République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITE Dr YAHIA FARES DE MEDEA



Faculté des Sciences et de la Technologie

Département du Génie des Procédés et Environnement

Mémoire de Fin d'Etudes pour l'obtention du diplôme de Master en
Génie des Procédés

Spécialité : **Procédés Pharmaceutiques**

THEME

Prédiction neuronale de la toxicité des pesticides

Proposé par :
Pr. HANINI Salah

Dirigé par :
Mr. HAMADACHE Mabrouk
M^elle^ KHAOUANE Latifa

Etudié par :

M^elle^ CHABA MOUNA Siham.

M^elle^ BOUDERGUA Fatiha.

Promotion 2011 - 2012

Dédicace

Je dédie ce mémoire

Tout d'abord á ma mère

Merci de m'avoir offert ton lait pendant de très nombreux

mois, qui en plus des bienfaits qu'il m'a apporté, a renforcé nos liens.

Tu m'as toujours encouragé et cru en moi.

A la mémoire de mon père et mon frère

À mon cher mari RABAH et À ma belle-famille ''Abri ''

de m'avoir soutenue et encouragée tout au long de ce travail.

À mes sœurs, SOUAAD et son mari Abdallah et leurs enfants

Youcef et Adam, et RATIBA et son mari Ibrahim et leurs enfants

Salim, Youcef et Siham

À mes frères MOHAMED et sa Femme HAMIDA et leurs enfants

Ahlam et Zouina et Riyad et Inass, MOURAD et sa femme

Zahra Sans oublier SHAFIKA et son mari et les petites

Marawa et Salma

Á toute la famille BOUDERGUA

Á ma chère amie et binôme CHABA MOUNA Siham pour sa patience et son sérieux

Á toutes mes amies surtout B. Manel, D. Amel, B. Wardia, C. Nabila,

M. Amina, B. Samira, G. Chahrazed, W. Dalila, B. Lamia, H. Dallel, Z. Mouna.

Á M. NAIMA, A. YAMINA

Á la deuxième promotion Master 2011-2012 (LMD) du Génie des Procédés Pharmaceutiques Et environnement

Et Merci à tous ceux que j'ai oublié…

Dédicace

Je dédie ce modeste travail à mes très chers et aimables parents C.Abdlkader, B.Rokia qui sont toute ma fierté et à qui Je dois tout ce que je suis aujourd'hui et, à ceux que j'aime tant et qui sont ma joie de vivre qu'ALLAH vous protège et ne me privera jamais de vous.

A mes chères sœurs Iman, Hayat,Khadidja, Manal.

A mon frère Mustafa et Rachid et aux petits Hadil et Abdallah.

A mon cher fiancé Imad, pour son soutien et sa présence et la joie qu'il m'a donné

A ma meilleure amie avec qui j'ai passé Les meilleurs temps b.Fatiha

A mes adorables amies Fatiha, Manal, Amel, Zakia, wardia, sans oublier Ismail etBilel qui ont m'aidé durant mon cursus universitaire.

A mes copines Amina, Nabila, Lamia, Dalila, Samira, Dalel, Mouna, Chahrazed, Naima et yamina

A tous mes enseignants.

Remerciements

Nous tenons à remercier en premier lieu Dieu « ALLAH DJELLA WA AALA » qui nous a donné le courage et la volonté durant notre travail pour qu'il soit une réussite méritée.

Nous remercions vivement notre promoteur Mr. HANINI Salah, Professeur à L'UYFM, qui nous a extrêmement aidé et encouragé tout au long de ce travail.

Nous tenons à exprimer nos plus vifs remerciements à nos Co- rapporteurs Mr HAMADACHE Mabrouk & Melle KHAOUANE Latifa, enseignants à l'UYFM, pour leurs aides, leurs accompagnements et leurs soutiens. Nous leur devons les remerciements les plus sincères.

Nous remercions Dr. SI MOUSSA Cherif et Melle REZAZI Sarah qui n'ont pas cessé de nous encourager et de nous conseiller tout au long de la préparation de ce mémoire.

Nous adressons également nos vifs remerciements au Dr. BENKORTEBI Othmane chef de département du Génie des procédés et environnement.

Nous remercions tous les enseignants de l'UYFM qui nous ont guidés tout au long de notre formation universitaire.
Sans oublier la secrétaire du Département du Génie Eléctrique et de l'Informatique C.Iman pour nous avoir aidé pendant nos 5 années d'études.

A tous nos collègues et amies qui nous ont soutenu et encouragé pour terminer ce travail et à toute la promotion Master 2011-2012

Espérons à la fin que notre travail va satisfaire les membres du jury tout en acceptant leur évaluation et leurs remarques qui nous seront utiles.

Sommaire

Introduction Générale 1

Chapitre I. Les Pesticides

I.1. Introduction 2

I.2.Définition 2

I.3.Historique 3

I.4.Composition des pesticides 3

I.5.Classification des pesticides 5

I-6 généralisationde l'utilisation mondiale des pesticides 8

I-7 les pesticides dans les pays en développements 9

I-8 devenir des pesticides 10

I-9 toxicité des pesticides 11

I-10 impacts chroniques sur la santé humaine et ampleur du problème 12

I-10-a L'ampleur du problème 12

I-10-b Les pesticides à toxicité aiguë 13

I-10-c Impacts chroniques sur la santé 13

I-11 effets sur l'environnement 16

I-12 utilisation des pesticides en Algérie 18

I.12.1. Situation géographique 18

I.12.2. Agriculture 19

I.12.3.Les principaux pesticides utilisés en Algérie 19

I.13. conclusion 21

Chapitre II. QSAR

II.1.Introduction 22

II.2. Définition 22

II.3. Historique 23

II.4. Application du QSAR 24

II.4.1. En chimie 24

II.4.2. En biologie 24

II.4.3 Analyse de données 24

II.5. QSAR en toxicologie de l'environnement 26

II.6. QSAR pour la prédiction de la toxicité 26

II.7. Les systèmes experts pour la prédiction de la toxicité 28

II.8. Calcul des propriétés physico-chimique 29

II.9. SMILES (Simplified Molecular Input Line Entry Specification) 30

II.9.1. Notation SMILES 30

II.9.2. Principes de SMILES et théorie des graphes 31

II.10.Conclusion 33

Chapitre III. Les réseaux de Neurones Artificiels

III.1.Introduction 34

III.2.Intelligence artificielle 34

III.2.1. Définition 34

III.2.2.Historique 35

III.2.3. Domaines de l'intelligence Artificielle 35

III.2.4.Les modèles de l'intelligence artificielle 36

III.3.Les réseaux de neurones 36

III.3.1 Introduction aux réseaux de neurones 36

III.3.2 Définition 37

III.3.3.Historique 37

III.3.4 Applications des réseaux de neurones 38

III.3.5 Notions fondamentales sur les réseaux de neurones artificiels 38

III.3.5.1 Analogie entre neurone biologique et neurone artificiel 38

III.3.6 Le modèle mathématique d'un neurone artificiel 41

III.3.7. Fonctions d'activation 43

III.3.8 Architecture des réseaux 46

III.3.9.Apprentissage 48

- L'apprentissage supervisé 48
- L'apprentissage non supervisé 49

III.3.9.1 comment apprendre au réseau 49

III.3.9.2 Surapprentissage 50

III.3.9.3 Rétropropagation 50

III.3.9.4 Élagage 50

III.3.9.5 Précision 50

III.4. Conclusion 51

Chapitre IV. Modélisation Neuronale de la Toxicité des Insecticides et des Herbicides

IV.1. Introduction 52

IV.2. Méthodologie de calcul 52

IV.2.1. Elaboration de la base de données 52

IV.2.2. Choix d'une famille de descripteurs 64

IV.2.3. Prétraitement de la base de données 68

IV.2.4. Normalisation de la base de données partielle 69

IV.2.5. Mise en forme des données 69

IV.2.6. Conception et optimisation 70

IV.3. Structure du RNSO obtenu 70

IV.3.1. L'algorithme d'apprentissage et la fonction d'activation 70

IV.4. Analyse de sensibilité 72

IV.5. conclusion 77

Conclusion générale 78

Références bibliographiques 79

Liste des figures

Figure	Titre de la figure	page
Figure I.1	Structures chimiques des principales familles de pesticides**,**	8
Figure I.2	Le marché mondial des pesticides par région dumonde	9
Figure I.3	Le marché mondial des pesticides par catégorie de produit	9
Figure I.4	Mécanismes de transport des pesticides dans l'environnement	17
Figure I-5	Processus impliqué dans le devenir des pesticides dans les sols	18
Figure III.1	Schéma complet d'un neurone biologie	39
Figure III.2	Représentation d'un neurone formel	40
Figure III.3	Schéma de comparaison entre un neurone biologique et un neurone artificiel	41
Figure III.4	Modèle mathématique du neurone artificiel	42
Figure III.5	Schéma plus formel d'un perceptron	43
Figure III.6	Fonctions de transfert : (a) du neurone "seuil"; (b) du neurone "linéaire", et(c) du neurone "sigmoïde"	45
Figure III.7	Architecture générale d'un RNA	46
Figure III.8	Réseau Monocouche	46
Figure III.9	Réseau Multicouche	47
Figure III.10	Réseau à connexions complètes	47
Figure III.11	Réseau à connexions locales	48
Figure III.12	Schéma représentatif d'un apprentissage supervisé	48
Figure III.13	Schéma représentatif d'un apprentissage non supervisé	48

Figure III.14	Représentation d'apprentissage	49
Figure IV.1	Méthodologie de calcul de la toxicité par les réseaux de neurones artificiels.	53
Figure IV.2	Procédure d'élaboration de la base de données.	57
Figure IV.3	Dessin des molécules sur ChemDraw (1,2).	57-58
Figure IV.4	Généralisation de la notation smiles par chemDraw(1,2)	58-59
Figure IV.5	Les étapes pour le calcul des descripteurs par E Dragon1.0. (1..9)	59-64
Figure IV.6	Schéma simplifié de l'architecture du RNSO.	71
Figure IV.7	Validations des valeurs calculées par RNA avec les valeurs expérimentales.	72
Figure IV.8	Processus de calcul par la méthode de poids de connexion.	74
Figure IV.9	Influence des descripteurs sur la toxicité.	77

Abréviation

Abréviation	Signification
2D	**D**eux **D**imensions
2,4-D	**D**ichloro acide phénoxyacétique
3D	**T**rois **D**imensions
EPA	**E**nvironnemental **P**rotection **A**gency
F	**I**ndice de **F**isher
FAO	**F**ood and **A**griculture **O**rganization
HO	**H**idden-**O**utput (cachée-sortie)
IA	**I**ntelligence **A**rtificielle
IH	**I**nput-**H**idden (entrée –cachée)
IR	**I**mportance **R**elative
IUPAC	International Union of Pure and Applied Chemistry
MLP	**M**ulti **L**ayer **P**erceptron
OBD	**O**ptimal **B**rain **D**amage
OBS	**O**ptimal **B**rain **S**urgeon
OMS	**O**rganisation **M**ondiale de la **S**anté
PDB	**P**rotein **D**ata-**B**ank
QSAR	**Q**uantitative **S**tructure–**A**ctivity **R**elationship
RNA	**R**éseau de **N**eurone **A**rtificiel
RNSO	**R**éseau de **N**eurone **S**tatique **O**ptimal
SMILES	**S**implified **M**olecular Input **L**ine **E**ntry **S**pecification
VTH	**V**ariétés **T**olérantes à un **H**erbicide

Liste des Tableaux

Numéro	Titre de tableau	Page
Tableau I.1	Historique de l'évolution des trois plus grandes familles d'activité des années 1900 à nos jours	04
Tableau I.2	Classification des pesticides en catégories de danger	06
Tableau I.3	Classification des pesticides selon la dose létale	07
Tableau I.4	Classement toxicologique d'un produit	12
Tableau I.5	Les pesticides les plus utilisés en Algérie	19
Tableau I.6	Les pesticides utilisés dans la région de Médéa	20
Tableau II.1	Notation SMILES pour quelques molécules	32
Tableau III.1	Analogie entre le neurone biologique et le neurone formel	40
Tableau III.2	les fonctions de transfert les plus courantes	44
Tableau IV.1	Liste des pesticides utilisés dans cette étude.	54
Tableau IV.2	Liste des paramètres des descripteurs (constitutionnels et les propriétés moléculaires)	65
Tableau IV.3	Liste des paramètres des descripteurs (constitutionnels et les propriétés moléculaires) après prétraitement	68
Tableau IV.4	Structure des RNSO utilisés dans la modélisation de la toxicité	70
Tableau IV.5	poids de connexion entrée-cachée, cachée-sortie.	72
Tableau IV.6.	Produit de IH et HO.	74
Tableau IV.7	Importance individuelle.	76
Tableau IV.8	Importance relative.	76

Nomenclature

Lecture	Signification	Unité
b	Biais d'un neurone artificiel	(-)
CL50	Concentration Létale 50	[mmole/l]
DJA	Dose Journalière Admissible	[mg/kg/jour]
DL50	Dose Létale 50	[mg/kg]
NOAEL	Dose sans effet toxique observable	[mg/kg]
w	Poids synaptique des connexions interneuronales	(-)

ملخص:التنبؤ العصبوني للتأثير السمي لمبيدات

الهدف من هذا العمل هو إقامة علاقات بين البنية والنشاط من اجل تنبؤ سمية 72 مبيد مع حساب 25 معيار تحت شكل ممثلات تركيبية وخصائص جزيئية لكل جزيء.

استراتيجيه المحاولة والخطأ استعملت كمعيار توقف من اجل النمذجة العصبونية مع قدرة التعميم ، يليها تحليل الحساسية من اجل تحديد المعايير الأكثرتأثيرا على سمية مبيدات الحشرات والأعشاب .

الشبكة العصبونية المثالية المتحصل عليها مكونة من طبقة إدخال، طبقة خفية وطبقة خروج.

تحليل الحساسية اظهرأن عدد ذرات الكلور وعدد ذرات الفوسفور هما الأكثر تأثيرا على سمية مبيدات الأعشاب والحشرات.

الكلمات المفتاحية: تنبؤ، سمية، ممثلات، علاقات البنية النشاط والكمية، شبكة عصبونية، تحليل الحساسية.

Résumé : Prédiction neuronale de la toxicité des pesticides

Le but de ce travail est d'établir des relations structure-activité (SAR) pour la prédiction de la toxicité dans un ensemble de données de 72 composés pesticides. 25 paramètres ont été calculés sous forme de descripteurs QSAR de type constitutionnels et propriétés moléculaires pour chaque molécule.

La stratégie d'essai et erreur a été utilisée pour la modélisation neuronale avec la capacité de généralisation comme critère d'arrêt suivie d'une analyse de sensibilité afin de déterminer les paramètres les plus pertinents sur la toxicité des pesticides.

Le réseau de neurones optimal obtenu est un perceptron multicouche : une couche d'entrée, une couche cachée et une couche de sortie (25 ; 13 ; 1), L'analyse de sensibilité à révélé que le nombre d'atomes de chlore et le nombre des atomes de phosphore sont les paramètres les plus influents sur la toxicité des pesticides.

Mots clés : Prédiction, Toxicité, Descripteur, QSAR, Réseau de Neurones, Analyse de Sensibilité.

Abstract: Neuronal toxicity prediction of pesticides

The aim of this work is to establish structure-activity relationships (SAR) which can predict toxicity in a database of 72 pesticide compounds and 25 parameters calculated as QSAR descriptors of constitutional type and molecular proprieties for each molecule.

The test and error strategy was used for neuronal modeling with the capacity of generalization like criterion of stop, followed by a sensitivity analysis to determine the most relevant parameters on the toxicity of herbicides and insecticides.

The optimal neural network obtained is a multilayer perceptron: one input layer, one hidden layer and one output layer. The sensitivity analysis revealed that the number of chloride atoms and number of phosphor atoms parameters have the greatest influence on the toxicity of herbicides and insecticides.

Keywords: Prediction, Toxicity, descriptor, QSAR, Neural Networks, sensitivity analysis.

Introduction générale

Les pesticides (insecticides, raticides, fongicides, et herbicides) sont des composés chimiques dotés de propriétés toxicologiques, utilisés essentiellement par les agriculteurs pour lutter contre les animaux (insectes, rongeurs) ou les plantes (champignons, mauvaises herbes) considérés nuisibles aux plantations. Par ailleurs et dans une moindre proportion, ils sont utilisés pour un usage domestique.

Malgré le fait que les pesticides soient bénéfiques, leur emploi massif a pour conséquence la contamination généralisée de notre environnement. Les eaux de surfaces et les eaux souterraines sont très fréquemment polluées. Cette contamination généralisée de l'environnement et de la nourriture par les pesticides rend inévitable la contamination de l'être humain. Leurs effets secondaires nocifs ont été rapidement mis en évidence et leur toxicité, liée à leur structure moléculaire, ne se limite pas en effet aux seules espèces que l'on souhaite éliminer.

La tendance actuelle est le recours à l'interdiction de l'utilisation des pesticides reconnus toxiques. Cependant, pour arriver à statuer sur la toxicité ou non d'un pesticide, de longues et harassantes analyses sont effectuées sur des animaux pendant une très longue durée. D'un autre côté, le coût de ces analyses est très élevé.

Pour arriver à surmonter ce probléme de coût et de durée, les recherches actuelles sont orientées vers l'utilisation de l'intelligence artificielle. A ce titre, l'objectif de notre étude s'inscrit dans cette démarche. Par ce travail, nous envisageons de modéliser la prédiction de la toxicité des pesticides à partir de leur structure moléculaire et ce par utilisation des réseaux de neurones artificiels de type perceptron multicouche.

Le premier chapitre de ce mémoire a été consacré aux pesticides. C'est un ensemble de données relatant essentiellement leur importance, mais aussi leurs inconvénients.

Le second chapitre traite de la relation « structure-activité » en spécifiant son importance et ses nombreuses applications dans la prédiction des phénomènes physico-chimiques.

Le troisième chapitre concerne les réseaux de neurones artificiels (R.N.A). On y a présenté les notions fondamentales inhérentes à ces réseaux.

Le quatrième chapitre est quant lui aborde l'étape expérimentale de notre travail. Il regroupe successivement le calcul des descripteurs et la méthodologie de choix des plus pertinents, la modélisation par les réseaux de neurones artificiels ainsi que les résultats obtenus pour la prédiction de la toxicité des pesticides.

En dernier, une conclusion où sont mis en évidence les résultats obtenus et les perspectives qui peuvent être envisagées dans le futur.

Chapitre I

Les Pesticides

I-1 INTRODUCTION

Les pesticides constituent un enjeu important pour la qualité de notre alimentation et de notre environnement. On peut constater que l'agriculture s'est grandement transformée suite au recours à l'utilisation intensive des pesticides. Malheureusement cette montée du secteur agricole n'a pas que des répercussions positives, elle détériore grandement l'environnement et elle est très néfaste pour la santé humaine et surtout pour celle des agriculteurs qui sont les plus exposés aux effets des pesticides **[1].**

Utilisés pour lutter contre les insectes nuisibles (insecticides), les mauvaises herbes (herbicides) ou les champignons nuisibles (fongicides), les pesticides contaminent les eaux de surface ainsi que les eaux souterraines. Selon l'OMS, les pesticides seraient responsables du décès de 20.000 personnes environ chaque année, dans le monde. Actuellement, 25 groupes de pesticides, dont la plupart sont utilisés en Algérie, ont été déclarés substances cancérigènes. Il est extrêmement important que les utilisateurs de pesticides doivent comprendre les processus impliqués dans la contamination des eaux souterraines et de surface par les pesticides. Dans notre pays, l'usage des insecticides, des fertilisants, des engrais, des détergents et autres produits phytosanitaires, se répand de plus en plus avec le développement de l'agriculture, mais aussi dans le cadre des actions de lutte contre les vecteurs nuisibles **[2].**

I-2 DEFINITION

Les pesticides sont des produits destinés à lutter contre les organismes nuisibles, en particulier les mauvaises herbes (herbicides), les animaux (insecticides, acaricides, *etc*.) ou les maladies (fongicides, bactéricides, *etc*.) . D'un point de vue réglementaire, on distingue ceux utilisés pour la protection des végétaux, appelés produits phytosanitaires ou phytopharmaceutiques, de ceux utilisés pour préserver la santé humaine et animale, appelés biocides. Ainsi, un insecticide (qui est un pesticide) sera considéré comme un produit phytosanitaire s'il est utilisé sur du blé, mais comme un biocide s'il est utilisé sur du bois de charpente **[3]**. Les pesticides possèdent aussi une définition juridique. Deux termes désignent ces produits : « produits antiparasitaires » et « pesticides » **[4].**

L'utilisation des pesticides est destinée a :

- ✓ protéger les végétaux ou les produits végétaux contre tous les organismes nuisibles ou à prévenir leur action.
- ✓ exercer une action sur les processus vitaux des végétaux pour autant qu'il ne s'agisse pas de substances nutritives.

- ✓ assurer la conservation des produits végétaux pour autant que les substances ou produits ne fassent pas l'objet de dispositions particulières du conseil ou de la commission concernant les agents conservateurs.
- ✓ détruire les végétaux indésirables.
- ✓ Détruire les parties de végétaux, freiner ou prévenir une croissance indésirables des végétaux » [5].

I-3 HISTORIQUE

Avant la seconde Guerre Mondiale, les pesticides employés en agriculture étaient des dérivés de composés minéraux ou de plantes : arsenic, cuivre, zinc, manganèse, plomb, pyrèthre, roténone, sulfate de nicotine... que l'on retrouve en partie dans les cigarettes actuelles.

Les armes chimiques de la première Guerre Mondiale comme le gaz moutarde composé de chlore assura un nouveau débouché industriel pour les pesticides, une fois le conflit terminé. Ainsi, les organochlorés firent leur apparition avec de nombreuses déclinaisons qui ont connu un énorme succès.

Depuis la mise au point des herbicides de synthèse dans les années 1950, l'élimination des adventices (plantes indésirables) repose principalement sur la sélectivité des herbicides, c'est à dire leur capacité à cibler les "mauvaises herbes" sans porter préjudice aux cultures. Cependant, face à la difficulté à découvrir de nouveaux modes d'action, les industriels ont cherché, dès les années 1980, à modifier les caractéristiques des variétés cultivées pour les rendre tolérantes à l'application d'un herbicide, on parle alors de variétés tolérantes à un herbicide (VTH). Pour ce faire, trois méthodes sont utilisées (INRA, mars 2012) :

- ✪ La sélection de spécimens naturellement résistants ;
- ✪ Sélection de spécimens par mutagénèse ;
- ✪ Introduction de gènes de tolérance par transgénèse (OGM) - 95% des VTH **[6].**

Sur le tableau I.1, est donné à titre d'illustration l'historique de l'évolution des trois plus grandes familles d'activité des années 1900 à nos jours.

I-4 COMPOSITION DES PESTICIDES

Un pesticide est constitué de nombreuses molécules comprenant :

- Une (ou plusieurs) substance active responsable de l'effet du pesticide.
- Un solvant. Ce sont souvent des huiles végétales dans le cas des liquides, de l'argile ou du talc dans le cas des solides.

- Des additifs n'ayant pas d'activité biologique, mais pouvant modifier les propriétés du pesticide et faciliter son utilisation. Il y a plus de 100000 pesticides commercialisés, fabriqués à partir de 900 substances actives différentes. Chaque année, entre 15 et 20 nouvelles substances actives sont créées **[6].**

Tableau I .1 Historique de l'évolution des trois plus grandes

Evolution des pesticides			
	Herbicides	**Fongicides**	**Insecticides**
Avant 1900	Sulfate de cuivre Sulfate de fer	Soufre Sel de cuivre	Nicotine
1900-1920	Acide sulfurique		Sels d'arcenic
1920-1940	Colorants nitrés		
1940-1950	Phytohormones		Organo-chlorés Organo-phosphorés
1950-1960	Triazines,urées Substituées Carbamates	Dithiocar-bamates, Phtalimides	Carbamates
1960-1970	Dipyridyles, Toluidines**...**	Benzimidazoles	
1970-1980	Amino Phosphonates, Propionate**...**	Triazoles Dicarboximides Amide, Phosphites Morholines	Pyréthrinoides Benzoyl-urées (régulateurs de croissance)
1980-1990	Sulfonyl urées...		
1990-2000		Phénylpyrroles Strobilurines	

I-5 CLASSIFICATION

I-5-a *Selon leur mode d'action*

a- 1 Les herbicides

Destinés à limiter l'installation d'espèces végétales adventices. Peuvent être sélectifs ou totaux. Les familles de substances les plus importantes sont les acides amino phosphoriques (glyphosate), les urées (diuron, isoproturon), les triazines (atrazine, simazine). En France, plus de 300 spécialités contenant du glyphosate sont commercialisées.

a-*2 Les insecticides*

Destinés à tuer les insectes ou à empêcher le déroulement normal de leur cycle de vie. Les familles les plus rencontrées sont les organophosphorés (malathion), les carbamates insecticides (carbaxyl), les pyréthrinoïdes (deltaméthrine) et les organochlorés (endosulfan).

a-3 Les fongicides

Destinés à éliminer les champignons. On distingue trois modes d'action différents. Les multisites s'attaquent aux spores des champignons. Ils sont donc préventifs. Les unisites attaquent la perméabilité membranaire des champignons. Les antimitotiques bloquent la division cellulaire. La famille la plus présente est celle des carbamates.

a-4 Les molluscicides et autres pesticides

Les molluscicides sont destinés à éliminer les escargots et les limaces. Ils sont épandus essentiellement sous forme de granulés. Les rodenticides agissent contre les rongeurs.
Les anticoagulants représentent 85% du marché. Quelques produits de gazage sont encore utilisés. Les nématicides agissent sur les nématodes **[8].**

I-5-b Selon leur structure chimique

Les principaux groupes chimiques sont, **[9]** :

- Les organochlorés.
- Les organophosphorés.
- Les carbamates.
- Les pyrethrynoides.
- Les triazines.
- Les urées substituées**.**

I-5-c Selon leurs usages

Les pesticides à usage agricole ou produits phytopharmaceutiques sont des substances chimiques minérales ou organiques, de synthèse ou naturelles. Elles sont utilisées pour la protection des végétaux contre les maladies et contre les organismes nuisibles aux cultures.

Les pesticides à usage non agricole ou biocides sont similaires aux premiers, mais utilisés par exemple en hygiène publique (lutte anti-vectorielle) et dans d'autres applications comme la conservation du bois, la désinfection, ou certains usages domestiques. Environ 300 à 350 matières actives sont autorisées au Maroc et entrent dans la composition de près de 1000 spécialités commerciales, dont plus de 80% homologuées pour utilisation agricole [**10].**

I-5-d Selon la catégorie de danger

Les produits phytopharmaceutiques et les pesticides à usage non agricoles sont classés en "Catégories de danger" et en "classes de toxicité". Mais, en raison du degré de dilution des substances actives dans les produits (préparations) à usage du grand public, la plupart des signes relatifs à cette classification n'apparaissent pas sur les produits vendus au particulier.

Ces catégories sont définies sur base de propriétés toxicologiques et physico-chimiques : catégories "très toxique", "toxique", "nocif"*:*
Cette classification est fonction de la toxicité aiguë et est basée principalement sur la DL50 orale ou dermale ou de la CL50 du rat ou sur des essais respiratoires pour les préparations pulvérulentes (voir tableau I-2) :

Tableau I.2 Classification des pesticides en catégories de danger [11].

	Catégories de danger	**Préparations Solides**	**Préparations liquides**	**Gaz**
DL50 orale en mg/kg de poids du corps	très toxique toxique nocif	jusque 5 5-50 50-500	jusque 25 25-200 200-2000	
DL50 dermale en mg/kg de poids du corps	très toxique toxique nocif	jusque 10 10-100 100-1000	jusque 50 50-400 400-4000	
CL50 en mg/1 d'air	Très toxique toxique nocif			jusque 0,5 0,5-2 2-20

I-5-e Selon l'organisation mondiale de la santé (l'OMS)

L'Organisation mondiale de la santé (OMS) classe les pesticides par dangerosité en se basant sur leur dose létale médiane orale ou cutanée (tableau I.3). L'OMS et la FAO recommandent de ne pas utiliser les pesticides des Classes Ia et Ib dans les pays en développement. Certains bailleurs de fonds ne financent pas de projets utilisant des pesticides des Classes Ia et Ib.
Les dangers et risques des pesticides sont généralement indiqués par les fabricants sur les emballages et étiquettes des pesticides. Les précautions d'emploi aussi **[12].**
Donc chaque pesticide est alors placé dans une des quatre classes :

- Ia, extrêmement dangereux;
- Ib, très dangereux;
- II, modérément dangereux;
- III, légèrement dangereux

Tableau I.3 Classification des pesticides selon la dose létale **[13].**

La classe de toxicité	**La dose létale**
extrêmement dangereux	< 5 mg/kg
très dangerous	5 - 50 mg/kg
modérément dangereux	50 - 500 mg/kg
légèrement dangereux	> 500 mg/kg

I-5-f Selon leur origine

Les pesticides peuvent être regroupés en pesticides organiques ou inorganiques

✪ Les pesticides organiques : contiennent du carbone et peuvent être divisés en 3 groupes :

1. pesticides de synthèse (développés en laboratoire et produits en usine).
2. pesticides naturels (d'origine animale, microbienne ou végétale).
3. micro-organismes.

✪ Les pesticides inorganiques : ne contiennent du carbone que sous forme de carbonate ou de cyanure. Ce sont des dérivés à base d'arsenic, de mercure, de fluor, de soufre et de cuivre, ainsi que des dérivés du cyanure. Les pesticides inorganiques sont dérivés essentiellement de minéraux **[13].**

I-6 GENERALISATION DE L'UTILISATION MONDIALE DES PESTICIDES

Le marché mondial (environ 40 milliards de dollars) est globalement stable depuis quelques années (2000). Il faut noter que certains évènements climatiques récents (chaleur et sécheresse en Europe, pluie en Océanie) influencent fortement ces chiffres.

En Europe et en Amérique du Nord, les herbicides représentent 70 à 80% des produits utilisés (notamment à cause de la forte augmentation des cultures de maïs) tandis que sous les tropiques, 50% des produits appliqués sont des insecticides. La diversification des cultures, avec l'amélioration du niveau de vie dans certains pays, modifie également cet équilibre, ainsi la Chine a converti l'équivalent de la surface de l'Angleterre de rizières en cultures maraîchères, entraînant une diversification des produits mis en œuvre.

Les usages non-agricoles représentent environ 12% du marché global (dont plus d'un tiers des usages pour les Etats-Unis). En France, les usages non agricoles représentent entre 6 et 10% de la consommation totale.

Les pesticides et les engrais ont permis, depuis les années 60, de multiplier la productivité par 3, les pertes occasionnées aux cultures représentent pourtant encore près de 30% en Europe pour le maïs contre 50% en Afrique ; moins de 30% en Asie pour le riz contre plus de 50% en Afrique. Malgré les efforts considérables réalisés, il reste encore dans le Monde près d'un milliard de personnes mal nourries, et il faudra d'ici à 2050 nourrir une population qui aura progressée de près de 50% **[14].**

Les figures I.1et I.2 illustrent respectivement Le marché mondial des pesticides par région du monde et celui des pesticides par catégorie de produit

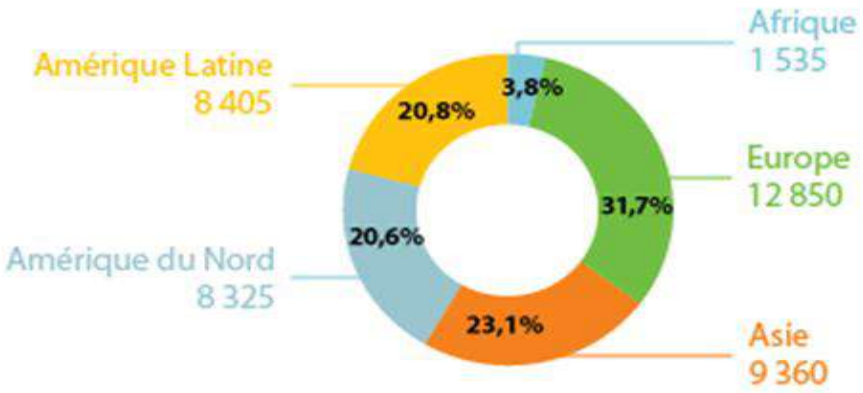


Figure I.1 Le marché mondial des pesticides par région du monde **[14].**

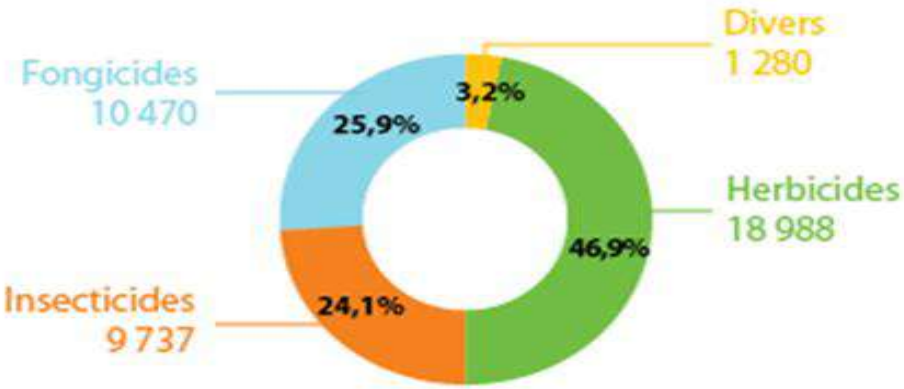


Figure I.2 Le marché mondial des pesticides par catégorie de produit **[14]**.

I-7 LES PESTICIDES DANS LES PAYS EN DEVELOPPEMENT

Au moment où l'Amérique du Nord, l'Europe occidentale et d'autres régions fortement industrialisées étaient en train de mettre en place des régimes nationaux de réglementation des pesticides élaborés en réponse à la pression locale du public, des sociétés basées dans ces régions ont activement promu les exportations de pesticides vers les pays en développement, notamment vers les pays avec des régimes de réglementation faible ou inexistante. Les bénéficiaires disposent d'une main d'œuvre agricole peu au fait des dangers sanitaire et écologique potentiels que peuvent induire les pesticides.

Un certain nombre d'organisations philanthropiques ont dans un premier temps soutenu l'exportation de pesticides dans les pays en développement. Cela s'inscrit dans le cadre des efforts destinés à accroître la productivité agricole et à lutter contre la famine et la faim. Quant à la plupart de ceux qui soutiennent l'utilisation accrue de pesticides, ils sont, cependant, principalement motivés par le profit. Beaucoup de grands propriétaires terriens et les exploitants de plantations ont commencé à utiliser de grandes quantités de pesticides avant de se rendre compte des méfaits que les pesticides peuvent occasionner. Beaucoup, cependant, ont fait montre d'un manque d'intérêt notoire même après s'être clairement rendu compte de leurs méfaits. Ces deux cas sont vrais pour les fabricants de pesticides, les exportateurs et les négociants.

Dans beaucoup de pays, de nombreux petits agriculteurs et les paysans sont également devenus de plus en plus dépendants des produits agrochimiques. Souvent, avant cette dépendance, ces petits producteurs n'étaient pas pleinement intégrés dans l'économie de marché. Ensuite, leur besoin d'argent est considérablement allé crescendo. Parfois, l'utilisation des pesticides a aidé les paysans à accroître leurs rendements afin de couvrir suffisamment ces nouveaux coûts de production, même si ce n'est pas toujours le cas. En outre, les petits agriculteurs qui sont devenus dépendants d'intrants extérieurs coûteux sont

devenus de plus en plus vulnérables aux forces du marché qu'ils ne pouvaient pas contrôler **[15].**

Dans les pays sous-développés, même les produits très toxiques, dont l'usage a été interdit dans les pays riches, sont encore largement utilisés, et avec beaucoup moins de précautions. Selon un communiqué de presse de la FAO (1er février 2001), environ 30% des pesticides commercialisés dans les pays en voie de développement ne sont pas conformes aux standards de qualité internationaux, car ils contiennent beaucoup d'impuretés très toxiques **[16].**

I-8 DEVENIR DES PESTICIDES

Lors d'un épandage, une grande partie du produit n'atteint pas l'organisme ciblé. Par exemple, sur une surface inclinée, le ruissellement des eaux de pluies peut entraîner le pesticide loin de sa cible. Il y a alors gaspillage du produit chimique, diminution de l'effet et risque de pollution du sol et des eaux.

De nombreux phénomènes influencent le devenir des pesticides dans l'environnement:

- Phénomènes de transfert

L'adsorption est un phénomène de surface par lequel les molécules se fixent aux particules du sol. La quantité de pesticide adsorbé varie selon le type de pesticide, la nature du sol, le pH du sol…. Les pesticides s'adsorbent facilement sur des sols riches en argile ou en matière organique. Les pesticides adsorbés sont moins susceptibles de se vaporiser ou de migrer dans le sol. Ils sont aussi plus difficilement captés par les plantes **[17].**

La volatilisation est la transformation des solides ou des liquides en gaz. Ce processus peut disperser une grande partie du produit épandu dans l'atmosphère. Ce mouvement est appelé dérive gazeuse. La dérive gazeuse de certains herbicides peut endommager des cultures voisines. Un temps chaud ou venteux accélère le phénomène de volatilisation. La dérive à l'épandage est le déplacement aérien des gouttelettes de produit phytosanitaire vaporisées lors du traitement des cultures. Ce phénomène peut entraîner les pesticides sur de grandes distances. La dérive à l'épandage dépend de : la taille des gouttelettes (plus celles-ci sont petites, plus elles seront entraînées), la vitesse du vent, la distance entre la buse de vaporisation et la plante. La dérive peut contaminer de grandes zones autour des sites traités **[17].**

Le ruissellement est l'entraînement des pesticides par l'eau sur des surfaces inclinées. Les pesticides étant soit mélangés à l'eau soit adsorbés sur les particules du sol érodé. L'infiltration est l'entraînement des pesticides par l'eau dans le sol. Elle peut se faire vers le bas, le haut ou horizontalement **[17].**

L'absorption est l'assimilation des pesticides par les plantes et les microorganismes. Une fois absorbé, le composé peut être dégradé ou peut subsister dans l'organisme et être relâché dans l'environnement lorsque l'animal meurt ou que la plante se décompose. De plus, lors des récoltes les résidus de pesticides peuvent également être déplacés **[17]**.

- Phénomènes de dégradation

Les pesticides sont dégradés par les microorganismes, par des réactions chimiques et par le rayonnement solaire. La dégradation peut prendre de quelques heures à plusieurs années selon les conditions environnementales et les caractéristiques physico-chimiques du pesticide. Les pesticides qui se dégradent le plus rapidement sont ceux présentant le moins de risques pour l'environnement.

La dégradation microorganique se fait par les champignons et les bactéries. La dégradation chimique se fait par réactions chimiques dans le sol, l'eau et l'air. La photodégradation quant à elle se fait par le rayonnement solaire. La vitesse de dégradation varie selon l'intensité lumineuse et les propriétés du pesticide. Les pesticides appliqués sur les feuillages sont plus exposés au rayonnement que ceux incorporés au sol **[17].**

I-9 TOXICITE DES PESTICIDES

La toxicité aigüe d'une substance chimique est estimée par des séries de tests réalisés sur des animaux de laboratoire. La valeur la plus utilisée est la dose létale 50 (DL50) qui correspond à la quantité de substance active qui, administrée en une fois, par ingestion, inhalation ou voie cutanée, entraîne la mort de 50% des animaux testés. Plus la DL 50 est basse, plus le produit est dangereux.

La toxicité chronique d'une substance chimique administrée en doses répétées est également estimée par des tests normalisés sur des animaux de laboratoire. L'effet carcinogène et l'effet sur la reproduction sont ainsi estimés. Par ces études, une dose sans effet (NOAEL) peut être déterminée. Cela correspond à la dose maximale n'entraînant pas d'effet statistiquement significatif sur le groupe testé. Tous ces tests permettent la détermination d'une Dose Journalière Acceptable (DJA ou ADI en anglais) qui indique la quantité de substance qu'un humain peut ingérer chaque jour, durant toute sa vie, sans danger pour sa santé **[17].** Sur le tableau I.4 est donné le classement toxicologique des pesticides.

Tableau I.4 Classement toxicologique d'un produit **[35]**.

Valeur de référence	**Ingestion DL50 (rat, orale) (mg/kg)**		**Pénétration cutanée DL50 (rat, cutanée) (mg/kg)**		**Inhalation CL50 (mg/l)**
formulation	liquide	Solide	liquide	solide	gazeux
Très toxique (T+)	<25	<5	<50	<10	<0.5
Toxique (T)	>25 et <200	>5 et <50	>50 et <400	>10 et <100	>0.5 et <2
Nocif (Xn)	>200 et < 2000	>50 et <500	>400 et 4000	>100 et <1000	>2 et <20

I-10 IMPACTS CHRONIQUES SUR LA SANTE HUMAINE ET AMPLEUR DU PROBLEME

Les pesticides dangereux ne sont pas seulement un problème du monde en développement. Les pays fortement industrialisés continuent d'utiliser de grandes quantités de pesticides dangereux. Ces derniers continuent encore de causer des agressions sur la santé et l'environnement. Pratiquement tous les pays ont besoin de réformes supplémentaires pour réduire et éliminer les préjudices causés par l'exposition aux pesticides. Néanmoins, l'exposition à des pesticides dangereux est un problème particulièrement grave dans beaucoup de pays en développement. Dans de nombreux pays, l'exposition aux pesticides toxiques est un grave problème de santé avec des proportions épidémiques **[15].**

I-10-a L'ampleur du problème

Plus d'un tiers de tous les travailleurs du monde entier sont dans l'agriculture, mais le pourcentage varie considérablement selon les régions. Dans des pays fortement industrialisés, l'agriculture occupe généralement moins de 5% de la main-d'œuvre. En Amérique latine, au Moyen-Orient et dans les pays de l'ex-Union soviétique, près d'un cinquième (20%) des travailleurs sont employés dans l'agriculture. En Afrique du Nord et l'Asie de l'Est, le nombre s'élève à plus d'un tiers. En Asie du Sud et du Sud-est, près de la moitié de la main-d'œuvre est dans l'agriculture. Et en Afrique sub-saharienne, les deux tiers de tous les travailleurs sont utilisés dans les activités agricoles **[18].**

Un récent rapport de la Banque mondiale estime que 355.000 personnes dans le monde meurent chaque année d'empoisonnement involontaire dû à des pesticides **[19]**. Une étude plus ancienne faisant autorité **[20].** Estime qu'il y a peut-être un million de cas

d'intoxications involontaires graves causés par les pesticides chaque année, et deux millions de cas de personnes supplémentaires hospitalisées pour cause de tentatives de suicide avec des pesticides. L'auteur note que cela ne reflète forcément qu'une infime partie du véritable problème et estime qu'il pourrait y avoir jusqu'à 25 millions de travailleurs agricoles dans les pays en développement souffrant d'une forme d'intoxication au travail chaque année à cause de pesticides, bien que la plupart des incidents ne soient pas enregistrés et la plupart des patients ne cherchent pas des soins médicaux. Une des conclusions de cet auteur est que les intoxications par les pesticides peuvent constituer dans certains pays en développement un grave problème de santé publique comme le sont les maladies contagieuses **[21].**

I-10-b Les pesticides à toxicité aiguë

Le pesticide aldicarbe qui a été pris dans l'exemple ci-dessus est l'un des nombreux pesticides qui sont classés comme très dangereux. Cette classification remonte à une décision de 1973 de l'Organisation mondiale de la Santé de mettre au point un système de classification des pesticides en fonction des risques qu'ils posent. Dans cette décision, le terme «pesticides dangereux» est défini comme un pesticide qui fait peser un grave risque sur la santé, c'est-à-dire un pesticide qui peut causer des préjudices graves à la suite d'une exposition pendant une courte période.

L'OMS tient une liste à jour des pesticides très dangereux qui sont très toxiques dans sa publication: *The Recommended Classification of Pesticides Hazard*. **[22]** L'OMS utilise comme principale méthode pour identifier les pesticides qui posent un risque grave pour la santé un test dans lequel les rats sont nourris avec le pesticide ou exposés à travers la peau. Ce test produit une valeur appelée DL50 (dose létale 50%), qui est une estimation statistique du nombre de milligrammes de la substance toxique par kilogramme de poids corporel qui va tuer 50% d'une population importante de cobayes. Dans ce système, l'OMS considère un pesticide extrêmement dangereux (classe Ia) ou très dangereux (classe Ib) si elle a une valeur DL 50 faible **[23].**

I-10-c Impacts chroniques sur la santé

Au cours des dernières années, le terme pesticide extrêmement dangereux a été élargi et est maintenant utilisé pour décrire non seulement les pesticides très toxiques, mais aussi les pesticides qui causent de graves effets chroniques sur la santé. Il est généralement beaucoup plus difficile de démontrer les effets chroniques sur la santé que de démontrer des effets toxiques aigus, mais d'importantes recherches sont en cours. Quand il y a une preuve sérieuse qui établit des liens entre un pesticide et un effet chronique grave sur la santé, ce pesticide est ainsi considéré comme un pesticide extrêmement dangereux. Les effets chroniques sur la santé liés aux pesticides concernent les cancers et tumeurs, les troubles du système nerveux, des problèmes de reproduction, les effets sur le système immunitaire, la perturbation du système endocrinien **[24].**

i. Le cancer

Il est connu que beaucoup matières actives des pesticides sont à l'origine ou sont soupçonnés d'être à l'origine du cancer. Les pesticides sont associés, soit par des éléments de preuve en laboratoire ou des études épidémiologiques, à une longue liste de cancers, y compris le myélome multiple, les sarcomes des tissus mous, le sarcome d'Ewing, le lymphome, le lymphome non hodgkinien, la leucémie, le mélanome, le neuroblastome ou de la tumeur de Wilm, les tumeurs à cellules germinales, le rétinopathie (tumeur de l'oeil) et le cancer de l'oesophage, de l'estomac, de la prostate, du testicule, du sein, de l'ovaire, du col de l'utérus, de la vessie thyroïde, des poumons, du cerveau, des reins, du pancréas, du foie, du colon et du rectum **[25 - 26].**

ii. Troubles du système nerveux

L'exposition aux pesticides a été associée à l'altération du développement du système nerveux qui peut entraîner une baisse de l'intelligence et des troubles du comportement **[27].** Il existe des preuves reliant divers pesticides aux effets sur le système nerveux central, le système nerveux périphérique et le développement prénatal du cerveau, notamment :

- Un développement plus lent des aptitudes et une augmentation de l'agressivité chez les enfants.
- Des effets dépressifs qui peuvent entrainer des suicides. Une neuropathie retardée, impliquant une dégénérescence des nerfs périphériques des membres avec des maux et des douleurs musculaires et des symptômes grippaux.
- Le changement de personnalité, les troubles de concentration et de mémoire, les troubles du langage, le sens de l'odorat très développé, la détérioration de l'écriture, une altération de la tolérance à l'exercice et les déficits neuromusculaires.
- La maladie de Parkinson et le parkinsonisme - des troubles avec des symptômes tels que la maladie de Parkinson qui peuvent toutefois être réversibles **[26].**

iii. Les problèmes de la reproduction

Certains pesticides peuvent être liés à un certain nombre de problèmes de reproduction, y compris des anomalies congénitales, la stérilité, les grossesses prolongées, l'avortement spontané, les mort-nés, les naissances prématurées, le retard de croissance intra-utérin, mortalité périnatale, l'endométriose, et la baisse de spermatozoïdes. Des études épidémiologiques ont établi un lien entre l'exposition des parents à certains pesticides et des anomalies du tube neural, des malformations congénitales cardiaques, la fente labiale et la fente palatine, les malformations congénitales musculo-squelettiques, anomalies du tractus urinaire et anomalies des organes génitaux masculins **[21].**

iv. Les effets sur le système immunitaire

L'exposition aux pesticides peut compromettre le système immunitaire ce qui augmente le risque de maladies infectieuses et de cancer. Cela devient particulièrement préoccupant quand le système immunitaire est déjà compromis par d'autres facteurs, telle que la malnutrition **[26].** Certains pesticides comme le thymus et la rate altèrent le développement

Des organes du système immunitaire et portent atteinte à la capacité des globules blancs et des lymphocytes à tuer les bactéries, les virus et les cellules cancéreuses **[26].**

v. Perturbation du système endocrinien

Certains pesticides, avec de très petites doses, peuvent imiter ou bloquer les hormones ou peuvent déclencher une activité hormonale inappropriée. Cela peut en rajouter aux effets nocifs sur la santé telle que la stérilité, la baisse du nombre de spermatozoïdes et le cancer du sein **[27].** D'autres effets indésirables éventuels sur la santé incluent l'avortement spontané, le sex-ratio biaisé dans la descendance des communautés exposées, les anomalies du tractus reproducteur mâle et femelle, y compris des malformations génitales et d'autres anomalies congénitales, la puberté précoce, le syndrome des ovaires polykystiques, une atteinte de la fonction immunitaire ainsi qu'une variété de cancers. Une étude a identifié 127 pesticides soupçonnés d'avoir des effets perturbateurs sur le système endocrinien **[27].**

vi. Le suicide

L'auto-intoxication aux pesticides dangereux est probablement la méthode de suicide la plus fréquemment utilisée dans le monde et constitue une pratique largement confinée dans les pays en développement **[28].** Les recherches montrent que beaucoup d'individus qui tentent de se suicider ne veulent pas mourir.

Dans les pays très industrialisés, une forme courante de tentative de suicide réside dans la surdose de somnifères ou de drogues similaires. Le taux de mortalité résultant de telles tentatives peut baisser jusqu'à 0,5%, et les études menées dans les pays industrialisés constatent que seulement 2% se suicident au cours des 12 mois qui suivent la première tentative. D'autre part, une tentative de suicide par ingestion de pesticides agricoles peut entrainer des taux de mortalité au-delà de l'ordre de 10% à 20% **[29].**

Selon les statistiques de l'OMS, 18 millions de personnes tentent de se suicider chaque année et cela se traduit par plus de 800.000 morts **[30].** Selon un dépliant de l'OMS sur la prévention du suicide :

> *«On trouve le taux le plus élevé de comportement suicidaire chez les femme dans les zones rurales des pays asiatiques, où le facteur de risque principal est la disponibilité des pesticides, qui transforme les tentatives de suicide impulsives par empoisonnement en actes meurtriers en moins de 3 heures»* **[21].**

Ce dépliant liste les interventions qui ont démontré leur efficacité dans la prévention des comportements suicidaires. Le premier point de la liste est : «Contrôle de l'accessibilité des substances toxiques (en particulier les pesticides dans les zones rurales des pays asiatiques**) [21].**

I-11 EFFETS SUR L'ENVIRONNEMENT

Les pesticides employés pour lutter contre les organismes nuisibles peuvent se retrouver dans l'environnement. Ils risquent alors d'engendrer une contamination ponctuelle ou diffuse. Cette pollution diffuse semble être la source dominante d'apport de pesticides vers les eaux de surface et souterraines .Celle-ci peut être associée à plusieurs mécanismes de transport qui permettent aux pesticides de quitter la parcelle agricole et de se retrouver dans l'environnement. La figure I.3 illustre ces mécanismes.

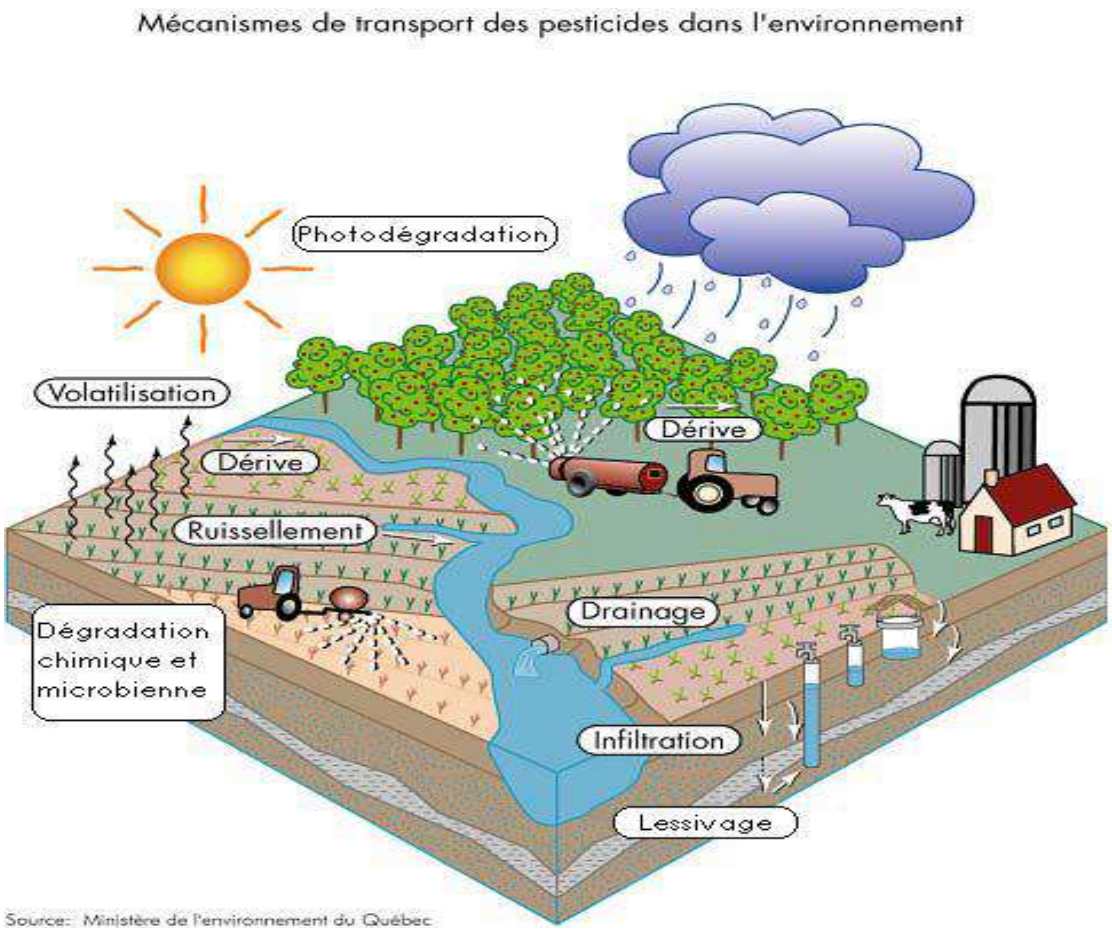


Figure I.3 Mécanismes de transport des pesticides dans l'environnement **[31]**

Les principaux mécanismes de transport des pesticides vers les eaux de surface et souterraines sont le ruissellement (pesticide sous forme dissoute ou lié à des particules de sols) et le lessivage. Les caractéristiques physicochimiques propres à chacun des pesticides - telle la persistance, l'adsorption, la pression de vapeur et la solubilité du pesticide - influencent sa tendance à quitter la parcelle. Les caractéristiques du sol et de la nappe d'eau souterraine, les précipitations, le travail du sol, les méthodes d'application, les pratiques culturales et le choix du pesticide jouent également un rôle important sur les pertes de pesticides par ruissellement et lessivage. Le drainage du surplus d'eau qui s'infiltre dans les sols agricoles par les canaux de drainage souterrain entraîne aussi les pesticides vers les eaux de surface.

Les pesticides peuvent aussi être transportés par l'air. Lors de l'application, la dérive des gouttelettes de pesticides fait en sorte que des produits peuvent se retrouver à l'extérieur de la parcelle. Le phénomène de la dérive est influencé, entre autres, par la taille des gouttelettes pulvérisées, la vitesse du vent, le type d'équipement utilisé et la hauteur de la pulvérisation.

Les pesticides peuvent également se retrouver dans l'air sous forme de vapeurs, un phénomène appelé la volatilisation. Après l'application de certains pesticides, ceux-ci sont libérés sous forme de gaz ou adhèrent à des particules comme les poussières. Les résidus de pesticides sous forme gazeuse dans l'atmosphère ont souvent une demi-vie de courte durée, car ils sont susceptibles à la photodégradation et sont transportés sur de courtes distances. Par contre, les pesticides moins sujets à la photodégradation et qui ont une demi-vie beaucoup plus longue. Par la suite, ces résidus de pesticides sous forme gazeuse peuvent être soustraits de l'atmosphère par condensation et retourner dans l'eau ou sur le sol par l'effet des précipitations ou par le dépôt de fines particules solides.

La volatilisation du pesticide à partir des surfaces traitées est principalement fonction de la chaleur, de la pression de vapeur associée au pesticide et de l'humidité relative. La présence de certains pesticides (atrazine, métolachlore, EPTC) dans l'eau de pluie. Une fois dispersés dans l'environnement, certains pesticides peuvent être dégradés par la lumière ou les micro-organismes, tandis que d'autres persistent et peuvent s'accumuler ou se transformer en d'autres contaminants parfois plus à risque que le produit d'origine **[31].** Sur la figure I.4, est illustré le processus impliqué dans le devenir des pesticides dans les sols conditionnant leur disponibilité et, par conséquent, leur efficacité phytosanitaire ou la manifestation de leur caractère polluant.

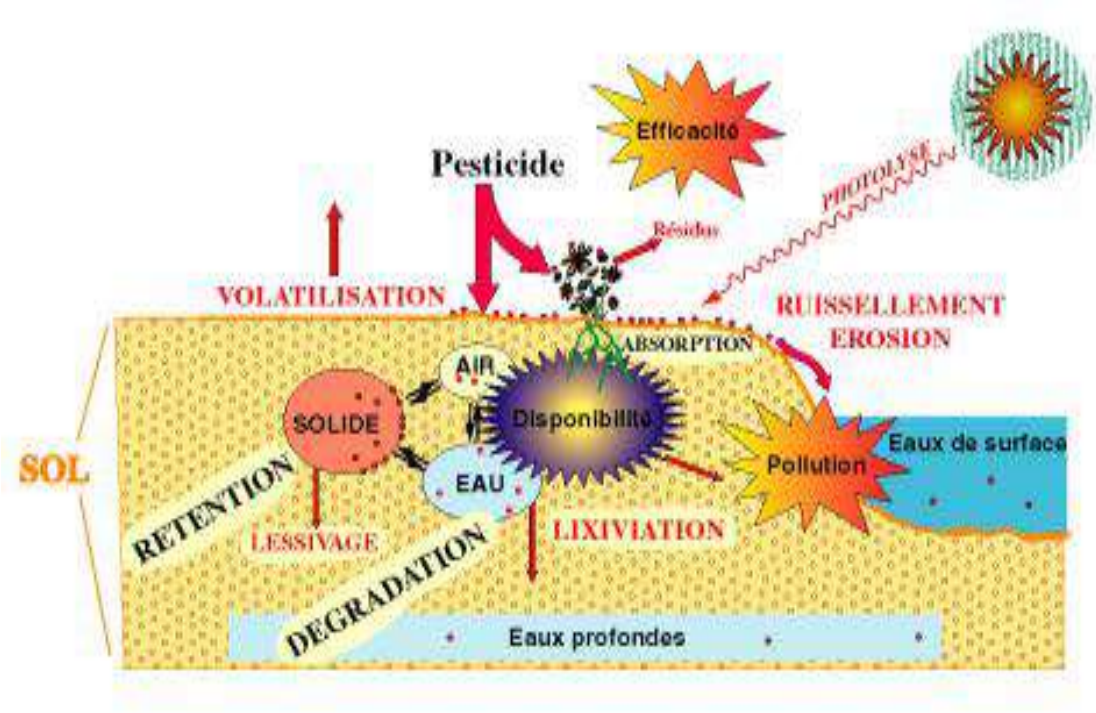


Figure I.4 Processus impliqué dans le devenir des pesticides dans les sols conditionnant leur disponibilité et, par conséquent, leur efficacité phytosanitaire ou la manifestation de leur caractère polluant **[32].**

Des études émanant de divers groupes de recherche ont été menées en France depuis la fin des années 80 et se poursuivent à l'heure actuelle. Les premières données permettent de constater la présence de pesticides dans toutes les phases atmosphériques, en concentrations variables dans le temps (caractère parfois saisonnier, lien avec les périodes d'application) et l'espace (proximité des sources) ; même des composés peu volatils ou interdits (lindane par exemple) sont parfois détectés.

I-12 UTILISATION DES PESTICIDES EN ALGERIE

I-12-1 Situation géographique

L'Algérie est bordée au nord par la mer Méditerranée, à l'est par la Tunisie et la Libye, au sud-est par le Niger, au sud-ouest par le Mali et la Mauritanie, à l'ouest par le Maroc, et le Sahara Occidental.

L'Algérie est le premier pays d'Afrique par sa superficie (2.381.741 km²), dont les quatre cinquièmes sont occupés par le Sahara. Le pays comprend deux grands ensembles géographiques : les chaînes de l'Atlas, au nord, et le Sahara, au sud. Entre les massifs de L'Atlas tellien, ou Tell algérien (Kabylie, Mascara, Ouarsenis, Saïda, Tlemcen, etc.), s'insèrent des plaines étroites et discontinues en bordure d'une côte très découpée qui s'étire sur près de 1200 km **[33].**

I-12-2 Agriculture

Défavorisée par un climat aride - les terres arables ne constituent que 3% du territoire- l'agriculture souffre de la dégradation des sols et de la faiblesse de l'irrigation. Elle a en outre été délaissée dans les années 1960 en faveur de l'industrie. Aujourd'hui, le secteur primaire emploie environ 26% de la population active et représente 8,6% du PIB (2000). Les principales cultures sont les céréales, qui occupent 35% des terres arables, notamment le blé (2,60 millions de tonnes en 2004 mais avec une faible productivité : 794kg/ha) et l'orge (1.200.000 tonnes). Les autres productions agricoles importantes incluent la pomme de terre, la vigne, les agrumes, les olives et les dattes. L'élevage constitue un secteur actif et dynamique.

En 2001, le cheptel se compose d'environ 18,7 millions d'ovins (14e rang mondial), 3,2 millions de caprins, 1,6 million de bovins et de 115 million de volailles. En 2000, les forêts couvrent 0,9% de la superficie du pays. D'importants projets de reboisement ont été entrepris dans les années 1970. Les coupes restent modestes : 7,5 millions de m3 en 2003. La pêche est un secteur actif : en 2001, les prises représentaient 100.281 tonnes de poissons (sardines, anchois, thons et fruits de mer) **[33].**

I-12-3 Les principaux pesticides utilisés en Algérie [36].

Nous donnons sur les tableaux I.4 et I.5 ci-dessous quelques exemples sur chaque famille de pesticides utilisés respectivement en Algérie et dans la région de Médéa.

Tableau I.5 Les pesticides les plus utilisés en Algérie **[34].**

Nom commercial	type	Matiere active
Probiénébe	Fongicide	Probinébe
Manébe	Fongicide	Manébe
Lannate	Insecticide	Méthomyl
Karaté	Insecticide	Lamda-cyhalathrine
Decis	Insecticide	Delthamitrine
2,4 D	Herbicide	Acide dichloro-2,4 Phenoxyl

Tableau I.6 Les pesticides utilisés dans la région de Médéa **[36].**

Herbicides					
Nom de molécules	Teneur en substance active	usages	Doses en produit	Délais avant récolte	Toxicité classement
Alcaudon 50 wp	Linuron 50%	Mauvaises herbes: pomme de terre, carotte, fève,	1-2.5Kg /HA	7 jours	Nocif : dangereux pour les poissons
Mamba 360 SL	Glyphosate 360 g/l	• Graminées dicotylédones annuelles bisannuelles et vivaces (vergers, Arboriculture,..) • Adventices des terrains non cultivés, drains, seguias friches, abords de fermes, abords d'usines et sur les chaumes.	6-9 L /HA 8-10 L /HA	/	Nocif
Fongicide					
Quinolate 15	Oxyquinoleate de cuivre 15%	• Carie et septoriose du blé • Fusariose et charbon nu du blé • Charbon de l'avoine • Helminthosporiose, charbon nu et couvert de l'orge	200 grs / quintal	/	Nocif
Oxycure 50%	50% Oxychlorure de cuivre	• Mildiou de la vigne • Mildiou de la pomme de terre • Mildiou des cultures légumières • Anthrachnose et alternaria des cucurbitacées • Septoriose des tomates et poivrons	300-500 3-4 300-500 200-400 300-500	20 jours 20 jours 20 jours 20 jours 20 jours	Nocif
Insecticides					
Beida zit	Huile minérale 76%	• Cochenilles noires des oliviers • Araignée rouge et cochenilles des agrumes • Pou rouge des arbres fruitiers	1.5-2 L/HL 2 L/HL 1.5 L/HL	/	Dangereux
Fostan 5%	Chlorpyriphos 5%	• Vers blancs, vers gris et insectes du sol. -céréales. -pomme de terre. -tabac	30-40 KG/HA	21 jours	toxique

I-13 CONCLUSION

Comme nous venons de le voir, l'exposition de l'humanité et de son environnement aux pesticides est une source de graves problèmes de santé dans toutes les régions du monde, mais surtout dans les pays en développement. Aussi, tous les efforts doivent être conjugués afin de minimiser ou d'éliminer carrément l'utilisation de pesticides reconnus dangereux pour l'homme, la faune et la flore. D'un autre côté, les efforts des scientifiques axés sur la recherche de méthodes de prédiction de la toxicité des nouveaux pesticides doivent être soutenus et encouragés.

Chapitre IV
Modélisation Neuronale de la Toxicité des Pesticides

IV-1 INTRODUCTION

Dans ce chapitre nous présentons les résultats préliminaires de la modélisation de la toxicité des pesticides par les réseaux de neurones artificiels.
Ce chapitre est divisé en trois parties : la première partie concerne l'élaboration de la base de données. Dans la deuxième partie on s'intéressera à la modélisation de deux familles de descripteurs (descripteurs constitutionnels et propriétés moléculaire) par les réseaux de neurones de type perceptron multicouches. Dans la troisième partie on étudiera la contribution des variables d'entrée sur la toxicité (analyse de sensibilité). Enfin on conclura ce chapitre en analysant et interprétant les différents résultats obtenus.

IV-2 METHODOLOGIE DE CALCUL

La méthodologie de calcul par réseaux de neurones de la toxicité est illustrée par l'organigramme de la figure IV.1.

IV-2-1 Elaboration de la base de données

Dans cette étape nous avons sélectionné 72 molécules de pesticides (Tableau IV.1). L'élaboration de la base de données est décrite par le diagramme de la figure IV.2.

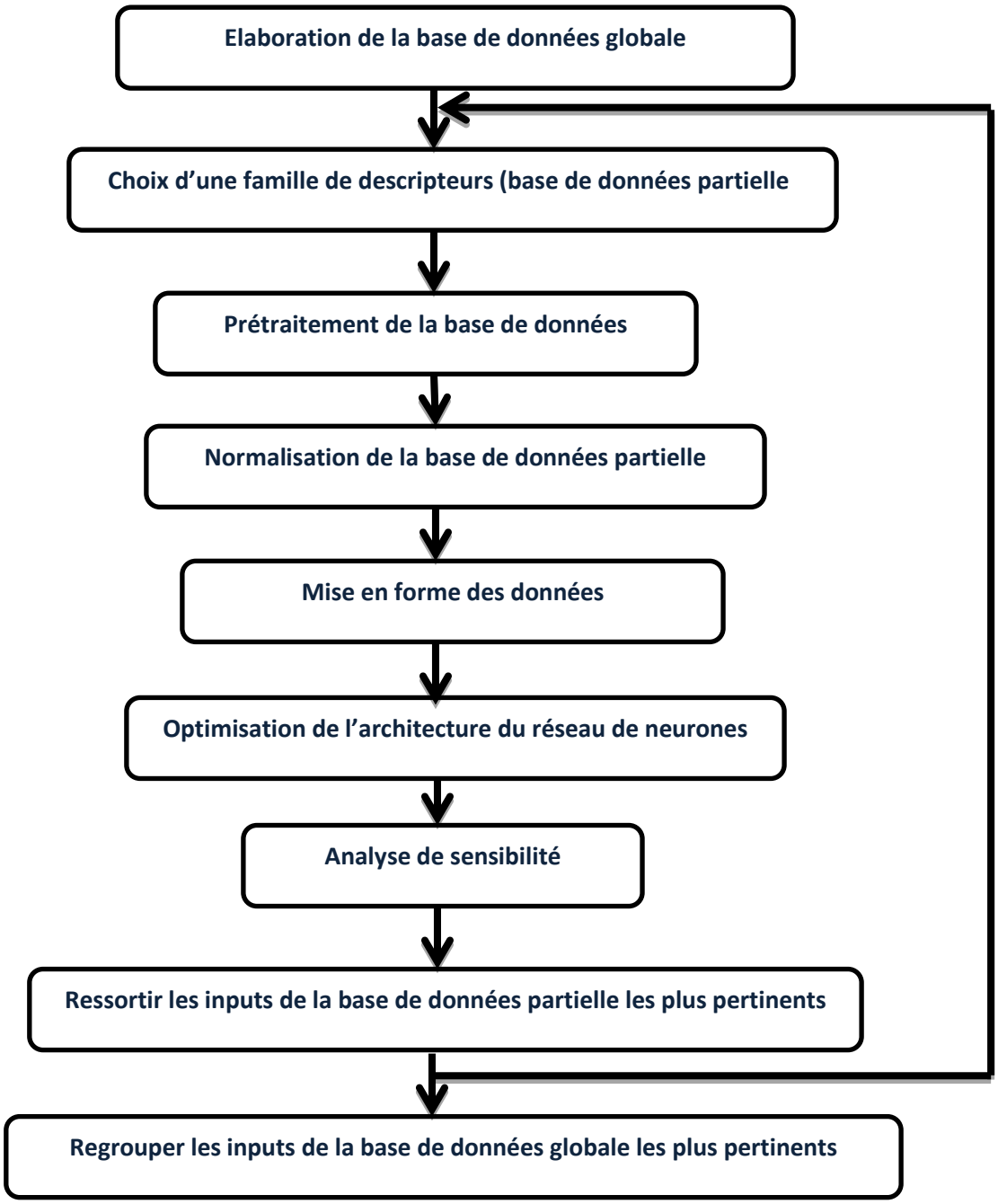


Figure IV.1 Méthodologie de calcul de la toxicité par les réseaux de neurones artificiels.

Tableau IV.1 Liste des pesticides utilisés dans cette étude.

Les pesticides		
PESTICIDES [79]	**NOTATION SMILE [80]**	**LogLD 50 [81]**
Organophosphate Insecticides		
o Acephate	O=C(NP(=O)(OC)SC)C	6,8512
o Azinphos-methyl	S=P(OC)(OC)SCN1\N=N/c2ccccc2C1=O	1,5261
o Bensulide	O=S(=O)(NCCSP(=S)(OC(C)C)OC(C)C)c1ccccc1	5,8861
o Bromophos	Clc1cc(OP(=S)(OC)OC)c(Cl)cc1Br	7,3778
o Carbophenothion	COP(=S)(OC)SCSC1=CC=C(C=C1)Cl	4,3745
o Chlorpyrifos	Clc1c(OP(=S)(OC)OC)nc(Cl)c(Cl)c1	5,4072
o Chlorthiophos	Clc1c(OP(=S)(OC)OC)ccc(SC)c1Cl	2,3702
o Cyanophos	S=P(Oc1ccc(C#N)cc1)(OC)OC	6,5653
o Demeton-S-methyl	O=P(OC)(SCCSCC)OC	3,6889
o Diazinon	S=P(OCC)(OCC)Oc1nc(nc(c1)C)C(C)C	5,5215
o Dichlorvos	O=P(O\C=C(/Cl)Cl)(OC)OC	4,382
o Dicrotophos	O=P(O/C(=C/C(=O)N(C)C)C)(OC)OC	3,0445
o Dimethoate	O=C(NC)CSP(=S)(OC)OC	5,8805
o Disulfoton	S=P(OCC)(SCCSCC)OCC	1,9169
o EPN	S=P(Oc1ccc(cc1)[N+]([O-])=O)(OCC)c2ccccc2	3,5835
o Ethoprophos	O=P(SCCC)(OCC)SCCC	4,1109
o Fenitrothion	S=P(Oc1cc(c(cc1)[N+]([O-])=O)C)(OC)OC	7,4384
o Fensulfothion	S=P(Oc1ccc(cc1)S(=O)C)(OCC)OCC	1,3863
o Isoxathion	S=P(Oc2noc(c1ccccc1)c2)(OCC)OCC	5,4889
o Leptophos	c1([P@@](Oc2c(cc(Br)c(c2)Cl)Cl)(OC)=S)ccccc1	3,9665
o Malathion	O=C(OCC)C(SP(=S)(OC)OC)CC(=O)OCC	1,7652
o Merphos	S(P(SCCCC)SCCCC)CCCC	0,3887
o Methamidophos	O=P(OC)(SC)N	3,2189
o Methidathion	O=C1SC(=N/N1CSP(=S)(OC)OC)\OC	3,434
o Methyl parathion+	[O][N+](C1=CC=C(OP(OC)(OC)=S)C=C1)=O	2,6391
o Monocrotophos	O=P(O/C(=C/C(=O)NC)C)(OC)OC	2,8904
o naled	ClC(Br)(Cl)C(Br)OP(=O)(OC)OC	5,5215
o Ronnel	ClC1=C(Cl)C=C(Cl)C(OP(OC)(OC)=S)=C1	7,1309
o Schradan+	P(N(C)C)(N(C)C)(OP(N(C)C)(N(C)C)=O)=O	2,6027
o Temephos	S=P(OC)(OC)Oc2ccc(Sc1ccc(OP(=S)(OC)OC)cc1)cc2	9,0595
o Terbufos	S=P(OCC)(SCSC(C)(C)C)OCC	0,6931
o Tetrachlorvinphos	Clc1cc(C(OP(=O)(OC)OC)=[C@H]Cl)c(Cl)cc1Cl	7,2998
o Trichlorfon	COP(=O)(C(C(Cl)(Cl)Cl)O)OC	6,4457

N-Methyl Carbamate Insecticides		
o Carbofuran	CC2(C)CC1=C(C(=CC=C1)OC(NC)=O)O2	2,0869
o Dicofol	Clc1ccc(cc1)C(O)(c2ccc(Cl)cc2)C(Cl)(Cl)Cl	6,3886
o Dieldrin	Cl\C1=C(/Cl)C2(Cl)C3C4CC(C3C1(Cl)C2(Cl)Cl)C5OC	3,8501
o Lindane	ClC1C(Cl)C(Cl)C(Cl)C(Cl)C1Cl	4,4773
o Methomyl	C/C(=N\OC(=O)NC)SC	2,8332
o Oxamyl	O=C(C(=N\OC(=O)NC)\SC)N(C)C	1,1314
o Propoxur	O=C(Oc1ccccc1OC(C)C)NC	4,4188
Solid Organochlorine Insecticides		
o Aldrin	Cl\C3=C(/Cl)C4(Cl)C2C(C1/C=C\C2C1)C3(Cl)C4(Cl)Cl	3,6636
o Endosulfan	Cl\C1=C(/Cl)C3(Cl)C(Cl)(Cl)C1(Cl)C2C3COS(=O)OC2	4,7005
Biologicals and Insecticides of Biological Origin		
o Pyrethrin II	COC(/C(C)=C/[C@H]1C(C)([C@]1([H])C(O[C@@H]2C(C)=C(C/C=C/C=C)C(C2)=O)=O)C)=O	7,7706
OtherInsecticides, Acaricides, and Repellents		
o Allethrin	O=C2\C(=C(\C)C(OC(=O)C1C(/C=C(/C)C)C1(C)C)C2)C\C=C	7,0031
o Chlorobenzilate	Clc1ccc(cc1)C(O)(c2ccc(Cl)cc2)C(=O)OCC	6,947
o Cypermethrin	Cl/C(Cl)=C/C3C(C(=O)OC(C#N)c2cccc(Oc1ccccc1)c2)C3(C)C	5,521
o Diethyltoluamide(DEET)	O=C(c1cc(ccc1)C)N(CC)CC	8,0064
o Fenvalerate	Clc1ccc(cc1)C(C(=O)OC(C#N)c3cccc(Oc2ccccc2)c3)C(C)C	6,8437
o Flucythrinate	FC(F)Oc1ccc(cc1)C(C(=O)OC(C#N)c3cc(Oc2ccccc2)ccc3)C(C)C	4,3944
o Fluvalinate	c1([C@@H](OC([C@@H](Nc2c(cc(C(F)(F)F)cc2)Cl)C(C)C)=O)C#N)cc(Oc2ccccc2)ccc1	5,6419
o Propargite	O=S(OCC#C)OC2CCCCC2Oc1ccc(cc1)C(C)(C)C	7,2998
o Resmethrin	O=C(OCc1cc(oc1)Cc2ccccc2)C3C(\C=C(/C)C)C3(C)C	7,1261

MISCELLANEOUS PESTICIDES		
○ piperonyl butoxide	CCCC1=CC2=C(OCO2)C=C1COCCOCCOCCCC	8.4273
Other Herbicides		
○ picloram	ClC1=C(N)C(Cl)=C(C(O)=O)N=C1Cl	6.8565
○ diuron	ClC1=C(Cl)C=C(NC(N(C)C)=O)C=C1	8.1315
○ dalapon	ClC(C)(C(O)=O)Cl	8.8715
○ bromacil	BrC1=C(C)NC(N(C(C)CC)C1=O)=O	8.5564
Paraquat and Diquat		
○ Paraquat	C[N+]1=CC=C(C2=CC=[N+](C)C=C2)C=C1	5.8406
Other herbicides		
○ molinate	CCSC(N1CCCCCC1)=O	5.9108
○ fluroxypyr	FC1=C(Cl)C(N)=C(Cl)C(OCC(O)=O)=N1	8.2263
○ dichlobenil	ClC1=C(C#N)C(Cl)=CC=C1	8.0583
○ chlorophacinone	O=C1C(C(C(C2=CC=CC=C2)C3=CC=C(Cl)C=C3)=O)C(C4=C1C=CC=C4)=O	-1.7148
○ alachlor	COCN(C(C Cl)=O)C1=C(CC)C=CC=C1CC	0.3001
Other pesticides		
○ hexachlorophene	OC1=C(Cl)C=C(Cl)C(Cl)=C1CC2=C(O)C(Cl)=CC(Cl)=C2Cl	1.7918
○ Glutaraldehyde	[H]C(CCCC([H])=O)=O	5.5053
○ ethylene dibromide	BrCCBr	4.9836
○ endothall	OC(C1C(C(O)=O)C2CCC1O2)=O	4.0431
○ o-phenylphenol	OC1=C(C2=CC=CC=C2)C=CC=C1	6.3818
Nitrophenolic and Nitrocresolic Herbicides		
○ dinoseb	OC1=C([N+]([O])=O)C=C([N+]([O])=O)C=C1C(CC)C	3.2958
Fumigants		
○ 1,3-dichloropropene	ClC/C=C/Cl	6.5695
○ acrolein	C=CC ([H])=O	2.3321
Rodenticides		
○ difenacoum	O=C(C(C1CC(C2=CC=C(C3=CC=CC=C3)C=C2)CC4=C1C=CC=C4)=C5O)OC6=C5C=CC=C6	0.5878

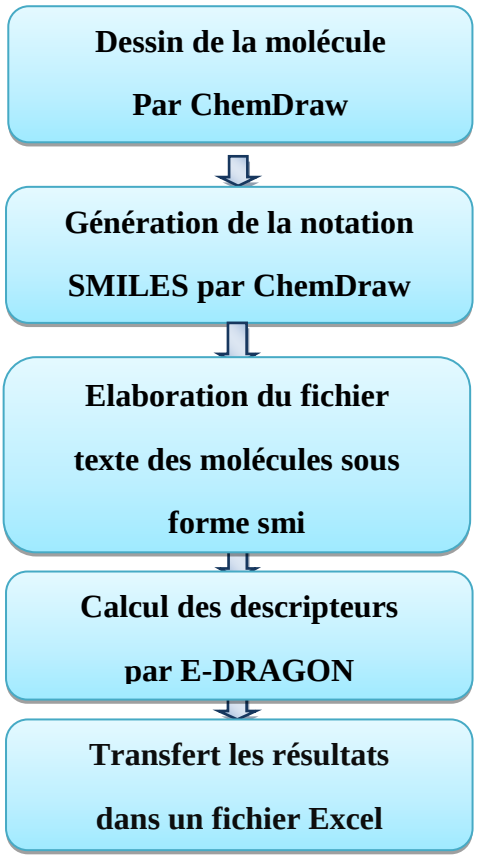


Figure IV.2 Procédure d'élaboration de la base de données.

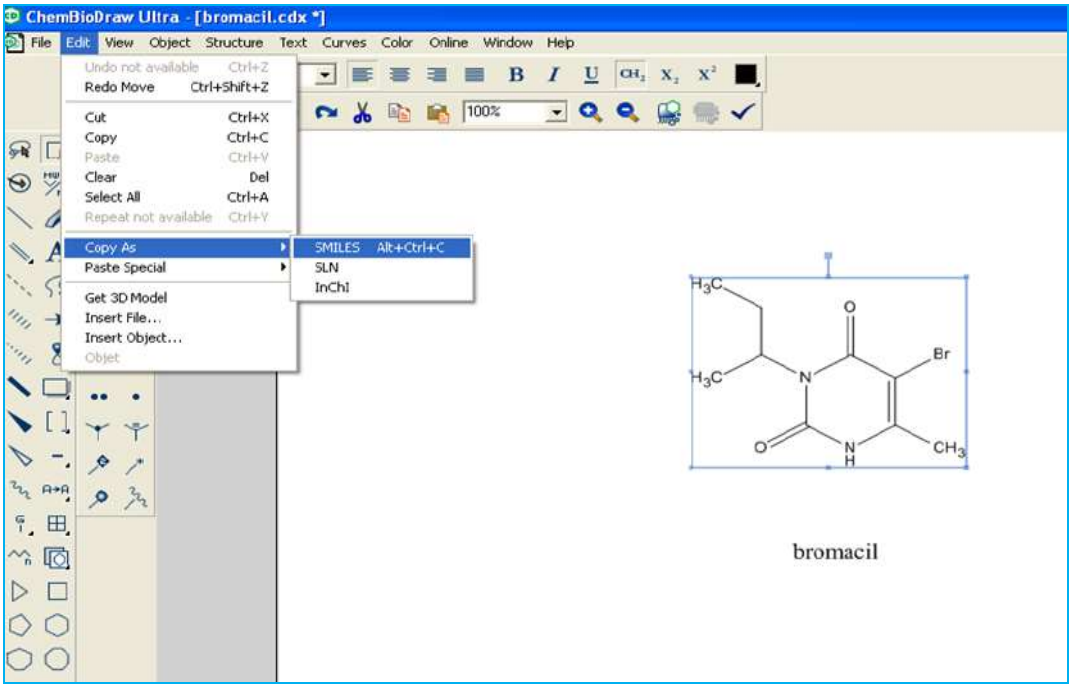


Figure IV.3 Dessin des molécules sur ChemDraw (1).

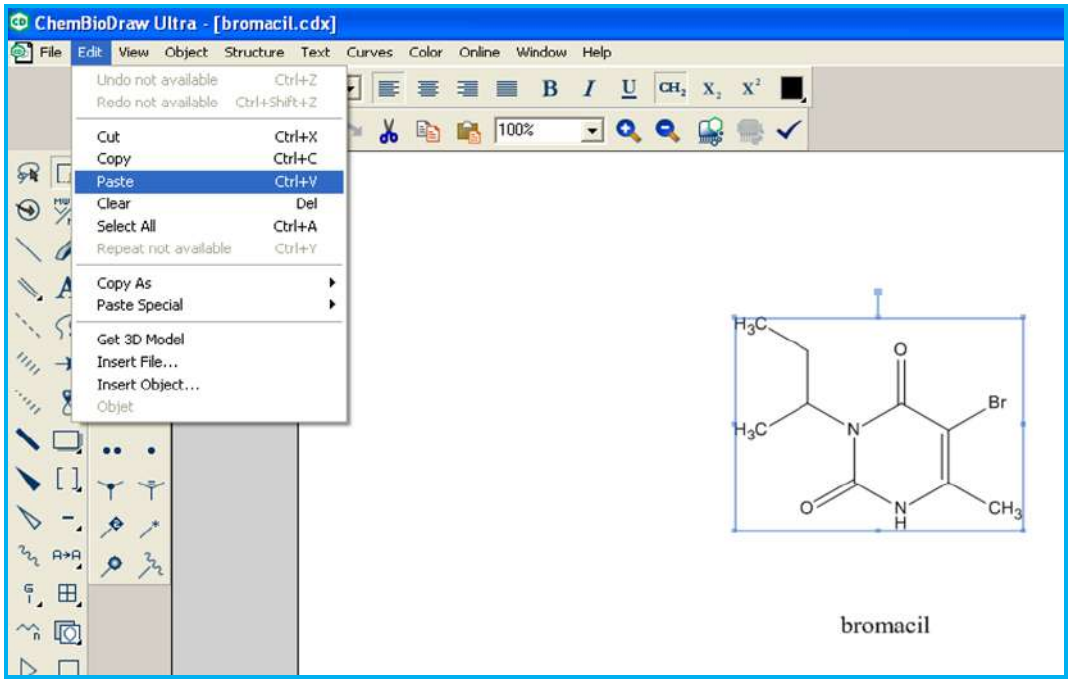


Figure IV.3 Dessin des molécules sur ChemDraw (2).

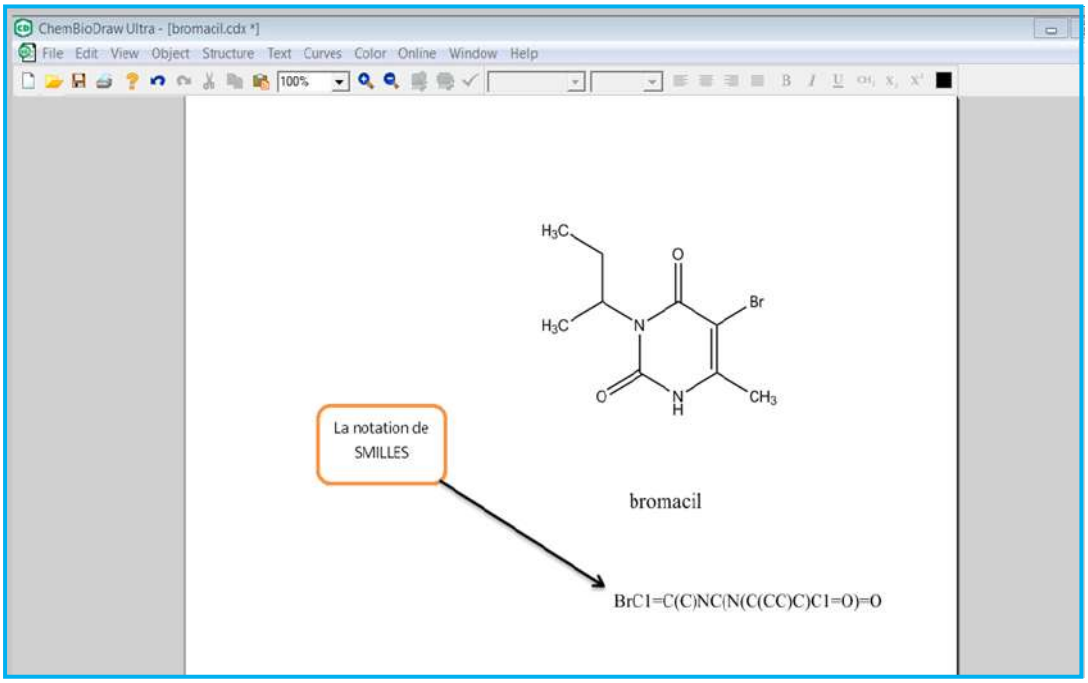


Figure IV.4 Génération de la notation SMILES par ChemDraw(1)

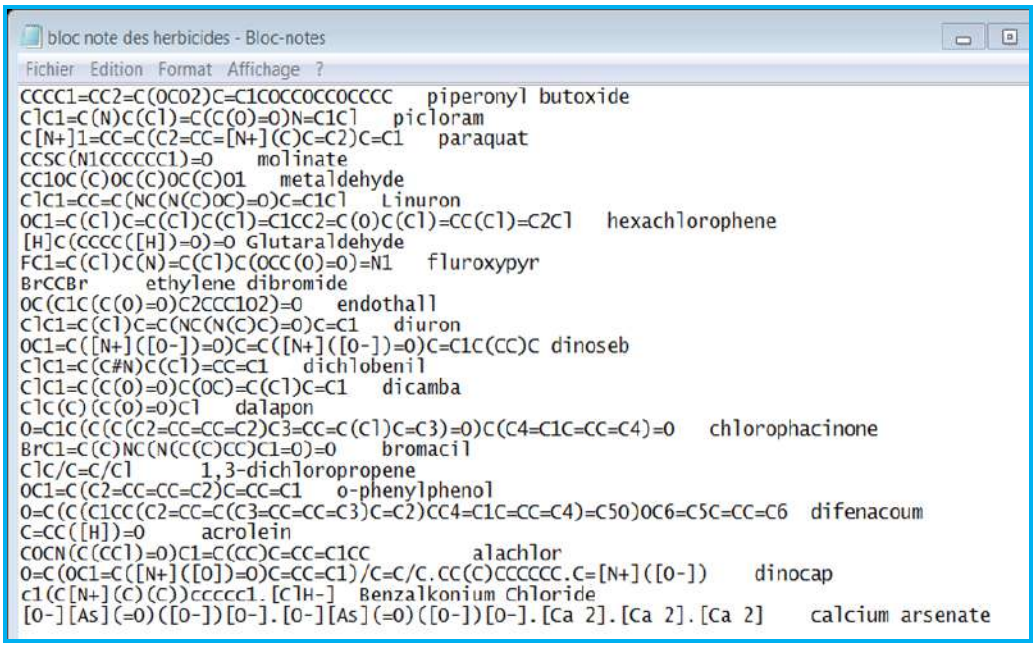


Figure IV.4 Génération de la notation SMILES par ChemDraw (2)



Figure IV.5 Première étape pour le calcul des descripteurs par E Dragon1.0. (1)



Figure IV.5 Deuxième étapes pour le calcul des descripteurs par E Dragon1.0. (2)

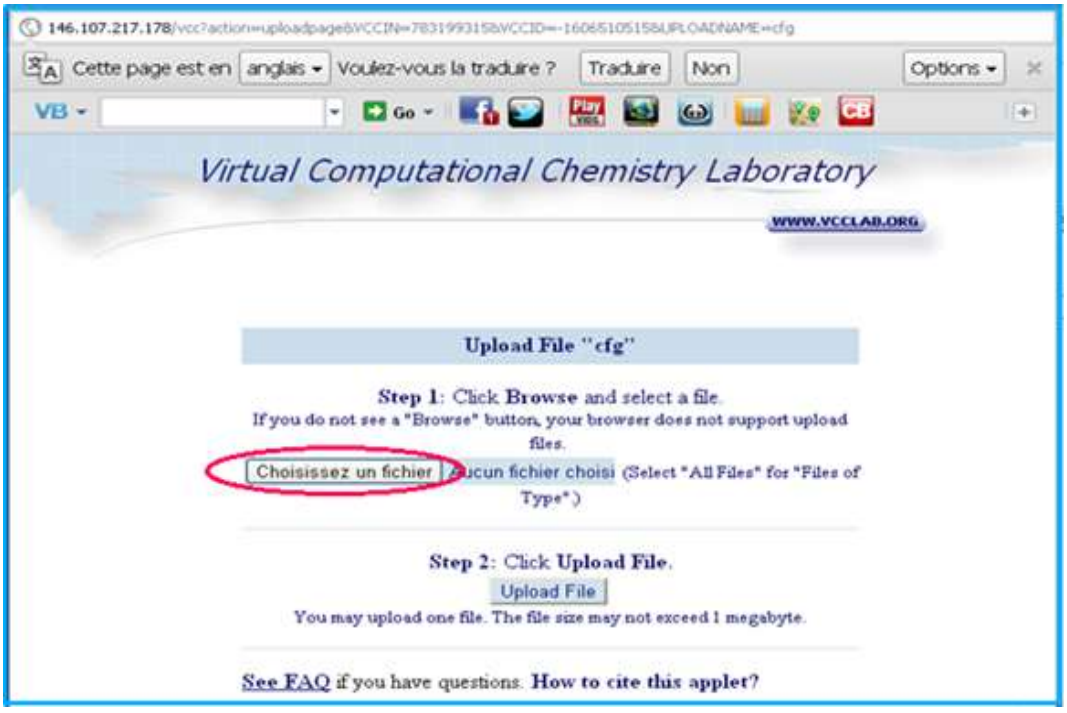


Figure IV.5 Troisième étape pour le calcul des descripteurs par E Dragon1.0. (3)

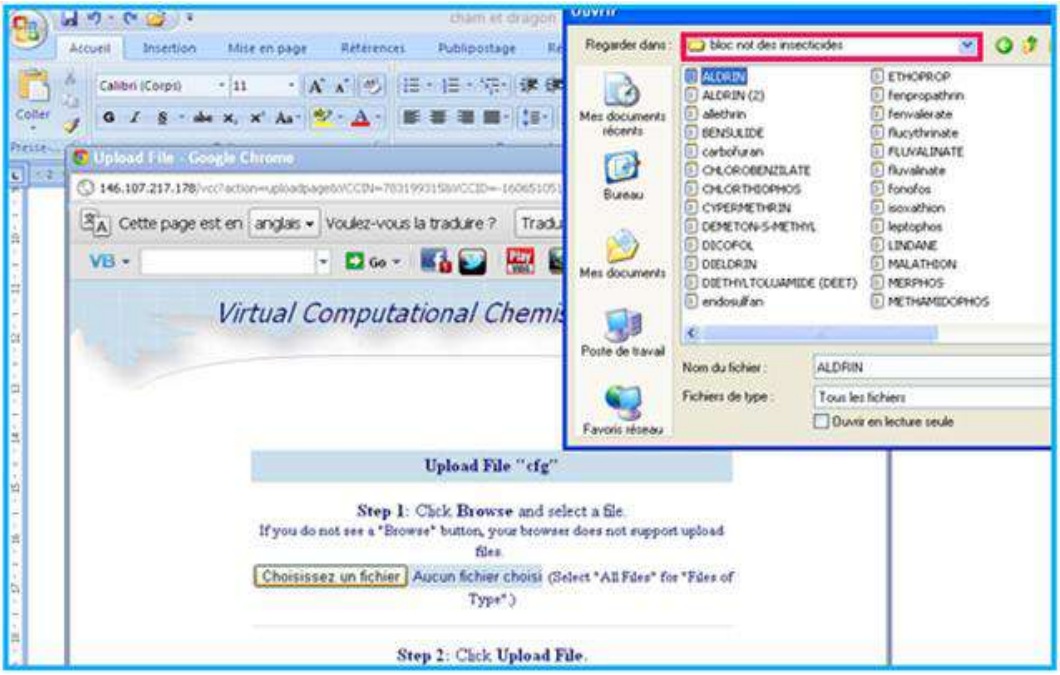


Figure IV.5 Quatrième étape pour le calcul des descripteurs par E Dragon1.0. (4)

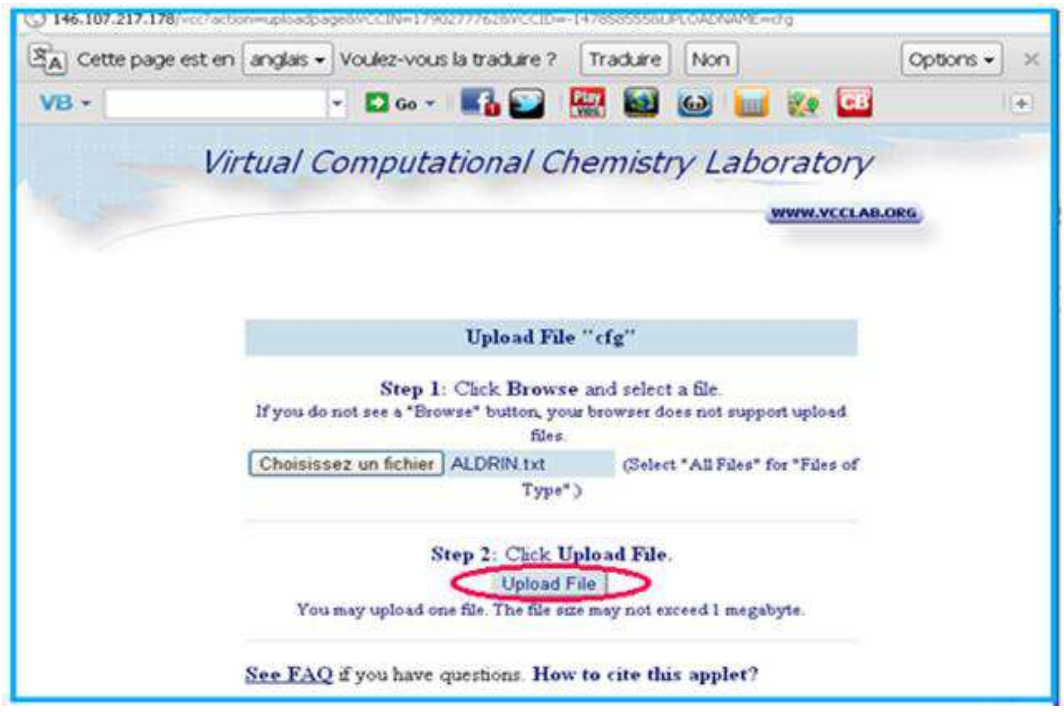


Figure IV.5 Cinquième étape pour le calcul des descripteurs par E Dragon1.0. (5)



Figure IV.5 Sixième étape pour le calcul des descripteurs par E Dragon1.0. (6)

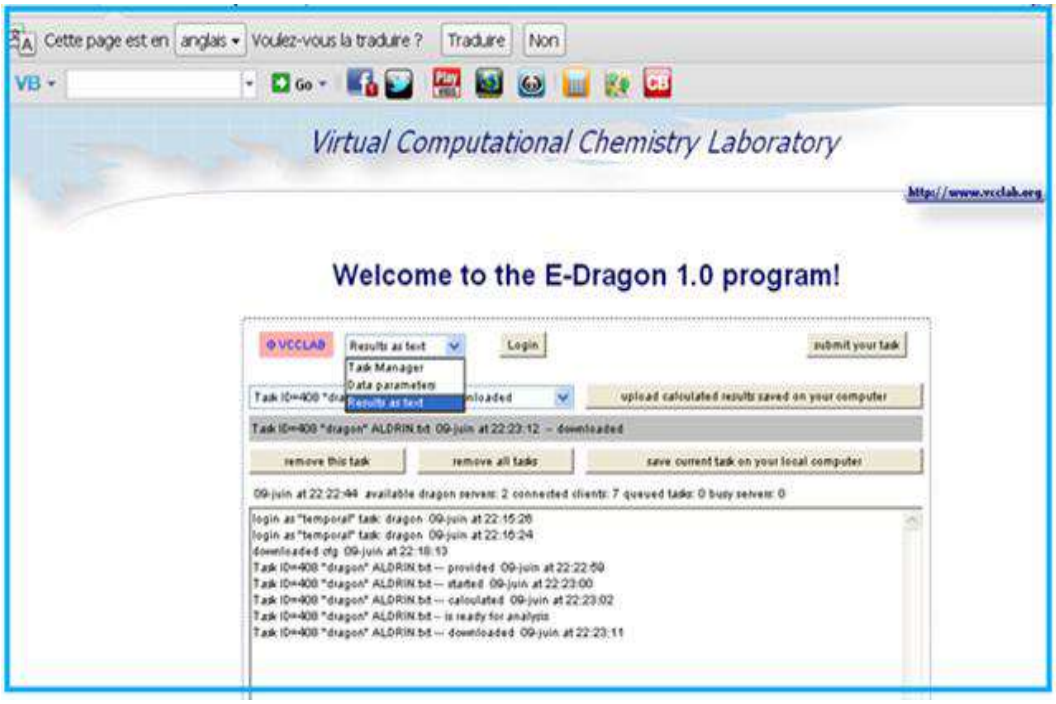


Figure IV. 5 Septième étape pour le calcul des descripteurs par E Dragon1.0. (6)



Figure IV.5 Huitième étape pour le calcul des descripteurs par E Dragon1.0. (7)

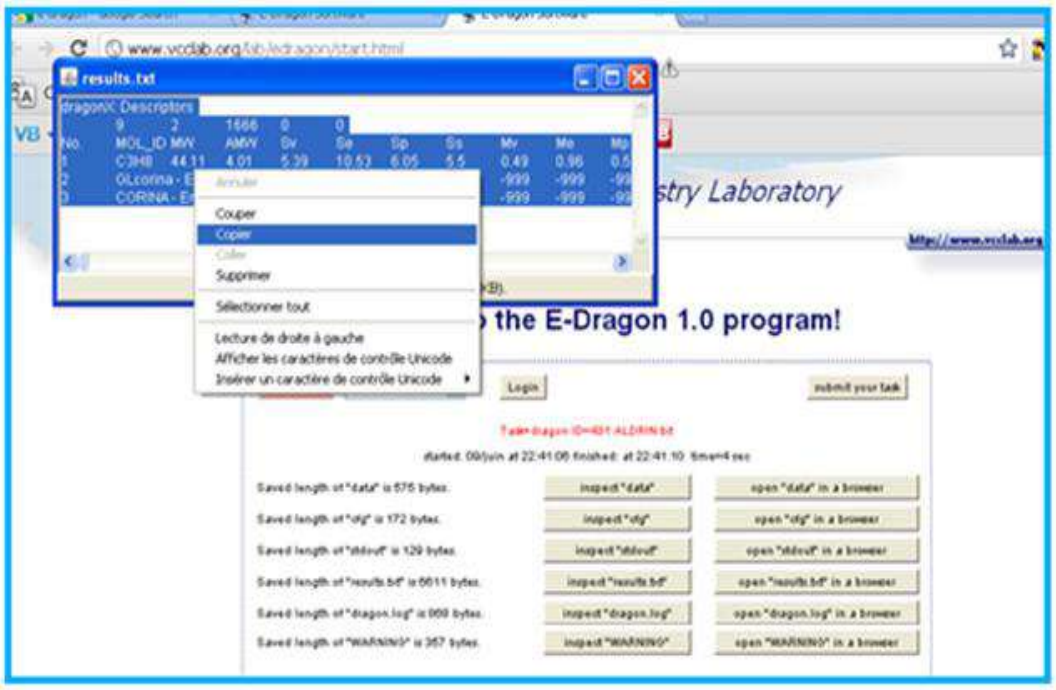


Figure IV.5 Neuvième étape pour le calcul des descripteurs par E Dragon1.0 (8)

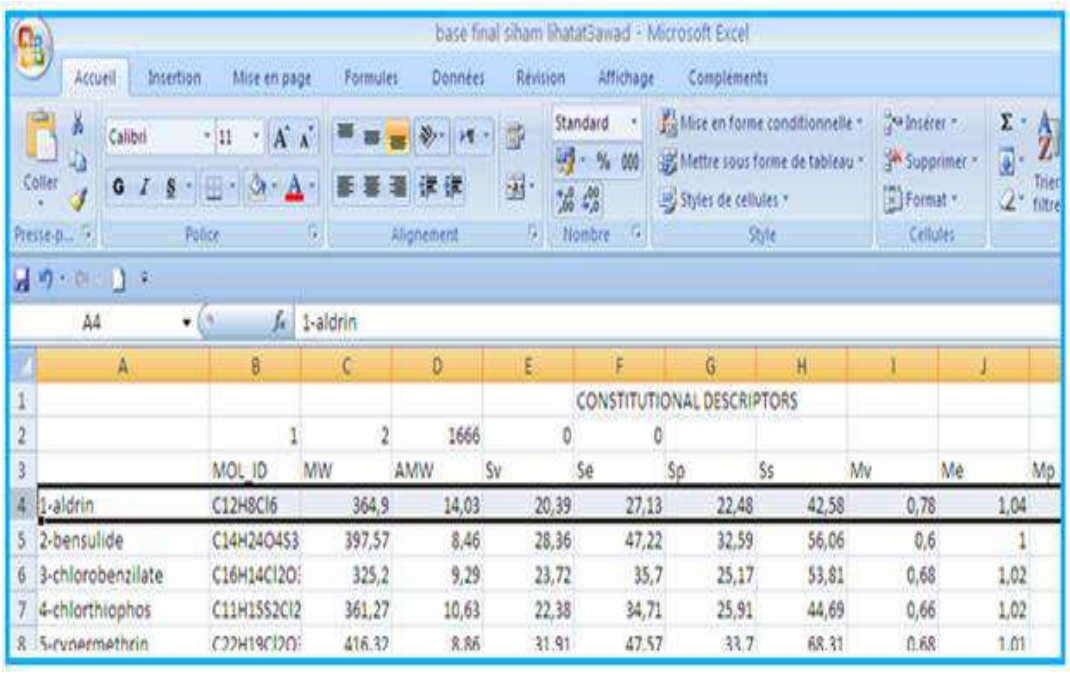


Figure IV.5 Dixième étape coller les descripteurs calculés par E Dragon1.0. (9)

La base de données de la figure VI.5 contient 72 lignes qui représentent les molécules, et 1667 colonnes qui représentent leurs propriétés (descripteurs) et une colonne réservée pour décrire la toxicité de chaque molécule.

IV-2-2 Choix d'une famille de descripteurs

Afin d'exploiter au maximum les informations contenues dans les structures moléculaires, celles-ci sont traduites en une série de grandeurs (en général scalaires) qui quantifient leurs caractéristiques physico-chimiques et structurelles. Ces grandeurs sont appelées descripteurs.

Depuis des décennies, de nombreux travaux ont été menés afin de développer des descripteurs capables de décrire de la manière la plus exhaustive possible les structures moléculaires. On en dénombre aujourd'hui des milliers **[82,83]**. Ils peuvent être calculés ou empiriques. Ces derniers nécessitant d'être mesurés, les descripteurs calculés seront privilégiés. Les détails les concernant sont disponibles dans l'ouvrage de Karelson **[84]**. On a choisi dans cette étude de travailler avec les descripteurs constitutionnels et les propriétés moléculaires illustrés dans le tableau IV.2.

Tableau IV.2 Liste des paramètres des descripteurs (constitutionnels et les propriétés moléculaires)

Molecular properties	**Ui**	**U**nsaturation **I**ndex
	Hy	**H**ydrophilic **F**actor
	AMR	**M**olarrefractivity
	TPSA(NO)	**T**opological **P**olar **S**urface **A**rea using N, O polar contributions
	TPSA (tot)	**T**opological **P**olar **S**urface **A**rea using N, O, S, P polar contributions
	MLogP	**M**origuchioctanol-water partition coefficient
	MLOGP2	**S**quared**M**origuchioctanol-water partition coeff. (logP^2)
	ALOGP	**G**hose-**C**rippenoctanol-water partition coeff. (logP)
	ALOGP2	**S**quaredGhose-Crippenoctanol-water partition coeff. (logP^2)
	LAI	**L**ipinskialert **I**ndex
	GVWAI-80	**G**hose-**V**iswanadhan-**W**endoloskidrug-like index **A**t 80%
	GVWAI-50	**G**hose-**V**iswanadhan-**W**endoloskidrug-like **I**ndex **A**t 50%
	Inflammat-80	**G**hose-**V**iswanadhan-**W**endoloskianti **I**nflammatory-**L**ike index at 80%
	Inflammat-50	Ghose-Viswanadhan-Wendoloskianti **I**nflammatory-like index at 50%
	Depressant-80	Ghose-Viswanadhan-Wendoloskianti **D**epressant-like index at 80%
	Depressant-50	Ghose-Viswanadhan-Wendoloskianti **D**epressant-like index at 50%
	Psychotic-80	Ghose-Viswanadhan-Wendoloskianti **P**sychotic-like index at 80%
	Psychotic-50	Ghose-Viswanadhan-Wendoloskianti **P**sychotic-like index at 50%
	Hypertens-80	Ghose-Viswanadhan-Wendoloskianti **H**ypertensive-like index at 80%
	Hypertens-50	Ghose-Viswanadhan-Wendoloskianti **H**ypertensive-like index at 50%
	Hypnotic-80	Ghose-Viswanadhan-Wendoloski **H**ypnotic-like index at 80%
	Hypnotic-50	Ghose-Viswanadhan-Wendoloski **H**ypnotic-like index at 50%

	Neoplastic-80	Ghose-Viswanadhan-Wendoloskianti **N**eoplastic-like index at 80%
	Neoplastic-50	Ghose-Viswanadhan-Wendoloskianti **N**eoplastic-like index at 50%
	Infective-80	Ghose-Viswan adhan-Wendoloskianti **I**nfective-like index at 80%
	Infective-50	Ghose-Viswanadhan-Wendoloskianti **I**nfective-like index at 50%
	BLTF96	Verhaar model of Fish **B**ase-**L**ine **T**oxicity for **F**ish (96h) from MLOGP (mmol/l)
	BLTD48	Verhaar model of Daphniabase-**L**ine **T**oxicity for **D**aphnia (48h) from MLOGP (mmol/l)
	BLTA96	Verhaar model of Algaebase-**L**ine **T**oxicity for Algae (96h) from MLOGP (mmol/l)
	MW	**M**olecular **W**eight
	AMW	**A**verage **M**olecular **W**eight
	Sv	**S**um of atomic van der Waals volumes (scaled on Carbonatom)
	Se	**S**um of atomic Sanderson electroneg ativities (scaled on Carbonatom)
	Sp	**S**um of atomic polarizabilities (scaled on Carbonatom)
	Ss	**S**um of Kier-Hall electro topological states
	Mv	**M**eanatomic van der Waals volume (scaled on Carbonatom)
	Me	**M**ean atomic Sanderson electronegativity (scaled on Carbonatom)
	Mp	**M**ean atomic polarizability (scaled on Carbonatom)
	Ms	**M**ean electro topological state
	nAT	**n**umber of **A**toms
	nSK	**n**umber of non-H atoms
	nBT	**n**umber of **B**onds
	nBO	**n**umber of non-H **B**onds
	nBM	**n**umber of **M**ultiple **B**onds
	SCBO	**S**um of **C**onventional **B**ond **O**rders (H-depleted)
	ARR	**AR**omatic **R**atio
	nCIC	**n**umber of rings

Consttutional descriptors	**nCIR**	**n**umber of **C**ircuits
	RBN	**N**umber of **R**otatable **B**onds
	RBF	**R**otatable **B**ond **F**raction
	nDB	**n**umber of **D**ouble **B**onds
	nTB	**n**umber of **T**riple **B**onds
	nAB	**n**umber of **A**romatic **B**onds
	nH	**n**umber of **H**ydrogenatoms
	nC	**n**umber of **C**arbonatoms
	nN	**n**umber of **N**itrogenatoms
	nO	**n**umber of **O**xygenatoms
	nP	**n**umber of **P**hosphorousatoms
	nS	**n**umber of **S**ulfuratoms
	nF	**n**umber of **F**luorine atoms
	nCL	**n**umber of **C**hlorineatoms
	nBR	**n**umber of **B**romineatoms
	nI	**n**umber of **I**odineatoms
	nB	**n**umber of **B**oronatoms
	nHM	**n**umber of **H**eavyatoms
	nX	**n**umber of halogenatoms
	nR03	**n**umber of **3**-membered **R**ings
	nR04	**n**umber of **4**-membered **R**ings
	nR05	**n**umber of **5**-membered **R**ings
	nR06	**n**umber of **6**-membered **R**ings
	nR07	**n**umber of **7**-membered **R**ings
	nR08	**n**umber of **8**-membered **R**ings
	nR09	**n**umber of **9**-membered **R**ings
	nR10	**n**umber of **10**-membered **R**ings
	nR11	**n**umber of **11**-membered **R**ings
	nR12	**n**umber of **12**-membered **R**ings

	nBnz	**n**umber of **B**enzene-like **R**ings

IV-2-3 Prétraitement de la base de données

1ère Étape : enlever les descripteurs avec des valeurs pratiquement constantes ou nulles.

2ème Étape : les descripteurs qui ont une valeur identique pour plus de 80% des échantillons sont éliminés.

3ème Étape : les descripteurs avec un écart type relatif inferieure à 0.05 sont éliminés.

$$RSD(i) = \frac{\text{std}(\text{data}(:,i))}{\text{mean}(\text{data}(:,i))} \qquad \textbf{IV.1}$$

4ème **Étape:** corréler tout les descripteurs restant avec log LD50. Corréler les descripteurs entre eux (par pair). Parmi chaque pair de descripteur avec |R|>0.8, celui qui a une |R| avec log LD50 élevée sera retenu.

Le **Tableau IV.3** représente la Liste des paramètres des descripteurs (constitutionnels et les propriétés moléculaires) qui reste après le prétraitement de la base de données

Tableau IV.3 Liste des paramètres des descripteurs (constitutionnels et les propriétés moléculaires) après prétraitement.

	Symbole	**Nom**
descripteurs constitutionnels.	**AMW**	**A**verage **M**olecular **W**eight
	Mv	**M**eanatomic van der Waals volume (scaled on Carbonatom)
	Ms	**M**ean electro topological state
	nDB	**n**umber of **D**ouble **B**onds
	nH	**n**umber of **H**ydrogenatoms
	nN	**n**umber of **N**itrogenatoms
	nO	**n**umber of **O**xygenatoms
	ARR	**AR**omatic **R**atio
	RBN	**N**umber of **R**otatable **B**onds
	RBF	**R**otatable **B**ond **F**raction
	nCIR	**n**umber of **C**ircuits
	nCL	**n**umber of **C**hlorineatoms
	nR06	**n**umber of **6**-membered **R**ings
	nP	**n**umber of **P**hosphorousatoms
	nS	**n**umber of **S**ulfuratoms

les propriétés moléculaires	**AMR**	**M**olarrefractivity
	TPSA(No)	**T**opological **P**olar **S**urface **A**rea using N, O polar contributions
	TPSA(Tot)	**T**opological **P**olar **S**urface **A**rea using N, O, S, P polar contributions
	GVWAI-80	**G**hose-**V**iswanadhan-**W**endoloskidrug-like index **A**t 80%
	Inflammat-80	**G**hose-**V**iswanadhan-**W**endoloskianti **I**nflammatory-**L**ike index at 80%
	Hypertens-80	Ghose-Viswanadhan-Wendoloskianti **H**ypertensive-like index at 80%
	Hypnotic-80	Ghose-Viswanadhan-Wendoloski **H**ypnotic-like index at 80%
	Neoplastic-80	Ghose-Viswanadhan-Wendoloskianti **N**eoplastic-like index at 80%
	Infective-80	Ghose-Viswan adhan-Wendoloskianti **I**nfective-like index at 80%
	ALOGPS_ logP	**G**hose-**C**rippenoctanol-water partition coeff. (logP)

IV-2-4 Normalisation de la base de données simplifier

La normalisation est une étape importante dans le processus de l'élaboration des données. Elle devient dans la plupart des cas, une exigence pour les matrices des inputs, du fait qu'on utilise souvent des fonctions de transfert bornées de type sigmoïdes dans les modélisations statiques.

Dans notre travail, on a utilisé la base de données qui est représentés en 26 lignes, correspondant respectivement aux descripteurs (constitutionnels, les propriétés moléculaires) et la toxicité. Les valeurs numériques ont été normalisées afin d'améliorer l'optimisation. La normalisation se fait par ligne (***mapminmax MATLAB version : 7.5***), chaque paramètre des entrées et de sortie (x_i) est normalisé (x_{in}) par l'expression suivante:

$$x_{in} = \frac{2\{x_i - \min(x_i)\}}{\{\max(x_i) - \min(x_i)\}} - 1 \qquad \textbf{IV 2}$$

IV-2-5 Mise en forme des données

Dans cette étape les données sont subdivisées en trois sous-ensembles :

- Un ensemble pour l'apprentissage (80%) ;
- Un ensemble pour la phase de validation (10%) ;
- Un ensemble pour le test (10%).

IV-2-6 Conception et optimisation

La conception et l'optimisation du réseau a était réalisé en élaborant un programme en MATLAB. Le nombre de neurones dans la couche cachée à été varié de 1 à 20.

On a choisi comme fonction de transfert la fonction tangente sigmoid « tansig » pour la couche cachée et la fonction identité « purelin » pour la couche de sortie et comme algorithme d'apprentissage l'algorithme de Levenberg-Marquardt « trainlm ».

IV-3 STRUCTURE DU RNSO OBTENU

La structure de base retenue est un perceptron multicouche à trois couches : une couche d'entrée, une couche cachée, une couche de sortie (Figure IV.6). Le tableau IV.4 récapitule les différents éléments structurels du réseau de neurones optimal.

IV-3-1 L'algorithme d'apprentissage et la fonction d'activation

L'algorithme d'apprentissage utilisé est Levenberg-Marquardt (la fonction trainlm de MATLAB®) avec l'option d'apprentissage atténué pour éviter le sur-apprentissage. La procédure commence par deux fonctions de transfert sigmoïde tangent hyperbolique (fonction tansig de MATLAB®) pour la couche cachée et la fonction linéaire (fonction purelin de MATLAB®) pour la couche de sortie.

Tableau IV.4 Structure des RNSO utilisés dans la modélisation de la toxicité.

Couche d'entrée	**Couche cachée**		**Couche de sortie**	
Nombre de neurones	Nombre de neurones	Fonction d'activation	Nombre de neurones	Fonction d'activation
25	13	Tangente sigmoïde « tansig »	1	Identité « purelin »

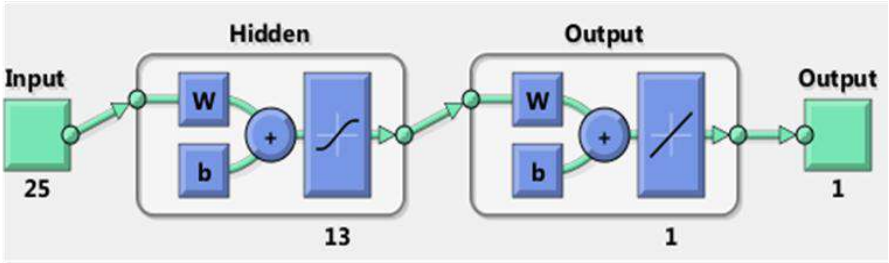


Figure IV.6 Schéma simplifié de l'architecture du RNSO.

La validation des valeurs calculées avec les valeurs expérimentales est schématisée dans la figure IV.7, elle a donné un coefficient de régression R égale à **0.94556**

Les différentes valeurs de l'erreur quadratique moyenne obtenues sont:

- ✓ EQM=0.039 pour la base d'apprentissage.
- ✓ EQM=3.225 pour la base de validation.
- ✓ EQM=3.092 pour la base de test.

La figure IV.7 présente la validation du modèle de réseau de neurone avec les résultats expérimentaux, on remarque bien que les points expérimentaux se rapprochent de la première bissectrice avec un coefficient de régression s'approchant de l'unité

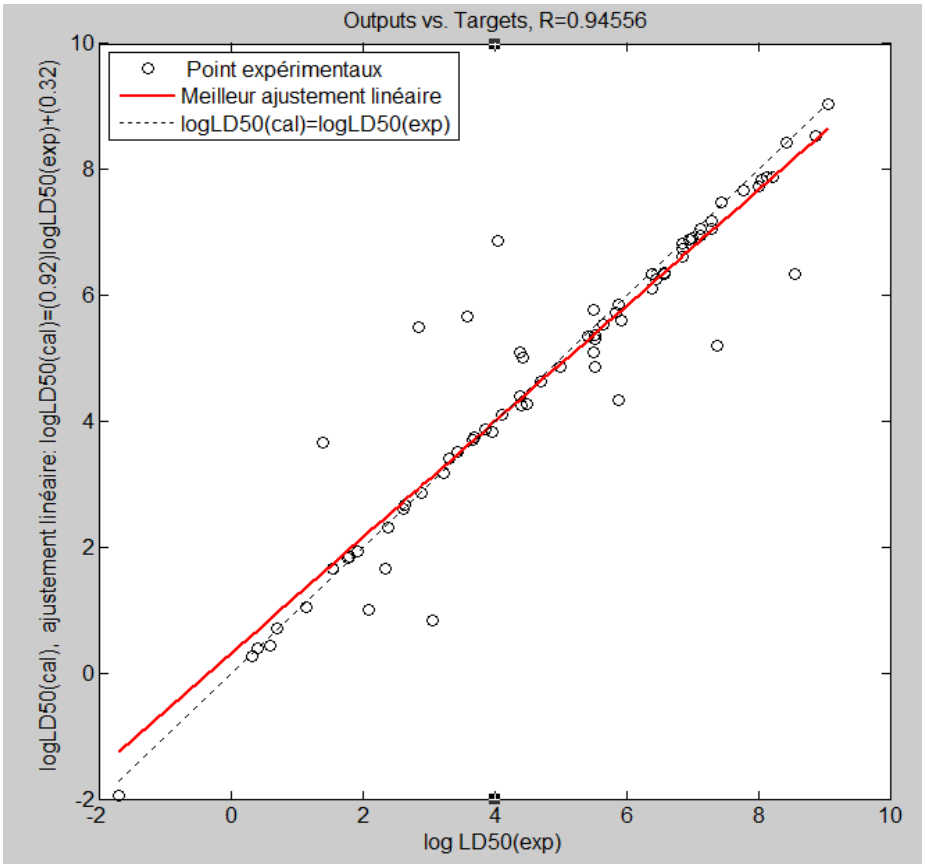


Figure IV.7 Validations des valeurs calculées par RNA avec les valeurs expérimentales.

IV-4 ANALYSE DE SENSIBILITE

Cette technique est très utilisée pour l'analyse de contribution des variables, elle consiste à éliminer les entrées peu ou moins influentes sur la sortie (toxicité). Elle se base sur le calcul de l'impotence relative des différentes variables d'entrées en utilisant la méthode des « poids de connexion » **[85]**.

Cette méthode nécessite la partition des poids de connexion en :

- Poids de connexion entrée –cachée (input-hidden) IH (Tableau IV.5).
- Poids de connexion cachée- sortie (hidden-output) HO (Tableau IV.5).

Tableau IV.5 Poids de connexion entrée-cachée, cachée-sortie.

	Hidden 1	Hidden 2	Hidden 3	Hidden 4	Hidden 5	Hidden 6	Hidden 7
Input1	0.18307389	0.38321715	0.05899789	0.41074862	-0.32265435	0.09647787	0.04837869
Input2	-0.09427062	0.47402015	0.55278846	0.00828555	-0.03664795	-0.27378181	0.57712135
Input3	-0.11946411	-0.04210346	-0.06174127	0.42088326	0.48322013	-0.57728008	-0.10471106
Input4	0.17985508	0.28366548	-0.13234376	0.40061573	-0.82247154	0.7459986	0.47586171
Input5	0.36858665	0.12663184	-0.39019042	0.21379782	0.21510672	-0.32929329	-0.43073334
Input6	0.14027305	-0.25609863	-0.4759361	-0.07563036	0.15278589	-0.06324011	-0.02678469
Input7	0.43609569	-0.76496426	0.14903849	0.60609968	-0.12813328	0.22865645	-0.2005316
Input8	0.65018758	-0.04399024	0.2168301	0.52768453	0.14762837	0.49502831	-0.68297071
Input9	0.29802152	-0.86619006	-0.3703924	0.11989962	-0.47188277	0.35063314	0.43042725
Input10	-1.02046119	-0.42373166	0.00338069	0.03341821	-1.02401898	-0.44325106	0.48815837
Input11	1.16217789	0.28436311	-0.19480141	-0.08212629	0.837305	0.03434225	-0.49745809
Input12	0.02962134	-0.05656543	0.71433092	-0.23902491	0.40192615	0.085819	-1.16272348
Input13	0.22813296	-0.09154574	0.01050523	0.12242734	0.21687498	0.31078713	-0.51986517
Input14	-0.24746605	0.17994849	-0.40919672	0.10999458	-1.05813947	0.51790712	0.40590057
Input15	0.51760475	-0.39399204	-0.21931163	0.46434338	0.24598314	-0.15308775	0.31426127
Input16	0.45259077	-0.16564926	-0.40457517	-0.21796747	-0.36203544	0.72765006	0.41571728
Input17	0.42383495	0.43133522	0.42998763	0.37306148	-0.2831209	0.39101367	0.48679517
Input18	0.40751951	-0.01659528	-0.31970387	0.09934875	-0.31901929	0.27271187	-0.20334128
Input19	0.67161529	0.85871775	0.32335808	0.03572362	0.84840029	-0.03116837	-0.95626381
Input20	0.94825693	-0.74066512	-0.68703424	0.21775123	-0.05142551	0.06789421	0.74240549
Input21	1.37996783	-0.56426696	-0.36582762	0.46260954	-0.44751057	-0.21665917	-0.17072448
Input22	-0.24338133	0.48549303	-0.58692031	0.04184041	0.22147466	0.14437138	0.12701232
Input23	-0.24606479	-1.18683034	-0.12403582	0.15261215	-0.11844856	-0.55630088	-0.04147872
Input24	-0.57876851	0.75502926	-0.01431199	-0.52317121	0.3826335	-0.11889671	0.58879919
Input25	0.18892127	0.48382568	0.49947736	0.09658996	-0.21104083	0.76233297	0.88394482
output	0.4035743	0.23752418	1.0684317	0.1488221	-0.44952741	-0.59188208	0.36886752

	Hidden 8	Hidden 9	Hidden 10	Hidden 11	Hidden 12	Hidden 13
Input1	-0.12466338	-0.14780332	-0.55115923	0.17753394	0.70362977	0.45179887
Input2	-0.02457772	-0.00137331	0.3428098	0.67352508	0.6665643	-0.12916283
Input3	0.07317812	-0.97606483	0.08585271	-0.69276858	1.6529059	-1.11887683
Input4	0.02144643	-0.41513108	0.52481191	0.55664171	-0.43891589	-0.27513543
Input5	-0.50390021	0.0462077	-0.30792741	0.18469464	-0.05765416	-0.22953942
Input6	0.36722975	-0.13706712	-0.03724479	-0.13544026	-0.08968147	0.12745136
Input7	0.13206437	-0.6691402	-0.03366225	-0.33788362	1.31516283	0.19575452
Input8	0.0278335	0.30747398	-0.57127567	-0.00086048	0.10928329	-0.17185417
Input9	0.05360956	-1.31038502	-0.02117799	-0.24567168	-1.59140121	0.44471412
Input10	0.0129034	0.11803728	-0.39630181	0.67182276	-0.46390073	-0.03512798
Input11	0.30855199	-0.46868193	0.27016932	0.58143203	0.6278042	0.05371312
Input12	0.11690714	0.18942623	-0.56304667	0.83573705	-0.64916784	0.46484035
Input13	-0.09636039	0.0824698	0.10827304	1.12255339	0.10308561	0.17502344
Input14	-0.46506935	-0.48767493	0.60474231	0.55874084	-0.36657812	1.41564768
Input15	-0.78086824	0.52368371	0.13470151	0.00205555	-0.26855382	-0.55137326
Input16	0.33757918	1.24099725	0.41424482	-0.3230471	0.08993218	0.01218563
Input17	-0.44001243	-0.85588954	0.73487516	0.21897665	1.20146652	0.36627993
Input18	0.07946456	-0.0196573	0.16894828	0.46133905	0.34219719	0.25456223
Input19	0.27428698	-0.2359868	0.36180934	0.19316144	0.05193001	0.28625787
Input20	-0.60071985	-0.73115787	-0.2492736	-0.1584228	-0.43696367	-0.1577776
Input21	0.17276974	0.50543195	-0.36211706	-0.66767742	-0.50982862	-0.25954262
Input22	-0.05032477	-0.56957631	-0.89109145	-0.86343318	-0.58107958	-0.51257665
Input23	0.11026762	0.28338727	-1.03028489	0.00498085	-0.22557978	0.07523667
Input24	-0.52266521	0.2355875	-0.55275068	0.41806756	0.10390007	-0.61302114
Input25	-0.44870025	0.45313816	-0.11162344	0.1736805	-0.23877959	-0.08442791
output	0.46353312	-0.84284873	0.25881836	0.66223296	-0.67893282	0.25556713

Le processus de calcul est comme suit :

1. Produit de IH et HO (tableau IV.6).
2. Calcul de l'importance de chaque neurone d'entrée (tableau IV.7).
3. Calcul de l'importance relative (tableau IV.8).

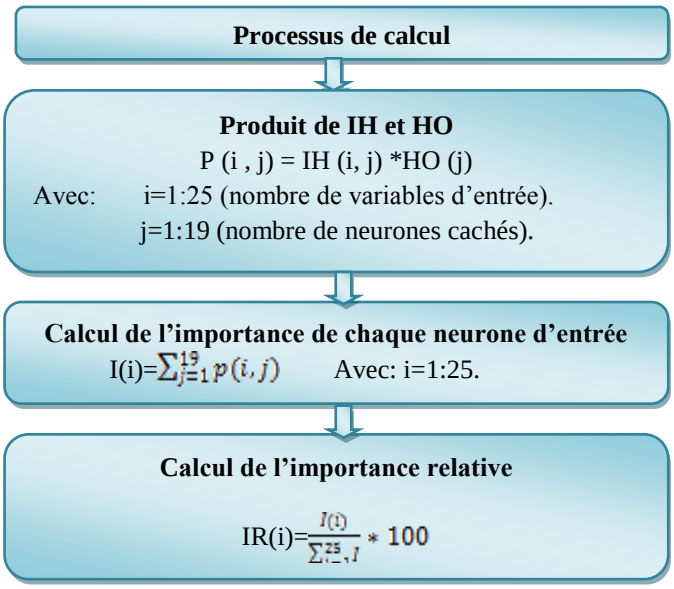


Figure IV.8 Processus de calcul par la méthode de poids de connexion.

Tableau IV.6 Produit de IH et HO.

	Hidden 1	Hidden 2	Hidden 3	Hidden 4	Hidden 5	Hidden 6	Hidden 7
Input1	0.07388391	0.09102334	0.06303522	0.06112847	0.14504197	-0.05710352	0.01784533
Input2	-0.0380452	0.11259125	0.59061672	0.00123307	0.01647426	0.16204655	0.21288132
Input3	-0.04821264	-0.01000059	-0.06596633	0.06263673	-0.2172207	0.34168174	-0.03862451
Input4	0.07258489	0.06737741	-0.14140027	0.05962047	0.3697235	-0.44154321	0.17552993
Input5	0.1487521	0.03007812	-0.41689182	0.03181784	-0.09669637	0.1949028	-0.15888354
Input6	0.0566106	-0.06082962	-0.50850522	-0.01125547	-0.06868144	0.03743069	-0.00988
Input7	0.17599701	-0.18169751	0.15923745	0.09020103	0.05759942	-0.13533766	-0.0739696
Input8	0.262399	-0.01044875	0.23166815	0.07853112	-0.066363	-0.29299839	-0.25192571
Input9	0.12027383	-0.20574108	-0.39573898	0.01784371	0.21212424	-0.20753347	0.15877063
Input10	-0.41183191	-0.10064651	0.00361204	0.00497337	0.4603246	0.26235236	0.18006577
Input11	0.46902512	0.06754312	-0.20813201	-0.01222221	-0.37639155	-0.02032656	-0.18349613
Input12	0.01195441	-0.01343566	0.76321381	-0.03557219	-0.18067682	-0.05079473	-0.42889093
Input13	0.0920686	-0.02174433	0.01122412	0.01821989	-0.09749125	-0.18394934	-0.19176138
Input14	-0.09987094	0.04274212	-0.43719875	0.01636962	0.4756627	-0.30653994	0.14972354
Input15	0.20889197	-0.09358264	-0.2343195	0.06910456	-0.11057617	0.0906099	0.11592078
Input16	0.182654	-0.0393457	-0.43226093	-0.03243838	0.16274485	-0.43068304	0.15334461
Input17	0.17104889	0.10245254	0.45941242	0.05551979	0.1272706	-0.23143399	0.17956293
Input18	0.1644644	-0.00394178	-0.34158175	0.01478529	0.14340792	-0.16141327	-0.07500599
Input19	0.27104667	0.20396623	0.34548603	0.00531646	-0.38137919	0.018448	-0.35273467
Input20	0.38269212	-0.17592588	-0.73404917	0.03240619	0.02311718	-0.04018537	0.27384927
Input21	0.55691955	-0.13402705	-0.39086182	0.06884652	0.20116827	0.12823668	-0.06297472
Input22	-0.09822245	0.11531633	-0.62708427	0.00622678	-0.09955893	-0.08545083	0.04685072
Input23	-0.09930543	-0.28190091	-0.1325238	0.02271206	0.05324588	0.32926452	-0.01530015
Input24	-0.2335761	0.17933771	-0.01529139	-0.07785944	-0.17200425	0.07037283	0.2171889
Input25	0.07624377	0.1149203	0.53365745	0.01437472	0.09486864	-0.45121123	0.32605854

	Hidden 8	Hidden 9	Hidden 10	Hidden 11	Hidden 12	Hidden 13
Input1	-0.0577856	0.12457584	-0.14265013	0.11756883	-0.47771734	0.11546494
Input2	-0.01139259	0.00115749	0.08872547	0.44603051	-0.45255238	-0.03300977
Input3	0.03392048	0.822675	0.02222026	-0.45877418	-1.12221205	-0.28594814
Input4	0.00994113	0.3498927	0.13583096	0.36862648	0.2979944	-0.07031557
Input5	-0.23357444	-0.0389461	-0.07969727	0.12231088	0.0391433	-0.05866273
Input6	0.17022315	0.11552685	-0.00963964	-0.08969301	0.0608877	0.03257238
Input7	0.06121621	0.56398396	-0.00871241	-0.22375767	-0.8929072	0.05002842
Input8	0.01290175	-0.25915405	-0.14785663	-0.00056984	-0.07419601	-0.04392028
Input9	0.0248498	1.10445635	-0.00548125	-0.16269188	1.0804545	0.11365431
Input10	0.00598115	-0.09948757	-0.10257018	0.44490317	0.31495743	-0.00897756
Input11	0.14302407	0.39502797	0.06992478	0.38504346	-0.42623688	0.01372731
Input12	0.05419033	-0.15965766	-0.14572681	0.55345262	0.44074135	0.11879792
Input13	-0.04466623	-0.06950956	0.02802305	0.74339185	-0.0699882	0.04473024
Input14	-0.21557505	0.41103619	0.15651841	0.3700166	0.24888191	0.36179302
Input15	-0.36195829	-0.44138615	0.03486322	0.00136126	0.18233	-0.14091288
Input16	0.15647913	-1.04597295	0.10721416	-0.21393244	-0.06105791	0.00311425
Input17	-0.20396033	0.72138541	0.19019918	0.14501356	-0.81571505	0.09360911
Input18	0.03683446	0.01656813	0.04372692	0.30551393	-0.2323289	0.06505774
Input19	0.1271411	0.19890117	0.0936429	0.12791787	-0.03525699	0.0731581
Input20	-0.27845354	0.61625548	-0.06451658	-0.1049128	0.29666898	-0.04032277
Input21	0.0800845	-0.42600268	-0.09372254	-0.44215799	0.34613938	-0.06633056
Input22	-0.0233272	0.48006667	-0.23063082	-0.57179391	0.39451399	-0.13099775
Input23	0.05111269	-0.2388526	-0.26665664	0.00329848	0.15315351	0.01922802
Input24	-0.24227263	-0.19856462	-0.14306202	0.27685812	-0.07054117	-0.15666806
Input25	-0.20798743	-0.38192692	-0.0288902	0.11501695	0.1621153	-0.021577

Tableau IV.7 Importance individuelle

Importance individuelle	
Input1	0.1155
Input2	-0.033
Input3	-0.2859
Input4	-0.0703
Input5	-0.0587
Input6	0.0326
Input7	0.05
Input8	-0.0439
Input9	0.1137
Input10	-0.009
Input11	0.0137
Input12	0.1188
Input13	0.0447
Input14	0.3618
Input15	-0.1409
Input16	0.0031
Input17	0.0936
Input18	0.0651
Input19	0.0732
Input20	-0.0403
Input21	-0.0663
Input22	-0.131
Input23	0.0192
Input24	-0.1567
Input25	-0.0216

Tableau IV.8 Importance relative

Importance relative(%)	
Input1	24.4149668
Input2	-6.97988972
Input3	-60.4634991
Input4	-14.8681701
Input5	-12.4041861
Input6	6.88740256
Input7	10.5784686
Input8	-9.28690668
Input9	24.0321105
Input10	-1.89829695
Input11	2.90262798
Input12	25.1197212
Input13	9.45817201
Input14	76.5008357
Input15	-29.7959129
Input16	0.65850509
Input17	19.7935697
Input18	13.7564055
Input19	15.4692208
Input20	-8.526216
Input21	-14.0255431
Input22	-27.6993653
Input23	4.06574913
Input24	-33.1273311
Input25	-4.56243841

La figure IV.9 illustre l'importance des descripteurs constitutionnels et les propriétés moléculaires sur la toxicité, on constate que la variable 14 qui correspond aux nombre d'atomes de chlore et la variable 12 qui correspond aux nombre d'atomes de phosphore sont les plus influentes sur la toxicité.

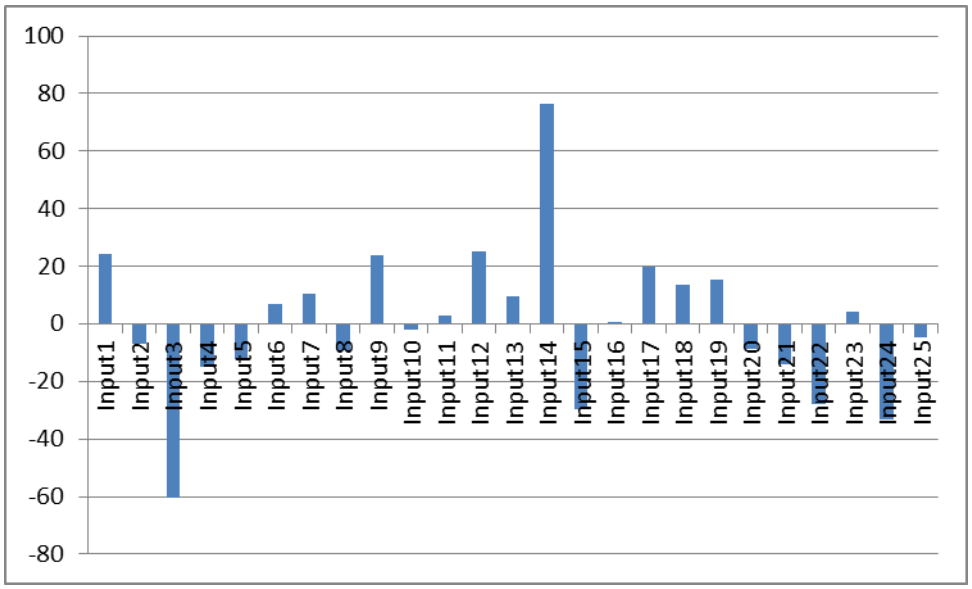


Figure IV.9 Influence des descripteurs sur la toxicité.

IV-5 CONCLUSION

Dans ce chapitre nous avons abordé le problème de prédiction de la toxicité des pesticides, en utilisant des programmes MATLAB® élaborés dans le cadre de ce travail.

Nous avons montré dans ce travail qu'il est possible de relier la toxicité à la topologie de la molécule après le calcul des descripteurs constitutionnels et les propriétés moléculaires par le logiciel E-DRAGON.

L'analyse de sensibilité nous a permis de connaître les entrées (descripteurs) qui ont une influence sur la toxicité. Par l'utilisation de la méthode de poids de connexion on a trouvé que le nombre d'atomes de chlore et le nombre des atomes de phosphore sont les deux paramètres les plus influents sur la toxicité.

Chapitre II

La Relation Quantitative Structure-Activité

II-1 INTRODUCTION

Les études toxicologiques et médicales sont de plus en plus onéreuses et nécessitent des moyens sans cesse plus importants. Aussi, pour pallier à ce problème de coût et de durée dans le temps, les laboratoires ont recours à l'étude de la relation structure-activité.

On entend par étude de la relation quantitative structure-activité l'analyse de la structure moléculaire d'un produit chimique pour en tirer des informations prédictives sur ses propriétés essentielles telles que sa stabilité, sa distribution, sa captation, son absorption et sa toxicité. L'étude de la relation quantitative structure-activité est l'une des méthodes d'identification des produits chimiques potentiellement dangereux. Elle permet de prendre une décision à un stade précoce pour un produit chimique nouveau et peut, de ce fait, répondre aux attentes de l'industrie et des pouvoirs publics.

II-2 DEFINITION

Les relations quantitatives structure-activité ou **QSAR** ***(Quantitative Structure Activity Relationship)*** représentent une tentative de corréler des descripteurs structuraux de composés avec des activités comme résumé ci-contre :

Activité = f (propriétés physico-chimiques et/ou structurales)

Ces descripteurs physico-chimiques, qui incluent des paramètres pour tenir compte de l'hydrophobie, la topologie, les propriétés électroniques, et des effets stériques, sont déterminés de manière empirique ou, plus récemment, par des méthodes de calcul. Les activités utilisées dans **QSAR** incluent des mesures chimiques et des dosages biologiques.

Ainsi par exemple, l'activité biologique peut être exprimée de manière quantitative, comme pour la concentration de substance nécessaire pour obtenir une certaine réponse biologique. De plus lorsque les propriétés ou structures physicochimiques sont exprimées par des chiffres, on peut proposer une relation mathématique, ou relation quantitative structure à activité, entre les deux. L'expression mathématique obtenue peut alors être utilisée comme moyen prédictif de la réponse biologique pour des structures similaires **[37].**

QSAR sont actuellement appliquées dans de nombreuses disciplines, avec plusieurs applications relatives à la conception de médicaments et l'évaluation des risques environnementaux **[38].** A titre d'exemple, l'industrie pharmaceutique s'est beaucoup

servie des études de la relation structure-activité pour identifier les molécules ayant un potentiel thérapeutique.

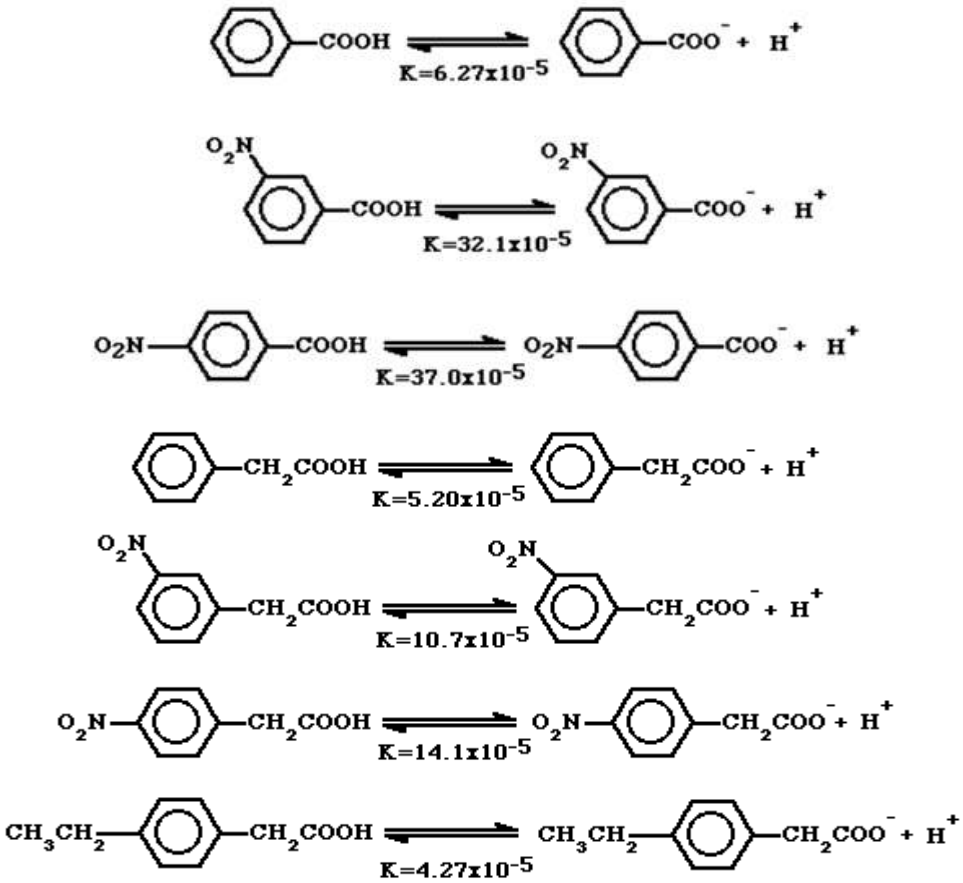
II-3 HISTORIQUE

•**19ème siècle** : Début de QSAR.

•**1863** : Des chercheurs appartenant à l'AFA0 Cros à l'Université de Strasbourg ont montrés que la diminution de la solubilité des alcools dans l'eau entrainait une augmentation de la toxicité de ces alcools pour les mammifères,

•**1890** : Hans Horst Meyer de l'Université de Marburg et Charles Ernest Overton de l'Université de Zurich, travaillant indépendamment, ont noté que la toxicité des composés organiques dépend de leur caractère lipophile.

•**1894-1987** : Louis Hammett à fait une corrélation entre les propriétés électroniques des acides et des bases organiques avec leurs constantes d'équilibre et leurs réactivité. En envisageant la dissociation de l'acide benzoïque (réaction 1), Hammett a observé que l'ajout de substituant sur le noyau aromatique de l'acide benzoïque a un effet ordonné et quantitatif sur la constante de dissociation (réactions 2-3). Hammett a également observé que les substituants ont un effet similaire sur la dissociation des acides organiques et d'autres bases (cas de la dissociation des acides phénylacétique schématisées par les réactions 4-7) **[38].**



II-4 APPLICATIONS DE LA QSAR

II-4-1 En chimie

Une des premières applications de la QSAR concerne la prédiction des points d'ébullition **[39].**

Il est bien connu par exemple que pour une famille de composés chimiques, particulièrement en chimie organique, il existe une corrélation forte entre la structure et les propriétés observées. On peut citer comme exemple simple, la relation entre le nombre de carbones dans les alcanes et leur point d'ébullition. Il existe une tendance nette à l'augmentation de la température d'ébullition avec le nombre d'atomes de carbone, ce qui sert de moyen prédictif pour les points d'ébullition des alcanes les plus lourds.Les méthodes de Hammett, de Taft et prédiction de pKa sont des applications particulièrement intéressantes **[40]**.

II-4-2 En biologie

L'activité biologique des molécules est mesurée habituellement au moyen d'essais afin d'établir le niveau d'inhibition d'une transduction de signal ou d'une voie métabolique particulière. Les produits chimiques peuvent être biologiquement actifs par leur toxicité. La recherche de médicament implique parfois l'utilisation de la QSAR afin d'identifier les structures chimiques pouvant présenter de bons effets inhibiteurs sur des cibles spécifiques et possèdent une faible toxicité (activité non-spécifique). Un intérêt particulier est la prédiction du coefficient de partition log P, qui est une mesure importante pour l'identification de la «similarité médicamenteuse » selon la règle des cinq de Lipinski.

Alors que de nombreuses analyses QSAR traitent des interactions d'une famille de molécules avec une enzyme ou un site récepteur, la QSAR peut aussi être utilisée pour étudier les interactions entre les domaines structuraux des protéines. Les interactions entre protéines peuvent être quantitativement analysées pour des variations structurales résultantes d'une mutagenèse **[41]** La réduction du risque de paradoxe RAS fait partie de l'apprentissage automatique, et particulièrement la prise en compte du fait que seul un nombre fini de données est disponible (voir aussi estimateur de minimum de variance non biaisé). En général, tous les problèmes QSAR peuvent être partagés entre une partie codage **[42].** Et une partie apprentissage **[43].**

II-4-3 Analyse de données

La première étape dans la modélisation par **QSAR** est la représentation des structures moléculaires par des nombres nommés des descripteurs moléculaires. Un nombre relativement important de caractéristiques ou de descripteurs moléculaires est

calculé pour le codage, qui peut cependant manquer de pertinence sur l'interprétation structurale. Il existe donc un problème de choix de variable, qui peut se résoudre par des méthodes d'apprentissage, appliquées comme étape de post-traitement ou de prétraitement. Parmi ces méthodes, les machines à vecteurs de support, les décisions, ou les réseaux de neurones peuvent être utilisées pour induire un modèle d'apprentissage prédictif **[38].**

Les descripteurs basés sur les propriétés pharmacophores mettent en relation les propriétés physicochimiques des atomes avec la topologie de la molécule. Ces propriétés pharmacophoriques sont essentiellement les atomes donneurs et accepteurs de liaisons hydrogène, les charges, les centres hydrophobes et les cycles aromatiques. La distance topologique entre deux éléments dupharmacophore donne une empreinte pharmacophorique qui pourra être utilisée pour la recherche de molécules de topologie similaire **[38].**

D'autres descripteurs de type QSAR (Quantitative Structure-Activity Relationship) sont des modèles prédictifs qui ont pour objectif d'estimer la capacité d'une molécule à devenir un médicament suffisamment stable dans l'organisme pour être actif et concernent par exemple la lipophilie d'une molécule **[44].**

Différents formats de fichiers ont été décrits pour représenter la structure des molécules. Les plus connus sont SMILES, SDF, mol2 (Tripos) et PDB (Protein Data-Bank). Le format SMILES (Simplified Molecular Input Line Entry System) permet de représenter les molécules sur une seule ligne de caractères. Les tirets simples ou doubles représentent des liaisons de même nature, et un système de parenthèses permet de repérer les embranchements au niveau de la molécule. Les chiffres permettent d'identifier les connections de cycle. Le problème de ce code est sa dégénérescence. En effet, à une même molécule peuvent correspondre plusieurs codes SMILES. Le code SMILES canonique a permis de résoudre ce problème. L'avantage principal de ce format est qu'il est léger, permettant le stockage des bases des molécules de grande taille dans un espace mémoire restreint. Cependant, d'autres informations sont souvent nécessaires, notamment la structure 3D de la molécule. Des formats comme PDB, SDF ou mol2 contiennent les coordonnées cartésiennes de chaque atome, et éventuellement leur charge partielle (mol2), ainsi que toutes les liaisons formant la molécule. Open Babel permet la conversion d'un format 2D vers un autre format 3D. La conversion d'un format 2D vers un format 3D nécessite de générer de l'information et est rendue possible grâce à des outils comme Omega, Corina, Catalyst, Concord ou Frog. Selon la molécule, différents états d'ionisation et stéréoisomères peuvent être générés et il est important de le vérifier avant **[38].**

II-5 LA QSAR EN TOXICOLOGIE DE L'ENVIRONNEMENT

La QSAR est un outil pour la prédiction de l'activité biologique, et se prête ainsi facilement à la prédiction de la toxicité pour l'environnement. Au cours des 20 dernières années, la QSAR dans le domaine environnemental n'a cessé de croître en importance, et Nendza **[45]**.a admirablement résumé ses réalisations. Il a maintenant atteint le stade où certains organismes de réglementation, tels que l'Agence américaine de protection de l'environnement, utilise régulièrement des QSAR pour prédire la toxicité à des fins réglementaires.Il est prévu que cette utilisation va augmenter fortement à l'avenir, à mesure que des assurances sont demandées sur la sécurité des produits chimiques d'une part et plus la pression publique est exercée contre l'utilisation des animaux dans les essais de toxicité d'autre part. Il convient de noter, toutefois, que les données expérimentales sur la toxicité sont nécessaires en premier lieu afin de développer une QSAR et il y'a encore une pénurie de données de bonne qualité dans de nombreux domaines.

II-6 LA QSAR POUR LA PREDICTION DE LA TOXICITE

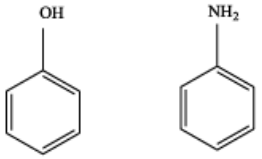
La plupart des données sur la toxicité de l'environnement ont été obtenues en utilisant des animaux aquatiques comme les poissons de diverses espèces, les daphnies, TetrahymenapyriformisVibriofischeri, et les algues. Cronin et Dearden **[46]** ont passé en revue la littérature concernant la modélisation QSAR de la toxicité aquatique. Huit modes d'action ont été identifiées chez les poissons, à savoir la narcose non polaire, narcose polaire, le découplage de la phosphorylation oxydative, l'irritation de la membrane respiratoire, l'inhibition de l'acétylcholinestérase, la saisie du système nerveux central, l'inhibition de la photosynthèse, et l'alkylation **[47]**. Toutefois, ceux-ci sont généralement plus regroupés comme suit: la narcose non polaire, narcose polaire, la réactivité sélective, et les mécanismes d'action spécifiques. Il est important, afin d'obtenir une prédiction RQSA correcte de la toxicité, que le mode d'action d'un produit chimique soit correctement identifié.

La corrélation QSAR première narcose non polaire a été développée par Könemann **[48].** Qui a corrélé la toxicité aiguë de divers produits chimiques industriels.

$$\text{Log1/LC}_{50}=0.87\text{logp-1.} \qquad \textbf{II.1}$$

Où LC50 = concentration (mmol/L) pour tuer 50% des poissons dans un délai déterminé.
Veith et Broderius **[49**] ont signalé que certains produits chimiques réactifs, tel que le phénol (2) et l'aniline (3) produits dérivés, qui produisent des effets toxiques compatibles avec narcose ont néanmoins été prévus plus toxiques que pour la narcose non polaire. Ce mode d'action est maintenant appelé narcose polaire, et peut résulter de la présence d'un groupe de liaisons hydrogène fortes dans une molécule. La corrélation trouvée par Veith et Broderius **[49].** De toxicité pour la tête de boule, Pimephales promelas, a été :

$$Log1/LC_{50}=0.65logp-0.71 \quad \textbf{II.2}$$



2 **3**

En règle générale, la toxicité des produits chimiques réactifs ne peut être modélisée que par l'inclusion d'un ou plusieurs descripteurs qui reflètent la réactivité. Généralement, une telle réactivité électrophile est dans la nature de la molécule, puisque les groupes nucléophiles tels que NH, OH et SH abondent dans les macromolécules biologiques. Cronin et Schultz **[50]** ont trouvé ce qui suit : la QSAR pour la toxicité des composés aromatiques par rapport à un cilié aquatiques, Tetrahymena pyrifomis :

$$Log1/IGC_{50}=0.603logp-0.330\,_{ELUMO}-1 \quad \textbf{II.3}$$

Avec n=239, r^2=0.8000, q^2=0.796, s=0.335, f=476etoù IGC50 = la concentration (mmol/L) pour inhiber la croissance de 50%, ELUMO = énergie de l'orbitale moléculaire inoccupée plus bas, Q = corrélation croisée validé coefficient (procédure de leave-one-out), et F = Fisher statistique.

Les statistiques fournies par l'équation II.3 sont les statistiques préférées pour la déclaration des QSAR. La corrélation de coefficient d'une validation croisée est une mesure de la productivité des QSAR (par opposition à sa capacité de seulement corrélative); F donne une mesure de probabilité que la corrélation n'a pas apparue par hasard.

Comme la plupart des QSAR sont développées à des fins prédictives, il est important que leur capacité prédictive soit évaluée. Bien que la validation croisée est une méthode d'évaluation de la capacité prédictive, il s'agit d'une méthode interne, et peut être critiquée comme étant simplement de donner une indication de la cohérence interne de l'ensemble des données utilisées pour élaborer les QSAR (l'ensemble d'apprentissage). Une meilleure méthode consiste à utiliser le modèle QSAR pour prédire la toxicité des produits chimiques non utilisés dans l'ensemble d'apprentissage (validation externe).

D'autres mesures de réactivité peuvent être utilisées dans QSAR impliquant des produits chimiques réactifs. Un exemple intéressant est donné par Mekenyan et Veith **[51]**. qui ont fait une corrélation entre la toxicité des produits chimiques avec trois modes d'action différents pour P. promelas par l'utilisation du descripteur orbitale moléculaire connue sous le nom de super de localisability:

$$\text{Log } 1/LC_{50}=0.56 \log p + 13.7\ s^{N}_{av} - 1.49 \qquad \textbf{II.4}$$

Avec n=114,r^2=0.81,s non donné

Où SavN = superdelocalisability moyenne nucléophile.

Les produits chimiques spécifiquement actifs sont caractérisés par les insecticides organophosphorés comme le malathion (4).Hermens et al **[52]** ont découvert qu'ils pouvaient modéliser la toxicité de ces composés à P.reticulata avec un QSAR à deux mandats

$$\text{Log } 1/\ LC_{50}=0.23\sum \pi + 0.80k + 2.77 \qquad \textbf{II.5}$$

Avec n=9,r^2=0.92

Où Sp = somme des constantes hydrophobes fragmentaire, et k = taux de réaction constante déterminée expérimentalement avec le 4-nitrobenzylpyridine.

II-7 LES SYSTEMES EXPERTS POUR LA PREDICTION DE LA TOXICITE

Un système expert a été défini **[53].** Comme « tout système formel, pas nécessairement sur ordinateur, qui permet à un utilisateur d'obtenir des prévisions rationnelles sur les propriétés des produits chimiques ». Greene **[54].** A récemment passé en revue les systèmes experts pour la prédiction de la toxicité.

Un certain nombre de ces systèmes sont disponibles dans le commerce, alors que la plupart d'entre eux s'occupent principalement des risques pour la santé de l'homme tels que la cancérogénicité, la tératogénicité et la sensibilisation de la peau. Certains comprennent des modules pour la prédiction écotoxicologique.
Parmi ces systèmes d'expert, on cite quelques uns :

- ✓ **TOPKA**

Ce système utilise une approche QSAR, qui repose en grande partie sur des descripteurs topologiques, pour prédire des valeurs CL50 pour Daphnia magna et P. promelas.

- ✓ **MULTICASE**

Ce système identifie des fragments structuraux qui sont responsables d'une toxicité donnée, y compris des valeurs CL50 chez les poissons.

✓ **HAZARDEXPERT**

Ce système intègre les connaissances d'experts de l'homme pour identifier les fragments structurels liés à la toxicité. Il fournit des prédictions à travers une série de niveaux trophiques avec les différents régimes de dosage.

✓ **ASTER**

Ce système a été développé par l'Agence américaine de protection de l'environnement. Les points de terminaison écotoxicologiques sont largement couverts, des valeurs de la CL50 nommément pour P. promelas, méné de-mouton (Cyprinodon) et D. magna. Il s'agit essentiellement d'une base de données, mais utilise QSAR pour prédire la toxicité lorsque des données expérimentales ne sont pas disponibles. Il utilise des modèles QSAR pour les différents types de produits chimiques (stupéfiants non-polaires, polaires et de stupéfiants ainsi de suite). En raison des restrictions de sécurité, ASTER n'est actuellement pas accessible au public.

✓ **ECOSAR**

Ce système est librement téléchargeable sur site. Il utilise QSAR hydrophobicité et prédit la toxicité pour les poissons, les daphnies et les algues vertes **[9].**

II-8 CALCUL DES PROPRIETES PHYSICO-CHIMIQUES

Un certain nombre de propriétés physico-chimiques sont d'une importance environnementale, que les QSAR données ci-dessus illustrent.

▪ **Coefficient de partage (P) :**

C'est sans doute le plus important d'entre eux, et il ya de nombreux logiciels disponibles pour le calcul de log P. Parmi les meilleurs, on cite les suivants :
Analyse interactive qui permet gratuitement de calculer en ligne ; Biobyte, SPARC**,** qui permet gratuitement de calculer en ligne, et le logiciel SRC KOWWIN, qui est téléchargeable gratuitement à partir de l'EPA **[9].**

▪ **Solubilité aqueuse :**

C'est une autre propriété importante de l'environnement. L'analyse interactive et de sites web SPARC permet le libre calcul en ligne de solubilité, et WSKOWWIN de la SRC est téléchargeable gratuitement à partir de l'EPA site web.

- **Point d'ébullition :**

Il est disponible à partir de SPARC (librement en ligne), MPBPVP de la SRC**,** ACD et ProPred **[9].**

- **pression de vapeur :**

Il est disponible à partir de SPARC (librement en ligne) et ACD.

- **point de fusion** :

Il est disponible auprès de MPBPVP et ProPred.

- **Le partage air-eau :**

Ce coefficient (constante de Henry) a des implications importantes pour la distribution de produits chimiques dans l'environnement. De nombreuses tentatives ont été faites pour le prédire ; Dearden et Schüürmann **[55].** Ont récemment revu la prévision QSAR de la constante d'Henry.

II-9 SMILES (Simplified Molecular Input Line Entry Specification)

II-9-1 Notation SMILES

The Simplified Molecular Input Line Entry Specification ou **SMILES** est un langage symbolique de description de la structure des moléculeschimiques sous forme de courtes chaînes de caractèresASCII. Les chaînes SMILES peuvent être importées dans la plupart des éditeurs de structure moléculaire pour être reconverties en représentation 2D ou en modèle 3D, et par conséquent, permettent de construire des bases de données informatiques **[38].**

La définition du langage SMILES a été initialement introduite par David Weininger à la fin des années 1980. Plusieurs extensions et modifications ont été ajoutées depuis, notamment par la société DaylightChemical Information Systems, dont Weininger est le fondateur. Il existe d'autres langages à ligne de texte pour représenter les molécules, dont notamment InChl, introduit par l'IUPAC. SMILES est cependant plus répandu, car plus facile à lire par l'utilisateur humain et surtout parce qu'il est implémenté dans de nombreux logiciels de chimie moléculaire **[38].**

II-9-2 Principes de SMILES et théorie des graphes

SMILES est bâti sur la représentation d'une molécule chimique sous forme d'un graphe, au sens mathématique du terme. Les atomes sont les sommets du graphe, et les liaisons chimiques en sont les arêtes. Chaque sommet du graphe est étiqueté avec le symbole de l'atome correspondant (C, N, O…) **[56].**

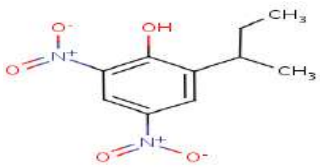
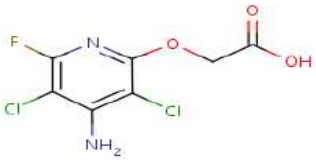
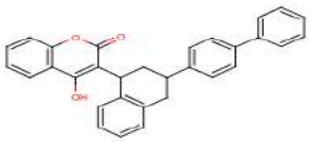
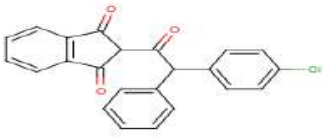
La chaîne de caractère SMILES correspondant à une molécule donnée est obtenue en parcourant progressivement le graphe suivant un algorithme de parcours en profondeur et en imprimant successivement le symbole de chaque sommet traversé **[56].**Le graphe de la molécule est préalablement simplifié en supprimant tous les atomes d'hydrogène (notation implicite) et en ouvrant les cycles pour transformer le graphe en arbre couvrant (spanningtree). À chaque fois qu'un cycle est ouvert, un suffixe numérique est ajouté pour indiquer la connexion des sommets correspondant à la liaison chimique supprimée. Des parenthèses sont utilisées pour indiquer les points de branchement sur l'arbre**, [56].**

De par la nature de ce langage, il existe, pour une même molécule organique constituée de plusieurs atomes autres que l'hydrogène, plusieurs écritures possibles, selon l'atome de départ et l'ordre de parcours de la structure moléculaire. Ainsi, CCO, OCC, C(C) O et C(O) C sont toutes les quatre, des représentations valides de la molécule d'éthanol CH_3–CH_2–OH **[56].**

Les principes théoriques du langage SMILES sont présentés de manière exhaustive sur le manuel du SMILES théorique proposé par Daylight Chemical Information Systems (en anglais). Le tableau II.1 illustre la notation SMILES pour quelques molécules **[56].**

Il existe des algorithmes informatiques pour transformer automatiquement le schéma d'une molécule donnée en chaîne SMILES. Comme il n'existe en général pas de description unique d'une molécule, il y a également plusieurs SMILES pouvant décrire la même structure chimique. Tout dépend de l'atome dont on part pour décrire la structure, de l'ordre dans lequel on décrit les branches du graphe moléculaire et de la manière dont on ouvre les cycles **[56].**

Tableau II.1 Notation SMILES pour quelques molécules

Nom	Formule brute	Notation de smille	Structure chimique
dinoseb	C10H12N2O5	OC1=C([N+]([O-])=O)C=C([N+]([O-])=O)C=C1C(CC)C	
Piperonyl butoxide	C19H30O5	CCCC1=CC2=C(OCO2)C=C1COCCOCCOCCCC	
fluroxypyr	C7H5Cl2FN2O3	FC1=C(Cl)C(N)=C(Cl)C(OCC(O)=O)=N1	
difenacoum	C31H24O3	O=C(C(C1CC(C2=CC=C(C3=CC=CC=C3)C=C2)CC4=C1C=CC=C4)=C5O)OC6=C5C=CC=C6	
chlorophacinone	C23H15ClO3	O=C1C(C(C(C2=CC=CC=C2)C3=CC=C(Cl)C=C3)=O)C(C4=C1C=CC=C4)=O	

Il existe plusieurs logiciels utilisés pour traduire la structure en formule SMILES, citons :

- **ChemOffice :** est une suite de logiciels et d'utilitaires offrant lenec plus ultra en matière de recherche. Elle se décline dans les versions Pro ou Ultra selon les besoins du scientifique ou du chercheur dans le domaine de la chimie **[57].**

- **ChemBioDrawUltra :** est un logiciel de représentation de structures pour spécialistes chevronnés, offrant des outils de prévision perfectionnés et une intégration Web complète par le biais du complément ChemBioDraw ActiveX / Plugin. La suite

ChemBioDraw Ultra comprend également ChemDraw / Excel, ChemBio3D Pro, ChemFinder Ultra avec BioViz, E-Notebook Pro et les bases de données ChemInfo **[57]**.

La méthode SMILES permet d'échanger numériquement des descripteurs de molécules arbitrairement complexes par le biais d'un fichier texte classique. Ces descripteurs peuvent êtres calculées par d'autres logiciels tels que DRAGON.Ce dernier est une application utilisée pour le calcul des descripteurs moléculaires à l'origine développé par la Chimio-métrie Milano et QSAR Research Group **[57]**.

Ces descripteurs peuvent être utilisés pour évaluer la structure moléculaire-activité ou structure-propriété des relations, ainsi que pour l'analyse de similitude et de dépistage des bases de données moléculaire. La représentation SMILES étant très compacte, elle permet de stocker simplement des fichiers de plusieurs centaines de milliers de molécules **[57]**.

II-10 CONCLUSION

La relation quantitative structure activité(QSAR) est depuis quelques années utilisée de façon systématique dans la conception de nouveaux médicaments et permet de prédire l'activité thérapeutique de molécules avant même de les avoir synthétisées. Cette prédiction, bien que non totale, a l'avantage d'économiser une partie de l'expérimentation initiale. La relation structure-activité est devenue partie intégrante d'un arbre de décision.

La même démarche a été utilisée pour prévoir une toxicité déterminée. Peut-être moins rationnelle ou plus approximative, cette méthode a donné dans le passé des résultats décevants en toxicologie. A partir de ce constat, des travaux récents impliquant des toxicologues, des chimistes et des informaticiens ont montré que le respect d'un certain nombre de règles simples permettait d'obtenir des résultats comparables à ceux obtenus lors du développement de molécules thérapeutiques.

SMARTS est une extension de SMILES qui permet, en plus des définitions classiques, de remplacer des atomes ou des liaisons par des jokers. Ceci est utilisé pour spécifier des motifs chimiques utilisés pour des recherches dans des banques de données de molécules (chimiothèques).
Cela permet en particulier la recherche de sous-structures chimiques communes à deux molécules. Les logiciels développés facilitent les calculs des descripteurs et les propriétés des molécules sans essais pratiques.

Chapitre III

Les Réseaux de Neurones Artificiels

III-1 INTRODUCTION

L'informatique a pris beaucoup d'ampleur avec l'arrivée de l'intelligence artificielle. De nombreuses expériences sont recensées dans cet annuaire **[58].** L'usage de l'expression«intelligence artificielle» s'est largement répanduau cours de l'histoire. L'intelligence artificielle, ou plus généralement noté I.A, estdueaux progrès techniques de l'informatique et à sa propagation dans les activités humaines. De nos jours, l'intelligence artificielle fait partie de notre culture comme en témoignent les films, les livres ou encore les jeux vidéo **[59].**

Parmi les applications de l'intelligence artificielle, on trouve le modèle des réseaux de neurones qui depuis une dizaine d'années a vu son utilisation se développer dans de nombreuses disciplines (sciences économiques, écologie et environnement, biologie et médecine…). Ils sont notamment appliqués pourrésoudre des problèmes de classification, de prédiction, de catégorisation, d'optimisation, de reconnaissance des formes et de mémoire associative **[60].**

III-2 L'INTELLIGENCE ARTIFICIELLE

III-2-1 Définition

Un système d'intelligence artificielle a pour but la modélisation ducomportement d'un expert humain, accomplissant une tâche de résolution deproblèmes pour laquelle on ne dispose d'aucun algorithme et ce dans un domaine bien précis **[61].**

Il existe un certain nombre de définitions différentes de l'IA. On cite entre autres :

- ✓ Système qui **pensecomme** les **hommes : [Modélisation cognitive** (GPS (Newel & Simon, 61))]
- ✓ Système qui **agit** comme les **hommes : [Système passant le test de Turing(1950)** Apprendre des K. (s'adapter)**,** Représenter des K. (mémoriser)**,** Résoudre des Pb. (raisonner) Comprendre (communiquer)]
- ✓ Système qui **pense rationnellement : Logicisme :** pensée logique Pascal [1623-1662] (machine à calculer), Leibnitz [1646-1716] (machine à raisonner), Babbage [1792-1871] (machine analytique)
- ✓ Système qui **agit rationnellement : Agent rationnel (199X)** agit selon ses croyances pour atteindre des objectifs (pas seulement logique) **[62].**

III-2-2 Historique

L'origine de l'intelligence artificielle se trouve probablement dans l'article d'Alan Turing **[63]** .où il explore le problème et propose une expérience maintenant connue sous le nom de test de Turing, dans une tentative de définition d'un standard permettant de qualifier une machine de « consciente ». Il développe cette idée dans plusieurs forums, dans la conférence « L'intelligence de la machine, une idée hérétique » **[64].** Dans la conférence qu'il donne à la BBC 3^{em} programme le 15 mai 1951 « Est-ce que les calculateurs numériques peuvent penser ? » d'où la discussion avec M.H.A. Newman, AMT, Sir Geoffrey Jefferson et R.B. Braithwaite le 14 et 23 Jan. 1952 sur le thème « Est-ce que les ordinateurs peuvent penser? » **[65].**

On considère que l'intelligence artificielle, en tant que domaine de recherche, a été créée à la conférence qui s'est tenue sur le campus de Dartmouth Collège pendant l'été 1956 **[66] .**à laquelle assistaient ceux qui vont marquer la discipline. Ensuite l'intelligence se développe surtout aux États-Unis à l' université Stanford sous l'impulsion de John McCarthy, au MIT sous celle de Marvin Minsky, à l' université Carnegie Mellon sous celle de Allen Newell et Herbert Simon et à l' université d'Édimbourg sous celle de Donald Michie. En France, l'un des pionniers est Jacques Pitrat.

III-2-3 Domaines de l'intelligence Artificielle

- **Tâches routinières**

  - ✓ Perception : – Vision
    – Parole

- **Langage naturel :** – Compréhension
  – Génération
  – Traduction
  - ✓ Sens commun
  - ✓ Robotique

- **Tâches formelles :**

  - ✓ Jeux : – Échecs
    – Backgammon
    – Dames
  - ✓ Mathématiques : – Géométrie
    – Logique
    – Calcul intégral
    – Preuves de théorèmes

- **Systèmes experts**

  - ✓ Ingénierie : – Conception
    – Réparation, déverminage
    – Planification industrielle
  - ✓ Analyse scientifique
  - ✓ Diagnostic médical
  - ✓ Analyse financière

- **Réseaux de neurones artificiels**

  - ✓ Classification
  - ✓ Apprentissage supervisé et non-supervisé : – Perceptron multicouche
    – Réseaux auto-organisés
    – Simulations biologiques
  - ✓ Imagerie et reconnaissance de formes **[67].**

II-2-4 Les modèles de l'intelligence artificielle

Il existe plusieurs modèles utilisables, parmi eux on cite :

- Les réseaux neuronaux artificiels
- Calcul évolutionnistes
- Intelligence D'Essaim
- Systèmes immunitaires Artificiels
- Systèmes Brouillés

Le modèle le plus utilisé dans les années précédentes est le réseau de neurone **[68].**

III-3 RESEAUX DE NEURONES

III-3-1 Introduction aux réseaux de neurones

Les réseaux de neurones ont d'abord été développés pour résoudre des problèmes de contrôle,de reconnaissance de formes ou de mots, de décision, de mémorisation comme une alternativeà l'intelligence artificielle, et en relation plus ou moins étroite avec la modélisation deprocessus cognitifs (capable de connaître ou faire connaître) réels et des réseaux de neurones biologiques **[69].**

III-3-2 Définition

Les réseaux de neurones sont composés d'éléments simples (ou neurones) fonctionnant en parallèle. Ces éléments ont été fortement inspirés par le système nerveux biologique. Comme dans la nature, le fonctionnement du réseau (de neurone) est fortement influencé par la connections des éléments entre eux. On peut entraîner un réseau de neurone pour une tâche spécifique (reconnaissance de caractères par exemple) en ajustant les valeurs des connections (ou poids) entre les éléments (neurone) **[70].**

III-3-3 Historique

En 1943, McCulloch et Pitts proposent le modèle du neurone formel inspiré des observationsneurophysiologiques et anatomiques du neurone biologique.Les neurones constituent des unités simples (automates linéaires à seuils) mais qui assemblés et sont capable de résoudre des fonctions logiques.

En 1949, Donald Hebb propose une règle d'apprentissage pour les réseaux de neurones artificiels. En 1962, Rosenblatt avec le perceptron et Widrow avec l'adaline proposent un modèle de réseau à une couche et une méthode pour modifier les connexions. Rosenblatt démontre la convergence d'un algorithme itératif d'adaptation des poids pour un perceptron monocouche pour apprendre certaines fonctions logiques.

En 1969, Minsky and Papert mettent en évidence les limitations du perceptron.Ils apportent un éclairage théorique sur certains problèmes impossibles à résoudre avec ce modèle, comme le OUexclusif (XOR).

Durant 16 ans, les chercheurs vont se détourner de l'approche connexionniste pour lui privilégier une approche symbolique. Ainsi, l'Intelligence Artificielle (IA) symbolique a connu ses heures de gloire et a permis de grandes avancées comme celles des systèmes experts. Cependant, les chercheurs se sont vite rendu compte des limites de l'IA symbolique pour modéliser le fonctionnement du cerveau, qui n'est manifestement pas basé sur une logique symbolique. Bien heureusement, certains chercheurs sont restés fidèles à l'approche connexionniste et ont poursuivi leurs recherches, notamment dans le domaine des mémoires associatives.

En 1982, Hopfield présente un modèle de réseau de neurones entièrement connecté et dotéd'un formalisme plus riche.Hopfield introduit la notion de fonction d'énergie et associe la mémoire à des attracteurs dans les systèmes dynamiques. Par ailleurs, ``la machine de Boltzmann" est le premier modèle à lever les limitations du perceptron par la méthode du recuit simulé.

Cependant, l'avancée la plus marquante est l'algorithme de rétropropagation découvert tout d'abord par Werbos en 1974, puis parallèlement par Rumelhart en 1986, Parker en 1985 et Le Cun en 1985. Cette approche, bien que non biologiquement plausible,

lève la limitation du perceptron de Rosenblatt. Cette méthode est destinée à des modèles multicouches statiques(MLP) (sans récurrences).Depuis, l'algorithme de rétropropagation a été étendu à des réseaux de neurones admettant desconnexions récurrentes.

Les algorithmes de rétropropagation dynamique permettent de modéliser les processusdynamiques tels que les séries temporelles. Les algorithmes de rétropropagation reposent sur leprincipe de minimisation d'une fonction de l'erreur commise en sortie du réseau et effectuent uncalcul et une descente du gradient pour trouver le minimum de cette fonction coût **[71].**

III-3-4 Applications des réseaux de neurones

L'étude des réseaux de neurones est une voie prometteuse de l'Intelligence Artificielle, qui a des applications dans de nombreux domaines :

• Aérospatial : pilotage automatique, simulation du vol…
• Automobile : système de guidage automatique,…
• Défense : guidage de missile, suivi de cible, reconnaissance du visage, radar, sonar, lidar, traitement du signal, compression de données, suppression du bruit…
• Electronique : prédiction de la séquence d'un code, vision machine, synthétiseur vocal, modèle non linéaire,…
• Finance : Prévision du coût de la vie
• Secteur médical : Analyse EEC et ECG
• Télécommunications : Compression de données … **[70].**

III-3-5 Notions fondamentales sur les réseaux de neurones artificiels

III-3-5-1 Analogie entre neurone biologique et neurone artificiel

L'idée de réseau de neurones artificiels a été inspirée de la manière par laquelleles neurones biologiques traitent les informations. Ce concept est utilisé pour mettre en œuvre des logiciels de simulations des processus massivement parallèles qui impliquent des éléments de traitement connectés dans une architecture de réseau. L'apprentissage dans le cerveau humain se fait à travers un réseau des neurones qui sont connectés par l'axone, les synapses et des dendrites **[72].**

❖ **Neurone biologique**

En biologie, un neurone est une cellule nerveuse dont la fonction est de transmettre un signal électriquedans certaines conditions. Il agit comme un relai entre une couche de neurones et celle qui la suit.

Le corps d'un neurone (figure III.1) est relié d'une part à un ensemble de dendrites (entrées du neurone) et d'autre part à un axone, partie étirée de la cellule, qui représentera

pour nous sa sortie. Le neurone étudié estconnecté aux neurones environnants : il reçoit au niveau de ses dendrites les signaux électriques desneurones "en amont", propagés par les axones de ces derniers. Les charges électriques s'accumulent dansle neurone jusqu'à dépasser un certain seuil : à ce moment la transmission du signal électrique se déclenchevia son axone vers d'autres neurones "en aval".

On remarque que les liaisons axone/dendrite entre deux neurones (connexions synaptiques) ne sont pas toutes de la même efficacité. Ainsi l'entrée associée à une certaine dendrite du neurone pourra avoir plus d'influence qu'une autre sur la valeur de sortie. On peut représenter la qualité de la liaison par un poids, sorte de coefficient s'appliquant au signal d'entrée. Le poids sera d'autant plus grand que la liaison est bonne. Un poids négatif aura tendance à inhiber une entrée, tandis qu'un poids positif viendra l'accentuer.

La structure d'un neurone se compose de trois parties :

- ***La somma*** : ou cellule d'activité nerveuse, au centre du neurone.
- ***L'axone*** : attaché au somma qui est électriquement actif, ce dernier conduitl'impulsion conduite par le neurone.
- ***Dendrites*** : électriquement passives, elles reçoivent les impulsions d'autres neurones **[73].**

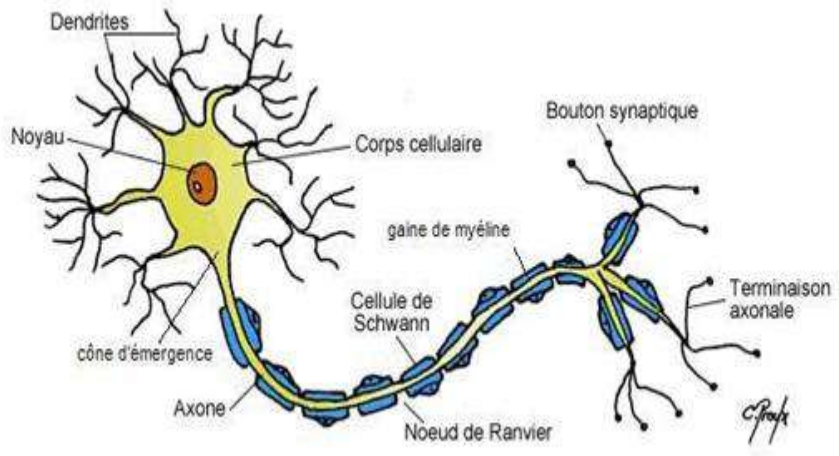


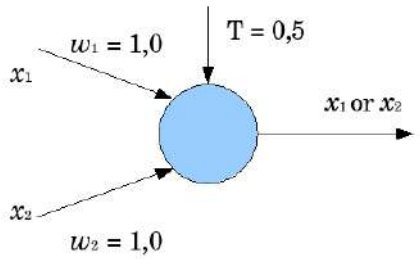
Figure III.1 Schéma complet d'un neurone biologie **[75].**

❖ Le neurone formel (artificiel)

Un neurone formel est une représentation mathématique et informatique d'un neurone biologique. Le neurone formel possède généralement plusieurs entrées et une sortie qui correspondent respectivement aux dendrites et au cône d'émergence du neurone biologique (point de départ de l'axone). Les actions excitatrices et inhibitrices des synapses sont représentées, la plupart du temps, par des coefficients numériques (les poids synaptiques) associés aux entrées. Les valeurs numériques de ces coefficients sont ajustées dans une

phase d'apprentissage. Dans sa version la plus simple, un neurone formel calcule la somme pondérée des entrées reçues, puis applique à cette valeur une fonction d'activation, généralement non linéaire. La valeur finale obtenue est la sortie du neurone

Le neurone formel (figure III.2) est l'unité élémentaire des réseaux de neurones artificiels dans lesquels il est associé à ses semblables pour calculer des fonctions arbitrairement complexes, utilisées pour diverses applications en intelligence artificielle **[74].** Mathématiquement, le neurone formel est une fonction à plusieurs variables et à valeurs réelles.



FigureIII.2 Représentation d'un neurone formel **[74].**

❖ Modélisation d'un neurone formel

Les réseaux de neurones formels sont à l'origine d'une tentative de modélisation mathématique du cerveau humain. Les premiers travaux datent de 1943 et sont l'œuvre deMM. Mac Culloch et Pitts. Ils présentent un modèle assez simple pour les neurones etexplorent les possibilités de ce modèle.

La modélisation consiste à mettre en œuvre un système de réseaux neuronaux sous un aspect non pas biologique mais artificiel, cela suppose que d'après le principe biologique, on aura une correspondance pour chaque élément composant le neurone biologique, donc une modélisation pour chacun d'entre eux.
On pourra résumer cette modélisation sur le tableau III.1, qui nous permettra de voir clairement la transition entre le neurone biologique et le neurone formel **[69].**

Tableau III.1 Analogie entre le neurone biologique et le neurone formel **[69].**

Neurone biologique	**Neurone artificiel**
Synapses	Poids de connexions
Axones	Signal de sortie
Dendrite	Signal d'entrée
Somma	Fonction d'activation

Nous avons vu, que les neurones sont des cellules organisées en « couche » : c'est à dire que lorsque un neurone reçoit un signal électrique par ses synapses, cela déclenche ou non un autre signal sur la couche suivante, etc. (cf. "2"). Ce fonctionnement est reproduit dans les réseaux de neurones artificiels **[75].**

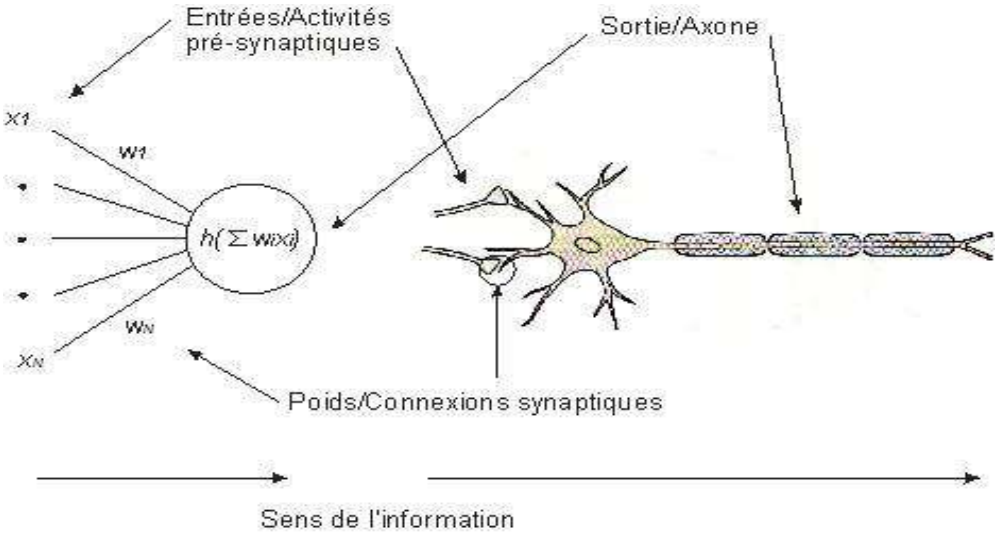


Figure III.3 Schéma de comparaison entre un neurone biologique et un neurone artificiel **[75].**

III-3-6 Le modèle mathématique d'un neurone artificiel

Le modèle mathématique d'un neurone artificiel (ou "perceptron") est illustré à la figure ci-dessous. Un neurone est essentiellement constitué d'un intégrateur qui effectue la somme pondérée de ses entrées (comme l'espérance statistique!). Le résultat n de cette somme ensuite transformée par une fonction de transfert *f* qui produit la sortie *a* du neurone.

Les *R* entrées du neurone correspondent au vecteur noté traditionnellement en ligne (mais en fait on prend la transposée d'où le *T* en suffixe)

$$\vec{p} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_R \end{pmatrix}$$

Alors que :

$$\vec{w} = \begin{pmatrix} w_{1,1} \\ w_{1,2} \\ \vdots \\ w_{1,R} \end{pmatrix}$$

Représente le vecteur des poids du neurone (nous les distinguons pour préparer le terrain à des neurones un peu plus complexes).

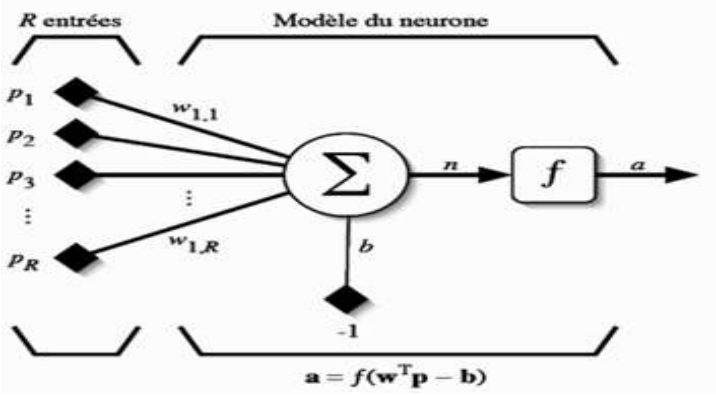


Figure III.4 Modèle mathématique du neurone artificiel **[76].**

La sortie *n* de l'intégrateur est définie par l'équation suivante :

$$n = \sum_{j=1}^{R} w_{1,j} p_j - b = w_{1,1} p_1 + w_{1,2} p_2 + \ldots + w_{1,R} p_R - b \qquad \textbf{III 1}$$

Que nous pouvons aussi écrire sous forme matricielle :

$$n = \vec{w}^T \vec{p} - b \qquad \textbf{III 2}$$

Cette sortie correspond à une somme pondérée des poids et des entrées moins ce que nous nommons "le biais *b* du neurone" (facteur correctif décidé par tâtonnement). Le résultat *n* de la somme pondérée s'appelle le "niveau d'activation du neurone". Le biais *b* s'appelle aussi le "seuil d'activation du neurone". Lorsque le niveau d'activation atteint ou dépasse le seuil *b*, alors l'argument de *f* devient évidemment positif (ou nul). Sinon, il est négatif.

En ajoutant la fonction d'activation *f* pour obtenir la sortie du neurone :

$$a = f(n) = f\left(\vec{w}^T \vec{p} - b\right) \qquad \textbf{III 3}$$

Il est temps maintenant de remplacer (pour alléger la notation) $\vec{w}^T$ par une matrice d'une seule ligne que nous noterons $W = \vec{w}^T$. Nous obtenons alors une forme générale que nous adopterons tout au long de notre étude :

$$a = f(n) = f\left(W\vec{p} - b\right) \qquad \textbf{III 4}$$

Cette équation nous amène à introduire un nouveau schéma plus formel de notre perception des choses :

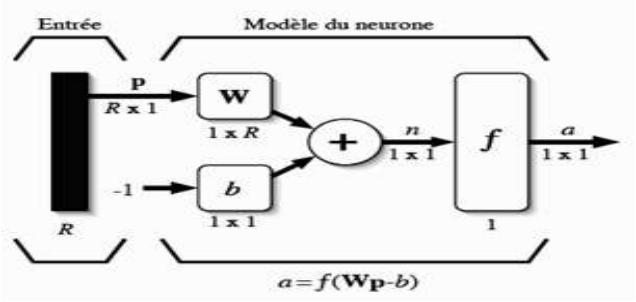


Figure III.5 Schéma plus formel d'un perceptron **[76].**

Nous y représentons les *R* entrées comme un rectangle noir (le nombre d'entrées est indiqué sous le rectangle). De ce rectangle sort le vecteur $\vec{p}$ dont la dimension matricielle est $R \times 1$. Ce vecteur est multiplié par une matrice *W* qui contient les poids (synaptiques) du neurone. Dans le cas d'un neurone simple, cette matrice possède la dimension $1 \times R$.

Le résultat de la multiplication correspond au niveau d'activation qui est ensuite comparé au seuil *b* (un scalaire) par soustraction. Finalement, la sortie du neurone est calculée par la fonction *f*. La sortie d'un neurone simple est alors toujours un scalaire **[76].**

III-3-7 Fonction d'activation

La fonction d'activation, ou fonction de transfert, est une fonction qui doit renvoyer un réel proche de 1 quand les "bonnes" informations d'entrée sont données et un réel proche de 0 quand elles sont "mauvaises". On utilise généralement des fonctions à valeurs dans l'intervalle réel [0,1]. Quand le réel est proche de 1, on dit que l'unité (le neurone) est **active** alors que quand le réel est proche de 0, on dit que l'unité est **inactive.** Le réel en question est appelé la **sortie** du neurone et sera noté *a*. Si la fonction d'activation est linéaire, le réseau de neurones se réduirait à une simple fonction linéaire.

En effet, si les fonctions d'activations sont linéaires, alors le réseau est l'équivalent d'une régression multi-linéaire (méthode utilisée en statistiques). L'utilisation du réseau de neurone est toutefois bien plus intéressante lorsque l'on utilise des fonctions d'activations non linéaires **[77].**

En notant *g* la fonction d'activation, on obtient donc la formule donnant la sortie d'un neurone:

$$a = g(in) = g(\sum_{i=0}^{k} w_i \times x_i) \qquad \textbf{III 5}$$

Remarquez que le coefficient de biais est inclus dans la somme, d'où la formule plus explicite :

$$a = g(in) = g((\sum_{i=1}^{k} w_i \times x_i) - w_0) \qquad \textbf{III 6}$$

Jusqu'à présent, nous n'avons pas spécifié la nature de la fonction d'activation $a = f(n)$ de notre modèle. Il se trouve que plusieurs possibilités existent et celles-ci sont quasiment empiriques et à adapter en fonction des situations. Les plus courantes et les plus citées dans la littérature sont énumérées dans le tableau III.2 ci-dessous :

TableauIII.2 Les fonctions de transfert les plus courantes **[76].**

Nom de la fonction	Relation d'entrée/sortie	Icône
seuil	$a = 0$ si $n < 0$ $a = 1$ si $n \geq 0$	
seuil symétrique	$a = -1$ si $n < 0$ $a = 1$ si $n \geq 0$	
linéaire	$a = n$	
linéaire saturée	$a = 0$ si $n < 0$ $a = n$ si $0 \leq n \leq 1$ $a = 1$ si $n > 1$	
linéaire saturée symétrique	$a = -1$ si $n < -1$ $a = n$ si $-1 \leq n \leq 1$ $a = 1$ si $n > 1$	
linéaire positive	$a = 0$ si $n < 0$ $a = n$ si $n \geq 0$	
sigmoïde	$a = \frac{1}{1+\exp^{-n}}$	
tangente hyperbolique	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$	
compétitive	$a = 1$ si n maximum $a = 0$ autrement	C

Les trois fonctions les plus utilisées dans le domaine de l'ingénierie sont les fonctions "seuil" (a) (en anglais "hard limite"), "linéaire" (b) et "sigmoïde" (c) comme représentées ci-dessous :

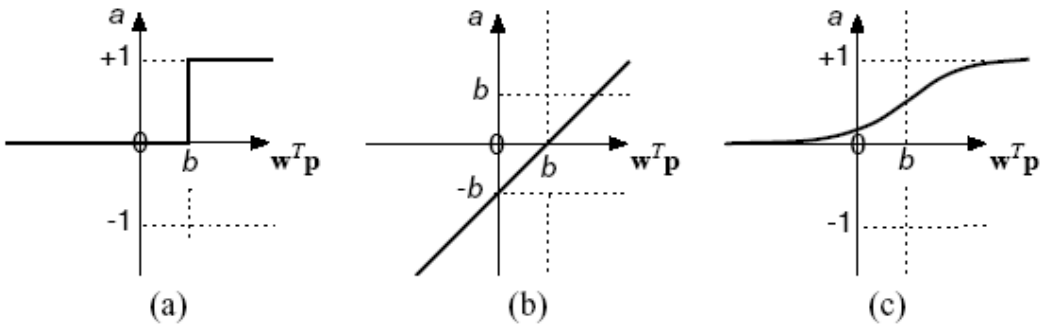


Figure III.6 Fonctions de transfert : (a) du neurone "seuil"; (b) du neurone "linéaire", et (c) du neurone "sigmoïde" **[9]**

Comme son nom l'indique, la fonction seuil applique un seuil sur son entrée.

Plus précisément, une entrée négative ne passe pas le seuil, la fonction retourne la valeur 0 (faux), alors qu'une entrée positive ou nulle dépasse le seuil, et la fonction retourne 1 (vrai). Il est évident que ce genre de fonction permet de prendre des décisions binaires (cette fonction peut aussi être assimilée à la fonction de Heaviside).

La fonction linéaire est quant à elle très simple, elle affecte directement son entrée à sa sortie selon la relation a=f(n)=n Il est évident que la sortie du neurone correspond alors à son niveau d'activation dont le passage à zéro (l'ordonnée à l'origine) se produit lorsque w.p=b.

La fonction de transfert sigmoïde est quant à elle définie par la relation mathématique :

$$a = \frac{1}{1+e^{-n}} \qquad \textbf{III 7}$$

Elle ressemble soit à la fonction seuil, soit à la fonction linéaire, selon que nous sommesrespectivement loin ou près de *b*. La fonction seuil est très non linéaire car il y a une discontinuité lorsque w.P=b. De son côté, la fonction linéaire est tout à fait linéaire. Elle ne comporte aucun changement de pente. La sigmoïde est un compromis intéressant entre les deux précédentes. Notons finalement que la fonction tangente hyperbolique est une version symétrique de la sigmoïde **[76].**

III-3-8 Architecture des réseaux

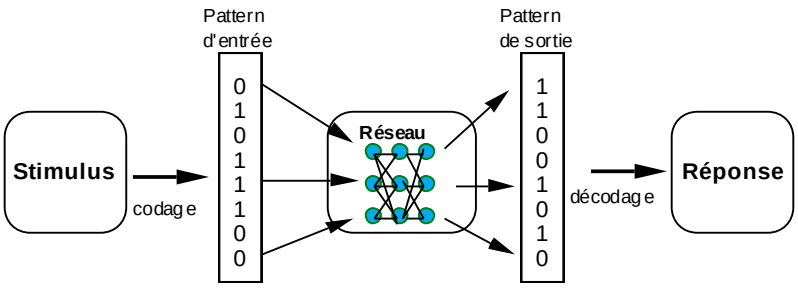
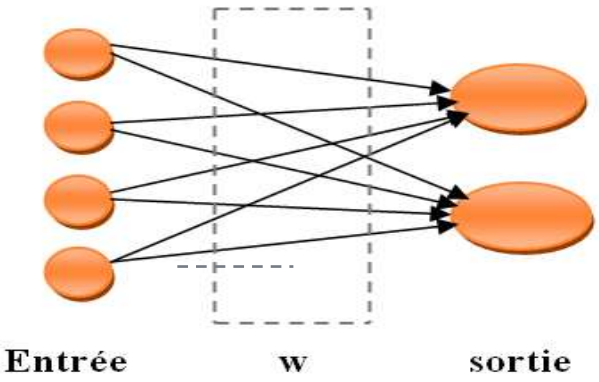


Figure III.7 Architecture générale d'un RNA **[78].**

Les connexions entre les neurones qui composent le réseau décrivent la topologie dumodèle. Elle peut être quelconque, mais le plus souvent il est possible de distinguer une certainerégularité (réseau à connexion complète) **[69].**

- **Réseau monocouche**

La structure d'un réseau monocouche est telle que des neurones organisés en entrée soiententièrement connectés à d'autres neurones organisés en sortie par une couche modifiable de poids (Figure III.8) **[69].**



FigureIII.8 Réseau Monocouche [12]

- **Réseau multicouche**

Les neurones sont arrangés par couche. Il n'y a pas de connexion entre neurones d'une même couche, et les connexions ne se font qu'avec les neurones de couches avales. Habituellement, chaque neurone d'une couche est connecté à tous les neurones de la couche suivante et celle-ci seulement. Ceci nous permet d'introduire la notion de sens de parcours de l'information (de l'activation) au sein d'un réseau et donc définir les concepts de neurone d'entrée, neurone de sortie. Par extension, on appelle couche d'entrée l'ensemble des neurones d'entrée, couche de sortie l'ensemble des neurones de sortie. Les couches intermédiaires n'ayant aucun contact avec l'extérieur sont appelées couches cachées (figure.III.9) **[69].**

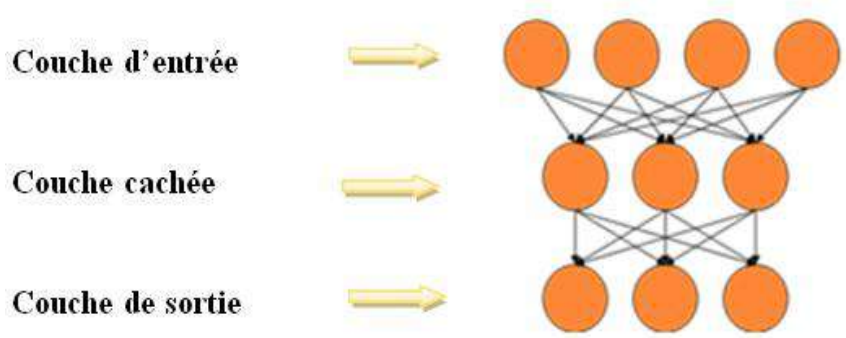


Figure III.9 Réseau Multicouche **[78].**

- **Réseau à connexion complète**

C'est la structure d'interconnexion la plus générale. Chaque neurone est connecté à tous lesneurones du réseau (et à lui-même) (Figure.III.10) **[69].**

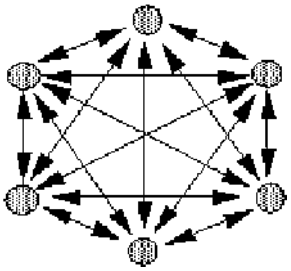


Figure III.10 Réseau à connexions complètes [78].

➢ **Réseau à connexions locales**

Il s'agit d'une structure multicouche, mais qui à l'image de la rétine conserve une certainetopologie. Chaque neurone entretient des relations avec un nombre réduit et localisé deneurones de la couche avale. Les connexions sont donc moins nombreuses que dans le casd'un réseau multicouche classique (figureIII.11) **[69].**

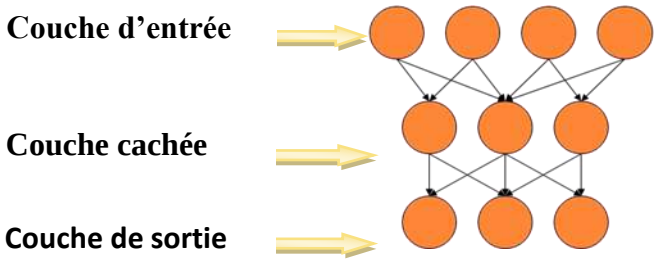


Figure III.11 Réseau à connexions locales [78].

III-3-9 Apprentissage

L'apprentissage est une phase du développement d'un réseau de neurones durant laquelle le comportement du réseau est modifié jusqu'à l'obtention du comportement désiré. On distingue deux grandes classes d'algorithmes d'apprentissage (figure III.12 et III.13) **[78].**

○ L'apprentissage supervisé

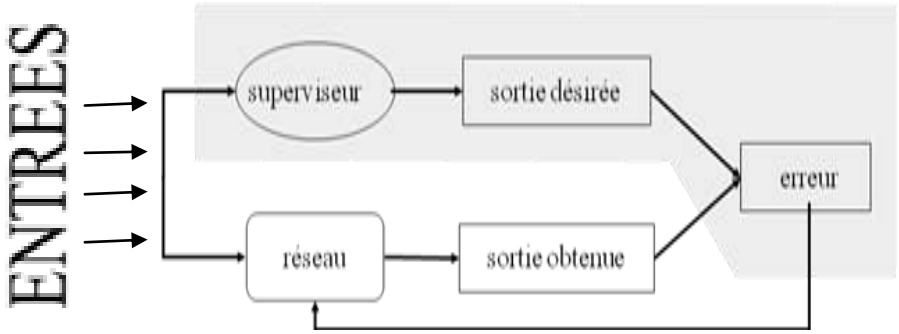


Figure III.12 Schéma représentatif d'un apprentissage supervisé **[78].**

- L'apprentissage non supervisé

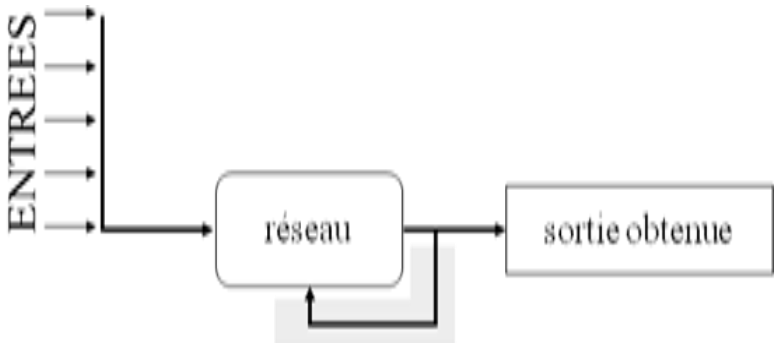


Figure III.13 Schéma représentatif d'un apprentissage non supervisé **[78].**

III-3-9-1 comment apprendre au réseau

Maintenant que l'on a créé un réseau, il faut lui faire apprendre quelque chose. Pour mieux comprendre cela, nous prendrons l'exemple de la reconnaissance de caractères : on présente des caractères au réseau et il a pour objectif de reconnaitre ce caractère. Par exemple, si on lui montre la lettre "a" dessinée, il doit nous répondre que cela est un "a".

Comment réaliser cela ? la méthode est en quelque sorte une imitation du cerveau :si la réponse est correcte,c'est bien, mais s'il y a une erreur, il faut modifier le réseau afin de ne pas réitérer l'erreur.

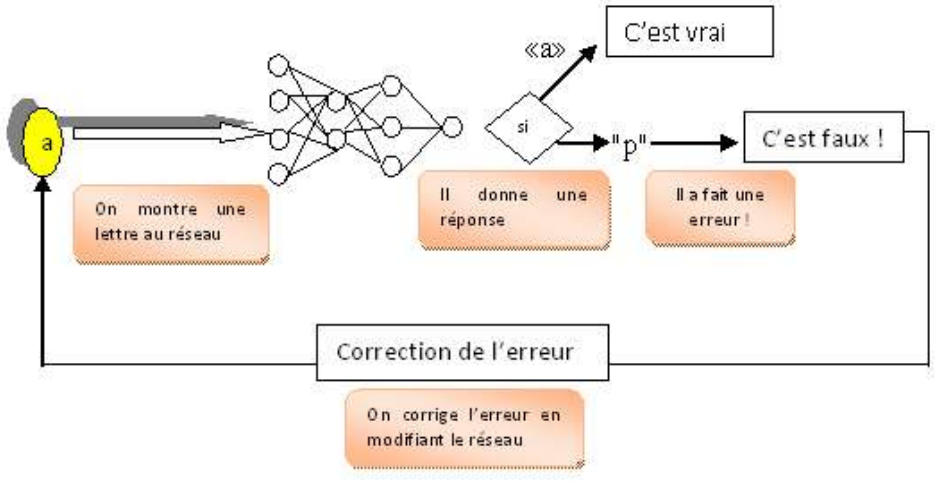


Figure III.14 Représentation d'apprentissage **[38]**

On recommence plusieurs centaines de fois l'opération, jusqu'à ce que le réseau ait bien appris sa leçon ! Le principe est donc relativement simple. On fait l'éduction du réseau comme on éduquerait un enfant : il faut lui répéter plusieurs fois qu'il a fait une erreur, et à la fin, il a appris **[68].**

III-3-9-2 Surapprentissage

Il arrive souvent que les exemples de la base d'apprentissage comportent des valeurs approximatives ou bruitées. Si on oblige le réseau à répondre de façon quasi parfaite relativement à ces exemples, on peut obtenir un réseau qui est biaisé par des valeurs erronées. Par exemple, imaginons qu'on présente au réseau des couples (xi, f (xi)) situés sur une droite d'équation y=ax+b, mais bruités de sorte que les points ne soient pas exactement sur la droite. S'il y a un bon apprentissage, le réseau répond ax+b pour toute valeur de x présentée. S'il y a surapprentissage, le réseau répond un peu plus queax+b ou un peu moins, car chaque couple (xi, f (xi) positionné en dehors de la droite va influencer la décision. Pour éviter le surapprentissage, il existe une méthode simple : il suffit de partager la base d'exemples en 2 sous-ensembles. Le premier sert à l'apprentissage et le 2eme sert à l'évaluation de l'apprentissage. Tant que l'erreur obtenue sur le 2eme ensemble diminue, on peut continuer l'apprentissage, sinon on arrête **[68].**

III-3-9-3 Rétro propagation

La rétro propagation consiste à rétro propager l'erreur commise par un neurone à ses synapses et aux neurones qui y sont reliés. Pour les réseaux de neurones, on utilise habituellement la rétro propagation du gradient de l'erreur, qui consiste à corriger les erreurs selon l'importance des éléments qui ont justement participé à la réalisation de ces erreurs : les poids synaptiques qui contribuent à engendrer une erreur importante se verront modifiés de manière plus significative que les poids qui ont engendré une erreur marginale **[68].**

III-3-9-4 Élagage

L'élagage ("pruning", en anglais) est une méthode qui permet d'éviter le surapprentissage tout en limitant la complexité du modèle. Elle consiste à supprimer des connexions (ou synapses), des entrées ou des neurones du réseau une fois l'apprentissage terminé.

III-3-9-5 Précision

En pratique, les éléments qui ont la plus petite influence sur l'erreur de sortie du réseau sont supprimés. Les deux algorithmes d'élagage les plus utilisés sont :

- Optimal Brain Damage (OBD) d'Y. LeCun et al.
- Optimal Brain Surgeon (OBS) de B. Hassibi et D.G. Stor **[38].**

III-4 CONCLUSION

L'intelligence artificielle est une discipline qui ne cesse d'évoluer depuis le milieu du vingtième siècle jusqu'à nos jours. De nombreux changements sont apparus pour perfectionner cette technologie à l'avenir prometteur dans le but d'atteindre le rêve ultime, de voir apparaître un jour une intelligence artificielle parfaite.

Il est difficile de déterminer si l'art inspire l'évolution scientifique .Cependant, on peut remarquer que les méthodes de développement de l'intelligence artificielle telles que les réseaux de neurones ou les algorithmes évolutionnistes sont une approche au fonctionnement de la logique humaine. Les scientifiques essayant d'approcher l'intelligence artificielle forte en simulant la logique humaine ont probablement les mêmes fantasmes que certains écrivains.

Un réseau de neurones est un outil d'analyse statistique permettant de construire un modèle de comportement a partir de données qui sont des exemples de ce comportement.Essentiellement utilisés pour faire de la classification, les réseaux de neurones peuvent être représentés par une boite noire à l'entrée de laquelle on présente un vecteur à « n » dimensions, représentant les données du problème, et à la sortie de laquelle on récupère un vecteur de dimension « m » qui représente la solution déterminée par le système .

Grace à leur grande capacité d'apprentissage automatique à partir de données modélisant le problème à résoudre, les réseaux de neurones permettent de remplacer efficacement des modèles mathématiques pouvant être extrêmement complexes. Dans ce chapitre nous avons donné une information sur l'intelligence artificielle et rappelé les propriétés principales des réseaux de neurones (les grands types, leurs applications…etc.) de façon générale.

Conclusion générale

Les résidus de pesticides constituent aujourd'hui un véritable problème de société. Depuis près de cinquante ans, ils ont été mis en évidence dans tous les compartiments de l'environnement, dans les eaux des rivières et des nappes phréatiques, dans l'air et dans les eaux de pluie. On les rencontre aussi dans les fruits, les légumes, les céréales et les produits d'origine animale (viande, poisson, lait et œuf). Ils existent sous leur forme initiale (substances actives) mais ils peuvent aussi existés sous une forme partiellement dégradée.

Ces pesticides regroupent un grand nombre de spécialités de toxicité variable pour l'hommeet l'environnement.

Cette étude traite la modélisation de la toxicité des pesticides par la méthode des réseaux de neurones artificiels qui fait partie des nouvelles techniques sous l'égide de l'intelligence computationnelle ou des sciences cognitives.

Nous avons présenté les étapes faites pour calculée les descripteurs QSAR et leur utilisation pour la prédiction de la toxicité qui contribue à la réduction de l'utilisation des animaux et en fournissant des données pour faire des classifications positive de la toxicité.

Le développement d'un modèle QSAR débute par la collecte de données expérimentales fiables. Il s'agit ensuite de développer une série de descripteurs qui caractérisent les structures moléculaires des composés de la base de données en vue de les relier à la propriété expérimentale étudiée (toxicité).

Parmi les nombreux descripteurs utilisés pour l'élaboration d'un modèle QSAR, nous avons choisi les descripteurs constitutionnels et les propriétés moléculaires pour initier cette étude.

Dans notre travail la modélisation de la toxicité repose sur l'utilisation d'un réseau de neurones de type perceptron multicouches (une couche d'entrée, une couche cachée, une couche de sortie)

Pour la recherche de cette topologie nous avons adopté la stratégie d'essai et erreur (trial and error), suggérée par plusieurs auteurs comme étant la stratégie la plus raisonnable. L'architecture du réseau de neurones optimal obtenu est de [25-13-1] : 25 neurones dans la couche d'entrée, 13 neurones dans la couche cachée, un neurone dans la couche de sortie. Un coefficient de régression s'approchant de l'unité a été obtenu (R=0.94556).

L'analyse de sensibilité a été réalisée comme une technique pour déterminer l'influence des descripteurs utilisés sur la toxicité en utilisant la méthode des poids de connexion.

Les descripteurs les plus pertinents sélectionnés par cette technique sont lenombre d'atomes de chloreet le nombre d'atomes de phosphore.

Références bibliographiques

[1] Chantal Gagné (1999) «L'utilisation des pesticides en milieu agricole» Mémoire présenté à la Commission sur l'avenir de l'agriculture et de l'agroalimentaire Québécois, Université du Québec à Rimouski.

[2] Mustapha Bounechada1, Mohamed Fenni, Abdellhamid Mekhlouf, Lamri tedjar (2011) «conséquences des pesticides dans les eaux et leur impact sur la santé des populations dans les hautes plaines setifiennes (nord-est de l'Algérie)» Université Ferhat Abbes, Sétif, Algérie.

[3] Direction générale de la santé (DGS). L'eau potable en France (2002-2004) « Guide technique Eau et santé » Juil. 2005, 52p. Disponible sur <http://www.sante.gouv.fr/> (Octobre 2005).

[4] http://www.mddep.gouv.qc.ca/pesticides/apropos.htm (Mars 2012).

[5] Nicolino F, François V (2007) «Pesticides, révélations sur un scandale français le Cavalier Bleu».

[6] http://www.notre-planete.info/ecologie/alimentation/pesticides_0.php(Novombre 2011)

[7] http://www.sénat.fr. (Mars 2011).

[8] Johan Mörner, Robert Bos et Marjon Fredrix (Geneve, 2002) «Réduire et éliminer l'usage des pesticides organiques persistants» Document d'orientation sur des stratégies alternatives de gestion durable des déprédateurs et des vecteurs, IOMC.

[9] Fergani Z, Mami N (Juin 2011) «Prédiction neuronale de la toxicité des pesticides». Thèse de Master. Université de Médéa. Algérie.

[10] OMS (2000-2002) «ipcs.the who recommended classification of pesticides by hazard and guide line to classification and guidelines to classification».

[11] Weinberg Jack (2009) « Un cadre d'action pour la protection de la santé humaine et de l'environnement contre les pesticides dangereux». Thèse de magistère.

[12] http://www.Dangers_ risques _pesticides _Final (Décembre 2011).

[13] Http://Www.Mddep.Gouv.Qc.Ca/Pesticides/Apropos.Htm (Janvier 2012).

[14] http:// www.observatoire-pesticides.gouv.frindex.phppageid=379 (Février 2012).

[15] Weinberg Jack (avril 2009) « un guide pour les ONG sur les pesticides dangereux et la SAICM » thèse de doctorat, International Society of Doctors for the Environment.

[16] http:// www.santemaghreb.com/algerie/index.asp (Février 2012).

[17] http://www.era-orleans.org/AFIP/FR/pesticides.html (Décembre 2011).

[18] http://www.ilo.org/public/english/employment/strat/kilm/download/kilm04 (Décembre 2011).

[19]http://siteresources.worldbank.org/INTWDR2008/Resources/WDR_00_book(Novembre2011).

[20] J. Jeyaratnam (1990) « Acute Pesticide Poisoning: A Major Global Health Problem» Associate Professor, Occupational and Family Medicine, National University of Singapore.

[21] http://fr.wiktionary.org/wiki/ibid (Juin 2012).

[22] http://www.who.int/ipcs / publications/pesticides_hazard_rev_3 (Février2012).

[23] http://en.wikipedia.org/wiki/Ibid. (May 2012).

[24] Watts MA (2009) Pesticides: Sowing Poison, Growing Hunger, Reaping Sorrow.

[25] http://www.unep.org/ceh/main01.html (Décembre 2011).

[26] Http://Www.Linkedin.Com/In/Stephanielynnwatts (Juin 2O12).

[27] http://www.inchem.org/documents/pds/pds/pest53_e.htm (Juin 2012).

[28] Bertolote J. M. (2006) « Deaths from pesticide poisoning, British Journal of Psychiatry (2006) ».

[29] Gunnell and Eddelston (2003) « Suicide by intentional ingestion of pesticides: a continuing tragedy in developing countries, International Journal of Epidemiology ».

[30] http://www.who.int/nmh/donorinfo/msd_ preventing_suicide. (Janvier 2010).

[31] Marlène Piché (1999) « La dérive des pesticides : prudence et solutions»Thèse de l'ingéniera, Université de Canada.

[32] Mamy L, Barriuso E, Gabrielle B (2008) «Evaluer les risques environnementaux des pesticides, Exemple du désherbage des cultures résistantes ou non au glyphosate».

[33] Projet POP's – Algérie GF/ALG/02/001(2006) Plan National de Mise en oeuvre (PNM) ALGERIE – Convention de Stockholm.

[34] Http://Images.Rechercher.Me/Fichiers/Anonyme-2001_Pdf_590115.Html (Février 2012)

[35] http://www.mce-info.org/Pdf/synthese_pesticides.pdf(Mars 2012)

[36] Document de l'Algérie phytosanitaire.

[37] http://fr.wikipedia.org/wiki/Relation_quantitative_structure_%C3%A0_activit%C3%A9 (Janvier 2012).

[38] Ami Y, Melzi N (Juin 2011) « Prédiction neuronale de l'effet toxicologique des fongicides». Thèse de master, université de Médéa, Algérie.

[39] Rouvray,D. H , Bonchev, Danail (1991) « Chemical graph theory: introduction and fundamentals » Tunbridge Wells, Kent, England: Abacus Press. ISBN 0-85626-454-7.

[40] Fraczkiewicz R (2007) «In Silico Prediction of Ionization» Comprehensive medicinal chemistryII. Amsterdam: Elsevier. ISBN 0-08-044518-7.

[41] Freyhult EK, Andersson K, Gustafsson MG (April 2003) «Structural modelin extends QSAR analysis of antibody-lysozyme interactions to 3D-QSAR » Biophysical Journal 84 (4): 2264–72. Doi: 10.1016/S0006-3495(03)75032-2. PMC1302793. PMID12668435.

[42] Hendrik T, Todeschini R, Viviana C, Raimund M, Hugo K (2002) « Handbook of Molecular Descriptors » Weinheim: Wiley-VCH. ISBN 3-527-29913-0.

[43] Strok, David G, Duda, Richard O,Hart, Peter W (2001). Pattern classification. Chichester: John Wiley & Sons. ISBN 0-471-05669-3.

[44] Corinne H. (2008) « Recherche D'inhibiteurs De Haute Affinite De L'adn Polymerase Beta Par Criblage Virtuel Et Rmn» Thèse de doctorat. Université De Toulouse (France).

[45] Nendza M (1998) «Structure-Activity Relationships in Environmental Sciences» Chapman & Hall: London.

[46] Cronin M.T.D,Dearden J.C(1995) « Quant. Struct.-Act. Relat».

[47] Bradbury S P,Henry T R, Carlson RW, Karcher W, Devillers J(1990) « In Practical Applications of Quantitative Structure-Activity Relationships (QSAR) in Environment Chemistry and Toxicology » Kluwer Academic Publishers: Dordrecht, Netherlands.

[48] Nendza Monica (1992) « report of the oecd workshop on quantitative structure activity relationships (QSAR) in aquatic effects assement» Paris.

[49] Veith, G.D, Broderius S.J (1987) «In QSAR in Environmental Toxicology II; Kaiser, K.L.E, ed, Reidel D: Dordrecht, Netherlands».

[50] CroninM.T.D, Schultz T.W (2001) « Chem. Res. Toxicol ».

[51] Mekenyan O.G, Veith G.D (1993)« SAR QSAR Environ».

[52] Hermens J, de Bruijn J, Pauly J, Seinen W(1987) « In QSAR in Environmental Toxicology».

[53] Dearden, J.C, Barratt, M.D, Benigni R, Bristol D.W,Combes R.D, Cronin M.T.D,Judson P.N, Payne M.P,Richard A.M, Tichý M,Worth A.P, Yourick J.J (1997) ATLA" Association Of Teachers Of Libreading To Adults".

[54] Greene Na(2002) «Drug Delivery».

[55] Dearden J.C, Schüürmann G (2001) « Environ. Toxicol. Chem » June.

[56] Weininger D (1988) « SMILES, a Chemical Language and Information System. 1. Introduction to Methodology and Encoding Rules» Journal Chemistry Information Computational Science.

[57] http://bkchem.zirael.org/index.html(Avril2012).

[58] Annuaire des Générateurs de systèmes experts et les Machines dédiées. Extrait du Tout sur l'intelligence artificielle(IA) et les systèmes experts(Mai 2009).

[59] http://ia-2011tpe.e-monsite.com/pages/l-intelligence/histoire-de-l-ia.html.(Maia2012).

[60] http://bmsap.revues.org/4463(Fèvrier 2012).

[61] Nils J.NILSSON«Systèmes d'intelligence artificielle»cours.

[62] Jean-Daniel Zucker (2001) « Introduction l'intelligence artificielle» Support de Cours IA .

[63] http://fr.wikipedia.org/wiki/Intelligence_artificielle#Histoire(Janvier 2012).

[64] http://www.turingarchive.org/viewer/?id=458&title=1(Fèvrier 2012).

[65] http://www.turingarchive.org/browse.php/B/6(Mai 2012).

[66] Crevier Daniel (1993) « AI: The Tumultuous Search for Artificial Intelligence » New York, NY:BasicBooks,ISBN 978-0-465-02997-6.

[67] «Réseaux de neurones et intelligence artificielle»Cours1 GPA-759.

[68] Marmouz H (2009-2010) « la modélisation neuronale de l'activité antibactérienne des molécules en utilisant les descripteurs QSAR» thèse d'ingéniera, université de Médéa, Algérie.

[69] www.ryounes.net/cours/chapitre%203%20RN.pdf(Juin 2012).

[70] wwwobs.univ-bpclermont.fr/.../cours_RN_2006.p(Juin 2012).

[71] Vincent Lemaire (2012)«Les réseaux de neurones».

[72] Rezazi S (2010-2011) « Modélisation de la prédiction de l'effet thérapeutique a partir de la structure moléculaire» thèseade magistère,université de Médéa,Algérie.

[73] Adrien Vergé (2009)«Réseaux de neurones artificiels TIPE»Lycée Michelet.

[74] http://fr.wikipedia.org/wiki/Neurone_formel(Avril 2012).

[75] https://sites.google.com/site/tpeiag5lnbahmv/.(Mai 2012).

[76] perso.pggj.net/...**Réseaux**%20de%20**neurones**/rapprto%20rn.doc (avril 2012)

[77] Feed Forward (2008) « Introduction aux Réseaux de Neurones Artificiels». Par Alp Mestan (Site perso de Alp)(Blog).

[78] Deptinfo.unice.fr/.../SystemesArtificielsComplexes(Mai 2012).

[79] Routt J, Reigart, James R Roberts (1999) «Recognition and management of pesticide poisonings»University of South Carolina,USA.

[80] Logiciel de chemBioDraw ultrat11.0.

[81] http://vsearch.nlm.nih.gov/vivisimo/cgibin/querymeta?v%3Asources=sisbundle&v%3Aproject=sis&querya(Marsa2012).

[82] Karelson M (2000)« Molecular Descriptors in QSAR/QSP»Wiley,New York.

[83] Todeschini R, Consonni V (2000) « Handbook of Molecular Descriptors »Wiley, Weinheim.

[84] Karelson M, Lobanov VS, Katritzky AR (1996) «Chem» Rev, 96, 1027-1044.

[85] Olden J.D,Joy M.K and Death R.G(2004) « An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data» Ecologie.