim et al 2008). La concentration d'alimentation, le colmatage ainsi que le phénomène de polarisation de concentration augmentent également avec le taux de conversion.

En ce qui concerne le colmatage, il est préférable de travailler à des faibles taux de conversion, l'avantage d'un taux de conversion élevé est l'optimisation énergétique (Diminution) (Maurel et al 1998).

I.6 Interactions membrane-soluté

Pour les molécules organiques, les propriétés de séparation ne sont pas faciles à prédire. En fait, il a été démontré qu'ils résultent de trois interactions qui influencent la rétention dans RO/NF : l'encombrement stérique (effet de tamisage) (Altalyan et al 2006), la répulsion électrostatique (effet de charge) et les interactions hydrophobes/ adsorption (Sagne 2008). Rappelons tout d'abord la structure du polyamide, un polymère principalement utilisé comme couche active dans les membranes d'osmose inverse et de nanofiltration (Figure I.9).

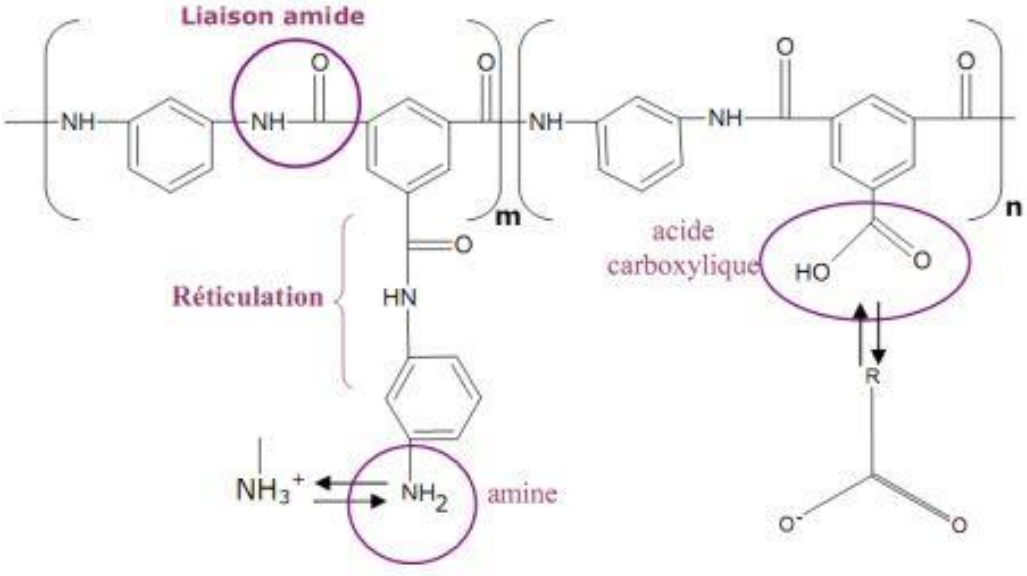


Figure I.9 Molécule de polyamide et ses équilibres acido-basiques (Sagne 2008).

Cette structure permet diverses interactions entre les molécules en solution et la surface de la membrane (Sagne 2008).

- ✓ Les groupes fonctionnels acide carboxylique libre et amine (non impliqués dans la réticulation) sont susceptibles d'être ionisés, c'est-à-dire présentent respectivement une charge négative ou positive, et peuvent former de fortes liaisons électrostatiques. Ils peuvent également participer aux liaisons hydrogène.
- ✓ Un noyau aromatique peut interagir avec le nuage électronique d'une autre molécule aromatique pour former un nuage électronique globalement plus stable.
- ✓ La liaison amide peut donner lieu à des interactions Lifshitz-van der Waals plus faibles via le doublet d'électrons libres des atomes d'azote et d'oxygène.

I.6.1 Encombrement stériques (effet du tamisage)

Elle est directement liée à la taille moléculaire du soluté d'une part et au nombre et à la forme des pores de la membrane d'autre part (Yangali-Quintanilla et al. 2010 ; Dolar et al 2011).a étudié l'importance des paramètres dimensionnels pour la rétention des molécules par les membranes.

D'autres auteurs ont également étudié la relation entre la rétention moléculaire parmi les solutés membranaires et les mécanismes d'exclusion de taille (Kiso et al. 2001 ; Ozaki et al. 2002 ; Schafer et al. 2003 ; Kimura et al. 2004 ; Nghiem et al. 2004 ; Yangali Quintanilla et al. 2009b-2010b, Khaouane et al. 2018, Ammi 2016, Ammi et al 2020, Jawad et al 2020) en sélectionnant des descripteurs de taille moléculaire.

I.6.1.1 État physique de la membrane

Les propriétés physiques de la membrane telles que le nombre de pores, la forme et la rugosité de la surface jouent un rôle majeur dans la rétention des solutés.

La porosité est exprimée par la taille, la distribution de taille et le nombre de pores disponibles dans la couche active. Le terme « pore » dans les membranes OI/NF correspond à des espaces vides au sein du polymère où le transport de liquide s'effectue sous l'influence d'une force motrice. Dans les membranes denses, les pores sont des vides de moins de 0,5 nm de diamètre qui se forment naturellement dans les polymères solides. Ces défauts de pores et de densité sont provoqués par des irrégularités dans l'enchevêtrement des chaînes polymères.

Cela peut également résulter d'un retrait local du polymère cristallin lors de la cristallisation. Ces pores peuvent être circulaires ou non, fermés ou ouverts, et forment un réseau continu de connexions. Il contient probablement de petites molécules qui interagissent fortement avec ses parois tout en étant transportées à travers ce pore (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.

- **Seuil de rétention "Sc ou MWCO"**

Les membranes d'OI/NF se caractérisent par un seuil de coupure. Il s'agit de la plus petite masse molaire d'une molécule qui sera bloquée à 90 % (ou 95 % selon le fournisseur) par la membrane dans des conditions opératoires spécifiques. Il est communément exprimé en Daltons (unités équivalentes à des grammes par mole). Bien que le seuil d'arrêt informe l'utilisateur du domaine d'utilisation de la membrane, sa valeur exacte ne doit être considérée qu'à titre indicatif. En fait, les molécules ont des formes différentes et les espèces ayant un poids moléculaire supérieur à la limite peuvent être retenues par la membrane, tout comme les espèces ayant un poids moléculaire légèrement inférieur peuvent traverser la membrane. Il est donc important de se rappeler que les limites sont définies par rapport à des solutés spécifiques. Par conséquent, la courbe de variation du taux de répulsion des solutés en fonction du poids moléculaire permet de caractériser la membrane d'OI/NF pour une séparation spécifique dans des conditions spécifiques (Figure 1.10) (Lanteri et al 2009).

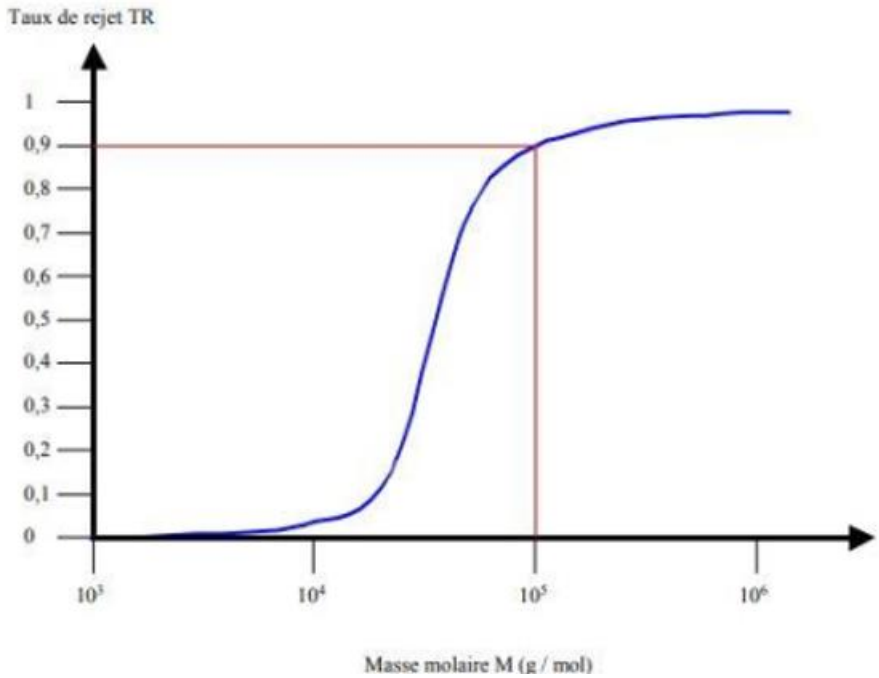


Figure I.10. Illustration du principe de détermination du seuil de coupure d'une membrane, la rétention en fonction de la masse molaire (défini ici pour un taux de rejet de 0.90) (Jean Marc et al 2002).

Plus la valeur Se de la membrane est faible, plus elle peut contenir de petites molécules ou colloïdes. Sc est principalement liée à la taille des pores de la membrane et est influencée par (Anselme et Jacobs 1996) :

- ✓ Forme de la molécule à filtrer ;
- ✓ Charge ;
- ✓ pH ;
- ✓ Pression de fonctionnement ;
- ✓ Débit de perméat ;
- ✓ Charge de la membrane.

❖ **Rejet 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.6.1.2 Taille moléculaire

La taille moléculaire est la géométrie de la molécule. Elle peut être représentée par différents paramètres obtenus à partir à partir de la structure de la molécule, tels que le poids moléculaire, la longueur, la largeur, la profondeur et la largeur équivalente. Van der Bruggen et al 1998) ont mis en évidence la relation entre la rétention des molécules organique et leurs paramètres de taille (Ammi 2016).

- **Poids moléculaire "Mw"**

Le poids moléculaire est un paramètre facilement disponible et ne fournit qu'une estimation approximative de la taille moléculaire. Cela permet d'avoir une première impression de la taille relative des molécules examinées. Il ne peut pas être utilisé pour la modélisation car il ne représente pas la géométrie de la molécule. Pour les molécules non chargées et moins polaires, la masse molaire est un indicateur précoce de rétention bien corrélé (Sadmani 2014). Cependant, il n'existe aucune

information concernant l'encombrement stérique de cette molécule. L'évaluation de la taille et de la forme moléculaires en combinaison avec la taille des pores peut être un meilleur descripteur que le poids moléculaire. En fait, les pesticides ayant des surfaces transversales plus faibles plutôt que des poids moléculaires plus faibles ont une rétention plus faible. Le poids moléculaire est considéré comme une variable de descriptive moins importante pour la modélisation de la rétention (Yangali-Quintanilla et al 2008).

Par ailleurs, Kratbi et al. 2023; Santos et al. 2006 ont utilisé un autre paramètre pour représenter géométriquement le poids moléculaire, ils ont introduit " diamètre effectif" qui est donné par la relation suivante.

$$d_c = 0.065 M_c^{0.438} \quad \text{(I.5)}$$

❖ **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).

❖ **Largueur 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 (Kiso et al 2010).

$$\text{Width} = \frac{1}{2}\sqrt{S_{\min}} \quad \text{(I.6)}$$

❖ **Profondeur moléculaire "Depth"**

Semblable à la largeur, elle est mesurée en remplaçant l'aire minimale de la molécule lorsqu'elle est projetée perpendiculairement au plan de l'axe L par l'aire d'un rectangle entourant l'aire maximale. Ils sont calculés à partir de la formule suivante (Kiso et al 1992 ; Santos et al 2006 et Ammi 2016) :

$$\text{Depdth} = \frac{1}{2}\sqrt{S_{max}} \quad \text{(I.7)}$$

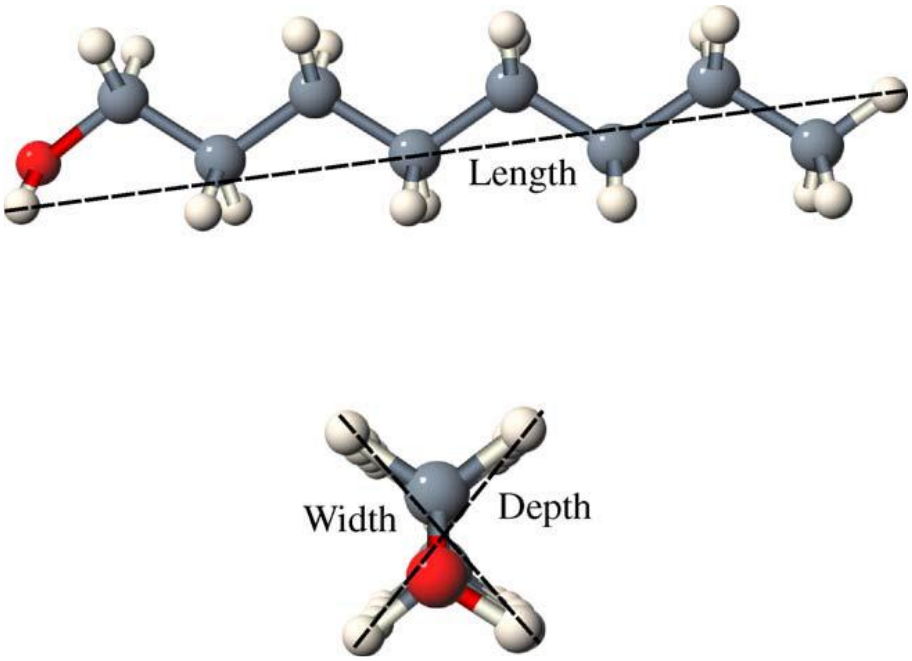


Figure I.11. Calcul des dimensions de l'octanol (Length, Width et Depth) de la structure 3D (Santos et al. 2006).

❖ **Largeur équivalente moléculaire "*EqWidth*"**

La largeur moléculaire équivalente est définie comme étant la moyenne géométrique de la largeur et de la profondeur (Santos et al 2006 ; Linares et al. 2011). On peut calculer la largeur équivalente par la formule suivante :

$$\text{EqWidth} = \sqrt{width * Depth} \quad \text{(I.8)}$$

I.6.2 Répulsions électrostatiques (effet de charge)

Il est bien connu que les interactions électrostatiques jouent un rôle important dans le mécanisme de rétention. Ils sont construits entre deux unités chargées : une molécule organique et une membrane (Van der Bruggen et al 1998).

I.6.2.1 Charge du soluté

Comme nous l'avons vu précédemment, les membranes d'osmose inverse sont généralement chargées négativement. Par conséquent, il est clair que les solutés chargés négativement sont mieux retenus par répulsion électrostatique que les solutés neutres. Respecter la neutralité des charges signifie que les composés chargés positivement se chargent également *(Shim et al 2002)*. Les molécules ionisables sont donc en général mieux retenues (Berg et al 1997 ; Ozaki et Li 2002). 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 charge opposée à celle de la membrane (Van der Bruggen et al 1999 ; Hu et al. 2003). Pour ces dernières, l'attraction électrostatique favorise leur passage.

I.6.2.2 Charge de la surface

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.6.2.3 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.6.3 Interaction hydrophobes/adsorption

La polarité de la molécule, sa capacité à former des liaisons hydrogène et son hydrophobie sont des paramètres importants pour l'adsorption.

I.6.3.1 Hydrophobicité de la membrane

L'hydrophobicité d'une membrane qui mesure sa mouillabilité par l'eau, a un impact sur ses

performances. 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 posée (Gumiet al 2003).

Avec la méthode de la goutte posée et l'eau (Figure I.12), si α est supérieur à 90°, la membrane est dite hydrophobe et s'il est inférieur à 90°, la membrane est dite hydrophile.

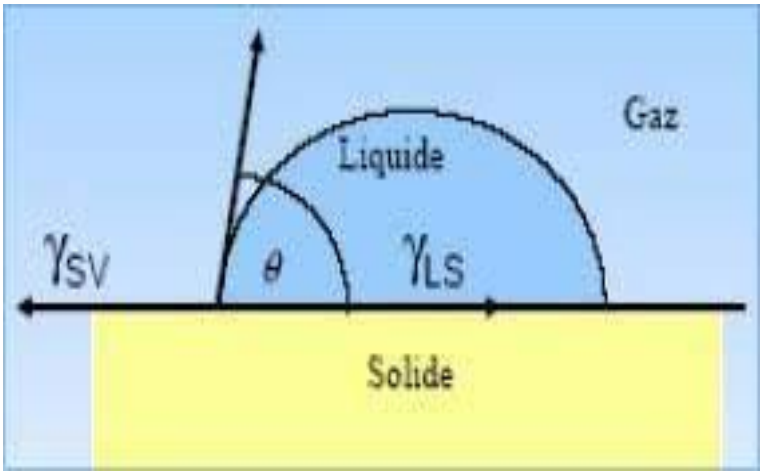


Figure I.12. Mesure de l'angle de contact par la méthode de la goutte déposée (http://www.lrs.upmc.fr/fr/l-unite-de-recherche/plateforme-technique/mesures-d-energie-de-surface.html).

I.6.3.2 Hydrophobicité des solutés

Le caractère hydrophobe ou hydrophile des solutés va donc jouer un rôle sur leur rétention. Elle est évaluée par le logarithme du coefficient de partage octanol-eau Kow, plus souvent exprimé sous forme logarithmique logKow. S'il est supérieur à 0, la molécule est hydrophobe ; s'il est inférieur, la molécule est hydrophile. On parle également de molécules polaires/apolaires. Les molécules hydrophobes sont en général apolaires.

Le logarithme du coefficient de partage n-octanol/eau, LogP, a été utilisé comme un indice hydrophobe, et est défini comme suit (Équation I.9)

$$\mathbf{Log}\ K_{OW} = \mathbf{log}(\frac{CO}{CW}) \qquad \text{(I.9)}$$

Où $C0$ et CWsont les concentrations des molécules dans les couches n-octanol et l'eau (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 $logkow$ du soluté (Équations I.13 et I.14) (Verliefde et al 2008).

Pour les acides :

$$\mathbf{Log}D_{(pH)} = \mathbf{log}k_{ow} - \mathbf{log}\,(1 + 10^{(pH-pKa)}) \qquad \text{(I.10)}$$

Pour les basses :

$$\mathbf{Log}D_{(pH)} = \mathbf{log}k_{ow} - \mathbf{log}\,(1 + 10^{(PKa-pH)}) \qquad \text{(I.11)}$$

I.7. Revue des travaux de la rétention des pesticides lors d'osmose et nanofiltration

De nombreux travaux réalisés au niveau international mettent en évidence la rétention des pesticides par d'osmose et nanofiltration. La base de données consultée est Science Direct. La figure I.13, reporte le nombre de publications de 2000 à 2024 pour les membranes de la nanofiltration, osmose inverse et osmose direct.

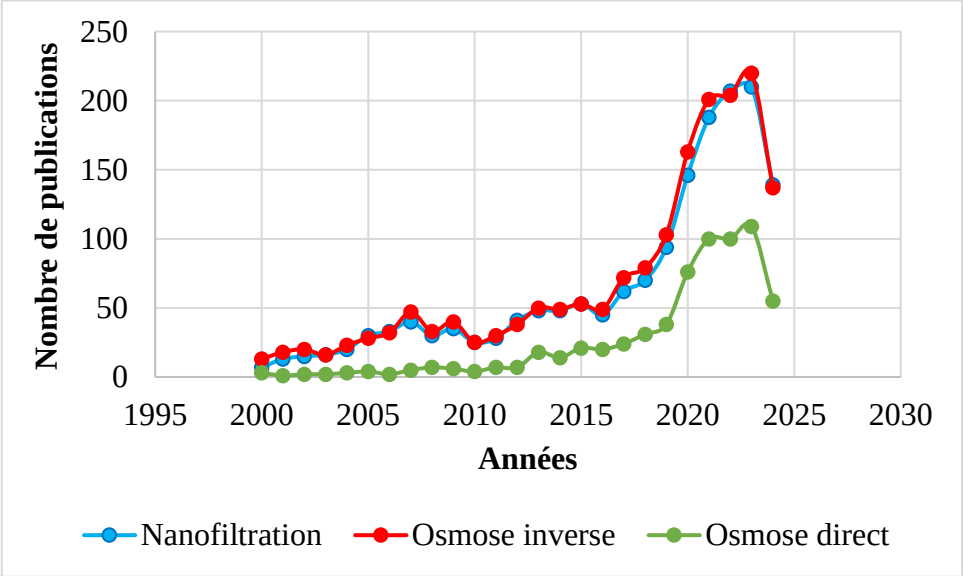


Figure I.13. Nombre de publications de 2000 à 2024 pour les membranes de la nanofiltration, osmose inverse et osmose direct.

I.8. Conclusion

La rétention des pesticides par les membranes de la nanofiltration, osmose inverse et osmose direct 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

QSPR-Réseaux de neurones articiels

II.1. Introduction

La Chemoinformatique est un nouveau domaine apparu il y a quelques années. Il concerne le développement, la création, l'organisation, le stockage, la diffusion, l'analyse, la visualisation et l'utilisation de l'information chimique, elle représente un domaine en évolution rapide. Actuellement, les méthodes spécifiques de la chemoin formatique sont devenues indispensables pour le développement de nouveaux composés, matériaux et processus.

La chemoin formatique est une discipline scientifique qui a beaucoup évoluée durant ces 40 dernières années. Cette science est à l'interface entre la chimie et l'informatique. Il a été rendu compte que dans beaucoup de domaines de la chimie moderne, le nombre important de données et d'informations fourni par la recherche en chimie peut seulement être analysé par des méthodes informatiques. De plus, les nouveaux problèmes rencontrés en chimie sont si complexes que l'utilisation de l'informatique a été rendue indispensable. C'est pourquoi des méthodes ont été développées pour prédire les propriétés chimiques, physiques et biologiques des matériaux et des composés chimiques. Ces méthodes ont aussi été développées pour la représentation des médicaments, l'élucidation des structures cristallines, dans la prédiction des réactions chimiques et dans la représentation des molécules organiques. La recherche et le développement en chemoinformatique sont essentiels pour :

- L'amélioration de la compréhension des phénomènes chimiques ;
- Permettre à l'industrie chimique de rester compétitive dans le marché mondial (http://www.chemspider.com/Chemical-Structure.6886.html?rid=0ae6e03a-77d9-40b1a7cd-44816ea08f94&page_num=0).

Ces dernières années, il y a eu plusieurs tentatives pour faire progresser l'utilisation des relations quantitatives structure-propriétés (QSPR) 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'étude sont également appliqué les modèles QSPR 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, Ammi et al. 2015, 2017, 2018, 2020, 2021, 2021).

Ce chapitre a été consacré à des modèles de réseaux de neurones artificiels basés sur la relation quantitative structure-activité "QSAR-RN". On parlera sur la méthode d'apprentissage statistique (relation quantitative structure-activité "QSAR" et les réseaux de neurones artificiels "RNA"). Une

conclusion clôturera ce chapitre et sera une synthèse par laquelle on justifiera le choix du thème de ce mémoire.

II.2. Méthodes d'apprentissages statistiques

La théorie de l'apprentissage statistique est un système d'apprentissage automatique à partir des domaines de la statistique et de l'analyse fonctionnelle. La théorie de l'apprentissage statistique traite du problème de la recherche d'une fonction prédictive basée sur des données. La théorie de l'apprentissage statistique a conduit à des applications dans des domaines tels que la vision par ordinateur, la reconnaissance de la parole, la bioinformatique(https://fr.wikipedia.org/wiki/Bio-informatique).

II.3. Relations quantitatives structure-Activités (QSAR)

Initiées sur la base de relations empiriques durant la seconde partie du 19ème siècle, les QSAR ont été mathématiquement formalisées dans les années 1930 et leur acceptation, en tant que discipline à part entière, a été définitivement acquise au début des années 1960 suites aux travaux de Hansch et Fujita (Devillers, 2013) qui ont proposé un modèle mathématique pour corréler l'activité biologique et la structure chimique. Depuis lors, un grand nombre de modèles a été conçu pour la prédiction de diverses activités d'intérêt biologique.

Il ya de nombreuses utilités dans l'utilisation des méthodes QSAR. Nous donnons ci-dessous, à titre d'exemple, quelques-unes :

- Prédire l'activité biologique et les propriétés physico-chimiques.
- Comprendre et rationaliser les mécanismes d'action dans une série de produits chimiques.
- Économiser le coût de développement de produits (par exemple dans l'industrie pharmaceutique, pesticides, produits d'hygiène personnelle,…etc.).
- Réduire ou même remplacer les tests longs et coûteux sur les animaux.

La capacité à prédire une activité biologique est utile dans un certain nombre de situations. Il existe un grand nombre d'applications de ces modèles QSAR au sein de l'industrie, du milieu universitaire et des organismes gouvernementaux chargés de la réglementation. Nous résumons ci-après quelques cas d'utilisations de ces modèles :

- l'optimisation de l'activité pharmacologique, biocide ou pesticide ;

- la conception rationnelle de nombreux autres produits tels que des agents tensio-actifs, des parfums, des colorants ou des produits chimiques ;
- l'identification des composés dangereux à des stades précoces de développement ;
- la conception de la toxicité et des effets secondaires pour les nouveaux composés ;
- la prédiction de la toxicité pour les humains pour les multiples expositions (délibérée, occasionnelle ou professionnelle) et de l'écotoxicité ;
- la prédiction d'une variété de propriétés physico- chimiques des molécules (qu'il s'agisse de produits pharmaceutiques, de pesticides, de produits d'hygiène personnelle, la chimie fine, etc.) (Hamadache 2014).

II.3.1. Outils de QSAR

Les modèles QSAR permettent de quantifier la relation entre la structure et l'activité biologique et d'apprécier l'impact de diverses propriétés physico-chimiques sur cette activité. L'élaboration de tels modèles QSAR repose sur l'utilisation de trois outils : une ou plusieurs données biologiques à prédire, des descripteurs moléculaires pour représenter les propriétés des molécules et des méthodes d'apprentissage statistique qu'on introduira dans une partie à part (Hamadache 2014).

II.3.2. Données biologiques

Les données biologiques sont habituellement exprimées sur une échelle logarithmique en raison de la relation linéaire entre la réponse et le logarithme de dose dans la région centrale de la courbe de log dose-réponse. Les logarithmes inverses de l'activité (log 1/C) sont également utilisés pour obtenir des valeurs mathématiques plus élevées lorsque les structures sont biologiquement très efficaces. Parmi les données biologiques utilisées dans l'analyse de QSAR, on trouve la constante de vitesse, la constante d'inhibition, … etc (Hamadache 2014).

II.3.3. Descripteurs moléculaires

De nombreuses recherches ont été menées, au cours des dernières décennies, pour trouver la meilleure façon de représenter l'information contenue dans la structure des molécules. Ainsi, les recherches ont abouti à la représentation de la structure moléculaire par un ensemble de valeurs numériques appelées descripteurs moléculaires. On en dénombre aujourd'hui plus de 3000 types, qui quantifient des caractéristiques physico-chimiques ou structurelles de molécules. Ils peuvent être obtenus de manière expérimentale, mais les descripteurs calculés, et non mesurés, sont à privilégier : ils permettent en effet d'effectuer des prédictions sans avoir à synthétiser les molécules, ce qui est un des objectifs de la modélisation. Les descripteurs moléculaires peuvent être divisés en deux grandes

catégories (Poezevara, 2011) : les descripteurs expérimentaux (par exemple : la lipophilie [logP], la polarisabilité, la densité électronique, le potentiel d'ionisation) et les descripteurs moléculaires théoriques. Ces derniers sont à nouveaux classés en trois catégories en fonctionde leur dimension (Hong 2008 et Hamadache 2014).

A. **Descripteurs 1D**

Ils sont accessibles à partir de la formule brute de la molécule et décrivent des propriétés globales du composé. Il s'agit par exemple de sa composition, c'est-à-dire les atomes qui le constituent, ou de sa masse molaire.

B. **Descripteurs 2D**

Ils sont calculés à partir de la formule développée de la molécule. Ils peuvent être de plusieurs types:

- Les indices constitutionnels caractérisent les différents composants de la molécule. Il s'agit par exemple du nombre de liaisons simples ou multiples, du nombre de cycles…
- Les indices topologiques peuvent être obtenus à partir de la structure 2D de la molécule, et donnent des informations sur sa taille, sa forme globale et ses ramifications.
- L'indice de Wiener par exemple, permet de caractériser le volume moléculaire et la ramification d'une molécule. Un autre indice, nommé "indice de Randic", est l'un des descripteurs les plus utilisés. Il peut être interprété comme une mesure de l'aire de la molécule accessible au solvant.

C. Descripteurs 3D

Ces descripteurs 3D reflètent les propriétés physiques dans la plupart des cas, mais sont insuffisants pour expliquer de façon satisfaisante certaines propriétés ou activités, telles les activités biologiques. Aussi, des descripteurs représentant l'information dérivée de la représentation en trois dimensions des molécules ont pu être calculés grâce au développement des techniques instrumentales. Ils sont évalués à partir des positions relatives de ses atomes dans l'espace, et décrivent des caractéristiques plus complexes. Ces descripteurs apportent davantage d'informations, et sont nécessaires à la modélisation de propriétés ou d'activités qui dépendent de la structure 3D. On distingue plusieurs familles importantes de descripteurs 3D :

- Les descripteurs géométriques les plus importants sont le volume moléculaire, la surface accessible au solvant, le moment principal d'inertie.

- Les descripteurs électroniques permettent de quantifier différents types d'interactions inter et intramoléculaires, de grande influence sur l'activité biologique de molécules. Par exemple, les énergies de la plus haute orbitale moléculaire occupée et de la plus basse vacante sont des descripteurs fréquemment sélectionnés. Par ailleurs, le moment dipolaire, le potentiel d'ionisation, et différentes énergies relatives à la molécule sont d'autres paramètres importants.
- Les descripteurs spectroscopiques: les molécules peuvent être caractérisées par des mesures spectroscopiques, par exemples par leurs fonctions d'onde vibrationnelles. En effet, les vibrations d'une molécule dépendent de la masse des atomes et des forces d'interaction entre ceux-ci ; ces vibrations fournissent donc des informations sur la structure de la molécule et sur sa conformation. Les spectres infrarouges peuvent être obtenus soit de manière expérimentale, soit par calcul théorique, après recherche de la géométrie optimale de la molécule.

Par exemple, les descripteurs de type MoRSE sont calculés à partir d'une simulation du spectre infrarouge. Par exemple, le programme appelé DRAGON permet de calculer différents groupes de descripteurs : auto correlations 2D, descripteurs géométriques, descripteurs RDF, descripteurs 3D-MoRSE,

D.Descripteurs 4-D

Ils correspondent à la mesure des propriétés 3D (potentiel électrostatique, d'hydrophobicité, de liaison hydrogène...) d'une molécule en tout point de l'espace. Ils permettent d'avoir l'information sur la structure de la cible (protéine). On pourra ainsi distinguer les descripteurs 4D qui nécessitent un alignement de la molécule guidé par l'étude des complexes ligand-cible (ou, au moins, par des contraintes visant d'optimiser le recouvrement spatial des champs électriques et stériques des ligands, faute d'information sur le vrai mode de fixation dans la cible) avant d'être calculés. Ces descripteurs sont obtenus par le calcul des champs d'interactions moléculaires (CoMFA, CoMSIA) entre une molécule et une sonde représentée par une autre molécule (eau, amide, ...) (Cramer et all 1988 ; Navajas et all 1996; Bonachera et all 2011 ; Goulon-Sigwalt-Abram et all 2008).

II.3.4. Formulation mathématique des QSAR

Les équations QSAR prennent la forme générale suivante :

$$QSAR\begin{cases} Activité\ biologique\ =\ fonction\ (Descripteurs) \\ Activité\ thérapeutique\ =\ fonction\ (Descripteurs) \end{cases}$$

$QSPR : Propriétés = f(Descripteurs)$

$QSTR : Toxicités = f(Descripteurs)$

Dans l'équation, l'activité biologique est normalement exprimée comme log [1/ (symbole de concentration)], où C'est généralement la concentration minimum requise pour causer une réponse biologique définie. La relation mathématique entre un paramètre ou plusieurs paramètres physico-chimiques et l'activité biologique d'un composé peut être exprimée à travers l'équation de Hansch :

$$\log\left(\frac{1}{c}\right) = b_0 + \sum_{i=1}^{j}(b_i D_i) \qquad \text{(II. 1)}$$

Dans cette équation, C'est la concentration de substance nécessaire pour obtenir une réponse biologique, $b_0$est une constante, b_i est un coefficient, $D_I$est un descripteur moléculaire et j est le nombre de descripteurs utilisés.

Les relations QSARs les plus connues ont été introduites par Hansh et Fujita (Poezevara, 2011). Une des équations les plus utilisées présume que les propriétés physiques responsable de l'activité biologique peuvent être séparées en trois catégories : (i) une composante hydrophobique (π), (ii) une composante électronique (σ) et (iii) une composante stérique (ES, RM) :

$$log\left(\frac{1}{c}\right) = a + b\pi + c\pi^2 + dE_s + eRM + f\sigma \text{(II. 2)}$$

a, b, c, d, e et f sont des coefficients déterminés durant l'étude QSAR, tandis que π, ES, RM et σ sont des paramètres fonctions des substituants. L'équation de Hansch est utilisée pour prévoir l'activité biologique de composés dont la structure est semblable à ceux utilisés pour former cette équation. Cependant, dans le cas où la valeur prévue est différente de celle déterminée par l'expérimentation, cela peut signifier que l'activité biologique réelle des composés n'est pas prise en compte par les paramètres mis en jeu dans le processus de construction de l'équation (Hamadache 2014).

II.4. Réseaux de neurones artificiels (RNA)

II.4.1. Histoire des réseaux de neurones artificiels

Un réseau de neurones artificiels est un système informatique matériel et / ou logiciel dont le fonctionnement est calqué sur celui des neurones du cerveau humain.Il s'agit là d'une variété de technologie Deep Learning (apprentissage profond), qui fait elle-même partie de la sous-catégorie d'intelligence artificielle du Machine Learning (apprentissage automatique).

Le concept des réseaux de neurones artificiels futinventé en 1943 par deux chercheurs de l'Université de Chicago : le neurophysicien Warren McCullough, et le mathématicien Walter Pitts. Dans un article publié dans le journal Brain Theory, les deux chercheurs présentent leur théorie selon laquelle l'activation de neurones est l'unité de base de l'activité cérébrale.

En 1957, le Perceptron fut inventé. Il s'agit du plus ancien algorithme de Machine Learning, conçu pour effectuer des tâches de reconnaissance de patterns complexes. C'est cet algorithme qui permettra plus tard aux machines d'apprendre à reconnaître des objets sur des images.

Malheureusement, à l'époque, les réseaux de neurones étaient limités par les ressources techniques. Par exemple, les ordinateurs n'étaient pas assez puissants pour traiter les données nécessaires au fonctionnement des réseaux de neurones. C'est la raison pour laquelle la recherche dans le domaine des Neural Networks est restée en sommeil durant de longues années.

Il aura fallu attendre le début des années 2010, avec l'essor du Big Data et du traitement massivement parallèle, pour que les Data Scientists disposent des données et de la puissance de calcul nécessaires pour exécuter des réseaux de neurones complexes. En 2012, lors d'une compétition organisée par Image Net, un Neural Network est parvenu pour la première fois à surpasser un humain dans la reconnaissance d'image.

C'est la raison pour laquell**e** cette technologie est de nouveau au cœur des préoccupations des scientifiques. A présent, les réseaux de neurones artificiels ne cessent de s'améliorer et d'évoluer de jour en jour (https://www.lebigdata.fr/reseau-de-neurones-artificiels definition#:~:text=(apprentissage%20automatique).,L'histoire%20des%20r%C3%A9seaux%20de%20neurones%20artificiels,et%20le%20math%C3%A9maticien%20Walter%20Pitts).

II.4.2. 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 une 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 sera méfié pour former les dendrites. C'est par les dendrites que l'information est acheminée de l'extérieur vers le soma (corps du neurone) (Ammi 2016).

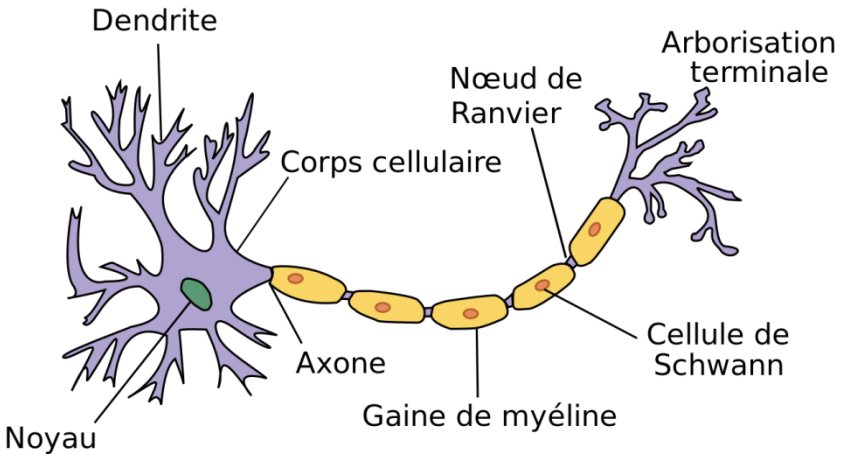


Figure II.1. Représentation d'un neurone biologique

https://fr.wikipedia.org/wiki/Neuromorphologie#/media/Fichier:Neurone_fran%C3%A7ais_sch%C3%A9ma.svg

Un neurone biologique reçoit des entrées ou signaux transmis par d'autres neurones (interaction dendrites- synapse). Au niveau du corps (soma), le neurone analyse et traite ces signaux en les sommant. Si le résultat obtenu est supérieur au seuil d'activation (ou d'excitabilité), il envoie une décharge alors nommé potentiel d'action le long de son axone vers d'autres neurones biologiques.

II.4.3. 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.

Entrées

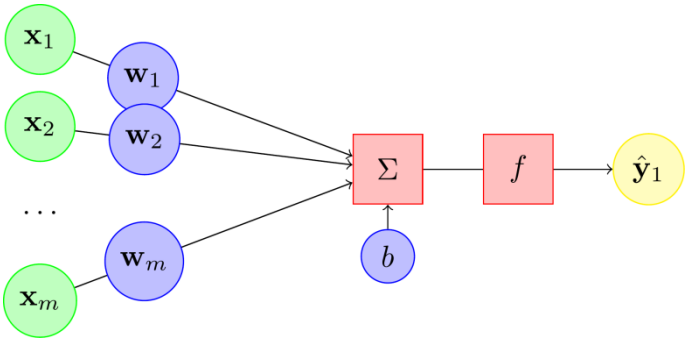


Figure II.2. Neurone formel (https://openclassrooms.com/fr/courses/5801891-initiez-vous-au-deep-learning/5801898-decouvrez-le-neurone-formel)

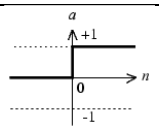
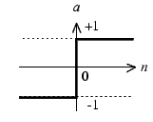
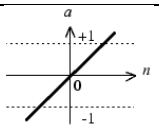
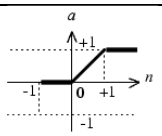
Un neurone formel est une fonction algébrique paramétrée, non linéaire en ses paramètres. Ses entrées peuvent être les sorties d'autres neurones ou des entrées de signaux extérieurs. Sa sortie est une fonction non linéaire *f* d'une combinaison linéaire *v* des entrées (x_i). Le potentiel *v* le plus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

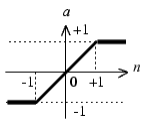
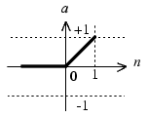
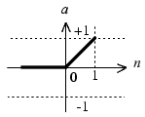
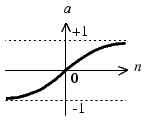
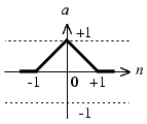
$$v = w_0 + \sum_{i=1}^{n} (w_i x_i) \qquad \text{(II. 3)}$$

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]. Un neurone formel, réalise donc l'opération suivante (Ammi2016) :

$$y = f(v) = f\left(w_0 + \sum_{i=1}^{n} (w_i x_i)\right) \qquad \text{(II. 4)}$$

Tableau II.1. Fonctions de transfert usuelles, a=f(n) (Khaouane 2013).

Nom de la fonction	**Relation entrée / sortie**	**Graphe de la fonction**	**Fonction MATLAB® Correspondante**
Seuil	a= 0 si n < 0 a=1 si $n \geq 0$		Hardlim
Seuil symétrique	a= −1 si n < 0 a=1 si $n \geq 0$		Hardlims
Linéaire	a= n		Purelin
Linéaire saturée	a= 0 si n <0 a= n si $0 \leq n \leq 1$ a=1 si n > 1		Satlin

Linéaire saturée Symétrique	$a = -1$ si $n < -1$ $a = n$ si $-1 \leq n \leq 1$ $a = 1$ si $n > 1$		Satlins
Linéaire positive	$a = n$ si $n \geq 0$ $a = 0$ si $n < 0$		Poslin
Logistique (Sigmoïde Logarithmique)	$a = \frac{1}{1 + e^{-n}}$		Logsig
Tangente hyperbolique	$a = \frac{e^{+n} - e^{-n}}{e^{+n} + e^{-n}}$		Tansig
Triangulaire de base	$a = 1 - \lvert n \rvert, -1 \leq n \leq 1$ $a = 0$ *autrement*		Tribas

Un neurone formel est une représentation artificielle et schématique d'un neurone biologique (Figure II.3) :

- Les synapses sont modélisées par des poids,
- Le soma ou corps cellulaire est modélisé par la fonction de transfert, appelé aussi fonction d'activation
- L'axone par l'élément de sortie (https://scilogs.fr/intelligence-mecanique/reprenons-bases-neurone-artificiel-neurone-biologique/).

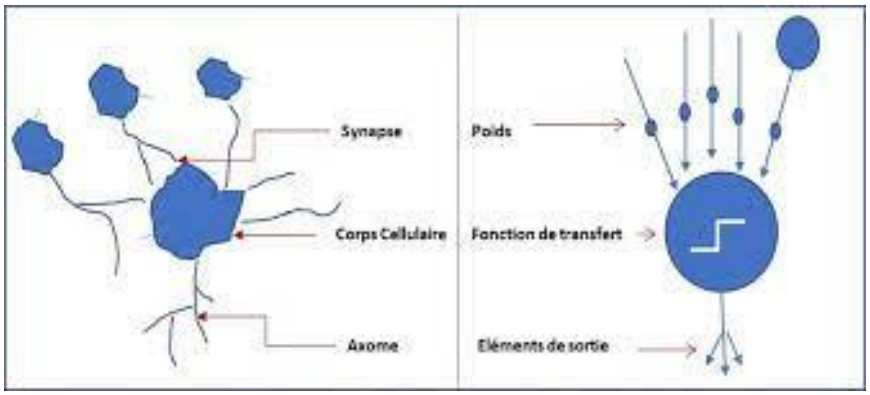


Figure II.3. Mise en correspondance neurone biologique / neurone artificiel (https://www.journaldunet.com/intelligence-artificielle/1494055-les-neurosciences-ont-inspire-l-intelligence-artificielle-le-modele-neuronal-et-les-concepts-2-2/)

II.4.4. 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 (Figure II.4) :

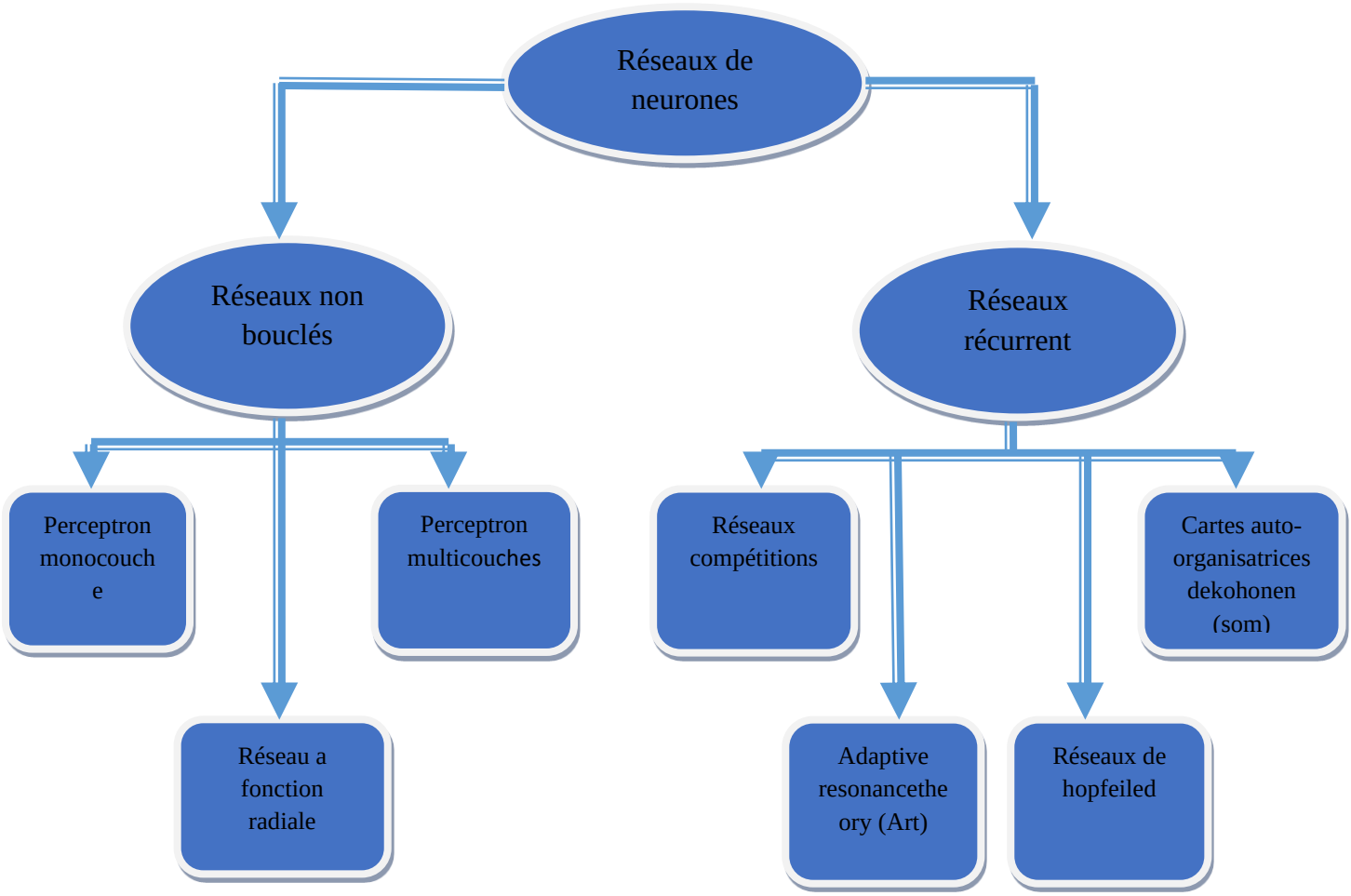


Figure II.4. Réseaux de neurones les plus utilisés.

- **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 (Figure II.5).

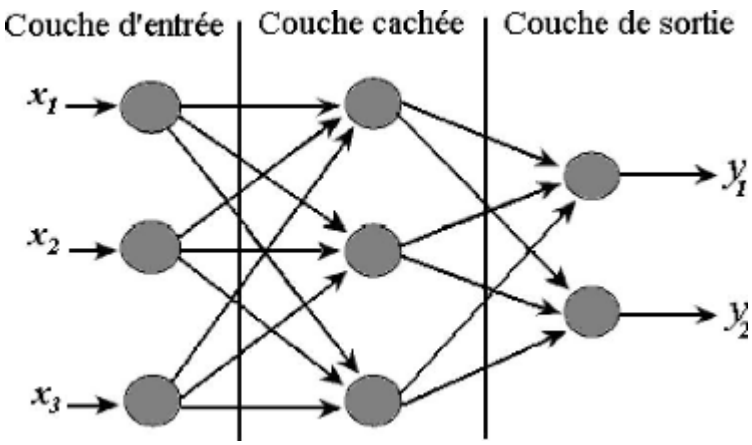


Figure II.5. Structure d'un Réseaux de neurones non bouclés (https://www.memoireonline.com/04/12/5750/m_Identification-et-commande-des-systemes-non-lineaires13.html)

- **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 Multicouches (PMC), illustré par la figure II.6 et dont les principes sont exposés dans ce qui suit (Ammi 2016).

II.4.5. Perceptrons multicouches

Le Perceptron multicouches "PMC" est un réseau de neurones non bouclé caractérisé par une structure bien particulière. La figure II.5 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 (Hamadache 2014).

II.4.6. Types d'apprentissage des réseaux de neurones

L'apprentissage est vraisemblablement la propriété la plus intéressante des réseaux de neurones. L'apprentissage est une phase du développement du réseau de neurones durant laquelle on calcule les poids des neurones de telle manière que les sorties du réseau soient aussi proches que possible des sorties désirées. Les procédures d'apprentissage sont divisées, elles aussi, en deux grandes catégories : apprentissage supervisé et apprentissage non supervisé (Khaouane 2013).

A. Apprentissage supervisé

L'apprentissage supervisé implique l'existence d'un professeur qui a pour rôle d'évaluer le succès (ou l'échec) du réseau quand il lui est présenté un stimulus connu (on dit que ce stimulus est un exemple appartenant à la base d'apprentissage). Cette supervision consiste à renvoyer au réseau une information lui permettant de faire évoluer ses connexions (parfois aussi sa propre architecture) afin de faire diminuer son taux d'échec. L'information peut être explicite, sous la forme d'une mesure de l'erreur commise, par exemple, ou globalement sur l'ensemble des exemples de la base. C'est ce qui se passe dans le cas du perceptron. Mais elle peut être plus implicite (apprentissage par renforcement), sous forme d'une simple appréciation (bon ou mauvais, punition ou récompense), sans mesure d'erreur et même être globale, sur l'ensemble des tâches que le réseau doit exécuter. Dans ce dernier cas, il est facile de comprendre que l'apprentissage est le plus difficile, la difficulté majeure consistant pour le réseau à identifier les étapes du processus qui sont responsables de l'échec ou du succès (Khaouane 2013).

B. Apprentissage non supervisé

L'apprentissage non supervisé implique la fourniture à un réseau autonome d'une quantité suffisante d'exemples contenant des répétitions (autrement dit de la redondance), telles que celui-ci en dégage les régularités automatiquement. Ces réseaux sont souvent appelés auto-organisateurs, ou encore à apprentissage compétitif. Dans l'apprentissage non supervisé, les données ne contiennent pas d'informations sur une sortie désirée, il n'y a pas de superviseur ou expert humain. Il s'agit de déterminer les paramètres du réseau de neurones suivant un critère à définir. Dans ce cas, les exemples

présentés à l'entrée provoquent une auto adaptation du réseau afin de produire des valeurs de sortie qui soient proches en réponse pour des valeurs d'entrées similaires. Ce type d'apprentissage possède souvent moins de complexité dans le calcul par rapport à l'apprentissage supervisé. Bien entendu, l'architecture du réseau, préalablement définie par son utilisateur, est une forme de supervision (Khaouane 2013).

II.4.7. Réseaux de neurones les plus utilisés

A. Perceptron monocouche (perceptron simple)

Le perceptron est le premier modèle des réseaux de neurones inventé en 1957 par Frank Rosenblatt (Rosenblatt 1958). C'est un réseau simple, puisqu'il ne se compose que d'une couche d'entrée et d'une couche de sortie. Il peut aussi être utilisé pour faire de la classification et pour résoudre des opérations logiques simples, il suit généralement un apprentissage supervisé (Talbi 2011).

B. Le perceptron multicouche (en anglais Multi Layer Perceptron)

C'est une extension du précédent, avec une ou plusieurs couches cachées entre l'entrée et la sortie (Keedwell et Narayanan 2005). Chaque neurone dans une couche est connecté à tous les neurones de la couche précédente et de la couche suivante. Il peut résoudre des problèmes non linéairement séparables et des problèmes logiques plus compliqués (Talbi 2011). Il suit un apprentissage supervisé, ce réseau a été choisi dans la suite de nos travaux.

C. Réseaux à fonction radiale

Les réseaux à fonction de bases radiales (RFR) ou plus simplement réseaux à bases radiales ont été proposés par Moody et Darken (Moody et Darken 1989). Les RFR seront donc employés dans les mêmes types de problèmes que les perceptrons multicouches à savoir, en classification et en approximation de fonctions, particulièrement. L'apprentissage le plus utilisé pour les RFR est le mode hybride (Talbi 2011).

D. Cartes auto-organisatrices de Kohonen (SOM)

Les cartes topologiques ou cartes auto organisatrices ont été introduites pour la première fois par Kohonen en 1982 (Kohonen 1982). Ce sont des réseaux à apprentissage non-supervisé qui établissent une carte discrète, ordonnée topo-logiquement, en fonction de patterns d'entrée. Le réseau forme ainsi une sorte de treillis dont chaque nœud est un neurone associé à un vecteur de poids. La

correspondance entre chaque vecteur de poids est calculée pour chaque entrée. Par la suite, le vecteur de poids ayant la meilleure corrélation, ainsi que certains de ses voisins, vont être modifiés afin d'augmenter encore cette corrélation(Talbi 2011).

E. Réseaux de Hopfield

Les réseaux de Hopfield (Hopfield 1982) sont des réseaux récurrents et entièrement connectés. Dans ce type de réseau, chaque neurone est connecté à chaque autre neurone et il n'y a aucune différence entre les neurones d'entrée et de sortie. Ils fonctionnent comme une mémoire associative non-linéaire et sont capables de trouver un objet stocké en fonction de représentations partielles ou bruitées. L'application principale des réseaux de Hopfield est l'entrepôt de connaissances, mais aussi la résolution de problèmes d'optimisation. Le mode d'apprentissage utilisé ici est le mode non-supervisé (Talbi 2011).

F. Réseaux ART

Les réseaux ART (en anglais Adaptive Resonance Theory) sont des réseaux à apprentissage par compétition (Carpenter et Grossberg1987 ; Carpenter et Grossberg 1988). Dans ce genre de réseau, les vecteurs de poids ne seront adaptés que si l'entrée fournie est suffisamment proche d'un prototype déjà connu par le réseau. Il est à noter qu'il existe deux principaux types de réseaux ART : les ART-1 pour des entrées binaires et les ART-2 pour des entrées continues. Le mode d'apprentissage des ART peut être supervisé ou non (Talbi 2011).

II.4.8. Conception d'un modèle neuronal

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

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

II.4.9. Détermination des entrées/sorties du réseau de neurones

Pour toute conception de modèle, la sélection des entrées doit prendre en compte deux points essentiels (Ammar 2007) :

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

II.4.10. Construction de la base de données et choix des échantillons

Le processus d'élaboration d'un réseau de neurones commence par la construction d'une base de données et le choix des échantillons. Il faut donc déterminer les grandeurs qui ont une influence significative sur le phénomène que l'on cherche à modéliser.

Cette étape est cruciale et aide à déterminer le type de réseau le plus approprié pour résoudre un problème donné. En particulier, la façon dont se présente l'échantillon conditionne le type de réseau, le nombre de cellules d'entrée et de sortie et la façon dont il faudra mener l'apprentissage, le test et la validation (Bishop 1995).

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

Pour mieux contrôler la phase d'apprentissage, il est souvent préférable de posséder une troisième base de données appelée "base de validation croisée" (Stone 1977). Les avantages liés à l'utilisation de cette troisième base de données seront exposés dans les paragraphes suivants. Notons qu'il n'y pas de règle pour déterminer ce partage d'une manière quantitative, néanmoins chaque base doit satisfaire aux contraintes de représentativité de chaque classe de données et doit généralement refléter la distribution réelle, c'est-à-dire la probabilité d'occurrence des diverses classes (Nascimento et al. 2000).

Généralement, les bases de données subissent un prétraitement qui consiste à effectuer une normalisation appropriée tenant compte de l'amplitude des valeurs acceptées par le réseau. Les valeurs d'entrées et de sortie sont normalisées dans un intervalle spécifique afin de donner à chaque

paramètre la même influence statistique. De plus, un problème additionnel peut survenir à cause d'une possible saturation de la sortie, si la somme pondérée des entrées du perceptron multicouches ($\sum(w_i x_i)$) est dans la partie asymptotique de la fonction d'activation (Dames 2005).

II.4.11. Définition de la structure du réseau

La détermination de l'architecture optimale est restée pendant longtemps un problème ouvert. Il existe actuellement diverses méthodes empiriques mettant notamment en jeu des tests statistiques pour une vaste classe de réseaux (Urbani et al. 1994). Comme nous l'avons mentionné, nous avons retenu le perceptron multicouche comme base du modèle. Nous structurons ce réseau en précisant le nombre de couches et de neurones cachés pour que le réseau soit en mesure de reproduire ce qui est déterministe dans les données.

A. Nombre de couches cachées

Mis à part les couches d'entrée et de sortie, il faut décider du nombre de couches intermédiaires ou cachées. Sans couche cachée, le réseau n'offre que de faibles possibilités d'adaptation. Néanmoins, il a été démontré qu'un perceptron multicouche avec une seule couche cachée pourvue d'un nombre suffisant de neurones, peut approximer n'importe quelle fonction avec la précision souhaitée (Hornik 1991).

B. Nombre de neurones cachés

Chaque neurone peut prendre en compte des profils spécifiques de neurones d'entrée. Un nombre plus important permet donc de mieux "coller" aux données présentées, mais diminue la capacité de généralisation du réseau. Il faut alors trouver le nombre adéquat de neurones cachés nécessaire pour obtenir une approximation satisfaisante. Le nombre de neurones cachés dépend du type de données dont on dispose. En général, ce nombre est déterminé par des règles empiriques. Suivant les auteurs, la taille de la couche cachée doit être (Dames 2005) :

- Soit égale à celle de la couche d'entrée ;
- Soit égale à 75 % de celle-ci ;
- Soit égale à la racine carrée du produit du nombre de neurones dans la couche d'entrée et de sortie.

Pour des cas plus sophistiqués, la recherche consisterait soit à :

- Estimer un réseau comportant de nombreux neurones puis à le simplifier par l'analyse des multi-colinéarités ou par une règle d'apprentissage éliminant les neurones inutiles (Dames 2005) ;
- Définir une architecture tenant compte de la structure des variables identifiées au préalable par une analyse en composantes principales (Bermejo et Cabestany 1999; Bermejo et Cabestany 2001)
- Augmenter progressivement la complexité du modèle à chaque itération, dite "Cascading" et réévaluer l'efficacité de généralisation à chaque addition d'un nouveau neurone (Dames 2005).

II.4.12. Apprentissage

Une fois l'architecture choisie, elle doit subir une phase d'apprentissage. Cette dernière consiste à calculer les coefficients synaptiques de telle manière que les sorties du réseau de neurones soient, pour les exemples utilisés lors de l'apprentissage, aussi proches que possibles des sorties "désirées". Celles-ci correspondent à la valeur de la fonction que l'on veut approcher ou à la sortie du processus que l'on veut modéliser (Dreyfus 2002).

L'apprentissage correspond donc à la phase du développement du réseau durant laquelle les coefficients synaptiques sont modifiés jusqu'à l'obtention du comportement désiré.

Dans notre étude, comme nous disposons des valeurs de sortie, nous avons choisi l'apprentissage supervisé, mettant en jeu un algorithme d'optimisation. Ce dernier cherche à minimiser, par des méthodes d'optimisation non linéaire, une fonction de coût qui constitue une mesure de l'écart entre les réponses réelles du réseau et ses réponses désirées (Dreyfus 2002).

Cette optimisation se fait de manière itérative, en modifiant les poids en fonction du gradient de la fonction de coût : le gradient est estimé par une méthode spécifique, dite la règle du gradient de l'erreur (en anglais delta rule). C'est une des règles les plus utilisées pour l'apprentissage supervisé de réseaux de neurones. Initialement développée pour résoudre des problèmes de traitements adaptatifs du signal, elle a conduit ensuite à un algorithme connu sous le nom de l'algorithme de rétro-propagation du gradient de l'erreur (en anglais Back-propagation) (Dames 2005). On distingue deux étapes dans la phase d'apprentissage (Dames 2005) :

- Une phase de propagation, qui consiste à présenter une configuration d'entrée au réseau, puis à la propager à celle de sortie en passant par les couches cachées.

- Une phase de rétro-propagation, qui consiste, après le processus de propagation, à minimiser l'erreur commise sur l'ensemble des exemples présentés, erreur considérée comme une fonction des poids synaptiques. Cette erreur représente la somme des différences au carré entre les réponses calculées et celles désirées pour tous les exemples contenus dans l'ensemble de l'apprentissage.

Les algorithmes de rétro-propagation implantés actuellement dans les logiciels de RNA appartiennent à une des cinq classes suivantes (Si-Moussa 2009):

- Descente du gradient
- Descente des gradients conjugués
- Quasi-Newton
- Levenberg-Marquardt
- Régularisation

Tableau II.2 résume les algorithmes d'apprentissage par rétro-propagation implanté dans la boite à outils des RN du programme MATLAB® (Demuth et Beale 2000).

Tableau II.2. Résumé des caractéristiques des algorithmes d'apprentissage de rétro-propagation implantés dans MATLAB® (Khaouane 2013).

Algorithme	**Caractéristiques**
Descente du gradient	
Descente du gradient simple (traingd)	Descente du gradient de base. Réponse lente; peut être utilisée pour apprentissage en mode séquentiel
Descente du gradient avec moment (traingdm)	Descente du gradient avec moment. Le moment permet à un réseau de répondre non seulement au gradient local, mais aussi aux tendances récentes dans la surface d'erreur. Sans moment un réseau peut être coincé dans un minimum local peu profond. Plus rapide que traingd, ne peut être utilisée que pour apprentissage en mode séquentiel
Descente du gradient avec taux d'apprentissage 9variable (traingda)	Taux d'apprentissage variable. Plus rapide que traingd, mais ne peut être utilisée que pour

	apprentissage en mode batch. Utilise des règles heuristiques pour la variation du taux d'apprentissage
Descente du gradient avec moment et taux d'apprentissage variable (traingdx)	Descente du gradient avec inertie (momentum) et taux d'apprentissage variable
Descente des gradients conjugués	
Descente des gradients conjugués avec la méthode de mise à jour de Fletcher-Reeves (traincgf)	Il possède la plus petite mémoire de stockage parmi les algorithmes de gradient conjugués.
Descente des gradients conjugués avec la méthode de mise à jour de Polak-Ribiére (traincgp)	Nécessite une mémoire de stockage légèrement plus grande que traincgf. Convergence plus rapide sur quelques problèmes.
Descente des gradients conjugués avec la méthode de reprises Powell-Beale (traincgb)	Nécessite une mémoire de stockage légèrement plus grande que traincgp. Convergence Généralement plus rapide.
Scaledconjugate gradient (trainscg)	Le seul algorithme de gradients conjugués qui n'exige aucune recherche de ligne. Très bon algorithme d'apprentissage en général.
Quasi-Newton	
BFGS quasi-Newton (trainbfg)	Il exige le stockage d'une approximation de la matrice Hessiènne et nécessite plus de calcul à chaque itération que les algorithmes de gradients conjugués, mais possède habituellement une convergence plus rapide.

II.4.13. Evaluation des performances du réseau de neurones

Enfin, la dernière étape doit permettre d'estimer la qualité du réseau obtenu en lui présentant des exemples qui ne font pas partie de l'ensemble d'apprentissage. Elle permet à la fois d'apprécier les performances du système neuronal et de détecter le type de données qui posent problème. Si les performances ne sont pas satisfaisantes, il faudra, soit modifié l'architecture du réseau, soit modifié la base d'apprentissage (Dames 2005).

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 pesticides lors des membranes d'osmose inverse et de nanofiltration.

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 pesticides lors d'osmose inverse et de nanofiltration en utilisant des réseaux de neurones artificiels (RNA) et les relations quantitatives structure-propriétés (QSPR). 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 neurones optimaux dans la couche cachée et pour le choix des fonctions de transfert.

Chapitre III

Modélisation de la rétention des pesticides lors d'osmose et nanofiltration par QSPR-RNA



III.1 Introduction

Les techniques de modélisation décrites dans le chapitre II ont été appliquées aux données expérimentales de la rétention des pesticides lors de nano filtration et osmose.

Ce chapitre comportera six axes. Dans le premier axe, on examinera les données collectées, prétraitement et analyse. La deuxième partie sera consacrée au développement des modèles neuronaux, nous discuterons les performances des modèles neuronaux dans la troisième partie, les expressions mathématiques de modèle neuronal optimal seront abordées dans la quatrième partie, nous traiterons dans la cinquième partie l'analyse de sensibilité par la méthode de poids et dans la dernière partie nous discuterons le diagnostic des points aberrants et domaine d'applicabilité du modèle neuronal optimise.

III. 2 Collecte des données, prétraitement et analyse

Dans cette étude, nous avons utilisé les données disponibles dans la littérature (Yoon et al. 2005 ; Verliefde et al. 2007, 2008 et 2009 ; Yangali-Quintanilla et al. 2008, 2009 et 2010 ; plakas et al 2008 ; Anna M et al .2008 et 2009; Chuyang et al. 2009 ; Huang et al. 2011 ; Konstantions et al. 2011 ; Riungu et al. 2012 ; Sadmani et al. 2014 ; Nikhbakht et al. 2019 ; Hai 2020 ; Jeong et al. 2021 ; Nuray et al. 2022).

La liste des 26 pesticides (20 herbicides, 5 insecticides et 7 fongicides) est répertoriée dans annexe A. La taille de la base de données et le nombre de pesticides sont montrés dans le tableau III.1.

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

Base de données "BD"	QSPR-RN	Taille de base de données	Nombre de pesticides
BD1 avec SR(MgSO4)	QSPR-RN1 QSPR-RN2 QSPR-RN3	516	26
BD2 avec SR(NaCl)	QSPR-RN4 QSPR-RN5 QSPR-RN6	522	19
BD3 avec SR	QSPR-RN7 QSPR-RN8 QSPR-RN9	574	26

Le choix des variables d'entrées et de sortie a été basé sur les interactions entre les propriétés des pesticides, les caractéristiques de la membrane et les conditions opératoires de filtration pour la rétention des pesticides par les membranes d'osmose et nano filtration. Les entrées considérées dans cette étude sont : les descripteurs moléculaires (Mw, de, LogD, le moment dipolaire, la longueur moléculaire "length", la surface minimale "Smin", la surface maximale "Smax", la largeur moléculaire "Width", la profondeur moléculaire "Depth" et la largeur équivalente moléculaire "Eqwidth"), les caractéristiques de la membrane (le poids moléculaire de coupure "MWCO" , le rejet de sel de sulfate de magnésium "SR(MgSO4)", le rejet de sel de chlorure de sodium "SR(NaCl)", la masse molaire de rejet de sel (Mw "SR") la charge de surface de la membrane "zêta potentiel" et l'hydrophobicité de la membrane "angle de contact") et les conditions opératoires (le pH, la pression, la température et le taux de récupération).

Les valeurs de LogD, moment dipolaire, surface minimale "Smin", surface maximale "Smax" ont été calculées par les logiciels qui sont mentionnés dans le tableau III.2. Sur l'annexe B, nous donnons la capture d'écran d'un site d'internet et deux logiciels utilisés.

Tableau III.2. Logiciels utilisés pour calculer LogD, le moment dipolaire, Smin, Smax et Length.

	Logiciels
LogD	Chemspider http://www.chemspider.com/
Moment dipolaire	Hyperchem
Smin	
Smax	
Length	ChemBio3D Ultra (Version 12 ; 2012)

Les valeurs de la largeur équivalente moléculaire (EqWidth) ont été calculées par l'équation I.8 (voir le premier chapitre). Le nombre total de variables (les propriétés des pesticides, les caractéristiques membranaires et les conditions d'opératoires) a été réduit lorsqu'il existait une relation entre elles.
Pour chaque base de données (BD1, BD2 et BD3), un modèle QSPR 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 QSPR-RNA.

	BD1	BD2	BD3
Entrées	-de -Log D -Moment dipolaire -Length -EqWidth -MWCO -SR (MgSO4) -Zeta potentiel -Angle de contact -pH -Pression -Taux de récupération -Température	-de -Log D -Moment dipolaire -Length -EqWidth -MWCO -SR (NaCl) -Zeta potentiel -Angle de contact -pH -Pression -Taux de récupération -Température	-de -Log D -Moment dipolaire -Length -EqWidth -MWCO -Mw (SR) -SR (MgSO4 ou NaCl) -Zeta potentiel -Angle de contact -pH -Pression -Taux de récupération -Température
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.1) "$\bar{x}$", du maximum "max" et de la déviation standard (écart type) "DS" (Équation III.2) (Kratbi 2023) sont présentées dans les tableaux III.4, III.5 et III.6.

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

$$DS = \left(\frac{1}{n}\sum_{k=0}^{n}(X_i - \bar{x})^2\right)^{\frac{1}{2}} \qquad \text{(III.2)}$$

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

	Min	Moyenne	Max	DS
de (g/mol)	0,2884	0,6924	0,8670	0,0628
LogD	-2,2300	2,0725	5,3300	1,5708
Moment dipolaire	0,4610	2,3140	5,5700	1,213
Length (nm)	0,1200	0,2129	1,7950	0,2838
EqWidth	0,5900	1,0046	1,0858	0,0661
MWCO (Dalton)	96,0000	202,7016	400,0000	83,6917
SR(MgSO4)	0,1000	0,9476	0,9920	0,1618
Zeta potentiel (mv)	-87,0000	-25,4129	-1,0000	21,2706
Angle de contact (°)	17,5000	49,2028	94,700	15,3854
Ph	3,3900	7,2530	10,0000	0,8091
Pression (KPa)	240,0000	940,2802	2500,0000	644,9555
Taux de récupération (%)	0,5000	41,2361	111,0000	30,365
Température (C°)	13,0000	22,9787	25,0000	2,6791
Rétention %	2,0000	80,8615	100,0000	20,4520

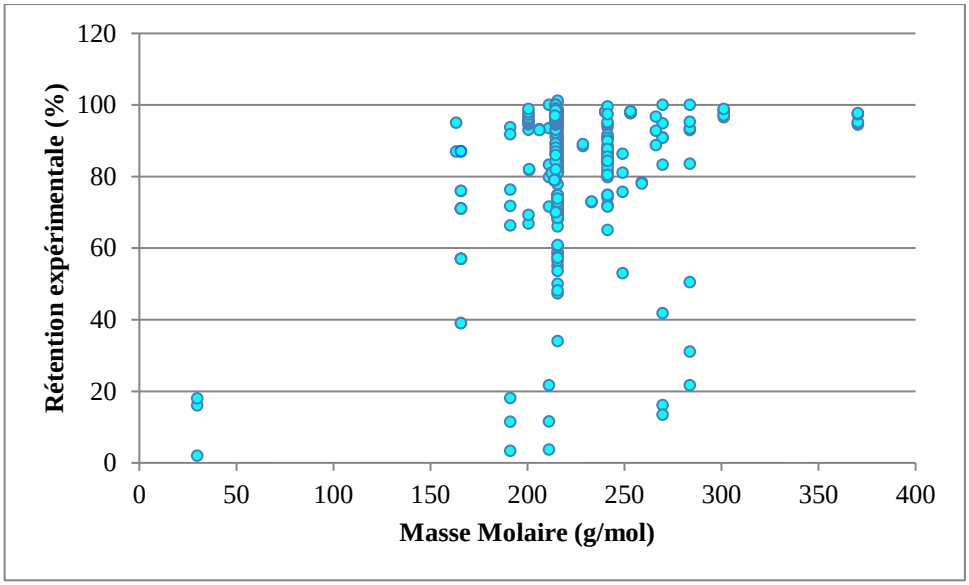
Tableau III.5. Analyse statistique des données d'entrée et de sortie (BD2).

	Min	Moyenne	Max	DS
de (g/mol)	0,6058	0,7018	0,8670	0,0520
LogD	-2,2300	2,1751	5,3300	1,6403
Moment dipolaire	0,4610	2,2319	5,5700	1,0950
Length (nm)	0,1500	0,1877	1,7950	0,2013
EqWidth	0,8200	1,0125	1,0858	0,0499
MWCO (Dalton)	96,0000	193,0920	400,0000	88,4598
SR (NaCl) (%)	0,5000	0,8744	0,9960	0,1631
Zeta potentiel (mv)	-87,0000	-27,0054	-1,0000	21,2953
Angle de contact (°)	17,5000	49,6177	94,7000	15,6033
pH	3,3900	7,2948	10,0000	0,8166
Pression (KPa)	240 ,0000	933,7534	2500,0000	611,2441
Taux de récupération (%)	0,5000	45,7948	122,0000	31,6525
Température (C°)	13 ,0000	23,2567	25,0000	2,588
Rétention %	3,3000	83,0537	100,0000	18,931

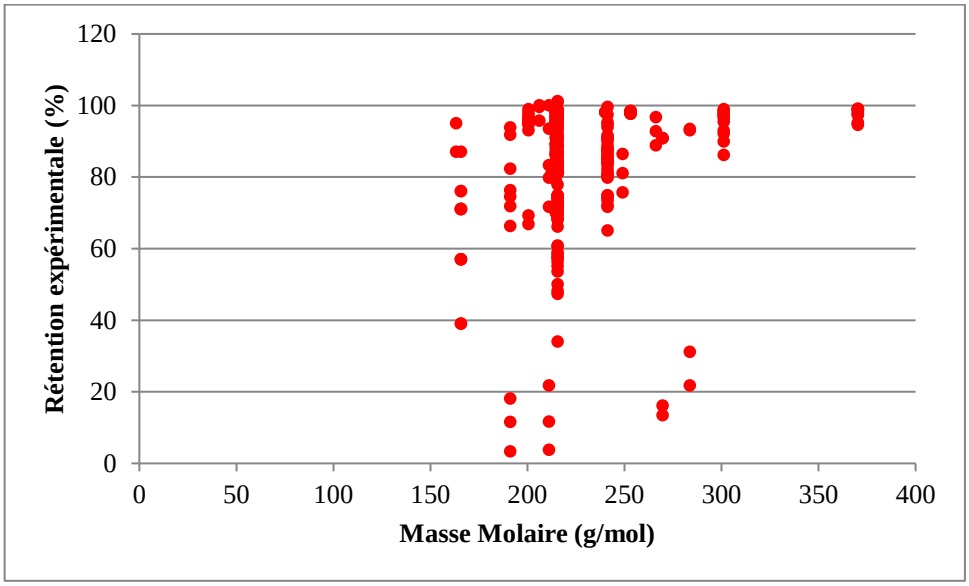
Tableau III.6. Analyse statistique des données d'entrée et de sortie (BD3).

	Min	Moyenne	Max	DS
de (g/mol)	0,2884	0,6986	0,8700	0,0700
LogD	-2,2300	2,1678	5,3300	1,5900
Moment dipolaire	0,4610	2,3613	5,5700	1,2000
Length (nm)	0,1200	0,2082	1,8000	0,2700
EqWidth	0,5900	1,0085	1,0900	0,0700
MWCO (Dalton)	96,000	193,369	400,0000	84,6500
Mw(SR) (g/mol)	58,4400	64,4816	120,3700	18,4100
SR (%)	0,1000	0,8805	1.0000	0,160
Zeta potentiel (mv)	-87,0000	-25,8636	-1,0000	20,6400
Angle de contact (°)	17,5000	49,7631	94,7000	14,950
pH	3,39000	7,2977	10,0000	0,79000
Pression (KPa)	240,0000	972,9696	2500,0000	645,1000
Taux de récupération (%)	0,5000	44,4632	122,0000	31,8000
Température (C°)	13,0000	23,1359	25,0000	2,6000
Rétention (%)	2,0000	82,1840	100,0000	19,9300

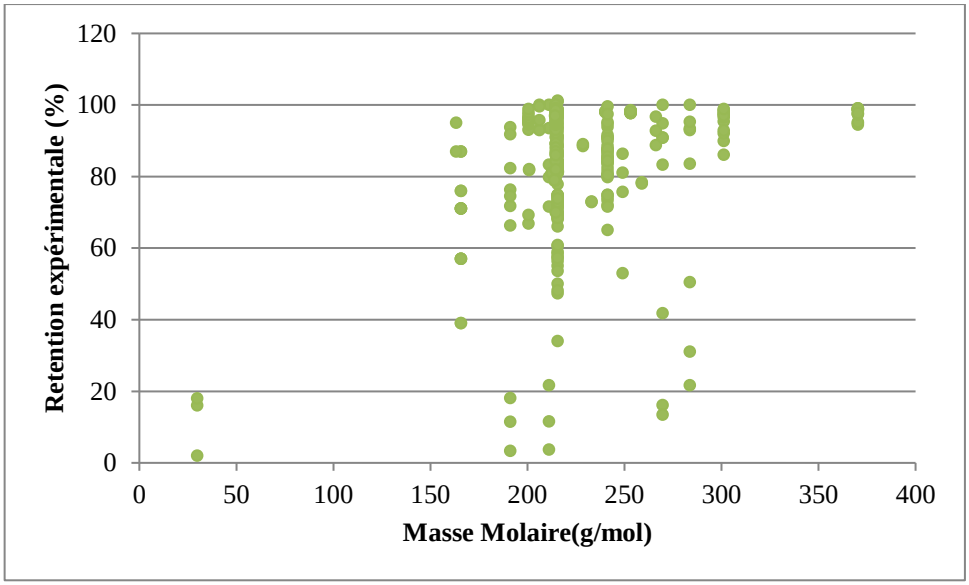
La figure III.1 représente les rétentions expérimentales en fonction de la masse molaire "M_W" pour les trois bases de données (les rétentions en fonction de la masse molaire des pesticides "M_W" sont indiquées par des cercles bleu pour BD1 "516 points", des cercles rouges pour BD2 "522 points" et des cercles verts pour BD3 "574 points"). Quand la masse molaire des pesticides "M_W" varié de 170 g/mol à 300g/mol, les rétentions expérimentales qui ont des pourcentages supérieurs à 60 % s'agglomèrent les unes aux autres.


Rétention expérimentale (%)
120
100
80
60
40
20
0
0
50
100
150
200
250
300
350
400
Masse Molaire (g/mol)

(a)


Rétention expérimentale (%)
120
100
80
60
40
20
0
0
50
100
150
200
250
300
350
400
Masse Molaire (g/mol)

(b)



(c)

Figure III.1. Rétentions expérimentales des pesticides en fonction de masse molaire : (a) BD1, (b) BD2, (c) BD3.

III.3. Développement des modèles QSPR-RNA

Cette étape concerne la création des modèles de réseaux de neurones artificiels basées sur la relation quantitative structure propriété "QSPR-RNA".

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 pesticides, les caractéristiques de la membrane et les conditions opératoires de filtration pour la rétention des pesticides par les membranes d'osmose et nano filtration, l'utilisation de la méthode RNA est nécessaire et est devenue une technique de modélisation importante. Elle est largement utilisée pour les études de QSPR 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 QSPR-RNA est décrite. Elle est basée, comme décrite sur l'organigramme ci-dessous.

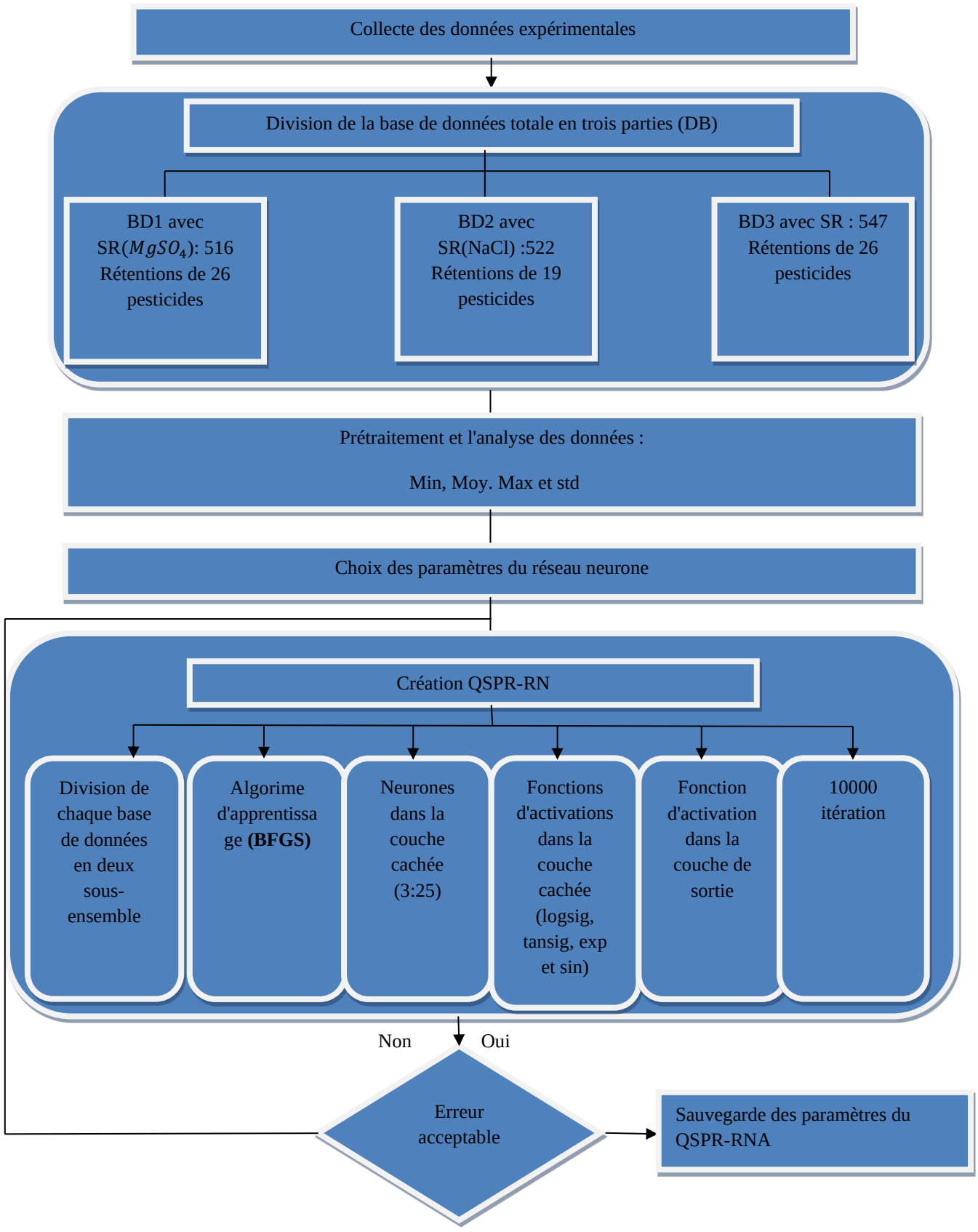


Figure III.2. Organigramme de développement et l'optimisation de l'architecture d'un QSPR-RNA.

L'algorithme d'apprentissage utilisé dans ce travail est le BFGS quasi-Newton (trainbfg). Chaque QSPR-RNA contient trois couches de neurones ou des nœuds : une couche d'entrée avec 13 pour QSPR-RNA1 et QSPR-RNA2 et 14 QSPR-RNA3neurones 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. 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), l'exponentielle et sin » 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-RNA décrite ci-dessus a été réalisée en utilisant STATISTICA (Annexe B). Ce logiciel est utilisé pour optimiser l'architecture des neuf modèles neuronaux QSAR-RNA employés pour la rétention des pesticides par les membranes d'osmose et nano filtration.

III.4 Performance des modèles neuronaux

Les résultats de performances de neuf modèles neuronaux "QSPR-RNA" en termes du coefficient de corrélation (R) et l'erreur quadratique moyenne (EQM) sont mentionnés dans le tableau suivant pour chaque division de base de données (division 01 "80% pour la base d'apprentissage et 20% pour la base de test", division 02 "70% pour la base d'apprentissage et 30% pour la base de test" et division 03 "60% pour la base d'apprentissage et 40% pour la base de test"). D'après ce tableau III.7, on constate que le septième modèle (QSPR-RNA07) est un modèle optimal ($QSPR\text{-}RNA_{optimal}$), il donne des meilleurs résultats par apport des autres modèles (R = 0.9627et MSE = 16.4937% pour la phase de test).

Tableau III.7. Comparaison entres les modèles neuronaux (QSPR-RNA).

		Pourcentage	**R**	**EQM (%)**
QSPR-RNA01	**Phase d'apprentissage : 413 points**	80%	0.9568	14.9540
	Phase de test : 103 points	20%	0.9217	29.9482
	Phase totale : 516 points	-	0.9487	-
QSPR-RNA02	**Phase d'apprentissage : 362 points**	70%	0.9348	22.5701
	Phase de test : 154 points	30%	0.9262	26.3163
	Phase totale : 516 points	-	0.9317	-
QSPR-RNA03	**Phase d'apprentissage : 310 points**	60%	0.9257	24.7331
	Phase de test : 206 points	40%	0.8639	52.5021
	Phase totale : 516 points	-	0.8972	-
QSPR-RNA04	**Phase d'apprentissage : 418 points**	80%	0.9574	15.0274
	Phase de test : 104 points	20%	0.9032	24.8090
	Phase totale : 522 points	-	0.9487	-
QSPR-RNA05	**Phase d'apprentissage : 366 points**	70%	0.9613	14.5083
	Phase de test : 156 points	30%	0.9055	22.4967
	Phase totale : 522 points	-	0.9492	-
QSPR-RNA06	**Phase d'apprentissage : 314 points**	60%	0.9328	24.4048
	Phase de test : 208 points	40%	0.9056	26.2334
	Phase totale : 522 points	-	0.9236	-
QSPR-RNA07	**Phase d'apprentissage : 460 points**	80%	0.9696	11.62995
	Phase de test : 114 points	20%	0.9627	16.4937
	Phase totale : 574 points	-	0.9680	-
QSPR-RNA08	**Phase d'apprentissage : 402 points**	70%	0.9747	9.4623
	Phase de test : 172 points	30%	0.9492	22.8343
	Phase totale : 574 points	-	0.9654	-
QSPR-RNA09	**Phase d'apprentissage : points**	60%	0.9295	24.2605
	Phase de test : 229 points	40%	0.8747	55.1151
	Phase totale : 574 points	-	0.9039	-

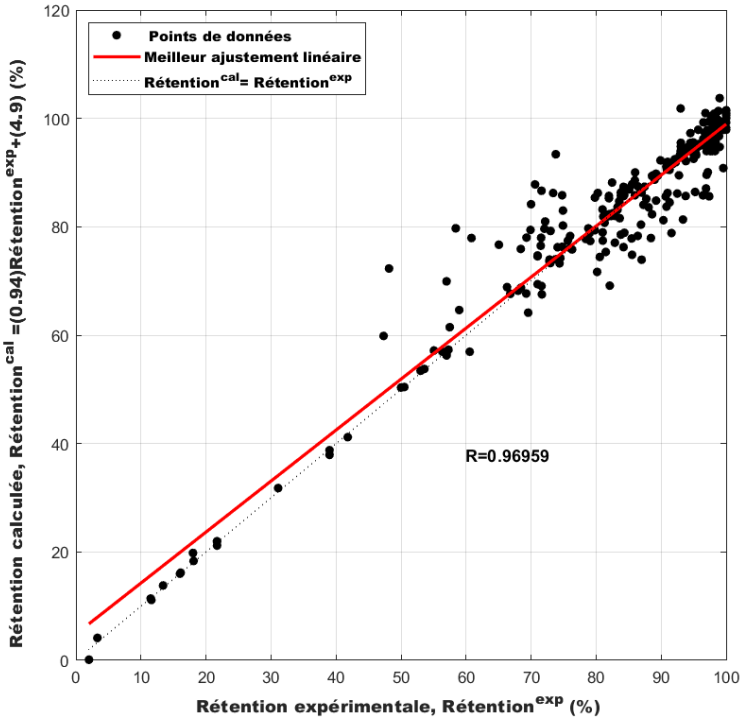
Le tableau III.8 montre la structure de modèle QSPR-RNA$_{optimal}$ optimisé.

Tableau III.8 Structure de modèle QSPR-RNA$_{optimal}$ optimisé.

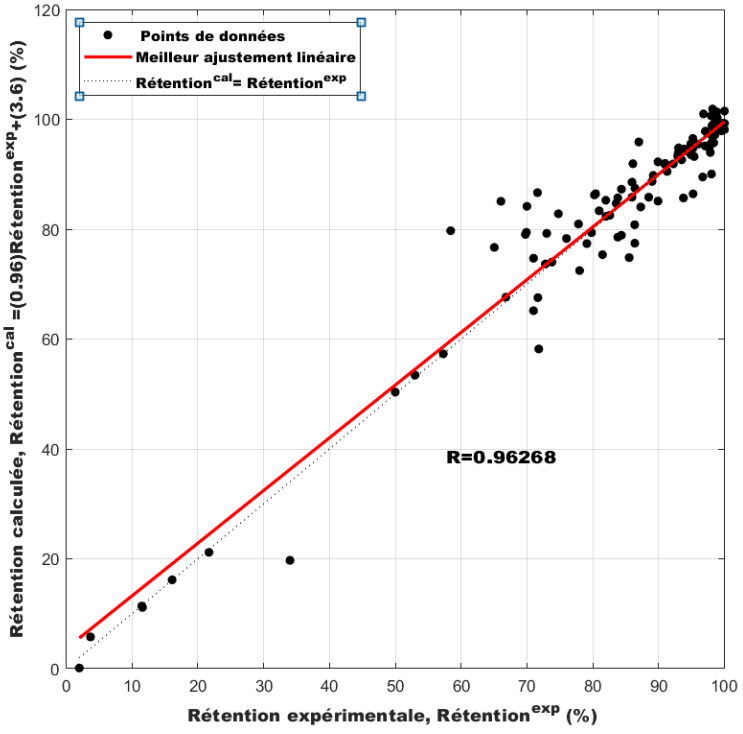
Algorithme d'apprentissage	**Couche entrée**	**Couche cachée**		**Couche sortie**	
	Nombre de neurone	**Nombre de neurone**	**Fonction d'activation**	**Nombre de neurone**	**Fonction d'activation**
BFGS quasi-Newton (trainbfg)	14	19	Exponentiel	1	Identité

Selon la discussion précédente, le modèle neuronal optimal (QSPR-RNA$_{optimal}$) a été développé dans le but de prédire la rétention des pesticides par les membranes d'osmose et nano filtration. L'intrigue et les paramètres de la régression linéaire sont franchement obtenus en utilisant la fonction MATLAB ***"postreg"***. Une valeur de coefficient de corrélation supérieure à 0.5 est généralement considérée comme satisfaisante, une valeur supérieure à 0.9 est excellente (Chabane et Koul2022).

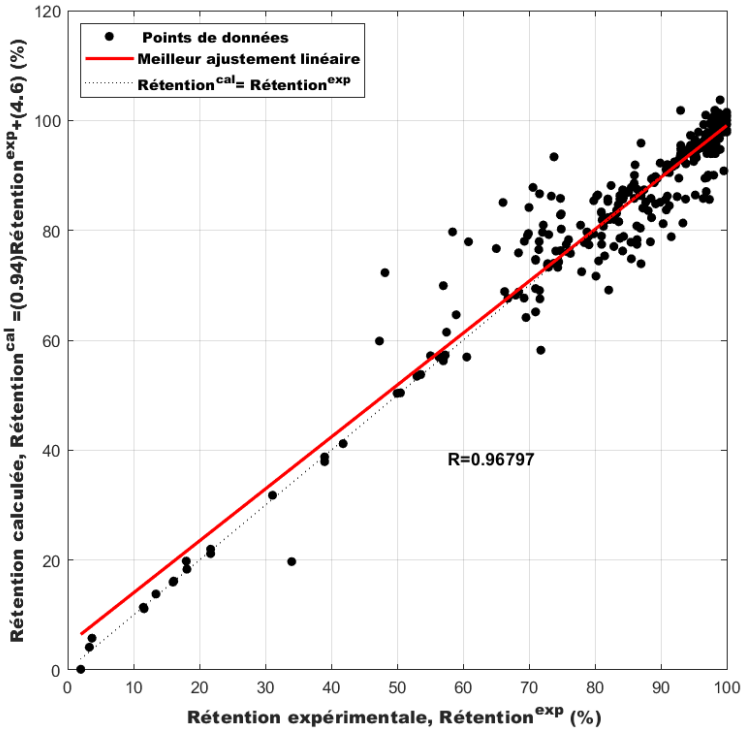
La figure III.3 montre les courbes de régression linéaire des taux de rétentions calculés par le modèle neuronal "QSPR-RNA$_{optimal}$" optimisé avec les taux de rétentions expérimentales avec des vecteurs de régression approchant de l'idéal ([α, β, R] = [0.9408, 4.8526, 0.9696] pour la phase d'apprentissage ; [α, β, R] = [0.9594, 3.6480, 0.9627] pour phase de test et [α, β, R] = [0.9450, 4.5769, 0.9680] pour 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, ce qui montre une bonne robustesse du modèle neuronal "QSAR-RNA$_{optimal}$" établi et la possibilité de la prédiction des différents paramètres qui caractérisent la rétention des pesticides lors d'osmose et nano filtration.


Points de données
Meilleur ajustement linéaire
Rétention^cal= Rétention^exp
R=0.96959
Rétention calculée, Rétention^cal =(0.94)Rétention^exp+(4.9) (%)
Rétention expérimentale, Rétention^exp (%)

(a)


Points de données
Meilleur ajustement linéaire
Rétention^cal= Rétention^exp
R=0.96268
Rétention calculée, Rétention^cal =(0.96)Rétention^exp+(3.6) (%)
Rétention expérimentale, Rétention^exp (%)

(b)



(c)

Figure III.3. Comparaison entre les valeurs expérimentales et les valeurs calculées pour QSPR-RNA$_{optimal}$: (a) la phase d'apprentissage, (b) la phase de test, (c) la phase totale.

La performance de modèle neuronal "QSPR-RNA$_{optimal}$" optimisé a été évaluée en termes de coefficient de corrélation (R) et la racine de l'erreur quadratique moyenne (REQM), l'erreur absolue moyenne (EAM) et l'erreur de standard de prédiction (ESP). REQM, EAM et ESP ont été utilisées pour déterminer l'erreur de modélisation entre les valeurs expérimentales et calculées. Les erreurs sont définies comme suit (Faisal et al. 2015 ; Kolluri et al. 2015).

$$EAM = \frac{1}{n}\sum_{i=1}^{n}\left|(y_{i,exp} - y_{i,cal})\right| \qquad (III.3)$$

$$REQM = \sqrt{\frac{\sum_{i=1}^{n}(Y_{i,exp} - Y_{i,cal})^2}{n}} \qquad (III.4)$$

$$ESP(\%) = \frac{REQM}{Y_e} * 100 \qquad (III.5)$$

Où n est le nombre total de données,$y_{i,exp}$est la valeurexpérimentale,$y_{i,cal}$représente la valeurcalculée à partir du modèle QSAR et $Y_e$estla valeur moyennedes données expérimentales.

Les vecteurs de régression linéaire pour le modèle neuronal "QSPR-RNA$_{optimal}$" optimiséet les erreurs (REQM, EAM, ESP) sont présentés dans le tableau III.9 et la figure III.4. Il est prouvé que l'approche neuronale proposée donne des résultats excellents 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 QSPR-RNA$_{optimal}$.

Table III.9 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 l'erreur quadratique moyenne ***(REQM***), l'erreur absolue moyenne ***(EAM)*** et l'erreur de standard de prédiction ***(ESP)***.

	α	**B**	**R**	**EAM**	**REQM**	**ESP**
Phase d'apprentissage	0.9408	4.8526	0.9696	2.9741	4.8229	5.8788
Phase de test	0.9594	3.6480	0.9627	3.7124	5.7435	6.9614
Totale	0.9450	4.5769	0.9680	3.1208	5.0191	6.1112

La figure III.4 représente la performance de modèle neuronal optimal. Elle montre que le modèle RNA$_{optimal}$ a donné des erreurs plus faibles. Ce qui indique que le modèle RNA$_{optimal}$ possède une capacité de modélisation très élevée.

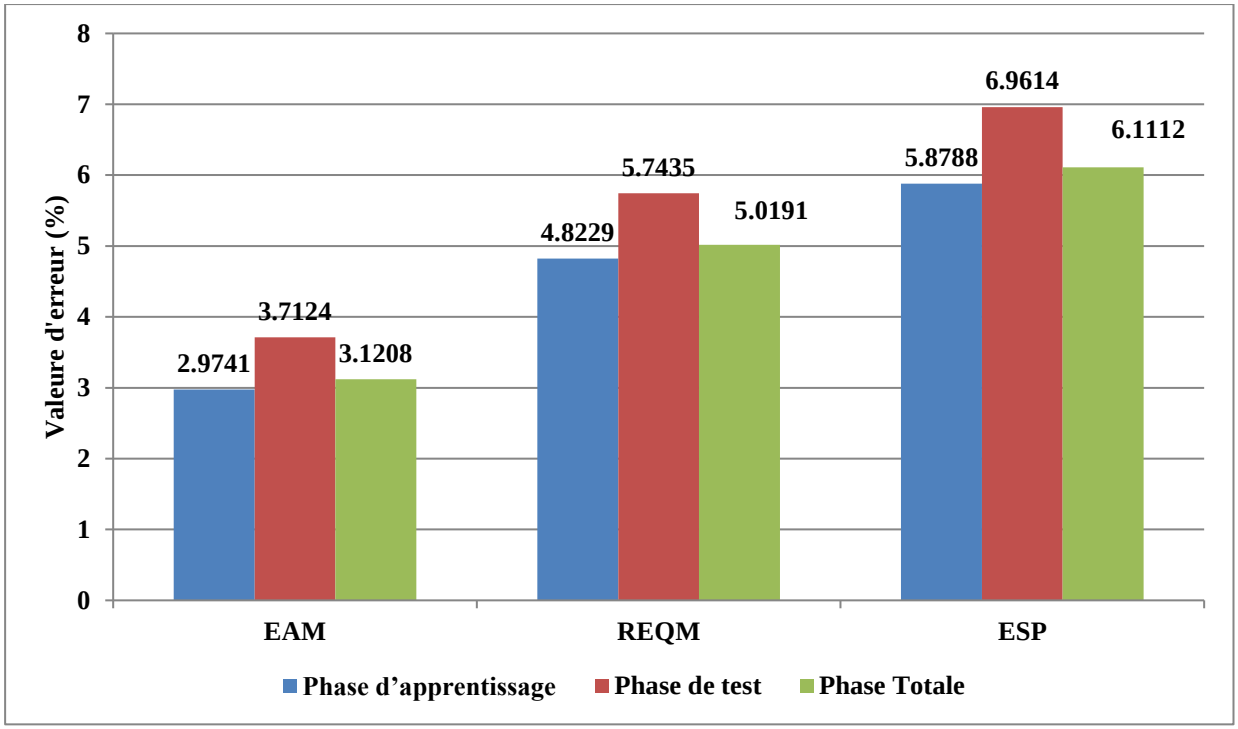


Figure III.4. Performances de modèle neuronal (*QSAR-RNA$_{optimal}$*) pour la phase d'apprentissage, la phase de test et la phase totale.

III.5 Expressions mathématiques des modèles neuronaux optimisés

À partir du modèle neuronal "QSPR-RNA$_{optimal}$" optimisé, représenté sur la figure III.5, on peut exprimer l'assimilation de la rétention des pesticides par les membranes d'OI/NF par un modèle mathématique y incorporant toutes les entrées Ei (de, LogD, moment dipolaire, Length, EqWidth, MWCO, M_W(SR), SR, angle de contact, potentiel zêta, pH, pression, taux de récupération et température) comme suit : Les sorties Zj de la couche cachée :

$$Z_j = f_H\left[\sum_{i=1}^{14} w_{ji}^I E_i + b_j^H\right] = exp^{-(\sum_{i=1}^{14} w_{ji}^I E_i + b_j^H)} \qquad (III.6)$$

j = 1,2,….19

La sortie "rétention des pesticides" :

$$Rétention\left(QSPR - RNA_{optimal}\right) = f_o\left[\sum_{j=1}^{19} w_{1j}^H Z_j + b_1^o\right] = \sum_{j=1}^{19} w_{1j}^H Z_j + b_1^0 \qquad (III.7)$$

La combinaison des équations (III.6) et (III.7) mène à la formule mathématique suivante qui décrit l'assimilation de rétention des pesticides lors de nano filtration et osmose en prenant en compte toutes

les entrées Ei(de, LogD, moment dipolaire, Length, EqWidth, MWCO, M_W(SR), SR, angle de contact, potentiel zêta, pH, pression, taux de récupération et température) :

$$Rétention\ (QSPR - RNA_{optimal}) = \sum_{j=1}^{19} w_{1j}^{H}\ exp^{-(\sum_{i=1}^{14} w_{ji}^{I} E_i + b_j^H)} + b_1^0 \qquad \text{(III.8)}$$

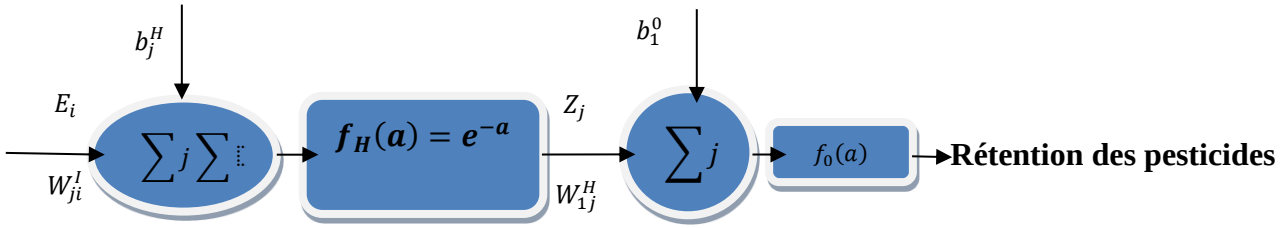


Figure III.5 Représentation schématique du modèle neuronal "QSPR-$RNA_{optimal}$" optimisé.

Ou W_{ji}^Ireprésente le poids synaptique de la connexion du neurone i de la couche d'entrée au neurone j de la couche cachée (19 lignes x 14 colonnes);W_{1j}^Hreprésente le poids synaptique de la connexion du neurone j de la couche cachée au neurone de la couche de sortie (19 lignes x 1 colonne); b_j^Hreprésente les biais du neurone j de la couche cachée (19 lignes x 1 colonne) et b_1^0représente les biais du neurone j de la couche de sortie (1 ligne). Le matrice des poids et biais de modèle neuronal "QSPR-$RN_{optimal}$" optimisé est donné dans le tableau III.10.

Tableau III.10. Poids et biais du QSPR-RNA$_{optimal}$.

Connexions														**Connexions Couche cachée-Couche de sortie**	
Couche d'entrée-Couche cachée															
Poids													**Biais**	**Poids**	**Biais**
$w_{j,1}^{l}$	$w_{j,2}^{l}$	$w_{j,3}^{l}$	$w_{j,4}^{l}$	$w_{j,5}^{l}$	$w_{j,6}^{l}$	$w_{j,7}^{l}$	$w_{j,8}^{l}$	$w_{j,9}^{l}$	$w_{j,10}^{l}$	$w_{j,11}^{l}$	$w_{j,12}^{l}$	$w_{j,13}^{l}$	$w_{j,14}^{l}$	b_j^H	$w_{1,j}^{H}$	b_1^Q
0,06159	1,82123	-2,00199	0,16463	0,16163	0,31159	-1,45076	-1,16419	1,19968	0,06061	-1,85762	1,31823	1,19646	-0,50368	-0,23531	-0,07706	
-1,46438	-2,07788	-1,15853	0,14835	-0,06386	-0,50774	1,26771	0,97671	-2,09667	-2,26886	2,47349	-2,15931	-0,17583	-0,3526	2,27126	0,96381	
1,98543	2,61385	0,05357	-0,68519	-0,84315	0,1887	2,78602	-3,84848	-0,04784	-0,28803	0,37582	4,04296	-1,77785	-0,10665	0,19392	0,1748	
0,13088	-1,93841	0,50836	0,70739	0,86716	-5,62429	0,89929	1,53737	-0,01889	0,00332	-1,22713	1,37484	3,81666	-0,14346	-0,37532	0,4655	0,05164
-0,32619	2,59047	-0,20862	1,42651	1,54748	7,5151	-0,00581	-1,19125	-0,41202	0,62045	1,56542	-0,01401	-1,74305	-0,5078	0,08374	-0,11412	
-2,66829	-1,52139	-1,36688	-0,93645	0,93398	2,95423	3,66273	1,47822	-3,31631	0,94949	0,46373	-0,04716	1,13295	0,54731	-0,06727	0,31229	
0,85118	-2,14302	0,20486	0,67055	0,15051	0,72911	-1,32228	2,07439	0,28347	-9,84749	-1,72769	1,88985	-6,00522	0,78144	-0,68293	-0,4978	
0,80539	1,61893	-0,15067	-0,48919	-0,31508	-3,13348	1,48158	-3,33821	0,62571	2,05299	1,88053	-3,85717	0,33095	-0,10216	1,47309	-0,27939	
-1,06206	-5,01662	-0,14842	-1,68134	-0,40604	0,83444	-3,59718	3,08559	-0,53021	0,38888	1,50749	3,08691	-0,76587	-0,24743	-1,53506	-0,13064	
0,70713	1,41877	-2,24878	0,14906	-0,4633	-3,02538	-1,30792	1,24979	0,0975	2,43296	0,68112	1,73042	-1,41464	1,83057	1,9077	-0,24039	
2,25271	0,45384	-1,83471	-1,34787	-0,98259	-0,28749	0,37604	0,1987	-0,578	3,46823	-0,29272	4,07685	0,21376	-2,27858	2,60809	-1,50637	
1,03135	1,41097	0,81200	-0,5828	-1,16417	-1,67744	-1,46161	-0,79448	-1,00525	-5,14285	2,40992	-5,93253	0,66113	0,78877	-0,10681	1,32124	
0,65329	-2,01412	-0,24239	0,01383	2,02554	2,29526	-0,65812	1,89513	1,14298	1,788	-0,45128	-2,29067	3,14834	-0,37061	-0,30956	0,11416	
-2,38551	0,60301	-1,34318	-1,5769	0,3159	1,38509	-0,69178	-0,46812	-0,69618	1,55069	2,32822	2,22806	-0,40951	-0,29252	-0,17635	-0,11299	
-5,03905	0,54484	0,84944	-0,29202	-1,34782	-1,34542	-1,47032	2,46091	-0,39923	-1,07529	-0,55787	-7,14679	-0,81394	-1,7314	-3,39704	0,13714	
3,98929	-1,00202	1,14308	1,47909	0,32455	2,10845	0,60077	-2,57459	0,33898	-0,36838	3,11974	-0,0812	-1,2341	1,1214	1,72445	-0,056	
1,28208	-2,31863	0,00998	0,14893	-2,0039	-1,77745	-1,79714	3,51158	0,18973	3,86295	-1,25728	1,84675	-0,26023	-0,21757	0,89825	0,03434	
0,15481	-0,21757	0,04789	0,10028	-1,43257	1,42774	0,04891	-6,62044	-0,53577	-3,33011	-3,27033	1,46853	-0,07163	0,69524	1,75612	-0,0553	
-0,03568	1,51437	0,57645	0,59548	2,52401	-1,09395	-0,56601	-4,47659	-1,37185	1,55637	2,11687	-0,88427	0,14716	-0,30392	-0,25537	0,08468	

L'équations mathématiques, dérivant le modèle neuronal "QSPR-RNA$_{optimal}$" optimisé, contient juste le degré exigé de complexité et incluent dans leurs formules les caractéristiques pertinentes de la rétention des pesticides par les membranes d'osmose et nanofiltration à savoir les propriétés des pesticides, 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 pesticides par les membranes d'osmose et nano filtration.

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

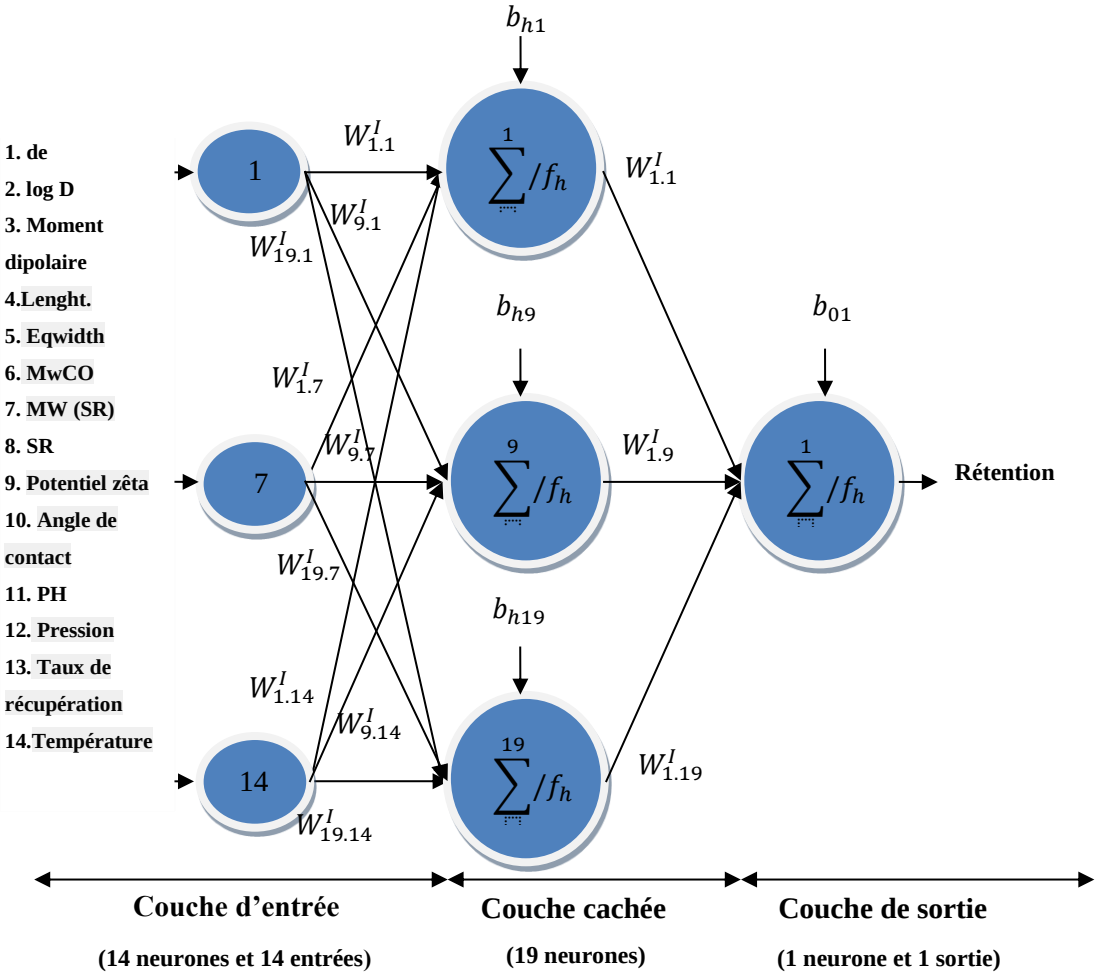


Figure III.6 Architecture finale du réseau de neurone optimal de type feedforward pour la prédiction de la rétention des pesticides par les membranes d'osmose et nano filtration.

III.6 Analyse de sensibilité par la méthode de "poids"

La contribution des variables d'entrées (de, *LogD*, moment dipolaire**,** *Lenght*, *EqWidth*, MWCO, M_W(SR), SR, angle de contact, potentiels zêta, pH, pression, taux de récupération et température) sur la sortie a été déterminée par une analyse de sensibilité en utilisant la méthode de "poids" pour chaque réseau de neurone QSPR-RNA$_{optimal}$.

Cette méthode, proposée d'abord par Garson(Garson 1991)et reprise par Goh (Goh 1995), donne une quantification de l'importance relative (*IR*) des différentes entrées sur la sortie du RNA(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_{ji}^{H}).

Le processus de calcul, compose de quatre étapes, il est mentionnée dans les équations suivantes (Chabane et Koul 2022):

$$P_{ji} = \left|W_{ji}^{I} W_{ji}^{H}\right| \quad \text{(III.9)}$$

$$Q_{ji} = \frac{P_{ji}}{\sum_{i=1}^{n_i} P_{ji}} \quad \text{(III.10)}$$

$$S_i = \sum_{j=1}^{n_J} Q_{ji} \quad \text{(III.11)}$$

$$IR = 100\frac{S_i}{\sum_{i=1}^{n_i} S_i} \quad \text{(III.12)}$$

Pour voir la contribution ainsi que le profil de variation de chaque variable d'entrée (de, LogD, Dipole moment, Lenght, EqWidth, MWCO, M_W(SR), SR, angle de contact, potentiels zêta, pH, pression, taux de récupération et température) sur la sortie de modèle neuronal "QSPR-RNA$_{optimal}$" optimisé, On a procédé à une analyse de sensibilité par méthode de "poids".

Nous avons appliqué cette méthode à modèle neuronal "QSPR-RNA$_{optimal}$" optimisé en utilisant leurs poids respectifs illustrés dans les tableaux III.10. Les résultats de contributions sont montrés sur la figure III.7. Les variables les plus pertinentes qui peuvent influer sur la rétention des pesticides par les membranes d'osmose et nano filtration sont : pression, SR, MWCO, angle de contact, logD, pH, de, Mw(SR), taux de récupération, Eqwidth, moment dipolaire, potentiel zeta, length, température.

Toutes les entrées exercent des contributions importantes (IR>5%) sur la rétention les pesticides par les membranes d'osmose et nanofiltration et les entrées (Moment dipolaire, Eqwidth, length, Zeta potentiel et température) exercent des contributions inférieures à 5%.

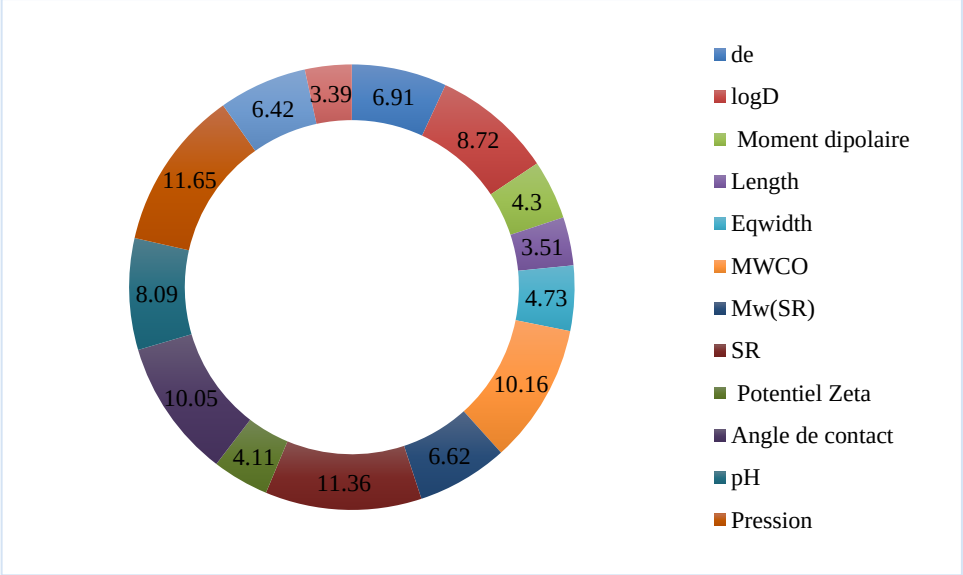


Figure III.7 Présentation des pourcentages de l'importance relative (*IR*) de chaque variable d'entrée sur la rétention des pesticides lors d'osmose et nano filtration.

III.7. Diagnostic des points aberrants et domaine d'applicabilité du modèle neuronal optimisé

Les diagrammes de dispersions des modèles optimisés ont montré un certain nombre des points éloignés de la première bissectrice que nous avons attribué à l'existence probable de points aberrants. Ceci nous a motivé à réaliser un diagnostic des données potentiellement suspectes selon la méthode décrite dans les deux références suivantes (Hosseinzadeh and Hemmati-Sarapardeh 2014) ; (Baghban, Mohammadi et al. 2017).

Les valeurs aberrantes sont bien connues en tant que points de données distincts de la tendance générale des principaux points de données. La détection des valeurs aberrantes, qui joue un rôle important dans la proposition de modèles, est utilisée pour reconnaître des données individuelles (ou des groupes de données) pouvant différer de la masse des données constituant un ensemble de données. À cette fin, il est essentiel de trouver des méthodes rigoureuses pour détecter les valeurs potentiellement suspectes afin d'éliminer des données expérimentales inexactes et d'améliorer la précision des modèles (Hosseinzadeh and Hemmati-Sarapardeh 2014) ; (Baghban, Mohammadi et al. 2017). Dans cette étude, la technique mathématique "Leverage" a été utilisée pour rechercher les valeurs aberrantes. Dans cette technique, les valeurs résiduelles sont d'abord calculées. Ensuite une matrice H de points de données d'entrée est créée sur la base de la formule suivante :

$$H = X(X^t X)^{-1} X^t \qquad \text{(III.13)}$$

Où X fait référence à la matrice m × n (m et n représentent respectivement le nombre d'échantillons et les paramètres (variables d'entrée) du modèle). Les valeurs de Hat sont obtenues à partir de la diagonale principale de la matrice de H.

$$Hat = diagonale(H) \quad \text{(III.14)}$$

Le diagramme de William qui représente les résidus normalisés en fonction du "Leverage" levier (Hat) peut détecter graphiquement les valeurs aberrantes. Dans ce graphique, le levier critique (H*) (seuil) est généralement fixé à la valeur donnée par l'équation suivante :

$$H^* = \frac{3(n+1)}{m} \quad \text{(III.15)}$$

Ou m et n représentent respectivement le nombre d'échantillons et les paramètres (variables d'entrée) du modèle.

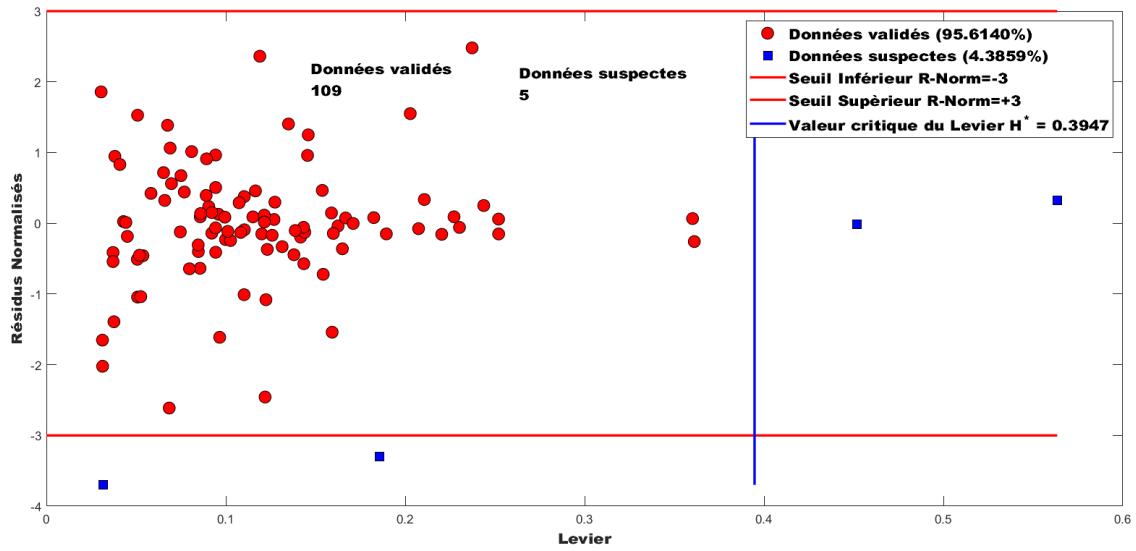
Les résidus normalisés sont calculés à partir des données expérimentales et celle calculée par le modèle.

$$(R_Norm)_i = \frac{\left(y_i^{exp} - y_i^{cal}\right)}{\sqrt{var\left(y^{exp} - y^{cal}\right)}} \quad i = 1, \dots m \quad \text{(III.16)}$$

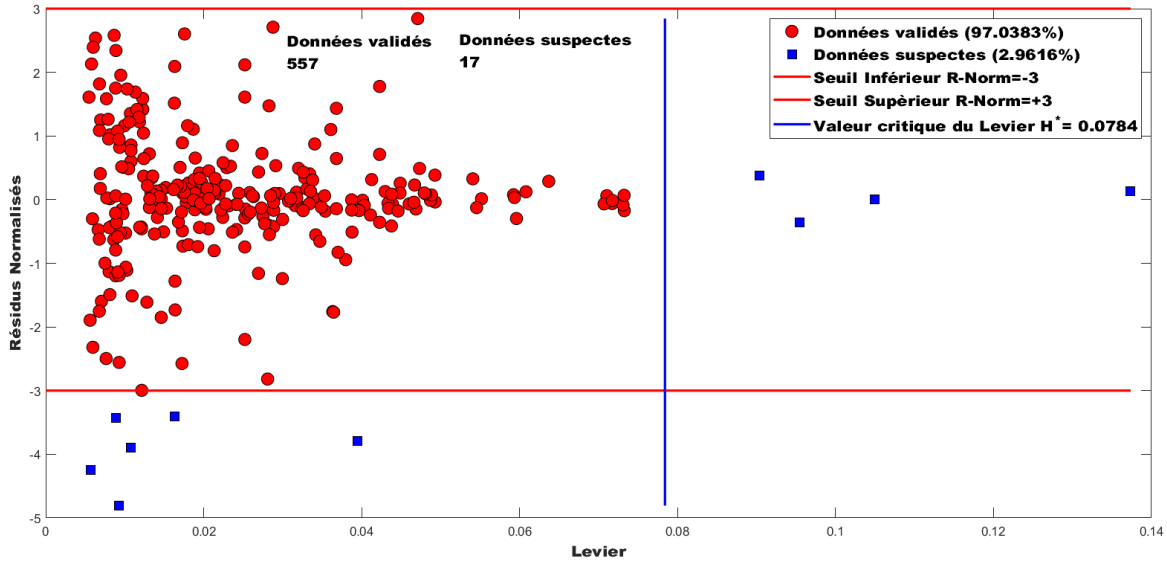
Le résidu normalisé égal à trois est normalement considéré comme une valeur "seuil" pour accepter les points de données dans les écarts types compris sur une plage de ± 3 par rapport à la moyenne (pour couvrir 99% des données normalement distribuées). Si la majorité des points de données se situent dans les plages de 0 ≤ Hat ≤ H* et de − 3 ≤ R_Norm ≤ 3, cela indique que le développement du modèle et ses prédictions sont effectués dans le domaine d'applicabilité, ce qui conduit à un modèle statistiquement valide. Ainsi on peut affirmer qu'il existe des points ‘’Good High Leverage’’ dans le domaine de 0 ≤ Hat ≤ H* et − 3 ≤ R_Norm ≤3. Les points situés dans la plage de R_Norm<−3 ou R_Norm>3 (qu'ils soient plus grands ou plus petits que la valeur H*) sont désignés comme points aberrants du modèle ou comme points "Bad High Leverage".

La figure III.8 Montre le diagramme de William pour le domaine d'applicabilité pour le modèle neuronal optimal "QSPR-RNA$_{optimal}$". Dans ce diagramme l'existence de **109/114 (95.6140%)** de points de données valides et **5/114 (4.3859%)** de points de données suspectes pour la phase test et **557/574 (97.0383%)** de points de données valides et **17/574(2.9616%)** de points de données suspectes pour la phase totale. La valeur critique de levier est $H^* = \frac{3(n+1)}{m} = \frac{3(14+1)}{114} =$

0.3947 pour la phase de test et $H^* = \frac{3(n+1)}{m} = \frac{3(14+1)}{574} = 0.0784$ pour la phase totale. Cela indique que le développement du modèle QSPR-RNA$_{optimal}$ et ses prédictions sont effectués dans le domaine d'applicabilité. Ce qui conduit à un modèle neural optimal statistiquement valide. Ainsi on peut affirmer qu'il existe des points "Good High Leverage" pour la phase validation.



(a)



(b)

Figure III.8 Diagnostic des données potentiellement suspectes et domaine d'applicabilité du Modèle neuronal optimal : (a) pour la phase test et (b) pour la phase totale.

III.8. Conclusion

Ce présent chapitre a donné une raison de plus pour l'utilisation généralisée des modèles neuronaux (QSPR-RNA) pour la modélisation de la rétention des pesticides lors d'osmose et nanofiltration. Le but de la présente étude était de développer un modèle RNA de type feed-forward.

Le modèle QSPR-$RNA_{optimal}$ optimisé développé dans ce travail a donné des erreurs plus faibles.

- Les erreurs absolues moyennes (EAM) étaient 3.1208% pour phase totale, 2.9741% pour la phase apprentissage et 3.7124% pour phase test.
- Les racines des erreurs quadratiques moyennes (REQM) étaient 5.0191 % pour phase totale, 4.8229% pour phase apprentissage et 5.7435% pour la phase test.
- Les erreurs standards de prédictions (ESP) étaient 6.1112% pour phase totale, 5.8788 % pour phase apprentissage et 6.9614% pour phase test.

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 pesticides sont : pression, SR, MWCO, angle de contact, logD, pH, de, Mw(SR), taux de récupération, Eqwidth, moment dipolaire, potentiel zeta, length, température. Toutes les entrées exercent des contributions plus importantes ($IR > 5\%$) sur la rétention des pesticides lors d'osmose et nanofiltration.

Analyse diagnostic des points aberrants et domaine d'applicabilité du modèle neuronal optimisé a montré que le développement du modèle QSPR-$RN_{optimal}$ et ses prédictions sont effectués dans le domaine d'applicabilité. Ce qui conduit à un modèle neural optimal statistiquement valide. Ainsi on peut affirmer qu'il existe des points "Good High Leverage" pour la phase de test et totale.

Conclusion générale et perspectives

Les réseaux de neurones possèdent une propriété notable qui ont prouvé leur contribution à la modélisation de processus non linéaires. Les réseaux de neurones sont des approximateurs généraux parcimonieux. Enfin, un modèle de réseau neuronal artificiel (RNA) a été appliqué pour prédire la séparation des micropolluants par des membranes d'osmose et nano filtration.

D'autres chercheurs ont abordé ce problème de manière réductible, notamment en termes de taille de la base de données et les paramètres calculés ont donné des résultats prometteurs.

Le but de ce mémoire a porté sur l'utilisation des réseaux de neurones pour la modélisation du processus de la rétention des pesticides par les membranes d'osmose de nano filtration. Un modèle neuronal optimal de type feedforward (QSPR-$RNA_{optimal}$) caractérisé par une structure (14 neurones dans la couche d'entrée, 19 neurones dans la couche cachée et un neurone dans la couche de sortie). Le modèle QSPR-$RNA_{optimal}$ développé dans ce travail a donné des erreurs faibles. Les erreurs absolues moyennes étaient 3.1208% pour phase totale, 2.9741% pour la phase apprentissage et 3.7124% pour phase test. Les racines des erreurs quadratiques moyennes étaient 5.0191% pour phase totale, 4.8229% pour phase apprentissage et 5.7435% pour la phase test. Les erreurs standards de prédictions étaient 6.1112% pour phase totale, 5.8788% pour phase apprentissage et 6.9614% pour phase test. Ce qui montre une bonne robustesse du modèle neuronal "QSPR-$RNA_{optimal}$" établi et la possibilité de la prédiction des différents paramètres qui caractérisent la rétention des pesticides lors d'osmose et nano filtration.

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 pesticides sont : pression, SR, MWCO, angle de contact, logD, pH, de, Mw(SR), taux de récupération, Eqwidth, moment dipolaire, potentiel zeta, length, température. Toutes les entrées exercent des contributions plus importantes (IR>5%) sur la rétention des pesticides lors d'osmose et nano filtration.

L'analyse diagnostique des valeurs aberrantes et de la couverture du modèle neuronal optimisé montre que le développement du modèle optimal QSPR-RNA et ses prédictions s'effectuent dans le cadre de la couverture. Il en résulte un modèle neuronal optimal statistiquement valide. Par conséquent, nous pouvons dire qu'il existe un « bon effet de levier élevé » dans la phase de test.

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

- Les réseaux de neurones, comme d'autres techniques statistiques, sont des outils statistiques qui nous permettent d'adapter une fonction non linéaire très générale à un ensemble de points.
- 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 et nanofiltration.

- La modélisation des réseaux de neurones est une méthode basée uniquement sur des données expérimentales et/ou semi-expérimentales (données issues de modèles analytiques, empiriques, numériques, etc). Par conséquent, la qualité et la quantité des données déterminent 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 extraites des modèles neuronaux ont des applications intéressantes dans la prédiction, l'optimisation, voire le contrôle des processus d'osmose et nano filtration.

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

- Cette perspective concerne spécifiquement la modélisation de la rétention des pesticides par les membranes d'osmose et nano filtration, mais en utilisant d'autres mesures de porosité ($MgCl_2$, KCl, Na_2SO_4, …).
- Cette perspective est particulièrement pertinente pour modéliser d'autres procédés (microfiltration, ultrafiltration) à des fins prédictives.
- Prédiction de la rétention des anticancers par nanofiltration à l'aide d'empreintes moléculaires.
- Nous pouvons aussi étendre l'idée de ce travail à d'autres méthodes de l'intelligence artificielle (Support vecteur machine, Réseaux de neurones multiples,…).

Références Bibliographique

Achilli A., Cath T. Y.et Childress A. E (2010). "Selection of inorganic-based draw solutions for forward osmosis applications". Journal of membrane science, 364(1-2), 233-241.

Achiou B., Elomari H., Ouammou M., Albizane A., Bennazha J., Younssi A. et AaddanS.A (2016)."Elaboration and characterization of flat ceramic microfiltration membrane made from natural Moroccan pozzolan" J. Mater. Environ. Sci. 7 (1) 196-204.

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.

Ahoulé (2016)."Performances comparatives des techniques de nanofiltration et d'osmose inverse pour le traitement d'eau de consommation contaminée à l'arsenic au Burkina Faso". Thèse de doctorat. Université de Montpellier. Montpellier. France.

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

Amirilargani M., Saljoughi E., Mohammadi T.et Moghbeli M.R. (2010). "Effects of coagulation bath temperature and polyvinyl pyrrolidone content on flat sheet asymmetric polyethersulfone membranes"Polym. Eng. Sci. 50, 885.

Ammar M. Y. (2007). "Mise en œuvre de réseaux de neurones pour la modélisation de cinétiques réactionnelles en vuede la transposition batch/continu". Thèse de Doctorat, Institut National Polytechnique de Toulouse. Toulouse France.

Ammi Y. (2016). "Modélisation de phénomènes de transport lors de l'osmose Inverse". Thèse de doctorat. Université deMédéa. Médéa. Algérie.

Ammi Y., Khaouane L. et Hanini S (2020). "A comparison of "neural networks and multiple linear regressions"models to describe the rejection of micropollutants by membranes". Kemija u industriji: 69 (3-4) 111-127.

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.

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

Aptel P., Moulin P. et Quemeneur F (2002)."Micro et Ultrafiltration : conduite des essais pilotes, traitements des eaux et effluents". Club Français des membranes, Lavoisier, Paris, 102.

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.

Bagheripour E ., Moghadassi A.R ., Hosseini S. M ., Van der Bruggen B. et Parvizian F (2018). **"**Novel composite graphene oxide/chitosan nanoplates incorporated into PES based nanofiltration membrane: Chromium removal and antifouling enhancement" , Journal of Industrial and Engineering Chemistry, 62 (2018) 311- 320.

Barbar R (2006)."Contribution à la mise au point de nouveaux matériaux membranaires pour la filtration de milieux complexes : Etude physico-chimique de membranes d'acétate de cellulose et de méthyl cellulose partiellement alkylée obtenues par immersion-précipitation". Thèse doctorat. Institut National Polytechnique de Lorraine de Nancy.Nancy France.

Bassyouni M., Abdel-Aziz M. H., Zoromba M. S., Abdel-Hamid S. M. S. et Drioli E. A(2019). "review of polymeric nanocomposite membranes for water purification", Journal of Industrial and Engineering Chemistry, 73 (2019) 19-46.

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.

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

Ben Frarès N., Taha S. et Dorange G. (2005). "Influence of the operating conditions on the elimination of zinc ions by nanofiltration" Journal of Membrane Science. 185(1–3) : 245-253.

Benhabiles O (2018)." Elaboration de membranes organiques photocatalytiques. Application pour la dépollution des eaux usées" .Thése de Doctorat. Ecole Nationale Polytechnique –Alger. Alger.Algéria.

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

Berland J.M .et Juery C(2002). "Les procédés membranaires pour le traitement de l'eau, Document technique FNDAE (fonds national pour le développement des adductions d'eau)", nº14.

Bermejo S. et Cabestany J (1999). "On-line gradient learning algorithms for K-nearest neighbor classifiers." Dans lelivre: Foundations and tools for neural modeling." Springer Berlin Heidelberg. 1606:546-555.

Bermejo S. et Cabestany J (2001). "Learning with nearest neighbour classifiers." Neural Processing Letters. 13(2):159-181.

Bessiere Y(2005)."Filtration frontale sur membrane mise en évidence du volume filtré critique pour l'anticipation et le contrôle du colmatage". Thèse de doctorat. Université paul sabatier, Toulouse III. Toulouse. France.

Bhatt P.P. (2004)."Osmotic drug delivery systems for poorly soluble drugs, PharmaVentures Ltd". Université of Oxford, UK. England.

Bishop C. M. (1995). "Neural networks for pattern recognition" Oxford University. Press. New York. USA.

Bonachera F(2011)**.** "Les triplets pharmacophoriquesous Développement et applications". Université de Lille 1.France.

Bouazizi A., Saja S., Achiou B., OuammouM., Calvo J.I.et Aaddane A (2016). "Elaboration and characterization of a new flat ceramic MF membrane made from natural Moroccan bentonite. Application to treatment of industrial wastewater"Applied ClayScience,132–133. 33-40.

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.

Boussahel R ; Baudu M ; Montiel A (2022)."Influence de la matière organique et inorganique de l'eau sur l'élimination des pesticides par nanofiltration " Revue des sciences de l'eau Journal of Water Science. 5 (4) 2002.

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

Carpenter G. A. et Grossberg S. (1988). "The ART of adaptive pattern recognition by a self-organizing neural network". Computer, 21(3), 77-88.

Carpenter, G. A. et Grossberg S (1987). "ART 2: self-organization of stable category recognition codes for analog input patterns". In Intelligent Robots and Computer Vision VI. 848 : 272-280.

Cath T. Y., Childress A. E. et Elimelech M (2006). "Forward osmosis: Principles, applications, and recent developments". Journal of membrane science, 281(1-2), 70-87.

Chabane et Koul (2022). "La prédiction de la rétention des pesticides lors de nanofiltration et d'osmose inverse par les modèles QSPR-RNA"Mémoire de Master.Université YAHIA FARES DE MEDEA.

Charcosset, C (2012)."1 - Principles on membrane and membrane processes" in Membrane Processes in Biotechnologies and Pharmaceutics". Amsterdam, Elsevier, 1–41.

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: 167-174.

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

Chuang W.Y., Young T. H., Chiu W.Y. et Lin C.Y(2000)."The effect of polymeric additives on the structure and permeability of poly(vinyl alcohol) asymmetric membranes". Polymer, 41, 5633-5641.

Cramer R. D., Patterson D. E.et Bunce J. D(1988). "Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins" Journal of the American Chemical Society, 110 (18) 5959–5967.

Cui Z. F., Jiang Y. et Field R. W (2010). "Fundamentals of pressure-driven membrane separation processes". In Membrane technology . Butterworth-Heinemann.

Dames M (2005). "Gestion de procèdes discontinus : Méthodologie de modélisation et d'optimisation d'opérations de dispersion liquide-liquide en cuve agitée." Thèse de Doctorat. Institut National Polytechnique de Toulouse. Toulouse. France

Demuth H. et Beale M (2000)." Neural network toolbox for use with Matlab. User's Guide" version 4. USA, The MathWorks Inc.

Desmukh S. S.et Childress A. E. (2001)."Zeta potential of commercial RO membranes: influence ofsource water type and chemistry". Desalination140: 87-95.

Devillers J (2013). "Methods for building QSARs". Computational Toxicology: Volume II, 3-27.

Dickson J. M .; Matsuura T ; Blais P.et Sourirajan S. (1975)."Reverse osmosis separations of some organic and inorganic solutes in aqueous solutions using aromatic polyamide membranes".Journal of AppliedPolymer Science 19: 801-819.

Diop S (2010). "Elimination partiel du fluor et de la salinité des eaux souterraines du bassin arachidier par nanofiltration et par adsorption sur argile" . Thèse de Doctorat. Université Cheikh Anta Diop de Dakar, 2010, 205.

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

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 (2008). "Réseaux de neurones: Méthodologie et applications".Eyrolles,Paris,France.

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, 1029-1036.

Fane A., Tang G. et Wang R (2011)."Membrane technology for water: microfiltration, ultrafiltration, nanofiltration, and reverse osmosis". Treatise on Water Science, 301-335.

Fell, C. J. D. (1995). "Reverse Osmosis in: Membrane Separations Technology. Principles and applications. Noble RD, Stern SA New York".

FengC.,Khulbe K.C., Matsuura T., Gopal R., Kaur S., Ramakrishna S. et KhayetM (2008)."Production of drinking water from saline water by air-gap membrane distillation using polyvinylidene fluoride nanofiber membrane". J. Membr.Sci, 3111.

Fleischer R.L., Price P.B. et Walker R.M (1975). "Nuclear tracks in solids, Principles and Applications", University of California Press, Berkeley.

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.

Gillespie L. J., et Beattie J. A (1930). "The thermodynamic treatment of chemical equilibria in systems composed of real gases. I. An approximate equation for the mass action function applied to the existing data on the Haber equilibrium". Physical Review, 36(4), 743.

Gittens G. J., Hitchcock P. A., Sammon, D. C. et Wakley G. E (1970). "The structure of cellulose acetate membranes for reverse osmosis. Part I. Membranes prepared from a dioxan based dope". Desalination, 8(3), 369-391.

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

Gopal R., Kaur S., Ma Z., Chan C., Ramakrishna S. et Matsuura T (2006). "Electrospun nanofibrous filtration membrane". Journal of Membrane Science, 281(1-2), 581-586.

Goulon-Sigwalt-Abram A (2008). "Une nouvelle méthode d'apprentissage de données structurées: applications à l'aide à la découverte de médicaments" .Thèse de Doctorat. Université Pierre et Marie Curie-Paris VI.

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". Journalof Membrane Science 212 (1-2): 123-134.

Haddada R., Ferjani E., Roudesli, M. S. et Deratani A (2004). "Properties of cellulose acetate nanofiltration membranes. Application to brackish water desalination". Desalination, 167, 403-409.

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.

Hamani (2016). "Récupération et séparation d'ions métalliques par la méthode de membranes liquides émulsionnées" Mémoire de master. Université A. Mira. Bejaia. Algérie.

HollowayW., Cath T.Y., Dennett K.E. et Childress A.E (2005). "Forward osmosis for concentration of anaerobic digester centrate, in: Proceedings of the Awwa Membrane Technology Conference and Exposition", Phoenix, AZ.

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.

Hopfield J.J (1982)."Neural networks and physical systems with emergent collective computational abilities." Proceedings of the National Academy of Sciences 79(8): 2554-2558.

Hornik K (1991). "Approximation capabilities of multilayer feedforward networks." Neural Networks 4(2): 251-257.

Hosseinzadeh M. et Hemmati-Sarapardeh, A (2014). "Toward a predictive model for estimating viscosity of ternary mixtures containing ionic liquids". Journal of Molecular Liquids, 200, 340-348.

http://www.lrs.upmc.fr/fr/l-unite-de-recherche/plateforme-technique/mesures-d-energie-de-surface.html (07/02/2024).

http://www.lrs.upmc.fr/fr/l-unite-de-recherche/plateforme-technique/mesures-d-energie-de-surface.htm l(26-02-2024).

https://fr.wikipedia.org/wiki/Bio-informatique (22/03/2024)

https://openclassrooms.com/fr/courses/5801891-initiez-vous-au-deep-learning/5801898-decouvrez-le-neurone-formel) (22/03/2024)

https://openmembranedatabase.org/membrane-science/) (05/02/2024)

https://scilogs.fr/intelligence-mecanique/reprenons-bases-neurone-artificiel-neurone-biologique/) (22/03/2024)

https://www.journaldunet.com/intelligence-artificielle/1494055-les-neurosciences-ont-inspire-l-intelligence-artificielle-le-modele-neuronal-et-les-concepts-2-2/) (24/03/2024)

https://www.lebigdata.fr/reseau-de-neurones-artificiels definition#:~:text=(apprentissage%20automatique).,L'histoire%20des%20r%C3%A9seaux%20de%20neurones%20artificiels,et%20le%20math%C3%A9maticien%20Walter%20Pitts) (28/02/2024).

https://www.memoireonline.com/04/12/5750/m_Identification-et-commande-des-systemes-non-lineaires13.html) (24/03/2024).

https://www.researchgate.net/figure/14-Representation-schematique-dune-membrane-symetrique-a-et-dune-membrane_fig12_319914244 (06/02/2024)
https://www.researchgate.net/figure/14-Representation-schematique-dune-membrane-symetrique-a-et-dune-membrane_fig12_319914244(10-02-2024)
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: 4801-4809.
Jawad J ., Hawari A. H. et Zaidi S (2020). "Modeling of forward osmosis process using artificial neural networks(ANN) to predict the permeate flux". Desalination, 484 114427.
Jean-Marc B. et Catherine J (2002). "Les procédés membranaires pour le traitement de l'eau". Direction de l'Espace Rural et de la Forêt.
Johnsen A.F., Kristensen S. K. et Dinesen D (1988). "Fluid dynamics in membrane filtration Systems", Sci. Environ. Eng., Vol, 11, Chap. 3, 95-112.
Judd S. et Judd C (2011). "Chapter 2-fundamentals". The MBR Book, 55-207.
Keedwell E. et Narayanan A (2005). "Intelligent bioinformatics: The application of artificial intelligence techniquesto bioinformatics problems" Wiley, Chichester, UK.
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
Khaouane L., Ammi Y. et Hanini S (2017)."Modeling the retention of organic compounds by nanofiltration andreverse osmosis membranes using bootstrap aggregated neural networks". Arabian Journal for Science and Engineering. 42(4), 1443-1453 (2017).
Kim J., Kim S. S., Park M .et Jang M(2008)."Effects of precursor properties on the preparation of polyethylene hollow fiber membranes by stretching". J. Membr. Sci.318 201-209.
Kim J., Kim S. S., Park M. et Jang M (2008). "Effects of precursor properties on the preparation of polyethylene hollow fiber membranes by stretching". Journal of Membrane Science, 318(1-2), 201-209.
Kimura K., Iwase T., Kita S. et Watanabe Y(2009)."Influence of residual organic macromolecules produced inbiological wastewater treatment processes on removal of pharmaceuticals by NF/RO membranes". Water Research. 43(15): 3751–3758.
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 nanofiltrationmembranes". Journal of Membrane Science. 182 (1-2): 205-214.
Kiso Y., Nishimura Y., Kitao T. et Nishimura K (2000). "Rejection properties of non-phenylic presticideswithnanofiltration membranes". Journal of Membrane Science. 171 (2): 229-237.
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.
Kohonen T (1982). "Self-organized formation of topologically correct feature maps." Biological Cybernetics 43(1): 59-69.
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.
Koschikowski J., Wieghaus M., Rommel M., Ortin V. S., Suarez B. P. et BetancortRodríguez J. R (2009). Experimental investigations on solar driven stand-alone membrane distillation systems for remote areas. Desalination, 248(1), 125–131.
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.
Kratbi F(2023)."Modelisation des phenomenes de transport a travers les membranes"diplôme de Doctorat.Université Yahia Fares de Médéa.

LafreniereL.Y., Talbot F. D. F., Matsuura T. et Sourirajan S(1987) ." Effect of poly (vinyl py-rolidone) additive on the performance of poly (ether sulfone) ultrafiltration membranes"., Ind. Eng. Chem. Res. 26 2385-2389.

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

Lanteri Y. (2009). "Transport à travers des membranes de nanofiltration: Caractérisation des propriétés électriques et diélectriques".Thèse de Doctorat. Universite de franche-comte: France.

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.

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

Liu F.,Hashim N.A., Liu Y.T., Abed M. R. M.et LiK(2011)**.** "Progress in the productionand modification of PVDF membranes", J. Membr. Sci. 375, 1-27.

Maurel A (1988). "Technique séparatives à membranes : Considérations théoriques, Technique de l'Ingénieur, traité Génie des procédés". J2 790 ,p1-23.

Maurel A (1998). "Osmose inverse in: Les séparations par membrane dans les procédés de l'industrie alimentaire". Lavoisier - Tec &Doc.

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.

Mona A. Abdel-Fatah (2018). "Nanofiltration systems and applications in waste water treatment", Ain Shams Engineering Journal,

Moody J. et Darken C. J (1989). "Fast learning in networks of locally-tuned processing units." Neural computation**1**(2): 281 - 294.

Nascimento C. A. O., Giudici R.et Guardani R (2000). "Neural network based approach for optimization of industrial chemical processes". Computers & Chemical Engineering, 24(9-10), 2303-2314.

Navajas C., Poso A., Tuppurainen K.et Gynther J (1996). "Comparative Molecular Field Analysis (CoMFA) of MX Compounds using different Semi-empirical Methods: LUMO Field and its Correlation with Mutagenic Activity". Quantitative Structure-Activity Relationships, 15(3), 189-193.

Nechad (2009). " Elaboration d'une membrane de filtration d'eau à base de polyéthylènebasse densité et de gypse de bouzggaia" mémoire de magister.Université Hassiba ben bouali.Chlef. Algérie.

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.

Nollet J.A. Leçons de physique expérimentale, Hippolyte-Louis Guérin and Louis-François Delatour (1748), Paris.

NorbergD., Hong S., Taylor J. et Zhao Y (2007). "Surface characterization and performance evaluation of commercial fouling resistant low-pressure RO membranes". Desalination202: 45-52.

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

Padaki, M., Isloor A. M., Wanichapichart P. et Ismail, A. F (2012)."Preparation and characterization of sulfonated polysulfone and N-phthloyl chitosan blend compositecation-exchange membrane for desalination". Desalination, 298, 42–48.

Poezevara G (2011). "Fouille de graphes pour la découverte de contrastes entre classes :application l'estimation de la toxicité des molecules". Thèse de Doctorat. Université deCaen.France.

Pontié M., Rumeau M., Letellier P., Sarrade S., Schrive L., Mandin, P. et Cassir M (2001). "Les techniques séparatives à membranes–Théorie, applications et perspectives". Paris: Union Internationale pour les applications de l'électricité, p212-214.

Prince J. A., Singh G., Rana D., Matsuura T., Anbharasi V.et Shanmugasundaram T. S (2012). "Preparation and characterization of highly hydrophobic poly (vinylidene fluoride)–Clay nanocomposite nanofiber membranes (PVDF–clay NNMs) for desalination using direct contact membrane distillation". Journal of membrane science, 397, 80-86.

Ramdani A., Deratani A. et Taleb S (2017). "Comparative study of the retention of monovalent salts by nanofiltration: application to natural water". Revue des Sciences de l'Eau: Journal of Water Science, 30(1), 65-69.

Rao A.P., Desai N.V. et Rangarajan R(1997)."Interracially synthesized thin film membrane for seawater desalination" J. Membr. Sci. 124 263.

Remigy J. C. et Desclaux S (2007). "Filtration membranaire (OI, NF, UF)-Présentation des membranes et modules".

Remise P(2006)."Etude des rétro lavages assistés par l'air et contrôle du colmatage résiduel :application à la production d'eau potable en filtration frontale", Thèse de doctorat. Université de Toulouse. Toulouse, France.

Reverchon E. et Cardea S(2004). "Formation of cellulose acetate membranes using a super critical fluid assisted process". J MembrSci 240 ,187–195.

Robert R., Fleischer L., Walker R. M. et Price, P. B (1975)." Nuclear tracks in solids: principles and applications". Univ of California Press.

Roy K. J., Anjali T. V. and Sujith A(2017)."Asymmetric Membranes Based on Poly (VinylChloride) : Effect of Molecular Weight of Additive and Solvent Power on theMorphology and Performance" . Journal of Materials Science, 52 , 5708-25.

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.

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

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.

Sahoo G.B. 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.

Saja S., Bouazizi A., Achiou B., Ouammou M., Albizane A., Bennazha J. et Younssi S. A (2018). "Elaboration and characterization of low-cost ceramic membrane made from natural Moroccan perlite for treatment of industrial wastewater". Journal of environmental chemical engineering, 6(1), 451-458.

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". Separationand Purification Technology. 50 (1): 122–131.

Savart T (2013). "design and realization of industrial ultrafiltration hollow fibers ofpolyvinylidene fluoride (PvDF) containing block copolymers". Thèse de doctorat, Toulouse .France.

Saxena A., Tripathi B. P., Kumar M.et Shahi, V. K (2009). "Membrane-based techniques for the separation and purification of proteins: An overview". Advances in colloid and interface science, 145(1-2), 1-22.

Schäfer A. I ., Nghiem L. D. et Waite T. D (2003). "Removal of the natural hormone estrone from aqueous solutionsusing 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 matterremoval by nanofiltration: effects of solution chemistry on retention of low molar mass acidsversus bulk organic matter". Journal of Membrane Science 242: 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: 3864- 3871.

Sigg L et al (2000). Chimie des milieux aquatique, chimie des eaux naturelles et desinterfaces dans l'environnement. Edition :Dunod. 567p,

Si-Moussa C (2009). "Modélisation des équilibres de phases a pressions élevées par les réseaux de neurones artificiels". Thèse de doctorat. École nationale supérieure polytechnique. Alger. Algérie.

Stone M (1977)."An asymptotic equivalence of choice of model by cross-validation and Akaike'scriterion." Journal of the Royal Statistical Society B 39(1): 44-47.

TaheranM; Brar S. K; Verma M; Surampalli R. 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.

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

Tamas A. P (2004). "Étude comparée du colmatage en nanofiltration et en ultrafiltration d'eau de surface". Thèse de Doctorat. Université Laval du Québec.Canada.

Tasselli F(2014). "Non-solvent Induced Phase Separation Process (NIPS) for Membrane Preparation". Encyclopedia of Membranes 1–3.

Tongwen X. et Weihua Y (2003). "A novel positively charged composite membranes for nanofiltration prepared from poly (2, 6-dimethyl-1, 4-phenylene oxide) by in situ amines crosslinking". Journal of membrane science, 215(1-2), 25-32.

Urbani D., Roussel-Ragot P., Personnaz L.et Dreyfus G. (1994). "The selection of neural models of nonlinear dynamical systems by statistical tests". In Proceedings of IEEE Workshop on Neural Networks for Signal Processing. 229-237.

Van der Bruggen B., Braeken L. et Vandecasteele C (2002a). "Flux decline in nanofiltration due to adsorption oforganic compounds". Separation and Purification Technology 29: 23-31.

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: 29-41.

Verliefde A. R. D; Cornelissen E. R; Heijman S. G. J; VerberkJ .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.

Wang H., Wang W., Wang L., Zhao B., Zhang Z., Xia X., Yang H., Xue Y. et Chang N(2017). "Enhancement of hydrophilicity and the resistance for irreversible fouling of polysulfone (PSF) membrane immobilized with graphene oxide (GO) through chloromethyl lated and quaternized reaction". Chemical Engineering Journal. 334: 2068-2078.

Wijmans J. G. et Baker R. W (1995). "The solution-diffusion model:areview". Journal of Membrane Science 107: 1- 21

Wu C., Li A., Li L., Zhang L., Wang H., Qi X. et Zhan, Q(2008)."Sodium alginate-polyvinylalcohol/polysulfone (SA-PVA/PSF) hollow fiber composite pervaporation membrane for dehydration of ethanol-water solution". Journal of Shanghai University (English Edition), 12(2), 163-170.

Xu Z., Liao J., Tang H., NanwenLi (2017)."Antifouling polysulfone ultra filtration membranes with pendent sulfonamide groups", Journal of Membrane Science, 548, 481–489.

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 ;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 Membrane Science. 342 (1-2): 251-262.

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.

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.

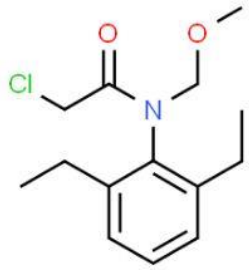
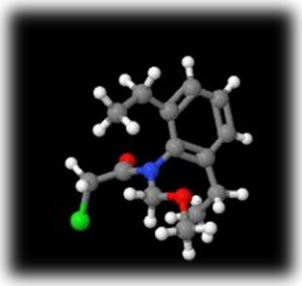
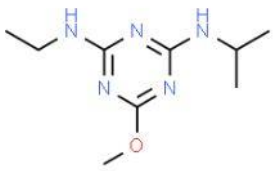
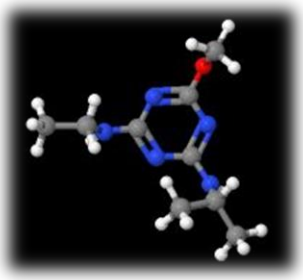
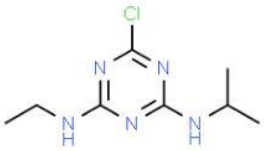
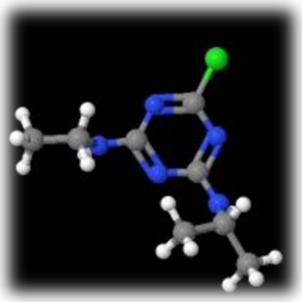
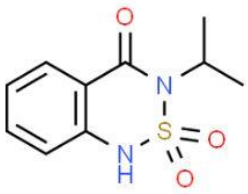
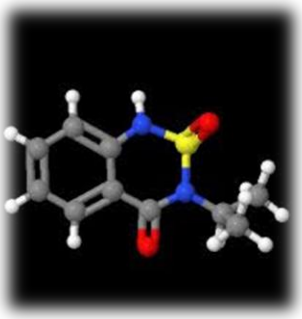
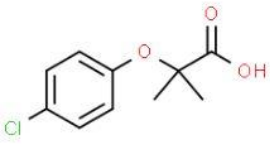
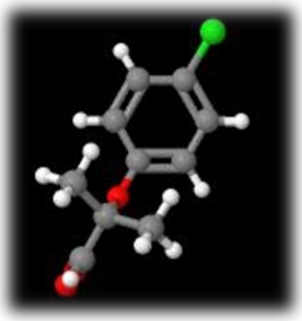
Yoo, J.E; Kim,J.H. et & Kim Y (2003)."Novel ultrafiltration membranes prepared from the new miscible blends of polysulfone with poly (1-vinylpyrrolidone-co-styrene) copolymers". Journal of membrane science, 216, 95-06.

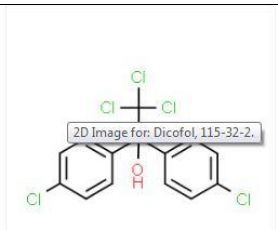
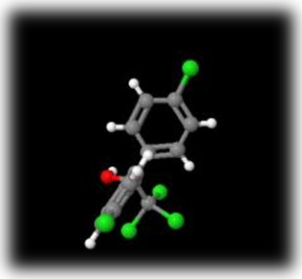
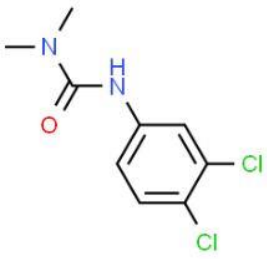
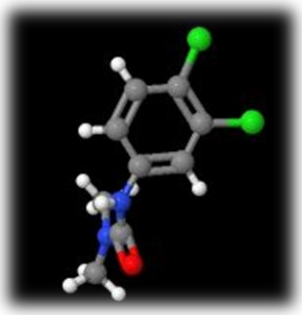
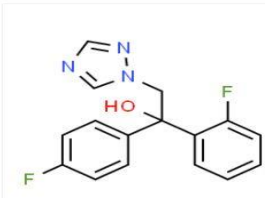
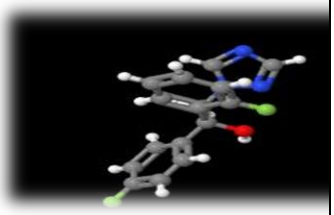
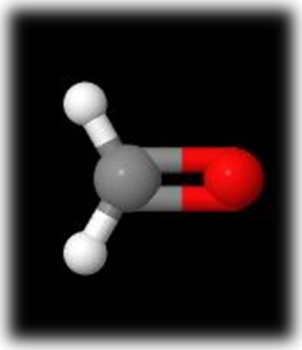
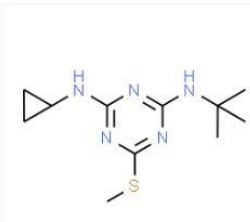
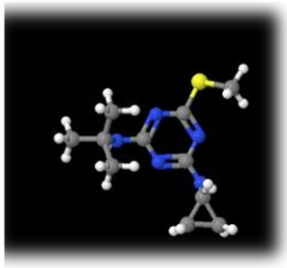
Zhao Z., Zheng J., Wang M., Zhang H. et Han C. C (2012). "High performance ultrafiltration membrane based on modified chitosan coating and electrospun nanofibrous PVDF scaffolds". Journal of membrane science, 394, 209-217.

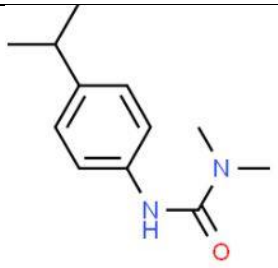
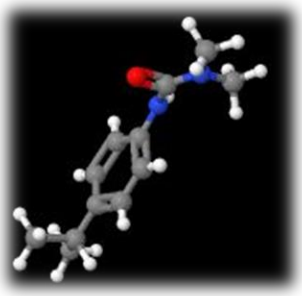
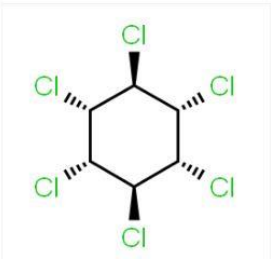
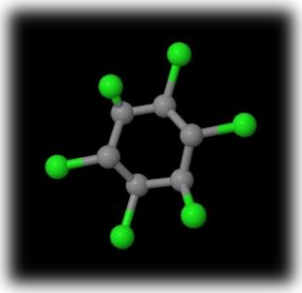
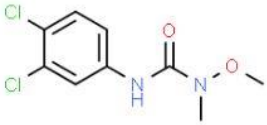
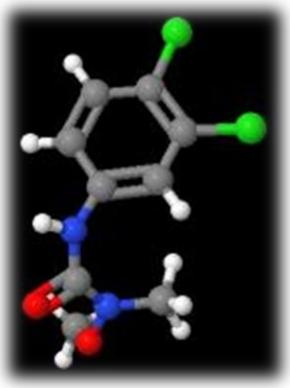
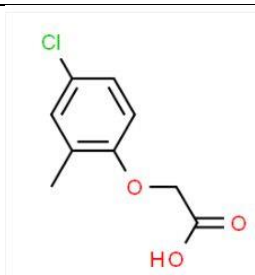
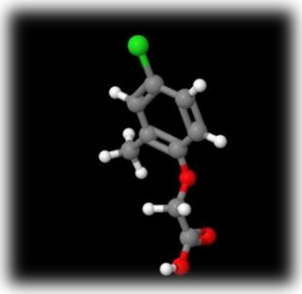
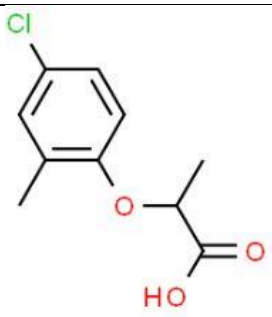
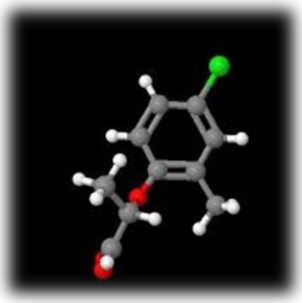
Zhou Y., Fang M-Y. et Fengsun L(2019)."Pore size tailoring from ultra filtration to nanofiltration with PVC-g-PDMA via rapid immersion thermal annealing". Journal of membrane science, 572, 401-409.

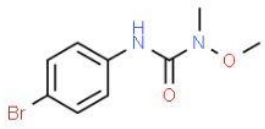
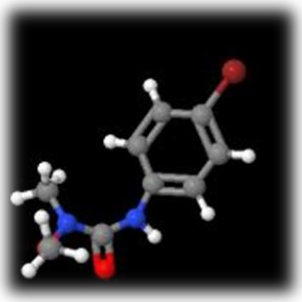
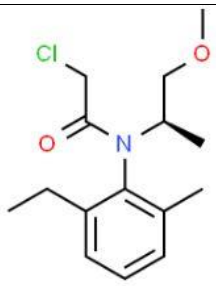
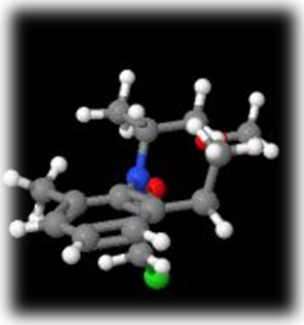
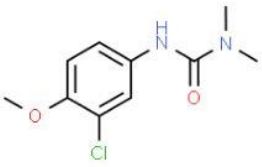
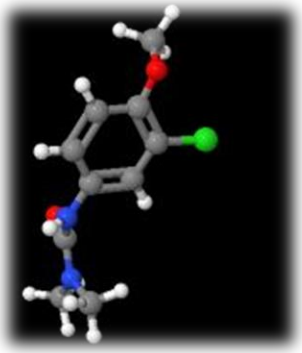
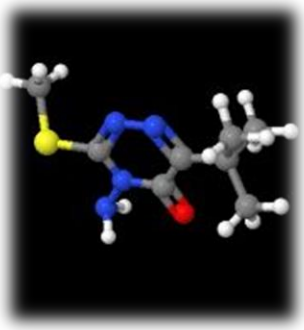
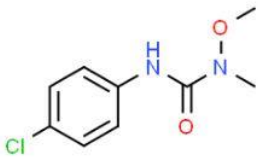
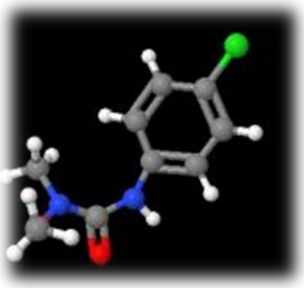
Zhu W., Zhang X., Zhao C; Wu W., Hou J. et Xu M (1996). "A novel polypropylene Microporous film".,Polym. Adv. Technol. 7(9). 743-748.

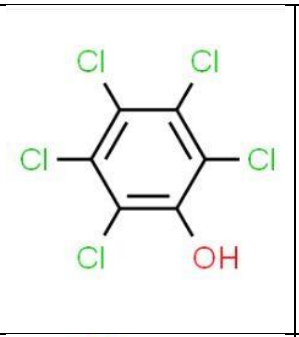
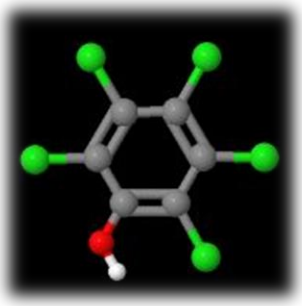
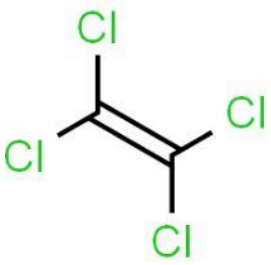
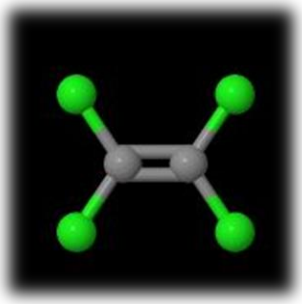
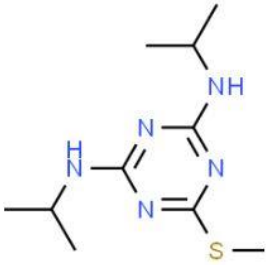
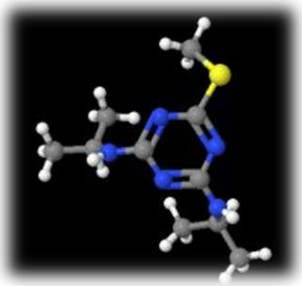
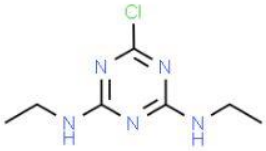
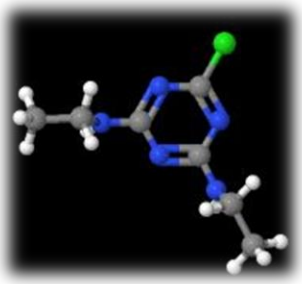
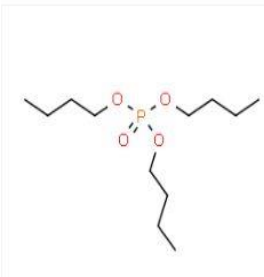
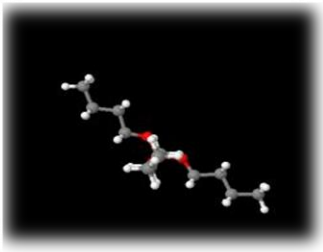
Annexe A

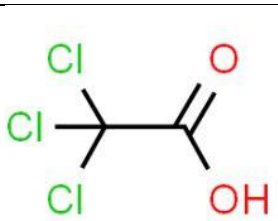
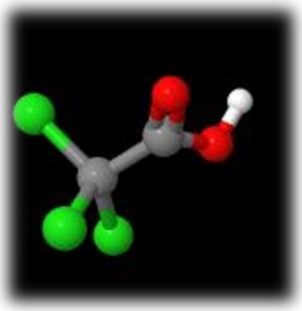
#	Nom de pesticides	Formulebrute	MW (g/mol)	Structure chimique 2D	Structure chimique 3D	Classe de pesticides
01	**Alachlor**	$C_{14}H_{20}ClNO_2$	269,77			Herbicide
02	**Atraton**	$C_9H_{17}N_5O$	211,26			Herbicide
03	**Atrazine**	$C_8H_{14}ClN_5$	215,68			Herbicide
04	**Bentazone**	$C_{10}H_{12}$	240,3			Herbicide
05	**Clofibric acid**	$C_{10}H_{11}$	214,64			Herbicide

06	**Dicofol**	$C_{14}H_9Cl$	370,49			Insecticide
07	**Diuron**	$C_9H_{10}Cl$	233,09			Herbicide
08	**Flutriafol**	$C_{16}H_{13}F$	301,29			Fongicide
09	**Formaldehyde**	CH_2O	30,02			Fongicide Insecticide
10	**Irgarol**	$C_{11}H_{19}N$	253,37			Herbicide

11	**Isoproturon**	$C_{12}H_{18}$	206,28			Herbicide
12	**Lindan**	$C_6H_6Cl_6$	290,83			Insecticide
13	**Linuron**	$C_9H_{10}Cl$	249,09			Herbicide
14	**Mcpa**	C_9H_9	200,62			Herbicide Fongicide
15	**Mcpp**	$C_{10}H_{11}$	214,65			Herbicide

16	**Metobromuron**	C_9H_{11}	259,1			Herbicide
17	**Metolachlor**	$C_{15}H_{22}$	283,79			Pesticide
18	**Metoxuron**	$C_{10}H_{13}$	228,67			Herbicide
19	**Metribuzin**	$C_8H_{14}N$	214,29			Herbicide
20	**Monolinuron**	C_9H_{11}	214,05			Herbicide

21	**Pentachlorophenol**	C_6HCl_5O	266,34			Fongicide Herbicide
22	**Perchloroethene**	C_2Cl_4	165,83			Fongicide
23	**Prometryn**	$C_{10}H_{19}N_5S$	241,36			Herbicide
24	**Simazin**	$C_7H_{12}ClN_5$	201,08			Herbicide
25	**Tributyl phosphate**	$C_{12}H_{27}$	266,31			Pesticide

26	**Trichloroacetic acid**	$C_2HCl_3O_2$	163,39			Herbicide

Annexe B

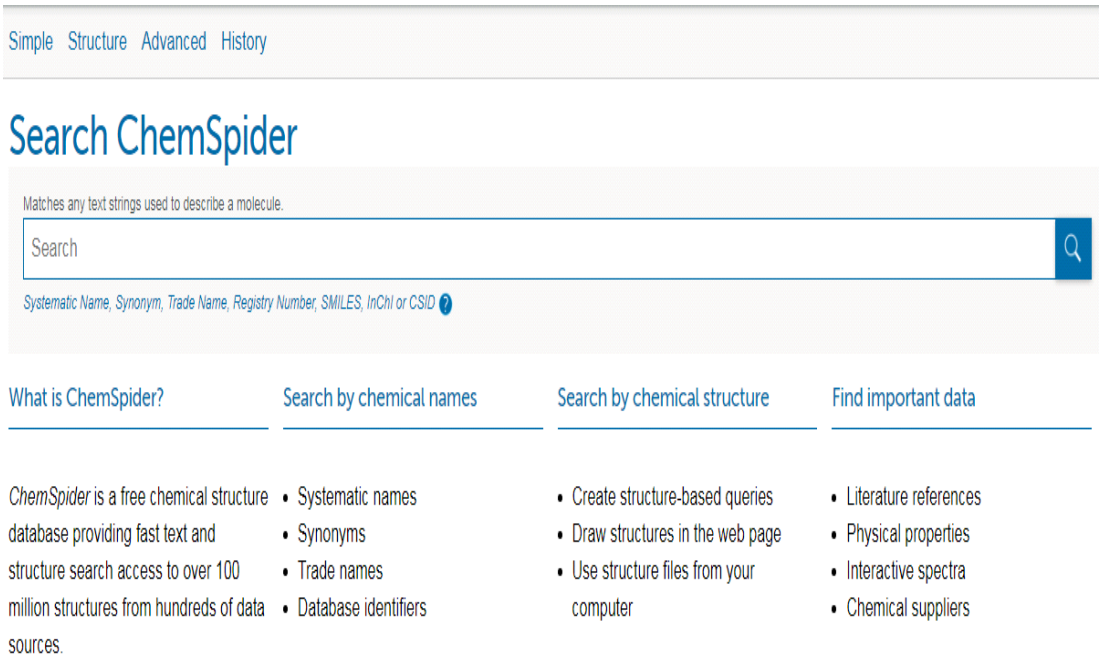
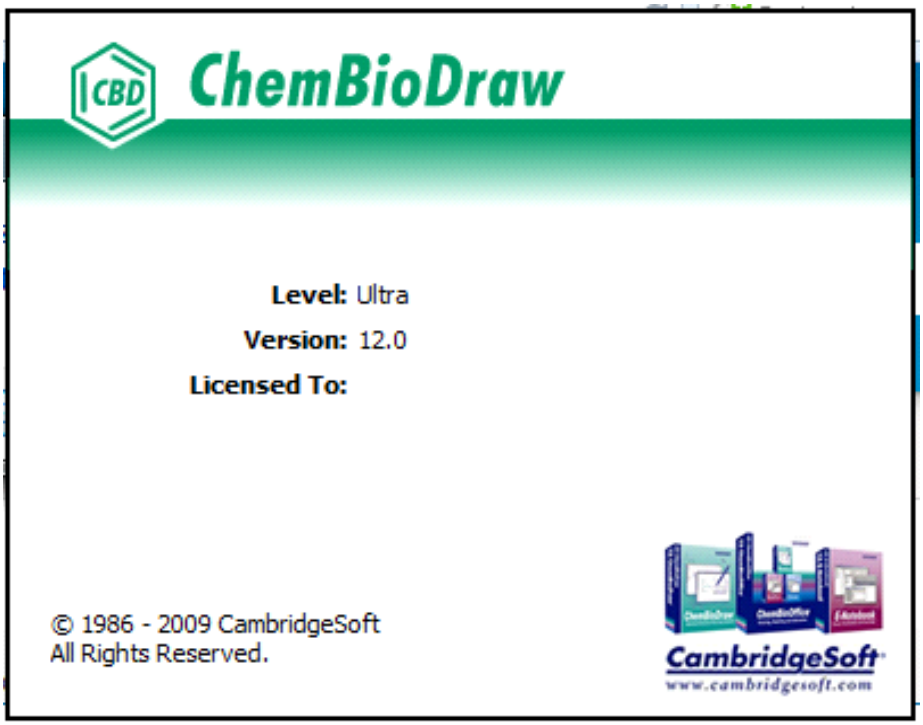
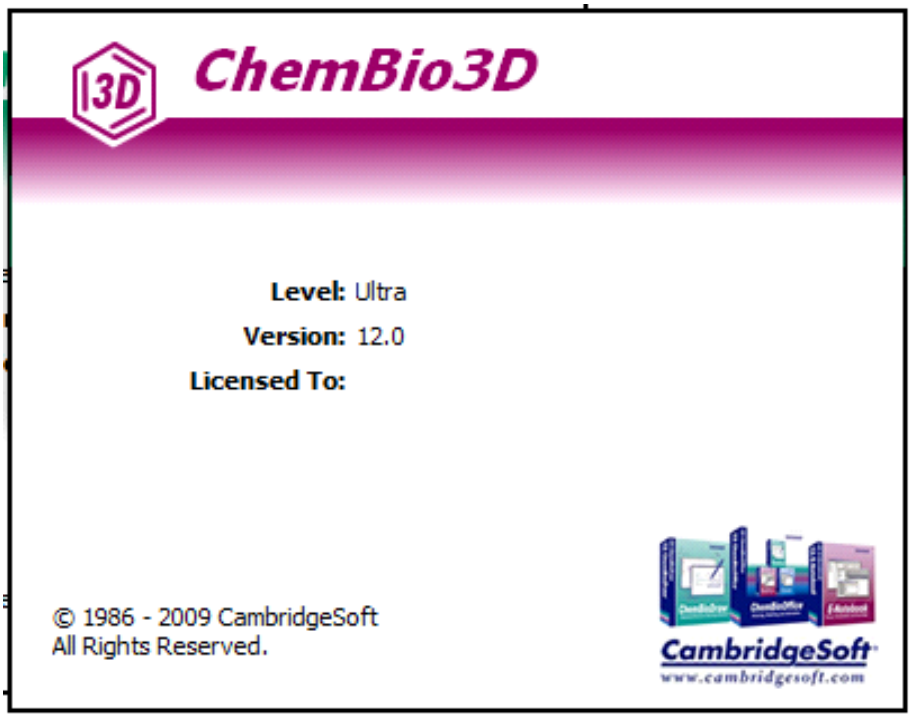


Figure 1. Capture d'écran du site intitulé Chemspider.



(a)



(b)

Figure 2. Capture d'écran des logiciels : (a) ChemBioDraw et (b) ChemBio3D.

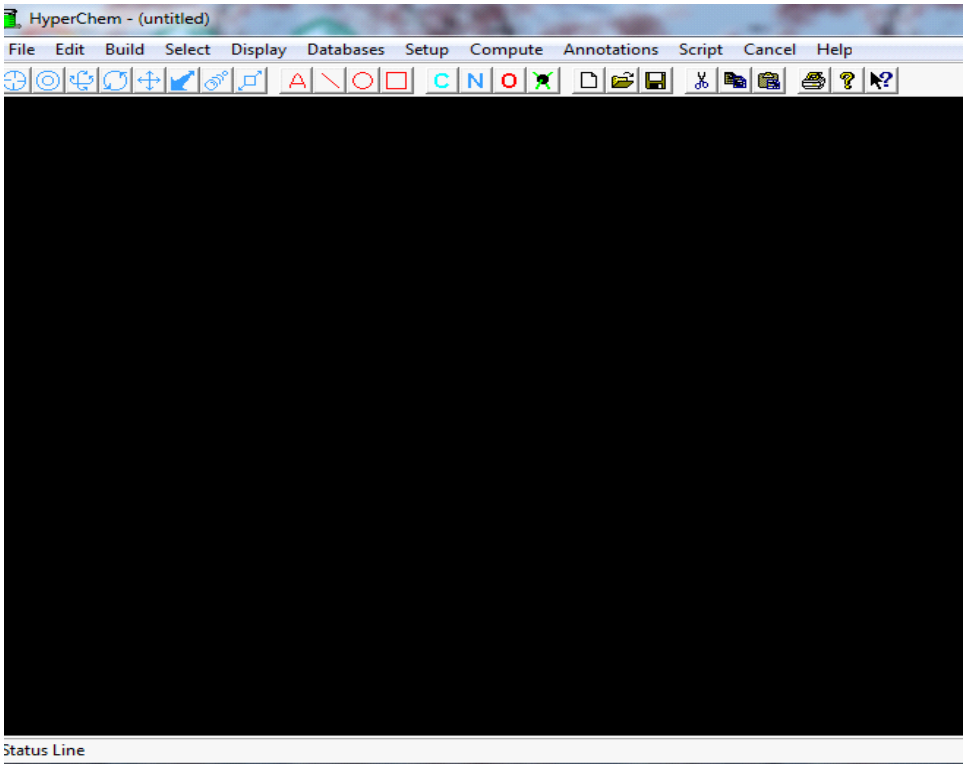


Figure 3. Capture d'écran de logiciel Hyperchem



Figure 4. Capture d'écran de logiciel STATISTICA

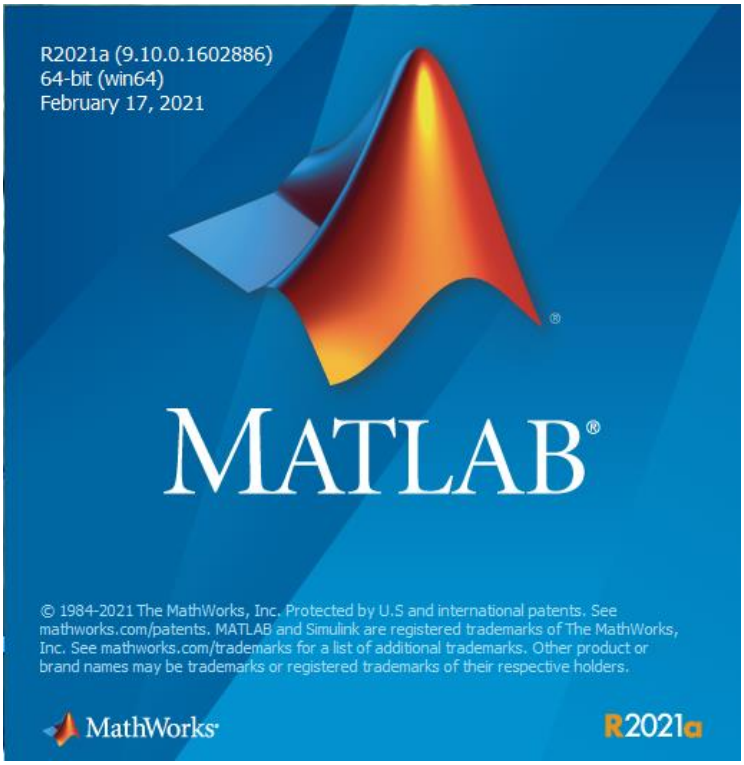


Figure 5. Capture d'écran de logiciel MATLAB.