



MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR  
ET DE LA RECHERCHE SCIENTIFIQUE  
UNIVERSITÉ «Abbés LAGHROUR» DE KHENCHELA  
FACULTÉ DES SCIENCES ET TECHNOLOGIE

Département des Sciences de la Matière

N° de série :....

## Mémoire de fin d'études

*Pour l'obtention du diplôme de Master (L.M.D)*

**Filière : Chimie.**

**Spécialité : Chimie Analytique et Environnement**

**Intitulé :**

**Analyse QSAR de la toxicité des  
pesticides aux poissons (LC50)**

*Réalisé par :*

-ACHOUR Soumia

-SELLAOUI Nadjat

*Dirigé par : Dr. BOUAKKADIA A*

*Membres de jury :*

<i>GERRAB Fahima</i>	<i>MAA</i>	<i>U.A.L. Khenchela</i>	<i>Président</i>
<i>BOUAKKADIA Amel</i>	<i>MCB</i>	<i>U.A.L. Khenchela</i>	<i>Encadreur</i>
<i>KERTIOU NourEddine</i>	<i>MCB</i>	<i>U.A.L. Khenchela</i>	<i>Examineur</i>

*Promotion : 2019/2020*

# *Remerciements*

Nous tenons tout d'abord à remercier **Allah** le tout puissant et philanthrope, qui nous a donné la force et la patience d'accomplir ce travail.

Nous tenons à remercier notre encadreur **Dr. BOUAKKADIA Amel**, pour ses précieux conseils et ses directives. Son aide était essentiel afin de réaliser ce travail

Nous tenons également à exprimer nos sincères remerciements aux membres de jury, **Dr.GERRAB Fahima** et **Dr.KERTIOU Nour Eddine** pour l'intérêt qu'ils ont porté à ce travail en acceptant de le rapporter.

Enfin, on remercie tous ceux qui, de près ou de loin, ont contribué à la réalisation de ce travail.

# *Dédicaces*

## *Je dédie ce travail à...*

*A la lumière de mes jours, la source de mes efforts, la flamme de mon cœur,  
ma vie et mon bonheur, qui n'a pas cessé de m'encourager de prier pour moi ;*

*Maman que j'adore.*

*A mon père, mon exemple éternel, mon soutien moral et matériel,  
tu représentes pour moi le symbole de la bonté par excellence, ce travail est  
le fruit de tes sacrifices que tu as consentis pour mon éducation et ma  
formation.*

*A mes sœurs : « Nasira, Salima, Ahlem ».*

*A mes frères : « Rafik, Bilal, Yahia, Azzou ».*

*Je vous dédie ce travail avec tous vœux de bonheur, de santé et de réussite.*

*A toute la famille : « Achour ».*

*A toute mes amis et collègues de promotion.*

*A mon amie, et mon binôme : « Nadjat ».*

*A tous ceux ou celles qui me sont chers et qui j'ai omis involontairement de  
citer. A tous qui aiment Soumia et tous que Soumia aime.*

*Soumia*

# *Dédicaces*

## *Je dédie ce travail à...*

*A la lumière de mes jours, la source de mes efforts, la flamme de mon cœur,  
ma vie et mon bonheur, qui n'a pas cessé de m'encourager de prier pour moi ;*

*Maman que j'adore.*

*A mon père, mon exemple éternel, mon soutien moral et matériel,  
tu représentes pour moi le symbole de la bonté par excellence, ce travail est  
le fruit de tes sacrifices que tu as consentis pour mon éducation et ma  
formation.*

*A mes sœurs à mes frères.*

*Je vous dédie ce travail avec tous vœux de bonheur, de santé et de réussite.*

*A toute la famille : « SELLAOUI ».*

*A toute mes amis et collègues de promotion.*

*A mon amie, et mon binôme : « Soumia ».*

*A tous ceux ou celles qui me sont chers et qui j'ai omis involontairement de  
citer. A tous qui aiment **Nadjat** et tous que **Nadjat** aime.*

*Nadjat*



## Sommaire

Introduction générale : ..... 2

### **PARTIE I: ETUDE BIBLIOGRAPHIQUE**

#### **Chapitre 01: LES PESTICIDES**

I - Généralité Sur Les Pesticides .....	6
I-1 Définition.....	6
I-2 Historique des pesticides .....	6
I-3Composition des pesticides (conception) .....	7
I-4 Classification des pesticides .....	8
I-4-1 Classification chimique : .....	8
I-4-2 Classification biologique :.....	8
I-4-3 Classification selon l'usage :.....	12
I-5 Effets des pesticides sur l'environnement .....	12
I-5-1Effet des pesticides sur les milieux.....	13
I-5-2Effet des pesticides sur les espèces.....	14
I-5-3Effets des pesticides sur les produits agricoles.....	15

#### **Chapitre 02: Modélisation QSAR**

II- Approche QSAR .....	18
II-1Définition : .....	18
II-2 Principe de QSAR.....	18
II-3 Paramètres biologiques.....	19
II.4.Les descripteurs QSAR.....	19
II-4-1Les descripteurs physico-chimiques.....	20
II-4-2Les descripteurs électroniques .....	20
II-4-3Les descripteurs topologiques .....	20
II-5Les applications de l'étude QSAR .....	20
II.6. Introduction à la régression linéaire multiple .....	21

### **PARTIE II: application**

I- Calcul du modèle : .....	32
I- 1- Choix du modèle : .....	32
I- 2- Choix d'une famille de descripteurs : .....	33
II- Développement du modèle : .....	35
II- 1- Matrice de corrélation : .....	35
II- 2- Equation et analyse de régression : .....	36
II- 3- Diagramme de Williams : .....	38

II- 4- Qualité de l'ajustement : .....	38
II- 5- Test de randomisation : .....	39
Conclusion générale .....	41
Références bibliographiques .....	43

## Liste des tableaux

Tableau I. Historique de l'évolution des trois plus grandes classes des Pesticides de 1900 à 2000 [1]. .....	11
Tableau II: les paramètres biologique. ....	19
Tableau III: Composés étudiés et leurs toxicités.....	23
Tableau IV : Descripteurs moléculaires intervenant dans la modélisation de pLC50 .....	33
Tableau V: Valeurs de pLC50 expérimentales, prédites, $h_{ij}$ , $e_{istd}$ .....	33
Tableau VI: Matrice de corrélation .....	35
Tableau VII: Analyse de régression .....	36
Tableau VIII: Les paramètres statistiques pour l'ensemble de calibration (43 composés) et celui de validation (18 composés).....	37



## Liste des figures

Figure 1: La composition d'un pesticide.....	7
Figure 2: Comportement des pesticides dans le sol .....	13
Figure 3: L'accumulation des pesticides.....	16
Figure 4: Modes d'exposition de l'homme et des milieux par les pesticides .....	16
Figure 5: Présentation de la méthodologie de QSAR. ....	18
Figure 6: Diagramme de Williams .....	38
Figure 7 : Qualité de l'ajustement .....	38
Figure 8: Test de randomisation.....	39

## Symboles et abréviation

AM1	Austin Model 1.
EQMC	Erreur quadratique moyenne sur l'ensemble de calibration.
EQMP	Erreur quadratique moyenne sur l'ensemble de prédiction.
EQMP <sub>ext</sub> :	Erreur quadratique moyenne sur l'ensemble de prédiction externe.
$e_i$	Résidu : différence entre les valeurs observées et estimées.
$e_{i \text{ std}}$ :	Résidu de prédiction standardisé.
F	Statistique de Fisher.
GA :	Algorithme génétique
H	Matrice de projection, ou matrice chapeau.
$h_{ii}$	Eléments diagonaux de la matrice chapeau.
LMO:	Validation croisée par omission d'un ensemble d'observations.
MP	Point de fusion (melting point)
N	Dimension de la population.
n-p	Nombre de degrés de liberté de la somme des carrés des résidus.
OCDE	Organisation de Coopération et de Développement Economiques
P	Nombre de descripteurs en comptant la constante (nombre de paramètres).
PRESS	Somme des carrés des erreurs de prédiction.
$Q^2$	Coefficient de prediction.
QSAR	Quantitative Structure/ Properties - Relationships.
$R^2$	Coefficient de détermination.
RLM :	Régression linéaire multiple.
S	Erreur standard.
SCE	Somme des carrés des écarts.

SCT	Somme des carrés totaux.
VIF :	Facteur d'inflation de la variance.

## ملخص:

سيكون التقدير التجريبي للخصائص دائما مطلوباً وسيظل ضرورياً لوصف المنتجات العضوية. في عملنا قمنا بنمذجة قيم السمية المتوقعة pLC50 لسلسلة من 61 مبيد للأعشاب. تم تقسيم سلسلة البيانات بشكل عشوائي إلى مجموعتين فرعيتين، مجموعة مكونة من 43 مركباً للبناء النموذجي ومجموعة من 18 مركباً للتحقق من الصحة. تم اختيار الوصفات الجزيئية الهامة بواسطة الخوارزمية الجينية باستخدام نهج QSAR. توضح القيم الإحصائية التي تم الحصول عليها لمجموعات المعايرة والتحقق (معامل التحديد والتنبؤ: R2 و Q2 والانحراف المعياري) جودة النماذج المحسوبة وأهميتها. تمت مناقشة النطاق باستخدام مخطط Williams.

**Résumé :**

L'estimation empirique des propriétés s'imposera toujours et demeurera incontournable pour caractériser les produits organiques.

Dans notre travail nous avons modélisés pLC50 d'une série de 61 herbicides. La série des données a été divisé aléatoirement en deux sous-ensembles, un ensemble de 43 composés pour la construction du modèle et un ensemble de 18 composés pour la validation.

La sélection des descripteurs moléculaires significatifs a été faite par algorithme génétique, en utilisant une approche QSAR

Les paramètres statistiques obtenus pour les ensembles de calibration et de validation (Coefficient de détermination et de prédiction :  $R^2$  ;  $Q^2$  et écart type s) ont ressortir la qualité et la pertinence des modèles calculés.

Le domaine d'application a été discuté à l'aide de diagramme de Williams.

**Mots clés :** Modélisation, Régression linéaire multiple, modèles QSAR, Pesticides, Pollution des eaux.

**Abstract:**

The empirical estimation of properties will always be essential and will remain essential to characterize organic products.

In our work, we modeled pLC50 from a series of 61 herbicides. The data series was randomly divided into two subsets, a set of 43 compounds for the construction of the model and a set of 18 compounds for validation.

The selection of significant molecular descriptors was made by genetic algorithm, using a QSAR approach

The statistical parameters obtained for the calibration and validation sets (Coefficient of determination and prediction:  $R^2$ ;  $Q^2$  and standard deviation  $s$ ) highlight the quality and relevance of the calculated models.

The applicability domain was discussed using a Williams's diagram.

**Key words :** Modélisation, Multiple linéaire régression, QSAR models, Pesticides, Pollution of water.

# *Introduction générale*

### Introduction générale :

Les facteurs qui déterminent la santé humaine et le développement des pathologies sont multiples, la qualité des milieux (air, eau, sol...), les contaminants (biologiques, chimiques, physiques) jouent un rôle fondamental. En effet, il apparaît que certaines pathologies sont aggravées, voire déterminées, par l'environnement et l'approfondissement des connaissances sur le rôle de l'environnement sur la santé constitue un enjeu scientifique majeur.

Depuis le début du XX<sup>ème</sup> siècle, l'environnement (atmosphérique, terrestre et aquatique) a été soumis à la pression croissante des activités industrielles et humaines. Les contaminants rejetés dans l'environnement finissent par se retrouver plus ou moins rapidement dans les milieux aquatiques, en particulier estuariens et côtiers où ils peuvent avoir des effets à court et à long terme [1].

La propagation de l'industrialisation, la naissance de nouvelles technologies, l'accroissement de la population, le développement de l'agriculture et l'obligation à améliorer ses productions agricoles dans le but de résoudre les problèmes de nutrition, sont tous liés à la consommation de quantités énormes de pesticides.

Les pesticides sont parmi les polluants les plus dangereux de l'environnement en raison de leurs stabilités, leurs mobilités, et les effets à long terme sur les organismes vivants [2].

Les pesticides permettent la protection et l'amélioration des produits agricoles que ce soit des fruits ou des légumes, mais d'après nos connaissances les pesticides sont des produits chimiques, et n'importe quelle matière qui contient une composition chimique est dangereuse pour la santé humaine [3].

Le travail que nous avons abordé se situe dans le cadre de protection de la santé humaine. Pour cela nous nous sommes proposé de faire une étude de La Relation Quantitative Structure et Activité (QSAR) et par conséquent l'analyse de la toxicité des pesticides aux poissons (LC50).

Ce travail est divisé en deux parties ; la première sacrifiée pour les informations bibliographiques basées sur les pesticides et le QSAR, qui est divisé en leur tour en deux chapitres, le premier concerne la généralité des pesticides et le deuxième concerne l'approche QSAR



## **Introduction générale**

La deuxième partie est sacrifiée pour les différentes méthodes statistiques utilisées dans notre travail et en termine par la présentation des résultats obtenus avec les différentes interprétations, et en fin une conclusion générale.

# *Etude bibliographique*

*Chapitre I*

*Les pesticides*

## I - Généralité Sur Les Pesticides

### I-1 Définition

L'étymologie du mot pesticide s'est construite à partir du suffixe «-cide-» qui signifie « tuer » et de la racine anglaise «Pest» (animal, insecte ou plante nuisible) provenant du latin «Pestis » qui désignait le fléau en général ; Étymologiquement « tueurs de fléaux » [4].

Le terme "pesticide" est une appellation générique couvrant toutes les substances (molécules) ou produits (formulations) qui éliminent les organismes nuisibles, qu'ils soient utilisés dans le secteur agricole ou dans d'autres applications. La substance ou le microorganisme qui détruit ou empêche les organismes nuisibles de s'installer sur les végétaux, parties de végétaux ou produits végétaux est dénommée matière active, à laquelle sont associés dans la préparation, un certain nombre de « formulant » (mouillants, solvants, anti-mousses, ...) qui la rendent utilisable par l'agriculteur (ACTA, 2000 ; FAO, 2015).

Les pesticides ont des risques sur la santé humaine par l'accumulation de ces derniers dans la chaîne alimentaire, et donc ils vont être consommé par l'être humain, d'une autre part ils ont un impact sur la pollution des eaux, le sol, la vie de la faune et la flore et aussi la santé des agriculteurs (Merghid, 2017) [5].

### I-2 Historique des pesticides

Selon Calvet (2005), la lutte contre les organismes nuisibles aux cultures a certainement été de tous temps une préoccupation de l'agriculteur. Pendant longtemps, l'essentiel des moyens étaient de nature physique : ramassage des larves, des œufs, des insectes adultes, destruction des plantes malades par le feu, désherbage manuel puis mécanique. L'utilisation des produits chimiques est malgré tout assez ancienne comme l'indique l'emploi du soufre et celle de l'arsenic. L'efficacité de ces moyens était souvent limitée et générait parfois un profond sentiment d'impuissance qui explique la place qui était occupée par les superstitions et autre diverses croyances comme le montre la pratique des rogations par exemple.

L'arsenic a été utilisé comme insecticide depuis la fin du XVII<sup>e</sup> Siècle ainsi que la nicotine dont les propriétés toxiques ont été découvertes par Jean de la Quintinie (1626-1688) qui en a recommandé l'usage. Cependant, c'est surtout au cours des XIX<sup>e</sup> et XX<sup>e</sup> Siècles que les propriétés biocides de nombreux produits chimiques ont été mises en évidence et ont donné lieu à de considérables développements des techniques de protection des plantes. Plusieurs

facteurs ont contribué à ce développement : l'apparition de graves épidémies (ex : phylloxéra, mildiou de la pomme de terre, doryphore), la nécessité de nourrir une population humaine croissante, les progrès considérables de la chimie organique de synthèse, et les innovations techniques [6].

### I-3 Composition des pesticides (conception)

Un pesticide est constitué de nombreuses molécules comprenant (figure 01) :

- **Une ou plusieurs matières actives (Ma)** : ce sont des éléments principales permettant l'efficacité du pesticide qui confèrent au produit l'effet poison désiré. La Ma peut également être reconnue grâce à un numéro de produit chimique, ainsi que grâce à un nom chimique.
- **Un diluant** : qui est une matière solide ou un liquide incorporé à une préparation et destiné principalement à diminuer la concentration de la matière active. Dans le cas d'une préparation liquide, il s'agira d'un solvant, d'argile ou de talc pour les préparations solides. Dans ce dernier cas le diluant est dénommé charge.
- **Un ou plusieurs adjuvants (adjuvants)** : ce sont des substances en théorie dépourvues d'activité biologique, mais qui sont susceptibles de modifier les propriétés du pesticide et d'en faciliter l'utilisation, l'application et le transport du produit permettant, par exemple, une meilleure pénétration dans le végétal. Ces adjuvants comprennent des stabilisants, des adhésifs, des colorants, des matières répulsives, des tensio-actifs, des émulsionnants et parfois des antidotes [7].

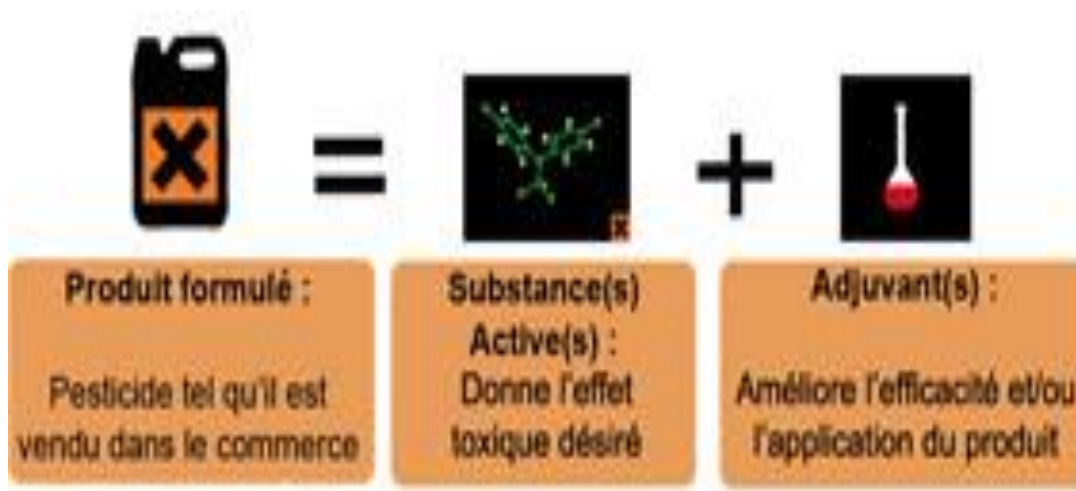


Figure 1: La composition d'un pesticide

## I-4 Classification des pesticides

Un pesticide est une substance chimique ; souvent utilisés en agriculture pour protéger les cultures et lutter contre les organismes considérés comme nuisibles, tel que les insectes (insecticides), les maladies fongiques (fongicides) et les mauvaises herbes (herbicides).

Il existe plusieurs types de pesticides, ces derniers sont classés selon le risque sur la santé et l'environnement

Un pesticide utilisé sous une forme différente de sa forme de mise en marché ; continue d'appartenir à la classe à laquelle il est rattaché, Cette classification permet d'ajuster les exigences réglementaires au niveau du risque, La plus simple et la plus globale d'entre elles consiste à les distinguer en fonction de leur cible en grandes catégories tel que Les fongicides, Les insecticides et Les herbicides [8].

### I-4-1 Classification chimique :

Il existe trois catégories de pesticides

- **Les pesticides inorganiques :** ils sont peu nombreux mais certains sont utilisés en grande quantités comme le soufre ou le cuivre. Ce sont pesticides très anciens dont l'emploi est apparu avant la chimie organique de synthèse. De cette époque ne subsiste qu'un seul herbicide employé en tant que désherbant total (chlorate de sodium) et quelques fongicides à base de soufre et cuivre.
- **Les pesticides organo-métalliques :** Ce sont des fongicides dont la molécule est constitué par un complexe fait d'un métal comme le zinc ou le manganèse et d'un anion organique dithioarbamate (exemple : mancozèbe avec le zinc, manèbe avec le manganèse).
- **Les pesticides organiques :** ils sont très nombreux et appartiennent à diverses familles chimiques (une famille chimique ou classe chimique correspond à l'ensemble des molécules dérivées d'un groupe d'atomes qui constituent une structure de base). Il existe actuellement plus de 80 familles ou classes chimiques dont les plus connues sont les Organophosphorés, Uraciles, Carbamates, Diphényl éther ...., mais leurs appellations sont parfois arbitraires et peuvent donc être différentes d'un ouvrage à un autre [9].

### I-4-2 Classification biologique :

Selon les organismes visés, trois grandes classes se distinguent :

- **Les insecticides** : Destinés à la lutte contre les insectes nuisibles, ils interviennent en les tuant ou en empêchant leur reproduction, ce sont souvent les plus toxiques.

Les trois plus grandes familles auxquelles appartiennent les insecticides organiques de synthèse sont :

- a- **Les organochlorés (OC)** : Ils contiennent du carbone, de l'hydrogène et des atomes de chlore. Il s'agit d'une famille où l'on trouve le DDT, la plupart des OC sont interdits vu leur persistance et leurs risques d'accumulation dans les sols, les tissus végétaux et les graisses animales.
- b- **Les organophosphorés (OP)** : Ce sont des esters de l'acide phosphorique. Les OP sont des pesticides utilisés en milieu agricole comme insecticides. Ils appartiennent à la famille chimique des anticholinestérasiques. Les organophosphorés sont très toxiques pour les vertébrés et la plupart des substances actives sont chimiquement instables.
- c- **Les carbamates** : Ce sont des dérivés de l'acide carbamique, qui agissent en inhibant l'activité enzymatique de l'acétylcholinestérase, inhibition qui peut être réversible dans certains cas. Le carbaryl est le carbamate le plus utilisé en raison de son spectre d'action très étendu pour les contrôles des insectes et en raison de sa faible toxicité chez les mammifères. Le carbofuran, l'aldicarbe, le carbosulfan ou encore le fénoxycarbe sont également des carbamates largement utilisés.
- d- **Les pyréthrinoïdes de synthèse** : Ils agissent sur le système nerveux central et périphérique des insectes en provoquant une excitation nerveuse répétée à travers des pompes à sodium. On les considère donc comme des poisons axoniques. Ce sont des insecticides particulièrement efficaces, dont l'utilisation est très répandue, tant dans le domaine agricole qu'en milieu domestique.
- **Les herbicides** : Ce sont les plus utilisés dans le monde en tonnage et en surface, Destinés à lutter contre certains végétaux (les mauvaises herbes ou les adventices), qui entrent en concurrence avec les plantes cultivées à protéger en ralentissant leur croissance.

Les herbicides seront traités en fonction de leur mode d'application et de leur action :

- a- **Les herbicides appliqués au niveau foliaire** : Existents sous différentes formes :
  - **Les régulateurs de croissance** : Les substances actives de ces composés affectent la croissance des plantes en agissant sur la synthèse des protéines et la division cellulaire. En fait ces herbicides vont entraîner une croissance

rapide des plantes pour arriver à leur sénescence. Les substances actives les plus connues et utilisées sont le 2,4-D, le dichloprope et le 2,4, 5-T par exemple.

- **Les inhibiteurs de la synthèse d'acides aminés** : Parmi lesquelles on retrouve le glyphosate qui est un herbicide de contact employé en post-levée sur les plantes annuelles, bisannuelles ou vivaces, les graminées.
- **Les destructeurs de la membrane cellulaire** : dont les deux principales familles sont les bypyridilium et les diphényl
- **Les inhibiteurs de la photosynthèse** : Les herbicides de la famille des triazines et des phénylurées agissent en interférant avec la photosynthèse.

**b- Les herbicides appliqués au niveau du sol** : On distingue :

- **Les inhibiteurs de la division cellulaire** : Agissent en inhibant les étapes de division cellulaire responsables de la séparation des chromosomes et de la formation de la paroi cellulaire au niveau des racines de la plante.
- **Les destructeurs de pigment** : agissent sur les plantes en détruisant la chlorophylle, ce qui empêche ensuite la plante de pouvoir réaliser correctement la photosynthèse.

Les herbicides sélectifs tuent certaines plantes avec peu ou pas de blessures à d'autres plantes. Habituellement, les types sélectifs tuent les plantes à feuilles larges ou les plantes herbeuses.

On distingue deux classes d'herbicides en fonction de leur mode d'action :

- **Les herbicides de contact** : Qui agissent au point d'impact du produit sur la plante.
- **Les herbicides systémiques** : Ils sont plus complexes ; assimilés par les feuilles ou les racines et transportés dans les différentes parties de la plante.

- **Les fongicides** :

Ils sont destinés à traiter les maladies fongiques des plantes, également les maladies bactériennes et virales 6, et à éliminer les moisissures et parasites (champignons) des plantes. Les fongicides les plus anciens sont le soufre, le cuivre et ses dérivés organiques, parmi ces derniers on trouve :

- **Les fongicides de contact** : empêchent les champignons de pénétrer dans la plante comme le **zinèbe** et le **captane**.
- **Les fongicides systémiques** : ils ont un rôle curatif comme le triadiméfon et la morpholine.



**Tableau I.** Historique de l'évolution des trois plus grandes classes des Pesticides de 1900 à 2000[1].

	<b>Herbicides</b>	<b>fongicides</b>	<b>insecticides</b>
Avant 1900	Sulfate de cuivre Sulfate de fer	Soufre Sels de cuivre	Nicotine
1900-1920	Acide sulfurique		Sels d'arsenic
1920-1940	Colorants nitrés		
1940-1950	Phytohormones...		Organochlorés Organophosphorés
1950-1960	Triazines, Urées substituées Carbamates	Dithiocarbamates phtalimides	Carbamates
1960-1970	Dipyridyles, Toluidines...	benzimidazoles	
1970-1980	Amino-phosphonates Propionates...	Triazoles Dicarboximides Amides, Phosphites morholines	Pyréthrinoides Benzoyl-urées (régulateurs de Croissance)
1980-1990	Sulfonyl urées...		
1990-2000		Phenylpyrroles Strobilurines	

- **Autres pesticides :** A côté de ces trois grandes catégories de pesticides, bien d'autres produits existent, pour lutter contre les limaces et les escargots (les molluscicides), contre les rongeurs (les rodenticides), contre les acariens (**acaricides**) et les nématodes (**nématocides**), les contre les oiseaux ravageurs tels que les corbeaux (les corvicides), pour désinfecter le sol (les fumigants), donc il existe des pesticides contre tout ce qui peut nuire à l'agriculture intensive **2** :
  - **Algicides :** Les algicides sont des substances qui permettent l'élimination des algues.

- **Nématicides** : Les nématicides sont des produits qui servent à éliminer les vers, notamment ceux qui s'attaquent aux racines, et Certains peuvent être très dangereux pour l'homme.
- **Rodenticides** : Ce sont des substances qui tuent les rongeurs (rats, souris, mulots...) après ingestion. Il peut s'agir par exemple d'anticoagulants, qui provoque des hémorragies internes chez l'animal qui l'a mangé.
- **Répulsifs** : Les répulsifs sont des produits destinés à éloigner les oiseaux, taupes et autres insectes nuisibles sans forcément chercher à les tuer [10].

#### I-4-3 Classification selon l'usage :

Les pesticides sont utilisés dans plusieurs domaines d'activités pour lutter contre des organismes vivants nuisibles. Il existe six catégories de pesticides selon leur destination de traitement, à savoir : les **cultures**, les **bâtiments d'élevage**, les **locaux de stockage des produits végétaux**, les **zones non agricoles**, les **bâtiments d'habitation**, **l'homme et les animaux** [11].

#### I-5 Effets des pesticides sur l'environnement

Les pesticides sont devenus un besoin pour les agriculteurs, car ils permettent l'intensification de l'agriculture afin de couvrir la demande exprimée dans les marchés sur les produits agricoles. En vérité les pesticides ont un avantage concernant la protection des cultures et l'augmentation de la production, mais malheureusement leurs risques ont dépassés l'utilité de ces derniers, parce qu'ils endommagent l'environnement.

Les pesticides ont contaminé presque toutes les parties de notre environnement, c'est vrai qu'ils tuent quelques espèces visés comme les insectes, champignons et les plantes indésirables.

Les substances et les molécules issues des pesticides sont susceptibles de se retrouver dans l'air, le sol, les eaux et les sédiments, ainsi que dans les aliments. Ces substances et molécules présentent, par leur migration entre les compartiments de l'environnement, des dangers importants pour l'homme et les écosystèmes, avec un impact à court ou à long terme.

### I-5-1 Effet des pesticides sur les milieux

Dans l'environnement, les pesticides se dégradent dans une variété d'autres substances ; à la suite d'interactions avec le sol, l'eau, la lumière du soleil et l'oxygène

- **Contamination du sol** : Le sol est un matériau à la fois minéral et organique. La partie minérale représente la fraction la plus importante :

- **Minéraux primaire** : Issus de l'altération du substrat géologique sous l'action conjuguée de la température, de l'air et de l'eau.

- **Minéraux secondaires** : produit d'altération comme les argiles, les oxydes et les hydroxydes.

Le sol joue un rôle fondamental dans le devenir des pesticides qui peuvent être appliqués. Un traitement important des sols avec des pesticides peut entraîner une baisse des populations de microorganismes du sol bénéfiques. Selon la scientifique du sol Elaine Ingham : " Si nous perdons les bactéries et les champignons, le sol se dégrade". L'utilisation excessive d'engrais chimiques et de pesticides a des effets sur les organismes du sol.

Le devenir des pesticides dans l'environnement c'est-à-dire, leur rétention, leur transformation et leur dégradation, dépend de leurs propriétés ainsi que celles des différents compartiments concernés, le sol, les eaux et l'atmosphère.

L'impact des pesticides sur le sol peut être expliqué par la figure suivante :

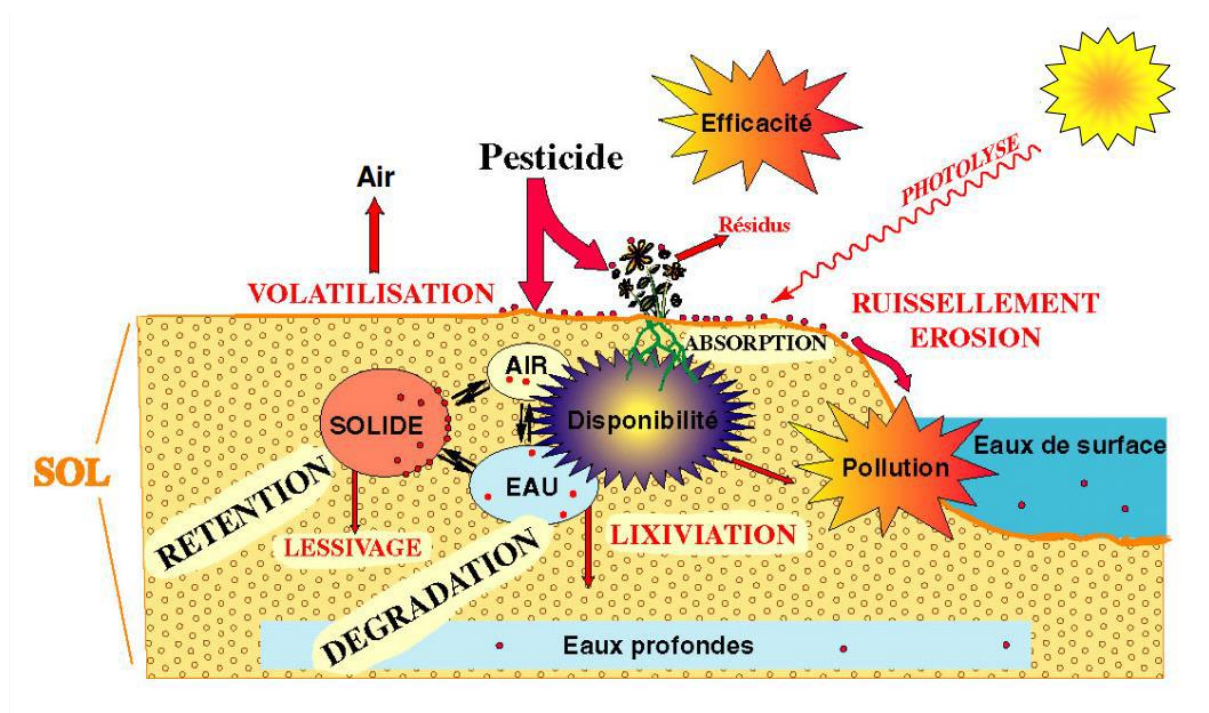


Figure 2: Comportement des pesticides dans le sol [12]

- **Contamination de l'air :** La dissémination des pesticides dans l'atmosphère se produit soit au moment de l'épandage, notamment lorsqu'ils sont pulvérisés, soit par évaporation, à partir des plantes sur lesquelles ils ont été répandu ou à partir du sol ou ils se sont déposé. Ainsi les procédés de pulvérisation et de traitement des produits jouent donc un rôle important dans la présence de résidus atmosphériques. Les pulvérisations aériennes sont les plus grandes sources de concentration de substances dans l'atmosphère. La présence des pesticides dans l'eau de pluie indique une contamination de l'atmosphère mais seuls les éléments solubles se retrouvent dans l'eau de pluie. Divers pesticides peuvent se trouver dans l'air à des concentrations parfois non négligeables et être transportés sur de grandes distances. La présence des pesticides dans l'air dépend des caractéristiques du produit, sa dégradabilité, type de surface, plantes ou animaux traités, des pratiques agricoles, du matériel de traitement, des conditions climatiques et pédologiques.
- **Contamination l'eau :** Les pesticides et leurs résidus se retrouvent dans les eaux de surfaces (cours d'eau et étendues d'eau) ainsi que dans les eaux souterraines et marines. La pollution des eaux souterraines due aux pesticides est un problème mondial. Une fois que les eaux souterraines sont polluées par des produits chimiques toxiques, la contamination peut prendre plusieurs années pour se dissiper ou être nettoyée. Le nettoyage peut également être très coûteux et complexe. Les eaux de surface destinées à la consommation ne contenaient que faibles concentrations des pesticides, rien ne semble indiquer que ces concentrations puissent présenter un danger significatif pour la santé. Les pesticides trouvés dans l'eau potable sont particulièrement préoccupants, car ils pourraient avoir des effets sur la santé et causé des maladies graves tel que le cancer et les maladies génétiques héréditaires. Un grand nombre d'insecticides et quelques herbicides et fongicides, peuvent avoir un effet toxique pour les organismes aquatiques, et peut avoir un effet nuisible sur le milieu naturel. La commercialisation du poisson contaminé, peut également représenter une menace pour la santé humaine.

### **I-5-2 Effet des pesticides sur les espèces**

De nombreux pesticides sont toxiques pour les insectes bénéfiques, les oiseaux, les mammifères, les amphibiens et les poissons. L'empoisonnement de la faune sauvage dépend de la toxicité du pesticide, la quantité appliquée, la fréquence, le moment et la méthode de

pulvérisation. Les animaux peuvent être intoxiqués par contact lors du traitement aérien, ou bien par ingestion. Les semences traitées par exemple, sont dangereuses pour la faune granivore et peuvent provoquer des mortalités importantes chez les populations d'oiseaux.

La faune est exposée aux pesticides en mangeant des aliments ou de l'eau contaminés, en respirant les vapeurs des pesticides ou par l'absorption de ces derniers à travers leur peau. Les prédateurs peuvent être empoisonnés en mangeant des animaux qui ont été exposés aux pesticides. Beaucoup d'insecticides affectent le système nerveux des animaux sauvages, ce qui peut interférer avec leur capacité à survivre ou à se reproduire.

Les pesticides peuvent passer à travers le placenta ou affecter les œufs des oiseaux ou des reptiles, ce qui cause des affaiblissements ou des défauts qui apparaissent plus tard dans la vie. Les herbicides, utilisés pour tuer les mauvaises herbes, affectent les plantes qui sont importantes pour la survie de la faune. Les poissons, les amphibiens et les insectes aquatiques sont spécialement susceptibles à la contamination de l'eau par les pesticides. Ceci affecte aussi les animaux comme les canetons qui dépendent de ces créatures pour survivre.

Les impacts précis des pesticides sur les populations de phytoplancton sont inconnus, mais parce qu'ils sont conçus pour tuer les plantes, les herbicides ont les impacts les plus importants, en outre, certains insecticides sont toxiques pour les plantes aquatiques. Les fongicides à base de cuivre sont hautement toxiques pour les organismes aquatiques, car le risque d'accumulation du cuivre dans les poissons et certains autres organismes aquatiques peut être élevé. Alors que Les insecticides chlorpyrifos et endosulfan ont la capacité de causer de sérieux préjudices aux amphibiens, à des concentrations présentes dans l'environnement inférieures aux conditions normales d'utilisation.

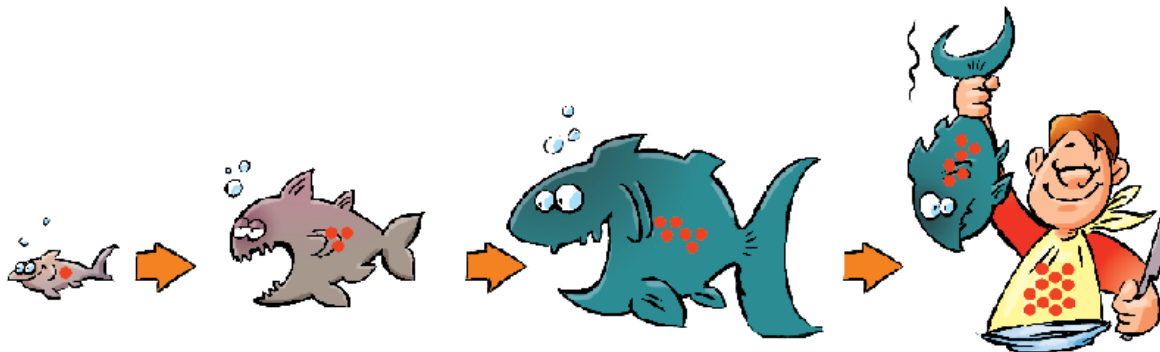
### **I-5-3 Effets des pesticides sur les produits agricoles**

L'association algérienne de la protection de l'environnement affirme que l'Algérie est un grand consommateur de pesticides, en effet 30 000 tonnes sont épandues chaque année. La moitié des fruits et des légumes - surtout les poivrons, piments, tomates, poireaux, laitues et épinards - vendus, contiendraient ces substances chimiques. Les fruits les plus touchés sont les fraises, les Mandarines et les raisins. Les résidus les plus fréquemment détectés sur les fruits et légumes sont les fongicides et les insecticides.

L'exposition de la population générale aux pesticides se produit principalement par la consommation d'aliments et l'eau potable contaminée par des résidus de pesticides, alors qu'une exposition importante peut également se produire dans ou à l'intérieur de la maison.

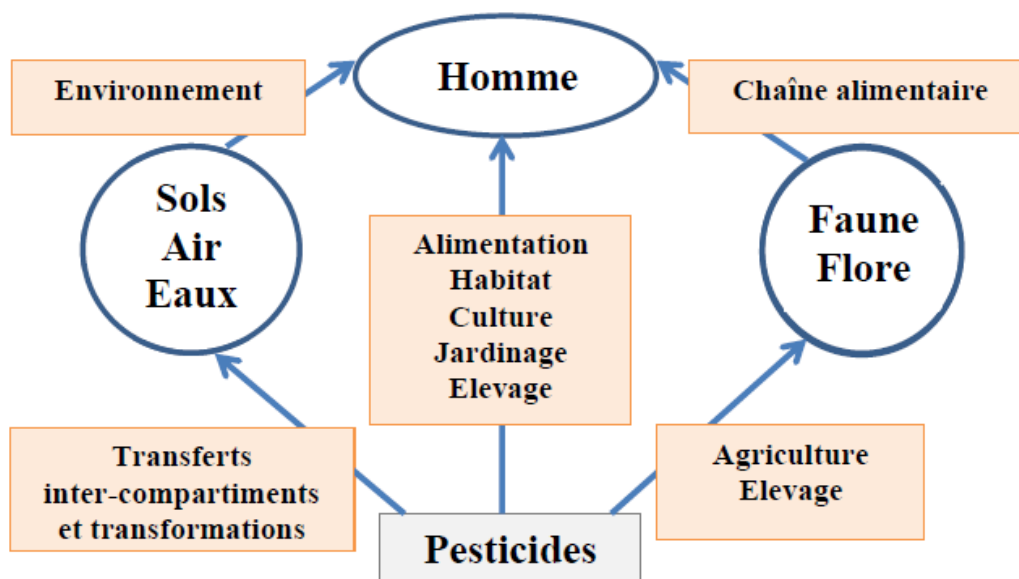
L'ingestion d'eau contaminée par les pesticides représente également un danger potentiel pour les nouveau-né. Ce problème est particulièrement important avec les aliments lactés qui doivent être préparés avec de l'eau.

Certaines espèces peuvent absorber et concentrer des pesticides sans en subir de conséquences apparentes, du moins jusqu'à un certain point. Les prédateurs qui s'en nourrissent risquent alors l'intoxication, Les pesticides se concentrent tout au long de la chaîne alimentaire, selon la figure suivante :



**Figure 3:** L'accumulation des pesticides

L'exposition aux pesticides se caractérise par plusieurs voies. L'exposition de la population générale se produit principalement par la consommation d'aliments et l'eau potable contaminée par des résidus de pesticides [12] comme l'indique la figure suivante :



**Figure 4:** Modes d'exposition de l'homme et des milieux par les pesticides

# *Chapitre II : Modélisation*

## *QSAR*

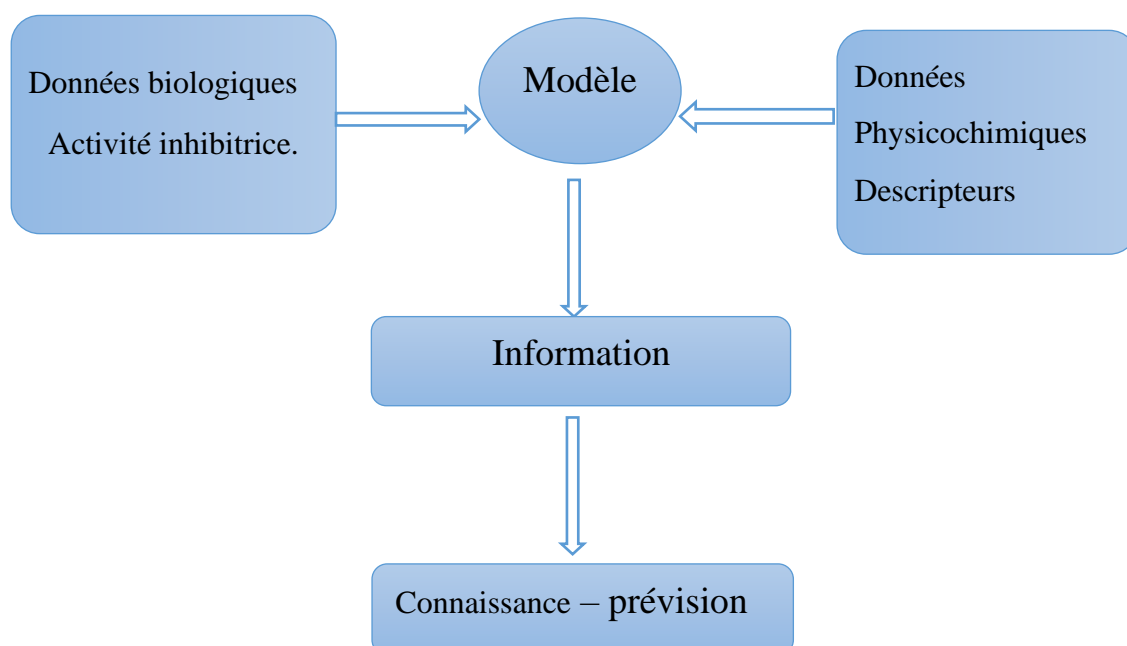
## II- Approche QSAR

### II-1 Définition :

La Relation Quantitative Structure et Activité (QSAR) est un outil que permet d'expliquer l'interaction des composés chimiques avec les sujets actuels, le principe des méthodes QSAR est comme leur nom l'indique, de mettre en place une relation mathématique reliant de manière quantitative des propriétés moléculaires, appelées descripteurs, avec une observable macroscopique (activité biologique, toxicité, propriété physico-chimique, etc.), pour une série des molécules biologiquement actives à l'aide des méthodes d'analyse de données [13].

### II-2 Principe de QSAR

Le principe des méthodes QSAR est, comme leur nom l'indique, de mettre en place une relation mathématique reliant de manière quantitative la structure moléculaire, codée par des propriétés moléculaires appelées descripteurs, avec une activité en utilisant des méthodes d'analyse de données. Par ces relations nous pouvons développer des modèles prédictifs de la forme générale suivante : **Activité = f (descripteurs moléculaires)** L'objectif de ces méthodes est alors d'analyser les données structurales afin de déterminer les facteurs influençant l'activité mesurée [14].



**Figure 5:** Présentation de la méthodologie de QSAR.



### II-3 Paramètres biologiques

Les données biologiques sont habituellement exprimées sur une échelle logarithmique en raison de la relation linéaire entre la réponse et le logarithme de dose dans la région centrale de la courbe de la relation linéaire entre la réponse. Les logarithmes inverses de l'activité ( $\log 1/C$ ) sont également utilisés pour obtenir des valeurs mathématiques plus élevées lorsque les structures sont biologiquement très efficaces. Des exemples de données biochimiques ou biologiques, utilisés dans l'analyse de QSAR, sont décrits dans le (tableau2) [15].

**Tableau II:** les paramètres biologique.

Source d'activité	Paramètres biologique
<b>1. Récepteurs isolés</b>	
Constante de vitesse	<b>log K</b>
Constante de Michaelis-Menten	<b>log 1/K<sub>m</sub></b>
Constante d'inhibition	<b>log 1/K<sub>i</sub></b>
<b>2. Systèmes cellulaires</b>	
Constante d'inhibition	<b>log 1/IC<sub>50</sub></b>
Résistance croisée	<b>log CR</b>
Donnée biologique in vitro	<b>log 1/C</b>
Mutation de gène	<b>log TA<sub>98</sub></b>
<b>3. Systèmes in vivo</b>	
Facteur de bioconcentration	<b>log BCF</b>
Vitesses de la réaction in vivo	<b>log I (introduction)</b>
Vitesses pharmacodynamiques	<b>log T (clairance totale)</b>

### II.4. Les descripteurs QSAR

Les descripteurs moléculaires sont déterminés pour chaque molécule ensuite liées mathématiquement à l'activité biologique mesurée. Ces descripteurs peuvent être des propriétés moléculaires simples telles que le coefficient de partage ( $\log p$ ) et le poids moléculaire, ou un descripteur plus complexe des champs d'interaction dans l'espace 3D. Dans ce dernier cas, les moléculaires sont superposés et alignés sur une grille, et les potentiels d'interaction sont déterminés pour chaque molécule au niveau de chaque point de grille [16].

### II-4-1 Les descripteurs physico-chimiques

Les descripteurs physico-chimiques caractérisent généralement la structure bidimensionnelle de la molécule. Parmi ces descripteurs, certains reflètent la composition moléculaire du composé, soit le type d'atomes et de liaisons présents dans la molécule, son nombre de cycle, son poids moléculaire .... : Ce sont les descripteurs constitutionnels. Certains représentent la surface accessible au solvant (nommée « Connolly-Accessible Surface »), le volume de solvant couvert par cette surface, le caractère hydrophile de la molécule généralement évalué à partir du coefficient de partage octanol/eau représenté par le log p. Ces descripteurs ne fournissent pas assez d'information sur la structure des molécules pour l'élaboration des modèles prédictifs plus complexes ; il est nécessaire d'ajouter d'autres types de descripteurs [16].

### II-4-2 Les descripteurs électroniques

Ces descripteurs caractérisent la distribution des charges des molécules (polarité des molécules) et les paramètres de la chimie quantique : le moment dipolaire, les énergies HOMO (orbitale moléculaire occupée de plus haute énergie) et LUMO (orbitale moléculaire non occupée de plus basse énergie) [16].

### II-4-3 Les descripteurs topologiques

Ces descripteurs décrivent les connectivités atomiques dans la molécule. Ils sont pour la plupart représentés par des indices comme ceux de Balaban, Wiener... et sont issus de la théorie des graphes. Cette théorie a été développée par Euler en 1736 avec les sept ponts de Königsberg. Un graphe est un ensemble de points, certains reliés par des lignes ; il permet de représenter la topologie des molécules sans se soucier de la géométrie spatiale exacte de cette dernière [16].

## II-5 Les applications de l'étude QSAR

Certaines études QSAR semblent être un peu plus d'études universitaires, il existe un grand nombre d'applications de ces modèles tels que :

- L'optimisation de l'activité pharmacologique.
- La conception rationnelle de nombreux autres produits tels que des agents tensioactifs, des parfums, des colorants et des produits chimiques fins.

- L'identification des composés dangereux dans les premiers stades de développement des produits ou la projection des stocks de composés existants.
- La prédiction de la toxicité et les effets secondaires de nouveaux composés.
- La prédiction de la toxicité pour les espèces.
- La prédiction de la toxicité pour les espèces environnementales.
- La sélection des composés ayant des propriétés pharmacocinétiques optimales, que ce soit la stabilité ou la disponibilité dans les systèmes biologiques.
- La prédiction d'une variété de propriétés physico-chimique des molécules.
- La prédiction du devenir des molécules qui sont libérées dans l'environnement.
- La prédiction des effets conjugués des molécules, que ce soit dans des mélanges ou des formulations [17].

## II.6. Introduction à la régression linéaire multiple

MLR est l'une des méthodes de modélisation les plus populaires grâce à sa simplicité d'utilisation et facilité d'interprétation. L'avantage important de la régression linéaire multiple est qu'elle est très transparente, puisque l'algorithme est disponible, et que les prédictions peuvent être réalisées facilement [18].

Le modèle de régression linéaire multiple est l'outil statistique le plus habituellement mis en œuvre pour l'étude de données multidimensionnelles. Cas particulier de modèle linéaire, il constitue la généralisation naturelle de la régression simple [19].

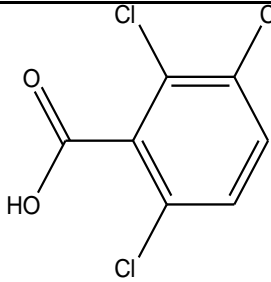
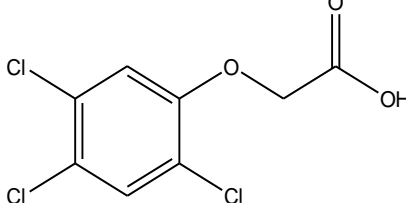
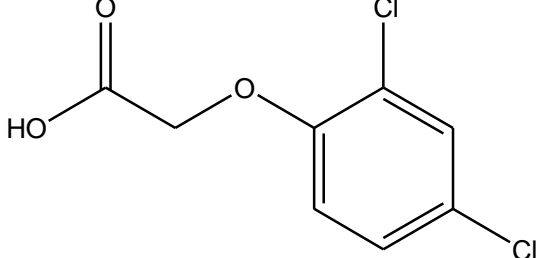
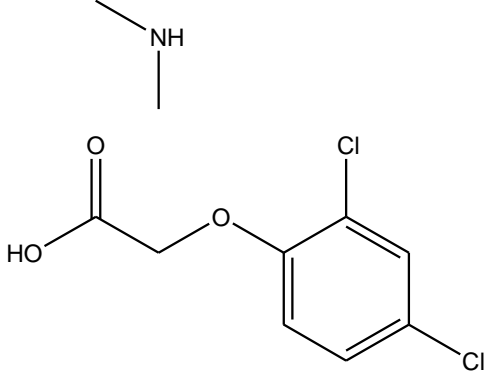
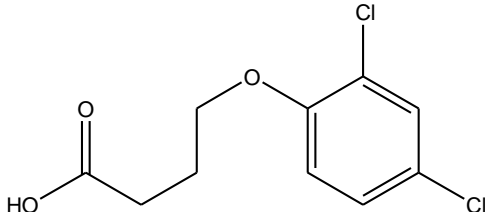
Une variable quantitative  $Y$  dite à expliquer (ou encore, réponse, exogène, dépendante) est mise en relation avec  $p$  variables quantitatives  $X^1, \dots, X^p$  dites explicatives (ou encore de contrôle, endogènes, indépendantes, régresseurs).

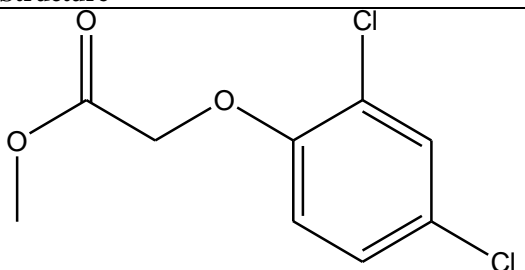
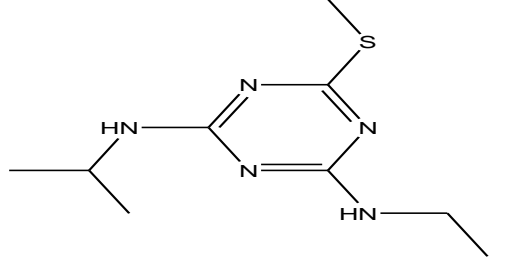
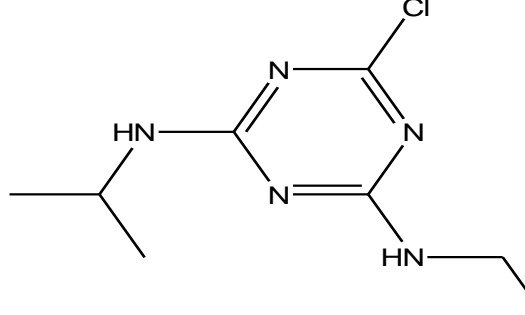
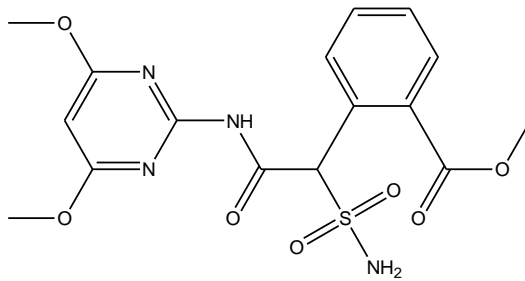
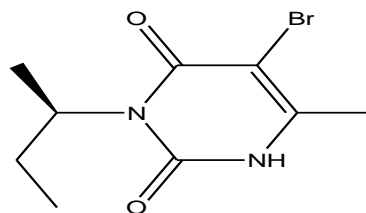
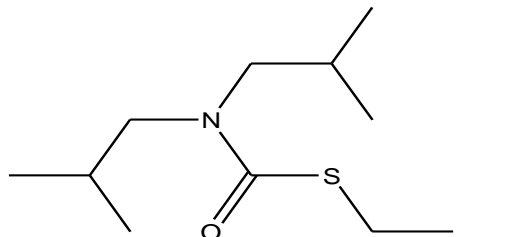
# ***APPLICATION***

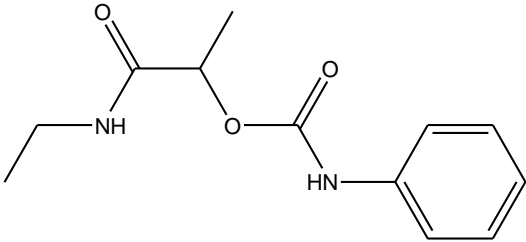
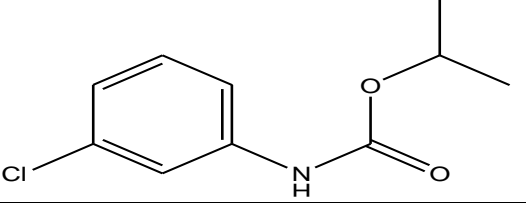
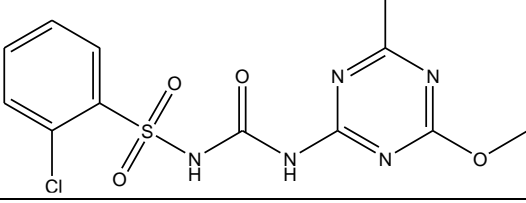
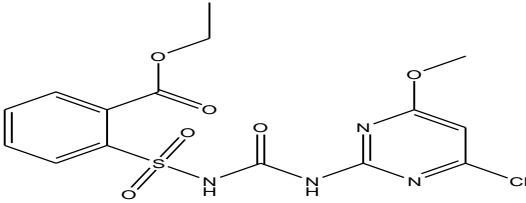
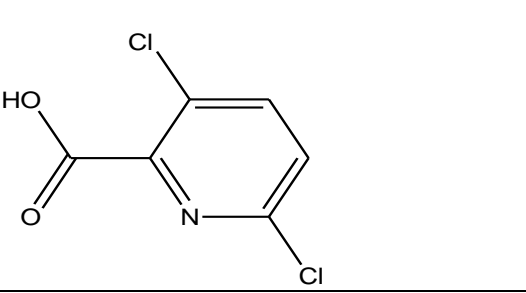
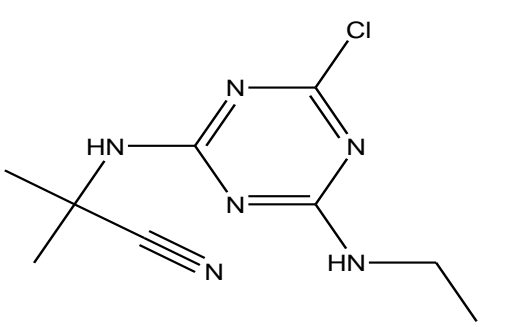
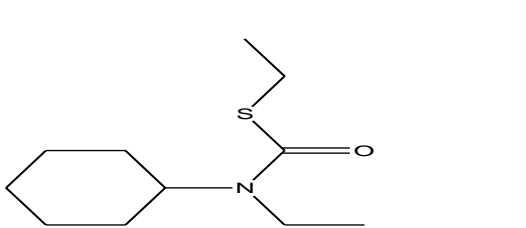
Dans cette partie expérimentale, nous allons déterminer la toxicité des pesticides aux poissons par modélisation d'un ensemble d'estimation. Le Meilleur modèle obtenue sera validé par l'approche hybride algorithme génétique et régression multilinéaire GA/MLR.

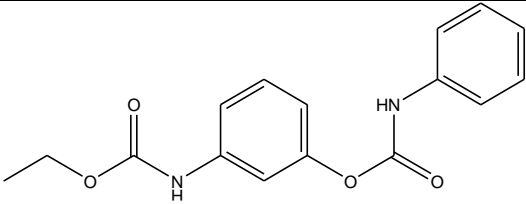
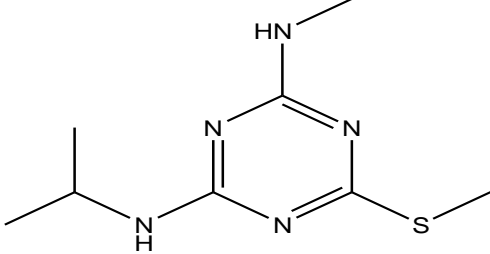
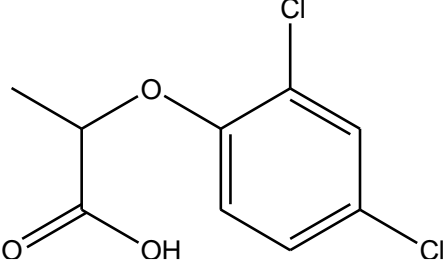
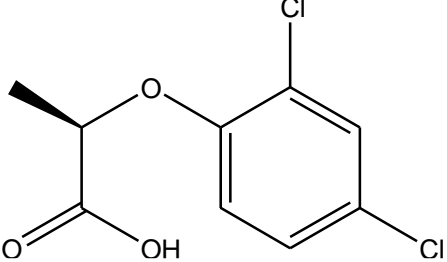
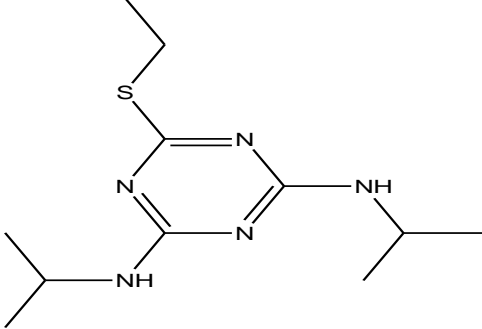
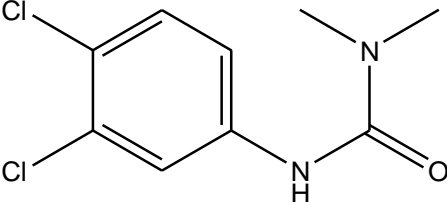
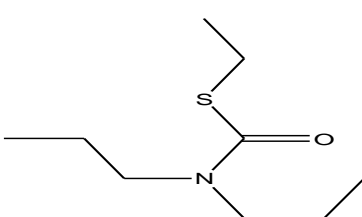
La liste des molécules étudiées dans notre travail se présente dans le Tableau III, cette base des données contient 61 composés, avec la propriété à étudier.

**Tableau III:** Composés étudiés et leurs toxicités

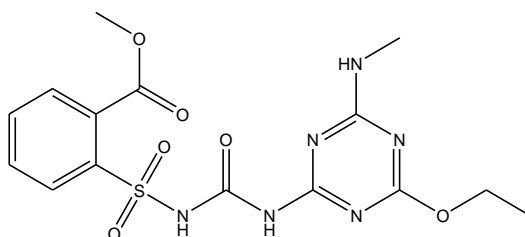
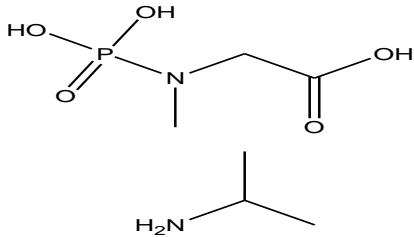
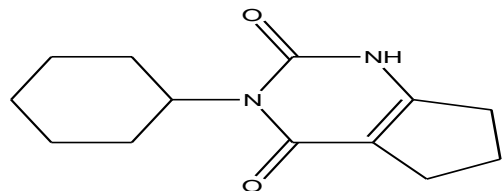
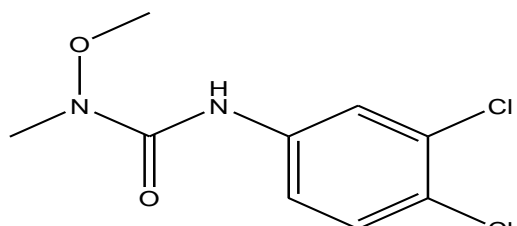
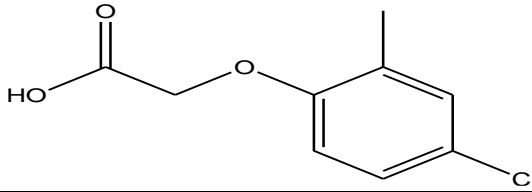
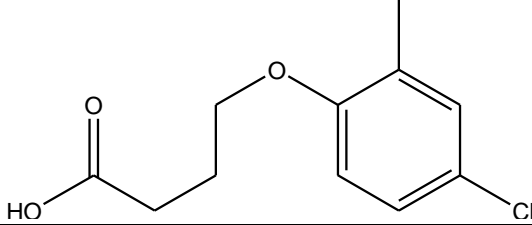
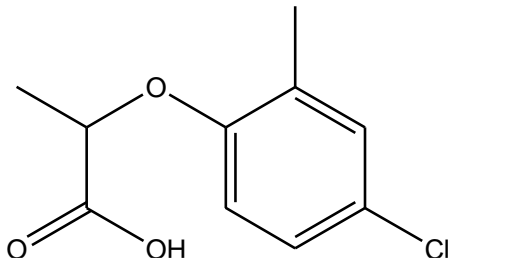
Compounds	Formule	Structure	pLC50
2, 3,6-TBA	$C_7H_3Cl_3O_2$		-2
2, 4,5-T	$C_8H_5Cl_3O_3$		-2.5440
2,4-D	$C_8H_6Cl_2O_3$		-0.04139269
2,4-Ddimethylammonium	$C_{10}H_{13}Cl_2NO_3$		-2
2,4-DB	$C_{10}H_{10}Cl_2O_3$		-0.60205999

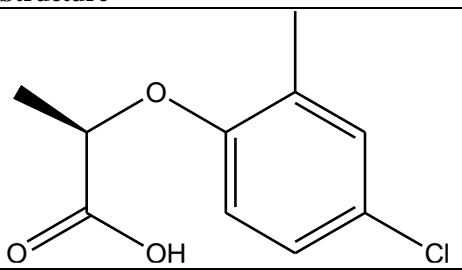
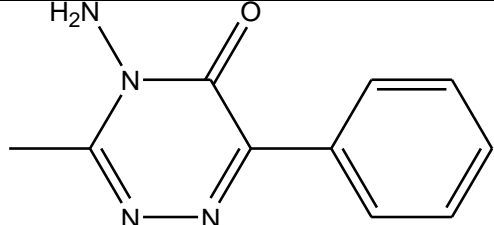
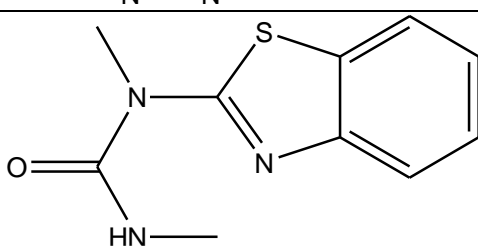
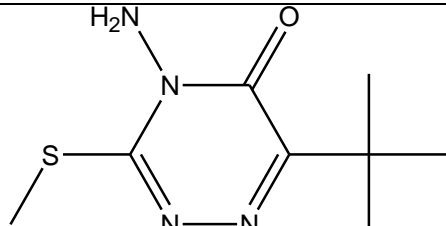
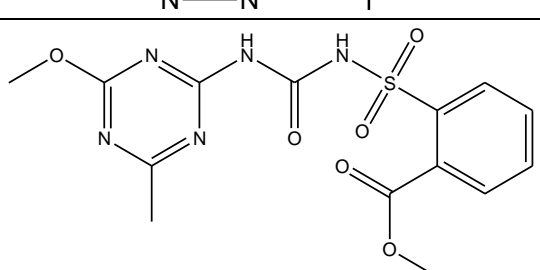
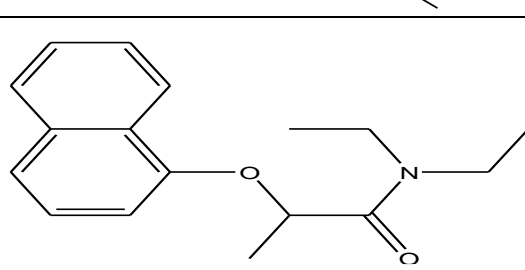
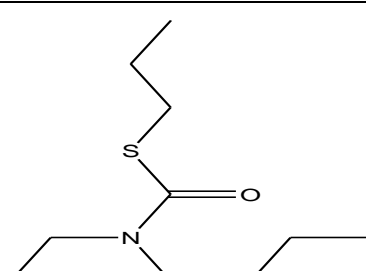
Compounds	Formule	Structure	pLC50
2,4-D-methyl	$C_9H_8Cl_2O_3$		0.04575749
Ametryn	$C_9H_{17}N_5S$		-0.69897
Atrazine	$C_8H_{14}ClN_5$		-0.65321251
bensulfuron methyl *	$C_{16}H_{18}N_4O_7S$		-2.17609126
Bromacil *	$C_9H_{13}BrN_2O_2$		-1.85125835
Butylate	$C_{11}H_{23}NOS$		-0.62324929

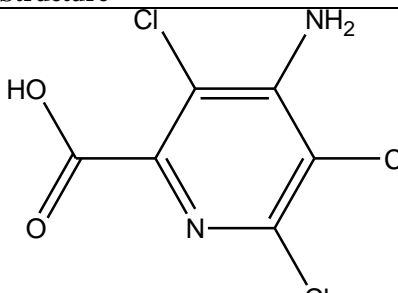
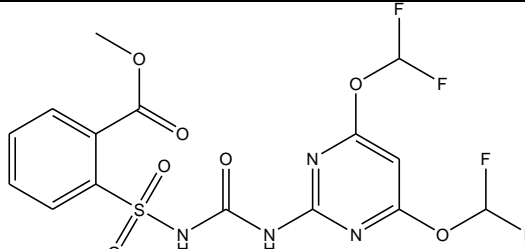
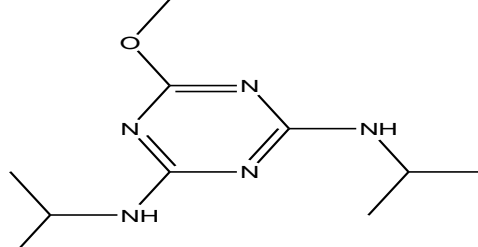
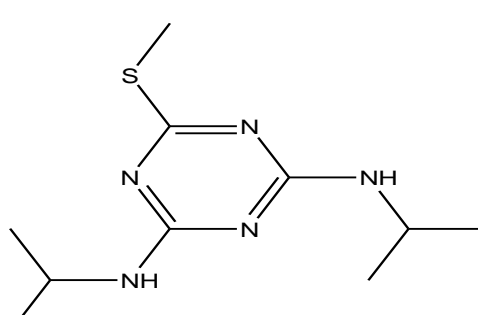
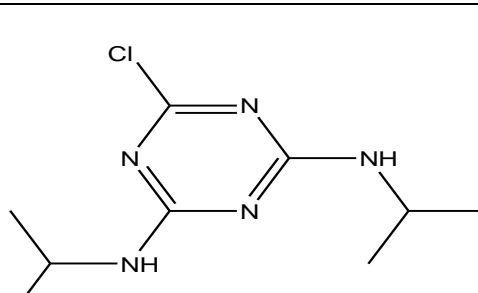
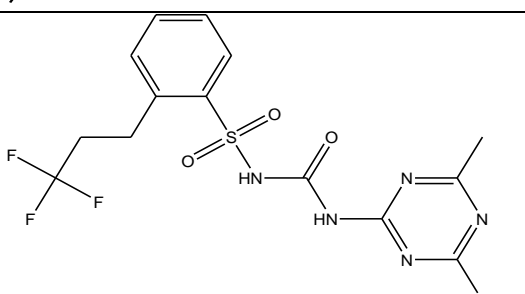
Compounds	Formule	Structure	pLC50
Carbetamide	$C_{12}H_{16}N_2O_3$		-2.54900326
Chlorpropham *	$C_{10}H_{12}ClNO_2$		-0.48000694
Chlorsulfuron *	$C_{12}H_{12}ClN_5O_4S$		-1.69897
chorimuron ethyl	$C_{15}H_{15}ClN_4O_6S$		-2
Clopyralide	$C_6H_3Cl_2NO_2$		-2.01494035
Cyanazine	$C_9H_{13}ClN_6$		-0.68124124
Cycloate *	$C_{11}H_{21}NOS$		-0.65321251

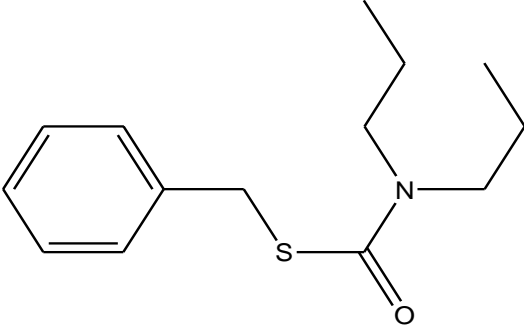
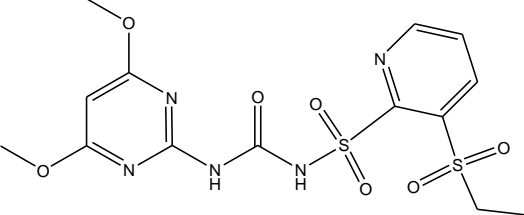
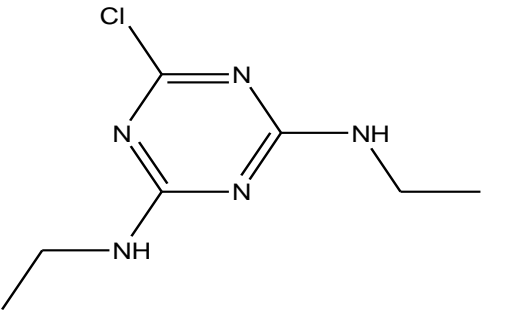
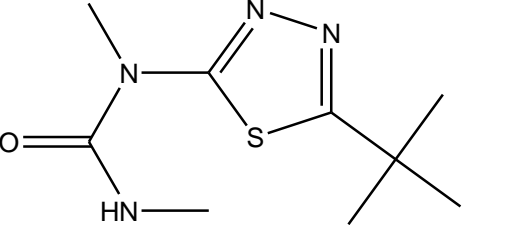
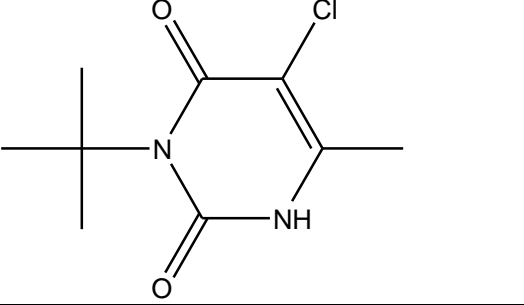
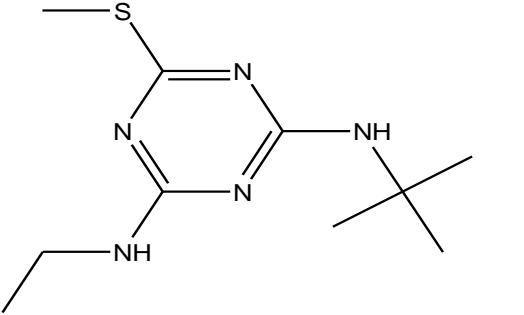
Compounds	Formule	Structure	pLC50
Desmedipham	$C_{16}H_{16}N_2O_4$		0.09691001
Desmetryn	$C_8H_{15}N_5S$		-0.34242268
Dichlorprop	$C_9H_8Cl_2O_3$		-2.63144377
Dichlorprop p *	$C_9H_8Cl_2O_3$		-2
Dipropetryn	$C_{11}H_{21}N_5S$		-0.20411998
Diuron	$C_9H_{10}Cl_2N_2O$		-0.77085201
EPTC *	$C_9H_{19}NOS$		-1



Compounds	Formule	Structure	pLC50
Ethametsulfuron methyl	$C_{15}H_{18}N_6O_6S$		-2.77815125
Glyphosate	$C_6H_{17}N_2O_5P$		-3
Lenacil	$C_{13}H_{18}N_2O_2$		-1
Linuron	$C_9H_{10}Cl_2N_2O_2$		0.05060999
MCPA	$C_9H_9ClO_3$		-1.98677173
MCPB *	$C_{11}H_{13}ClO_3$		-0.74818803
Mecoprop	$C_{10}H_{11}ClO_3$		-2

Compounds	Formule	Structure	pLC50
Mecoprop P	$C_{10}H_{11}ClO_3$		-2.17609126
Metamitron	$C_{10}H_{10}N_4O$		-2.5132176
Methabenzthiazuron	$C_{10}H_{11}N_3OS$		-1.20139712
Metribuzin *	$C_8H_{14}N_4OS$		-1.77085201
Metsulfuron methyl	$C_{14}H_{15}N_5O_6S$		-2.17609126
Napropamide *	$C_{17}H_{21}NO_2$		-1.22010809
Pebulate	$C_{10}H_{21}NOS$		-0.79588002

Compounds	Formule	Structure	pLC50
Picloram	$C_6H_3Cl_3N_2O_2$		-1.28555731
Primisulfuron methyl *	$C_{15}H_{12}F_4N_4O_7S$		-1.84509804
Prometon	$C_{10}H_{19}N_5O$		-1.07918125
Prometryn	$C_{10}H_{19}N_5S$		-0.39794001
Propazine	$C_9H_{16}ClN_5$		-1.24303805
Prosulfuron	$C_{15}H_{16}F_3N_5O_4S$		-2

Compounds	Formule	Structure	pLC50
Prosulocarb	$C_{14}H_{21}NOS$		-0.23044892
Rimsulfuron	$C_{14}H_{17}N_5O_7S_2$		-2.04139269
Simazine	$C_7H_{12}ClN_5$		-1.69019608
Tebuthiuron	$C_9H_{16}N_4OS$		-2.04921802
Terbacil	$C_9H_{13}ClN_2O_2$		-1.66464198
Terbutryn	$C_{10}H_{19}N_5S$		-0.47712125

Compounds	Formule	Structure	pLC50
Terbutylazine	$C_9H_{16}ClN_5$		-0.25527251
Thifensulfuron methyl	$C_{12}H_{13}N_5O_6S_2$		-2.5563025
Thiobencarb	$C_{12}H_{16}ClNOS$		-0.5563025
Triallate	$C_{10}H_{16}Cl_3NOS$		-0.07918125
Triasulfuron	$C_{14}H_{16}ClN_5O_5S$		-2
Tribenuron methyl	$C_{15}H_{17}N_5O_6S$		-3
Triclopyr	$C_7H_4Cl_3NO_3$		-2.06818586

Compounds	Formule	Structure	pLC50
Triflusulfuron methyl	$C_{17}H_{19}F_3N_6O_6S$		-2.86332286
Vernolate	$C_{10}H_{21}NOS$		-0.66275783
Vinclozolin	$C_{12}H_9Cl_2NO_3$		-1.51188336

\*Composés de validation

## I- Calcul du modèle :

### I- 1- Choix du modèle :

Les molécules sont dessinées par le logiciel ChemDraw (ChemDraw ultra 7.0) [20]. Puis elles sont optimisées en utilisant le logiciel HyperChem [21]. Les descripteurs moléculaires ont été calculés à l'aide du logiciel informatique Dragon [22], plus de 1600 descripteurs sont calculés.

L'ensemble des données a été décomposé en deux sous-ensembles aléatoirement, 70% de la totalité des composés pour la construction du modèle et 30% pour la validation externe.

D'après l'algorithme génétique dans la version MobyDigs [23], plusieurs modèles sont obtenues pour chaque jeu de groupe ; le choix a été opté pour le modèle qui conduit aux

meilleurs statistiques des 100 modèles générés par algorithmes génétiques. Sa validation est effectuée par 18 composés, alors que sa calibration par 43.

**I- 2- Choix d'une famille de descripteurs :**

Les descripteurs entrant dans le modèle obtenus par algorithme génétique, leurs classes et de brèves définitions sont donnés dans le tableau IV

**Tableau IV : Descripteurs moléculaires intervenant dans la modélisation de pLC50**

N°	Descripteur	Classe	Signification
1	ATS3m	2D autocorrelation	Broto-Moreau autocorrelation of a topological structure - lag 3 / weighted by atomic masses
2	BEHm2	Burden eigenvalue descriptors	highest eigenvalue n. 2 of Burden matrix / weighted by atomic masses
3	GGI2	Topological charge indices indices	topological charge index of order 2
4	Du	WHIM descriptors indices	D total accessibility index / unweighted
5	R5v+	GETAWAY descriptore	R maximal autocorrelation of lag 5 / weighted by atomic van der Waals volumes
6	TPSA (NO)	Molecular porperties	Topological polar surface area using N.O polar contributions

Les valeurs du pLC50expérimentales et prédites, les valeurs des Leviers, et les valeurs des erreurs standardisées sont présentées dans le tableau V au-dessous :

**Tableau V:** Valeurs de pLC50expérimentales, prédites,  $h_{ii}$ ,  $e_{istd}$

Composé	pLC50 <sub>exp</sub>	pLC50 <sub>pred</sub>	$h_{ii}$	$e_{istd}$
2, 3,6-TBA	-2	-1.49	0.215	1.63
2, 4,5-T	-2.54	-1.6	0.149	2.66
2,4-D	-0.04	-1.19	0.129	-3.13
2,4-D dimethylammonium	-2	-1.62	0.069	0.92

Composé	pLC50 <sub>exp</sub>	pLC50 <sub>pred</sub>	h <sub>ii</sub>	e <sub>istd</sub>
2,4-D-methyl	0.05	-0.68	0.083	-1.81
Ametryn	-0.7	-0.52	0.1	0.47
Atrazine	-0.65	-1.06	0.112	-1.08
Bromacil	-1.85	-2.32	0.418	-2.31
Butylate	-0.62	-0.83	0.096	-0.53
Carbetamide	-2.55	-1.98	0.109	1.49
Chlorpropham	-0.48	-0.62	0.069	-0.34
Chlorsulfuron	-1.7	-1.86	0.11	-0.43
Cyanazine	-0.68	-1.45	0.14	-2.14
Desmedipham	0.1	-0.54	0.158	-1.81
Dichlorprop	-2.63	-2.74	0.219	-0.35
Dipropetryn	-0.2	-0.06	0.153	0.4
Diuron	-0.77	-0.3	0.156	1.35
Ethametsulfuron methyl	-2.78	-2.92	0.212	-0.44
Glyphosate	-3	-3.08	0.433	-0.39
Lenacil	-1	-0.82	0.062	0.43
Linuron	0.05	-0.77	0.144	-2.28
Mecoprop	-2	-1.9	0.084	0.26
Mecoprop P	-2.18	-1.9	0.084	0.7
Metamitron	-2.51	-2.06	0.123	1.22
Metribuzin	-1.77	-1.97	0.088	-0.5
Metsulfuron methyl	-2.18	-2.09	0.198	0.28
Napropamide	-1.22	-1.15	0.148	0.19
Pebulate	-0.8	-0.99	0.094	-0.5
Primisulfuron methyl	-1.85	-1.93	0.167	-0.24
Prometon	-1.08	-0.85	0.128	0.63
Prometryn	-0.4	-0.16	0.122	0.64
Propazine	-1.24	-1.12	0.096	0.32
Prosulocarb	-0.23	0.04	0.159	0.78
Rimsulfuron	-2.04	-1.98	0.761	1.1
Simazine	-1.69	-1.18	0.111	1.36
Tebuthiuron	-2.05	-1.7	0.17	1.03
Terbacil	-1.66	-2.11	0.097	-1.15
Terbutryn	-0.48	-0.1	0.202	1.18
Triasulfuron	-2	-1.88	0.156	0.33
Triclopyr	-2.07	-2.11	0.185	-0.14
Triflusulfuron methyl	-2.86	-2.69	0.279	0.63
Vernolate	-0.66	-0.85	0.096	-0.49
Vinclozolin	-1.51	-1.37	0.116	0.36
2,4-DB *	-0.6	-0.25	0.133	0.82
bensulfuron methyl *	-2.18	-2.68	0.176	-1.23
chorimuron ethyl *	-2	-1.54	0.195	1.14
Clopyralide *	-2.01	-2.64	0.398	-1.78
Cycloate *	-0.65	-0.5	0.158	0.36



Composé	pLC50 <sub>exp</sub>	pLC50 <sub>pred</sub>	h <sub>ii</sub>	e <sub>istd</sub>
Desmetryn *	-0.34	-0.53	0.128	-0.45
Dichlorprop p *	-2	-2.74	0.214	-1.85
EPTC *	-1	-0.91	0.114	0.22
MCPA *	-1.99	-1.16	0.173	2.02
MCPB *	-0.75	-0.54	0.142	0.49
Methabenzthiazuron *	-1.2	-0.71	0.132	1.17
Picloram *	-1.29	-1.9	0.259	-1.58
Prosulfuron *	-2	-1.51	0.16	1.18
Terbutylazine *	-0.26	-0.97	0.113	-1.69
Thifensulfuron methyl *	-2.56	-2.42	0.222	0.34
Thiobencarb *	-0.56	-0.08	0.237	1.2
Triallate *	-0.08	0.4	0.395	1.35
Tribenuron methyl *	-3	-2.9	0.199	0.25

\*Composés de validation

## II- Développement du modèle :

### II- 1- Matrice de corrélation :

La matrice de corrélation entre la variable dépendante et les variables explicatives est présentée dans le tableau VI

**Tableau VI:** Matrice de corrélation

	pLC50	ATS3m	BEHm2	GGI2	Du	R5v+
ATS3m	-0.389 0.009					
BEHm2	-0.252 0.099	0.786 0.000				
GGI2	-0.568 0.000	0.836 0.000	0.686 0.000			
Du	0.095 0.541	-0.230 0.134	0.070 0.652	-0.028 0.856		
R5v+	-0.311 0.040	0.135 0.383	-0.085 0.582	-0.009 0.951	-0.645 0.000	
TPSA(NO)	-0.492 0.001	0.644 0.000	0.650 0.000	0.567 0.000	0.015 0.924	-0.251 0.100

Les descripteurs GGI2, ATS3m, R5v+, et TPSA(NO) sont fortement corrélés ( $p < 0.05$ ) avec la variable dépendante (pLC50), par contre les deux autres descripteurs sont moins corrélés avec pLC50, mais comme même ces descripteurs portent des informations complémentaires au pLC50. La suspension d'un descripteur d'entre eux entraîne une dégradation des capacités prédictives du modèle, ils portent des informations importantes au pLC50.

**II- 2- Equation et analyse de régression :**

L'équation de régression du modèle calculé est la suivante :

$$\mathbf{pLC50} = - 5.53 (\pm 1.963) + 1.66 (\pm 0.047) \mathbf{ATS3m} + 1.44 (\pm 0.644) \mathbf{BEHm2} - 0.937 (\pm 0.148) \mathbf{GGI2} - 5.54 (\pm 2.201) \mathbf{Du} - 72.4 (\pm 10.24) \mathbf{R5v+} - 0.0189 (\pm 0.002) \mathbf{TPSA(NO)} \quad (1)$$

**Tableau VII:** Analyse de régression

Régresseur	Coef	SE Coef	T	P	VIF
Constant	-5.525	1.963	-2.81	0.008	
ATS3m	1.6557	0.4702	3.52	0.001	6.890
BEHm2	1.4388	0.6445	2.23	0.032	3.371
GGI2	-0.9366	0.1482	-6.32	0.000	3.693
Du	-5.535	2.201	-2.52	0.016	2.101
R5v+	-72.37	10.24	-7.07	0.000	2.062
TPSA(NO)	-0.018926	0.002669	-7.09	0.000	2.277

Les valeurs de T des descripteurs sont proches se qui confirme la bonne homogénéité de la contribution des descripteurs dans le modèle.

Les valeurs des VIF inférieur à 5, sauf celle du descripteur ATS3m, ce qui traduit l'indépendance des cinq descripteurs intervenant dans le modèle.

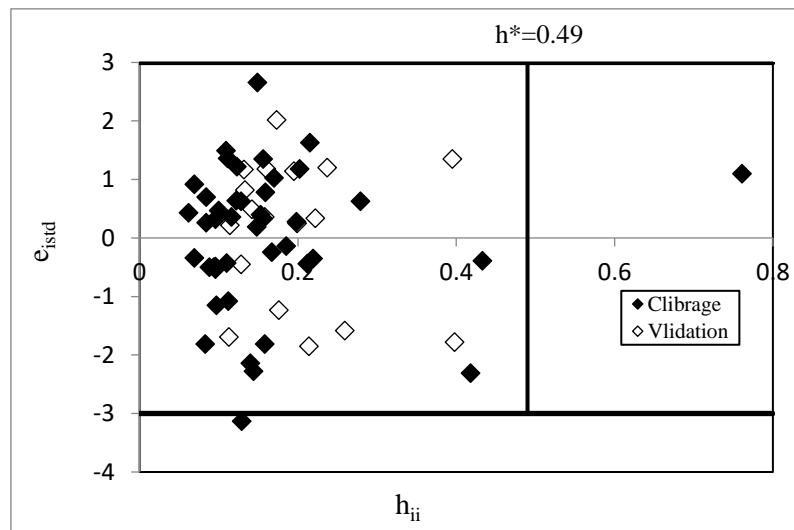
Les paramètres statistiques sont présentés dans le tableau suivant

**Tableau VIII:** Les paramètres statistiques pour l'ensemble de calibration (43 composés) et celui de validation (18 composés)

Les paramètres statistiques	
F	21.87
R <sup>2</sup>	78.48
R <sup>2</sup> <sub>adj</sub>	74.88
Q <sup>2</sup> <sub>Loo</sub>	70.00
Q <sup>2</sup> <sub>ext</sub>	70.75
S	0.453
EQMC	0.414
EQMP	0.489
EQMP <sub>ext</sub>	0.483

Les paramètres statistiques reproduit dans le tableau ci-dessus montrent une bonne explication de la variabilité pLC50, par les descripteurs choisis, de l'ordre de 78 pour le coefficient de détermination, et une robustesse du modèle due à la valeur de  $Q_{Loo}^2$ , alors que les valeurs de l'erreur quadratique moyenne de prédiction et de calcul sont petites et proches; en outre ce modèle est significatif avec une valeur du paramètre de Fisher: (F=21,87), Les statistiques montrent aussi que le modèle à une très grande capacité prédictive externe, ceci traduit par la grande valeur de  $Q_{ext}^2$  et la faible valeur de EQMP<sub>ext</sub>.

## II- 3- Diagramme de Williams :

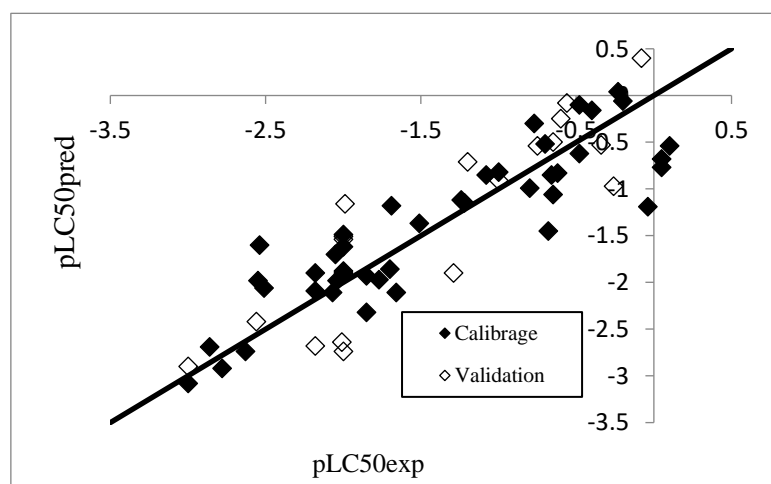


**Figure 6:** Diagramme de Williams

D'après le graphe on remarque que tous les erreurs standardisé sont comprise entre les limites  $\pm 3$ , a part d'un point de l'ensemble de calibration qui représente un point aberrant, on remarque aussi une présence d'un point influent qui a une valeur de levier supérieur à la valeur critique  $h^*=0.49$  ( $h^* = (3(k+1))/n$ ).

## II- 4- Qualité de l'ajustement :

La qualité de l'ajustement a été vérifiée par le graphe des valeurs calculées de l'ensemble de calibration et les valeurs prédites de l'ensemble de validation en fonction de celles expérimentales du pLC50.



**Figure 7 :** Qualité de l'ajustement

D'après le graphe on remarque une dispersion autour de la première bissectrice des deux ensembles de calibration et de validation. Ceci est confirmé par la valeur de  $Q^2_{\text{LOO}}$ , (égale à 70 %).

## II- 5- Test de randomisation :

Dans le but de montrer que le modèle obtenu n'est pas dû au hasard nous avons appliqué le test de randomisation de Y (Figure 08), dans cette figure on a présenté le graphe des coefficients statistiques  $Q^2$  et  $R^2$ .

Il est clair que les statistiques obtenues pour les vecteurs pLC 50 modifiés sont plus petites que celles du modèle réel, les cercles vides regroupés dans la région des valeurs négatives de  $Q^2$  ont des valeurs petites de  $R^2$ , seul le cercle noirci a des valeurs élevés et proches pour ces deux statistiques, il représente notre modèle calculé, donc il n'est pas dû au hasard.

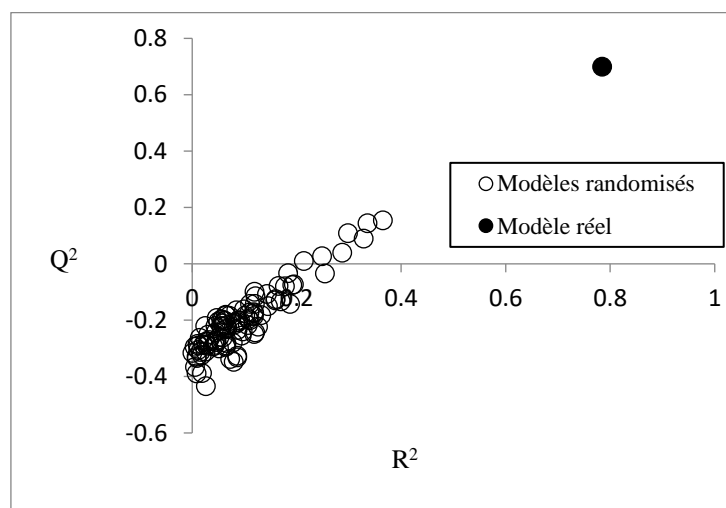


Figure 8: Test de randomisation

*Conclusion  
Générale*

### Conclusion générale

Nous avons utilisé la méthodologie QSAR pour relier la toxicité des poissons par un mélange hétérogène d'herbicides ayant des propriétés chimiques et des origines diverses, à des descripteurs moléculaires théoriques caractéristiques de la molécule entière ou de ses fragments, calculés à l'aide de logiciels spécialisés du commerce.

Nous avons recherché des corrélations linéaires entre la variable dépendante et variables explicatives sélectionnées par algorithme génétique, en utilisant la régression linéaire multiple.

Le modèle obtenu conduit aux bons résultats : qualité de l'ajustement, robustesse interne et externe, capacité prédictive ... ce qui prouve que la relation entre le pLC50 et les variables explicatives sélectionnées par algorithme génétique elle est fondamentalement linéaire.

*Références  
bibliographiques*



## **Références bibliographiques**

- [1] : ANIS, A. (2012). Évaluation de la toxicité de pesticides sur quatre niveaux trophiques marins : microalgues, échinoderme, bivalves et poisson. Thèse de Doctorat. Université de Bretagne occidentale –Brest. P10.
- [2] : AYAD-MOKHTARI, N. (2012). Identification et dosage des pesticides dans l'Agriculture et les problèmes d'Environnement liés. Mémoire de Magister. Université d'Oran. P13.
- [3] : MERGHID, M, DEBBACHE, M, FOUGHALI, I. (2017). Impacts des pesticides utilisés dans la plasticulture sur la santé humaine En Algérie- Etude de cas la wilaya de Constantine -. Mémoire de Master. Université des Frères Mentouri Constantine. P4.
- [4] : GHENNAM, S, HIMEUR, I, NAFAA, H. (2015). Étude de la génotoxicité du pesticide « Topik 80 » *in vivo* (*Allium cepa* test). Mémoire de Master. Université 8 MAI 1945 GUELMA.P2.
- [5] : MENNAI, Kh. (2019). Effets d'utilisation aléatoire des pesticides sur la valeur nutritionnelle des tomates (*Solanum lycopersicum* L). Mémoire de Master. Université Echahid Hamma Lakhdar -El OUED. P29.
- [6] : GUEDDOU, A, NEDJAA, Kh. (2017). Évaluation de la toxicité des pesticides par l'utilisation d'un biotest. Mémoire de Master. Université A. MIRA – Bejaia. P3.
- [7] : GHENNAM, S, HIMEUR, I, NAFAA, H. (2015). Étude de la génotoxicité du pesticide « Topik 80 » *in vivo* (*Allium cepa* test). Mémoire de Master. Université 8 MAI 1945 GUELMA. P3.
- [8] : MERGHID, M, DEBBACHE, M, FOUGHALI, I. (2017). Impacts des pesticides utilisés dans la plasticulture sur la santé humaine En Algérie- Etude de cas la wilaya de Constantine -. Mémoire de Master. Université des Frères Mentouri Constantine. P20.
- [9] : TIDAFI, N, BEKHTI, S. (2019). Evaluation du stress chez des poissons d'eau douce : Tilapia rouge (*Oreochromis* sp) exposés à des résidus de pesticide à base de (L'oxyfluorène). Mémoire de Master. Université Blida 1. P14.

- [10]: MERGHID, M, DEBBACHE, M, FOUGHALI, I. (2017). Impacts des pesticides utilisés dans la plasticulture sur la santé humaine En Algérie- Etude de cas la wilaya de Constantine -. Mémoire de Master. Université des Frères Mentouri Constantine. PP20-28.
- [11] : TIDAFI, N, BEKHTI, S. (2019). Evaluation du stress chez des poissons d'eau douce : Tilapia rouge (*Oreochromis sp*) exposés à des résidus de pesticide à base de (L'oxyfluorène). Mémoire de Master. Université Blida 1. P15.
- [12] : MERGHID, M, DEBBACHE, M, FOUGHALI, I. (2017). Impacts des pesticides utilisés dans la plasticulture sur la santé humaine En Algérie- Etude de cas la wilaya de Constantine -. Mémoire de Master. Université des Frères Mentouri Constantine. PP32-41
- [13] : ZERGUI, A. (2019). Application de la méthode DFT dans l'étude quantitative des relations structure-activité d'un ensemble de composés à visées thérapeutique. Mémoire de Master. Université Mohamed Khider de Biskra. P36.
- [14] : SAIHI, Y. (2015). Etude de la relation quantitative structure-activité inhibitrice des enzymes hydrolytiques : cas des alpha-glucosidases. Thèse Doctorant. Université badji Mokhtar – Annaba. P55.
- [15] : MELKEMI, N. (2013). Etude des corrélations entre les paramètres structuraux et les propriétés bioactives de quelques composés organiques. Thèse Doctorant. Université Mohamed Khider de Biskra. P108.
- [16] : ZERGUI, A. (2019). Application de la méthode DFT dans l'étude quantitative des relations structure-activité d'un ensemble de composés à visées thérapeutique. Mémoire de Master. Université Mohamed Khider de Biskra. PP37-38.
- [17] : ZERGUI, A. (2019). Application de la méthode DFT dans l'étude quantitative des relations structure-activité d'un ensemble de composés à visées thérapeutique. Mémoire de Master. Université Mohamed Khider de Biskra. P39.
- [18] : ROY, K, KAR, S, NARAYAN Das, R. (2015). Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment. Chapter 6- Selected Statistical Methods in QSAR|. Academic Press. Boston, PP 191–229.
- [19] : BESSE, PH. (2003).Thème Pratique de la modélisation Statistique. Université Paul SABATIER Toulouse Version. PP 9-21.
- [20]: ChemDraw Ultra 7.0 for windows,

## *Références bibliographiques*

[21]: Hyperchem<sup>TM</sup> Release 6.03 for windows, Molecular Modeling system, 2000.

[22] : TODESCHINI, R., CONSONNI, V, PAVAN, M. (2005). DRAGON, Software for the calculation of Molecular Descriptors. Release 5.3 for windows, Milano.

[23]: TODESCHINI, R, BALLABIO, D, CONSONNI, V, MAURI, A, PAVAN, M. Milano (2009). MOBYDIGS Software for Multilinear Regression Analysis and variable Subset Selection by Genetic Algorithm. Release I.1 for Windows.