



MINISTRE DE L'ENSEIGNEMENT SUPERIEUR  
ET DE LA RECHERCHE SCIENTIFIQUE  
UNIVERSITE «Abbès LAGHROUR» DE KHENCHELA  
FACULTE 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 :**

*Prédiction de la constante de Henery d'une  
série des pesticides utilisant une méthode  
linéaire.*

*Réalisé par :*

-MEBARKICHAHRA

*Dirigé par : Dr. BOUAKKADIA A*

-SEGHIRI NABILA

*Membres de jury :*

RADJEHI LAMIA	MCB	UAL Khenchela	Président
BOUAKKADIA Amel	MCB	UA L Khenchela	Encadreur
KERTIOU Nour Eddine	MCB	UA L Khenchela	Examineur

Promotion : 2018/2019

# **Dedicase**

*Je dédie ce modeste travail A mes très chers parents et ma grande famille.*

*A tous les enseignants de département des Sciences de la Matière de l'université de  
khenchela.*

*A l'enseignant ABAR MOHAMED.*

*A mes très chers frères.*

*A mes amis, AHLEM, SAMIA, HASSIBA, NAIMA, et*

*A Tous mes amis.*

*A tous qui m'aiment et*

*A tous que je les aime.*

*NABILA....*

# *Dédicace*

*Je dédie ce travail*

*A deux personnes que j'ai tant aimés, mon adorable*

*Mère qui m'a beaucoup donné, et mon cher père.*

*A tous les enseignants de département des Sciences de la Matière de l'université de  
khenchela.*

*A mes trois chers frères Oussama, Abd Aldjalil et Mouatez.*

*A mes chères sœurs Hanane, Hanna et Sadjida.*

*A tous mes amis surtout Samia, Ahlem, Fatma et Loubna.*

*A tous et à toutes les personnes qui j'aime.*

**CHAHRA ZED.....**

# Remerciement

Avant tout, nous tenons à remercier«**Allah**» le tous puissant, pour nous avoir donné la force

Et la patience.

J'exprime mes remerciements à notre encadreur Dr: **BOUAKKADIA AMEL**

Pour la grande bienveillance qu'elle nous a accordé et pour leurs efficaces et précieuses aides,

Tout le long d'achèvement de ce mémoire.

Je remercie également tous les membres du jury d'avoir accepté de juger ce travail.

Et enfin je remercié l'ensemble d'enseignants et collègues de notre promotion.

Un remerciement A toutes et à tous qui, de loin ou de près, ont contribué à la réalisation de ce

Mémoire.

**Merci.....**



## SOMMAIRE

Résumé	
Abstract	
Liste des tableaux	
Liste des figures	
Symboles et abréviation	
Introduction générale	2
<b>Chapitre I : Généralité sur les pesticides</b>	
1- Introduction	4
I- 2- Définition des pesticides	4
I.3- Les familles	5
I 3-1 Les fongicides	5
a-Les fongicides minéraux	6
b-Les fongicides organiques	6
I 3-2- Les insecticides	7
I 3-2-1- Les principales familles d'insecticides	7
a- Les organochlorés (OC)	8
b- Les organophosphorés (OP)	8
c- Les carbamates	8
d- Les pyréthrinoïdes de synthèse	8
I 3- 3- Les herbicides	9
a-Les herbicides appliqués au niveau foliaire	9
b-Les herbicides appliqués au niveau du sol	10
I 4- Effets des pesticides sur l'environnement	10
I 4-1- Effet des pesticides sur les milieux	11
I 4-1-1- La Contamination des pesticides dans le sol	11
I 4-1-2- La contamination des pesticides dans l'air	11
I 4-1-3- La Contamination des pesticides dans l'eau	12
I-5- Les propriétés physico-chimiques	13
<b>Chapitre II : Les méthodes</b>	
II-1 –Introduction	15
II- 2 –Historique	15
II- 3 - Définition	17
II- 4 –Principe	17
II - 5- Méthodologie générale d'une étude QSPR/QSAR	19

II-6-Base de données	20
II- 7-Descripteurs moléculaires	21
II-7-1 : Types de descripteurs	21
II-8- Applications des méthodes QSPR/QSAR	21
II- 9 - Introduction à la régression linéaire multiple	22
Application	
I-Calcul du modèle	28
I-1- Choix du modèle	28
I- 2- Choix d'une famille de descripteurs	30
II-Développement du modèle	31
II-1- Test de normalité:	31
II- 2- Matrice de corrélation	33
II- 3- Equation et analyse de régression	34
II- 4- Diagramme de Williams	35
II- 5- Qualité de l'ajustement	35
II- 6- Test de randomisation	36
Conclusion générale	39
Références bibliographiques	40

## Liste des tableaux

<b>Tableau I</b> : les composés étudiés et leurs solubilités.....	24
<b>Tableau II</b> : Descripteurs moléculaires intervenant dans la modélisation du (log P). .....	30
<b>Tableau III</b> : Valeurs de log S, IVDE, H6p, R5u .....	30
<b>Tableau IV</b> : Matrice de corrélations:(log Cte H; IVDE; H6p; R5u).....	33
<b>Tableau V</b> : Analyse de régression.....	34
<b>Tableau VI</b> : Les paramètres statistiques pour l'ensemble de calibration.....	34

## Liste des figures

Figure N° 0 1 : Structure chimique de l'organochlorés.....	8
Figure N° 02 : Structure commune des Organophosphorés.....	8
Figure N° 0 3 : Structure générale de carbamate.....	9
Figure N° 04 : Les insecticides de la classe des pyréthrinoïdes.....	9
Figure N° 0 5 : Processus et transfert du pesticide dans le sol.....	12
Figure N° 06: QSPR / QSAR relatives structures moléculaires avec critères d'évaluation...	15
Figure N° 07 : Modèle de l'étude de relation structure activité/propriété.....	18
Figure N° 08 : Diagramme de Williams.....	35
Figure N° 09: Qualité de l'ajustement.....	36
Figure N° 10 : Test de randomisation.....	37

## SYMBOLES ET ABREVIATION

ADEME : Agence de l'Environnement et de la Maîtrise de l'Energie.
AG : Algorithme Génétique.
AM1 : Austin Model 1.
COMFA : Comparative Molecular.
COMSIA : Comparative Molecular Similarity Indices Analysis.
$e_i$ : Résidu : différence entre les valeurs observées et estimées.
$e_{i \text{ std}}$ : Résidu de prédiction standardisé.
EQMC : Erreur quadratique moyenne sur l'ensemble de calibration.
EQMP : Erreur quadratique moyenne sur l'ensemble de prédiction.
EQMPext: Erreur quadratique moyenne sur l'ensemble de prédiction externe.
F : Statistique de Fisher.
$H_{ii}$ : Eléments diagonaux de la matrice chapeau
IFEN : Institut de Formation d'Educateur de Normandie.
KH : La constante de Henry.
Kx : La corrélation.
MLR : Multiple Linear Regression.
OC : Les organochlorés.
OCDE : Organisation de Coopération et de Développement Economique.
OP : Les organophosphorés.
PLS : Régressions aux moindres carrées partielles.
$Q^2$ : Coefficient de prédiction.
QSAR : Quantitative Structure Activity Relationship.

QSPR : Quantitative Structure-Property Relationships.
$R^2$ : Coefficient de détermination.
S : Erreur standard.
SSR : la somme des carrés de la régression.
SST: la somme totale des carrés (total sum of squares).
T : La contribution de descripteur dans le modèle.
VIF : Facteur d'inflation de la variance.
VLCT : Valeur limite d'exposition à court terme.
VME : Valeur limite moyenne d'exposition.

## ملخص :

سيكون التقدير التجريبي للخصائص دائماً مطلوباً وسيظل ضرورياً لوصف المنتجات العضوية.

في عملنا، قمنا بتصميم نموذج Henery الثابت لسلسلة من 22 مبيد حشري.

تم تقسيم سلسلة البيانات بشكل عشوائي إلى مجموعتين فرعيتين، مجموعة مكونة من 16 مركباً للبناء النموذجي ومجموعة من 6 مركبات للتحقق من الصحة.

تم اختيار الواصفات الجزيئية الهامة بواسطة الخوارزمية الجينية، باستخدام نهج QSPR

توضح المعلومات 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 la constante de Henry d'une série de 22 insecticides. La série des données a été divisé aléatoirement en deux sous-ensembles, un ensemble de 16 composés pour la construction du modèle et un ensemble de 6 composés pour la validation.

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

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  $s$ ) font ressortir la qualité et la pertinence des modèles calculés.

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

**Abstract :**

Empirical estimation of properties will always be required and will remain essential to characterize organic products.

In our work we have modeled the Henry constant of a series of 22 insecticides. The data series was randomly divided into two subsets, a set of 16 compounds for model construction and a set of 6 compounds for validation.

The selection of significant molecular descriptors was done by genetic algorithm, using a QSPR 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 models calculated.

The scope was discussed using Williams Diagram.



*Introduction*

*Générale*

## Introduction générale :

La propagation de l'industrialisation, la naissance de nouvelles technologies, l'accroissement de la population, le développement de l'agriculture et l'obligation de l'Algérie 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.

D'après l'Institut Nationale de Protection des Végétaux, plus de 480 pesticides est enregistrés en Algérie, dans le domaine de l'agriculture les autorités algériennes emploient l'expression d'usage <produits phytosanitaires à usage agricole>.

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. Le devenir des pesticides concerne tout le milieu naturel dans son ensemble (sol, eau et air) mais le sol reste un compartiment clé car une grande proportion des pesticides appliqués lors du traitement des cultures arrive au sol, par application directe et /ou par lessivage du feuillage Leur comportement dans les sols conditionne donc leurs impacts sur d'autres compartiments de l'environnement. C'est pourquoi il est crucial d'étudier les sols et le devenir des pesticides sur ces différents sols en vue de comprendre, ou mieux de prédire leur répartition ultérieure dans l'environnement et les risques de contamination des eaux [1].

Les pesticides peuvent être groupés de différentes manières: par famille chimique, par le type de formulation (liquide, granulé...), etc. Toute fois, la classification la plus courante est la classification de ces produits selon la nature de l'espèce nuisible sur laquelle ils doivent agir. On retrouve ainsi les termes: d'herbicides (contre les " mauvaises herbes "), d'insecticides (contre les insectes, leurs larves et leurs œufs), de fongicides (contre les champignons, virus, et bactéries), de molluscicides (contre les limaces et escargots), de rodenticides (contre les taupes et rongeurs), de corvicides (contre les oiseaux ravageurs), de nématicides (contre les vers), d'acaricides (contre les arachnides) [2].

Ainsi, l'objectif principal de ce travail est la Prédiction de constante de Henry d'une série des pesticides utilisant une méthode linéaire.

Ce mémoire est constitué de deux parties:

- ✓ A la suite d'une introduction générale, la première partie est une étude bibliographique qui consiste en une synthèse des connaissances sur les pesticides et aussi une étude consacré à différentes méthodes de traitement QSPR.
- ✓ La deuxième partie présente la partie application

Le mémoire est enfin complet par une conclusion générale.



*Chapitre I*

*Généralité sur les pesticides*

**I- 1- Introduction :**

Tout organisme qui interfère négativement avec les activités humains est considéré comme nuisible. Des insectes, des rongeurs, des bactéries, des champignons, des nématodes et de nombreux autres organismes entrent en compétition avec l'homme pour les ressources en nourriture.

D'autres organismes nuisibles sont vecteurs ou causes de maladies. La définition d'un organisme nuisible est sujette à caution. Un moustique peut être nuisible à l'homme mais constitue une proie pour les oiseaux ou les chauves-souris qui s'en nourrissent. L'homme a tenté et tente toujours de se débarrasser des nuisibles en réduisant la taille de leur population. Pour cela, l'utilisation de pesticides est le moyen le plus courant, particulièrement en agriculture. Les pesticides peuvent être classés en fonction des organismes-cibles qu'ils sont censés éliminer. Les insecticides éliminent les insectes, les herbicides les mauvaises herbes, les fongicides les champignons, les rodenticides les rongeurs comme rats et souris [3].

Les produits phytosanitaires ou pesticides sont par nature et définition, des molécules dotées d'une certaine toxicité. Lors d'études réalisées avant mise sur le marché, la toxicité « propre », la spécificité d'action et l'écotoxicité d'une substance sont déterminées. Ce dernier paramètre inclut la dégradation de la substance dans les différents compartiments de l'environnement, ainsi que l'évaluation des risques vis-à-vis de l'environnement. Afin d'évaluer ces risques, on fait appel à des « descripteurs » physico-chimiques. Dans le cas du risque de volatilisation, la constante de Henry, décrivant le partage entre les phases aqueuses environnementales et les phases gazeuses de l'atmosphère, ainsi que la pression de vapeur sont utilisées [4].

Les pesticides sont utilisés depuis plus d'un siècle, et on peut naïvement penser qu'il existe quelque intérêt à leur usage. Pour s'en convaincre, il suffit de donner une liste de leurs diverses applications qui concernent. D'ailleurs, une grande variété de domaines relatifs à plusieurs aspects des activités humaines. On peut assez commodément envisager trois groupes d'intérêts tout en observant que cette classification peut toujours être discutée bien qu'elle puisse être utile. On distinguera donc la protection des végétaux et de leurs produits. Les aspects sanitaires liés aux organismes nuisibles et l'entretien des espaces non agricoles [5].

**I- 2- Définition des pesticides :**

Les pesticides sont, en terme générique utilisé pour désigner toutes les substances naturelles ou de synthèse capables de contrôler, d'attirer, de repousser, de détruire ou de s'opposer au développement des organismes vivants (microbes, animaux ou végétaux) considérés comme indésirables pour l'agriculture, l'hygiène publique (par exemple les cafards dans les habitations), la santé publique (les insectes parasites (poux, puces) ou vecteurs de maladies telles que le paludisme et les bactéries pathogènes de l'eau détruites par la chloration, la santé vétérinaire, ou les surfaces non-agricoles (routes, aéroports, voies ferrées, réseaux électriques...), ils sont l'un des rares substances qui sont à la fois toxiques et délibérément rejetés dans l'environnement [6].

Les pesticides sont présents dans presque tous les milieux de l'environnement y compris les eaux de surface, les eaux souterraines, l'air ambiant, la poussière, le sol, le brouillard, la pluie, et la glace [6].

Les pesticides sont diffusés partout, car les résidus de pesticides ont été trouvés dans plus de 70 % des fruits et légumes, plus de 60 % des échantillons de blé et de 99% pour cent de lait, Ils sont également présents dans les organes de presque tous les adultes et les enfants [6].

Le terme « pesticide » couvre un champ plus vaste et général que les expressions « produit phytosanitaire » ou « produit phytopharmaceutique » car il englobe tous les produits destinés à lutter contre tous les dits nuisibles, ou indésirables (ex : les champignons qui pourraient attaquer une charpente) et les médicaments vétérinaires destinés à protéger les animaux domestiques, gibiers ou de compagnie (par exemple, le collier antipuces pour chien) [6].

**I.3- Les familles :**

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) [7].

Les principales familles de pesticides utilisées en agriculture fruitière et légumière sont les fongicides, les herbicides et les insecticides [8].

**I 3-1 Les fongicides :**

Ils sont destinés à traiter les maladies fongiques des plantes, également les maladies bactériennes et virales [7].

Ils sont utilisés pour lutter contre des champignons pathogènes pouvant infecter les plantes cultivées. Ils affecteront les processus respiratoires, pourront agir sur les microtubules (carbamates) [3].

On distingue deux grands groupes de fongicides : les fongicides minéraux et les fongicides organiques qui sont majoritairement des produits de synthèse [10].

**a- Les fongicides minéraux :**

Parmi les fongicides minéraux on trouve :

- Les fongicides à base de cuivre : Ce sont les plus nombreux et les plus fréquents.
- Les fongicides à base de soufre : La qualité des soufres employés a une importance dans l'efficacité du traitement. Il existe 3 types de fongicides soufrés :
- Les soufres triturés : utilisés sous forme de poudre.
- Les soufres sublimés : obtenus par condensation des vapeurs et employés en poudrage et les soufres mouillables utilisés en pulvérisations qui contiennent 80% de soufre à l'état pur et qui sont rendus miscibles à l'eau par l'adjonction d'agents mouillants.

Le plus connu est la bouillie bordelaise, c'est un mélange de sulfate de cuivre, de chaux et d'eau. Il existe également des préparations à base d'oxychlorure de cuivre.

- Les fongicides à base de permanganate de potassium : Ils agissent uniquement à titre curatif sur les oïdiums qu'ils détruisent par contact [8].

**b- Les fongicides organiques :**

Ils se sont rapidement développés après leurs apparitions dans les marchés en 1950. Très efficaces, ils possèdent un large spectre d'action. Les fongicides organiques représentent un groupe important de molécules dont la structure chimique est variée, et parmi les principaux fongicides on trouve 1: [1]

- Les carbamates
- les dérivés de l'acide carbamique et benzimidazolés

A partir des années 1970, l'introduction du groupe des benzimidazolés a révolutionné le traitement de nombreuses maladies des plantes. Ce groupe comprend le bénomyl, le thiophanate-méthyl et le carbendazime.

- les dérivés de l'acide thiocarbamique et thiocarbamique constituent une famille chimique très importante du point de vue phytosanitaire.

Ces fongicides ont en commun leur absence totale de phytotoxicité H, une polyvalence assez grande et une faible écotoxicité. On trouve dans cette famille : le mancozèbe, Le manèbe, le propinèbe, le zinèbe et le zirame.

Les dithiocarbamates ont une action inhibitrice sur la respiration de champignons. Ils agiraient également par le biais d'espèces réactives de l'oxygène à l'origine d'un stress oxydant.

- Les dérivés du phénol Ce groupe chimique, proche des dérivés du benzène est assez restreint. Son principal représentant est le dinocap qui agit sur la respiration. Le dinocap est phytotoxique si la température est supérieure à 35°C
- Les dicarboximides : Ces molécules ont une action sur la respiration du champignon. Les principaux représentants sont le captane et le folpel.
- Les amides et amines : Les fongicides de cette classe ont une action sur les Basidiomycètes.

Ils inhibent la respiration du champignon par inhibition de la succinate déhydrogénase.

- Les inhibiteurs de la biosynthèse des stérols H : Cette classe de fongicides agit sur les cellules du champignon en inhibant la synthèse des stérols. Ils ont un effet sur les attaques dues aux ascomycètes, aux basidiomycètes et aux champignons imparfaits mais ils n'ont pas d'activité sur les Oomycètes.

Ils peuvent être utilisés lors de phénomènes de résistances H aux benzimidazolés.

- Les anilinopyrimidines : Ce groupe comprend le cyprodinil et le pyriméthanil utilisé contre le développement de la pourriture grise au niveau des parties aériennes du fraisier.
- Les méthoxyacrylate et fongicides apparentés : Ces produits ont été développés à partir de substances naturelles secrétées par des champignons de forêts. Le mode d'action de ces antifongiques est l'inhibition de la chaîne respiratoire du cytochrome b/c1. Ce groupe est composé de l'azoxystrobine, du krésoxim-méthyl, de la picoxystrobine et de la trifloxystrobine [8].

### I 3-2- Les insecticides :

Les insecticides sont des substances actives ayant la propriété de tuer les insectes, leurs larves et/ou leurs œufs. Les insecticides organiques de synthèse sont des molécules carbonées, synthétisées, et se distinguent des insecticides inorganiques ou minéraux [10].

Sont les premiers pesticides utilisés et les plus utilisées en Algérie. Ils sont destinés à détruire les insectes nuisibles [1].

#### I 3-2-1- Les principales familles d'insecticides :

Les insecticides peuvent être classés selon leur structure chimique ou leur mode d'action [3].

Les trois plus grandes familles auxquelles appartiennent les insecticides organiques de synthèse sont : [8]

- 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 [11].

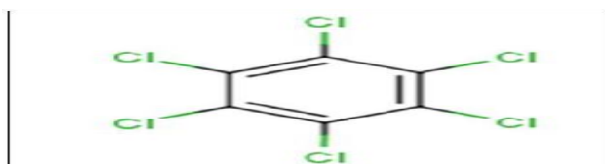


Figure N °0 1: Structure chimique de l'organochlorés.

- b- Les organophosphorés (OP):** Contient du phosphore. Ils ont été largement utilisés dès la seconde Guerre mondiale. Ils sont hautement toxiques pour les poissons, les abeilles et l'homme. ILS sont moins persistants que les organochlorés et les ont remplacés pour des usages agricoles [3].

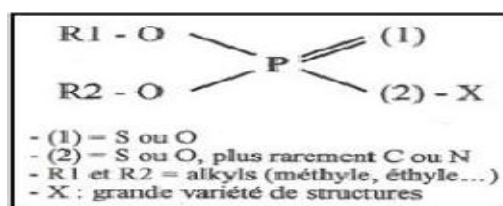
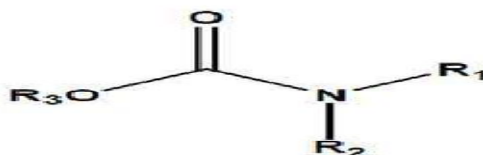


Figure N °0 2: Structure commune des Organophosphorés.

**c- Les carbamates:** Sont des dérivés de l'acide carbamique ; ils ont aussi un large spectre d'action mais sont moins toxiques que les précédents [3].

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.

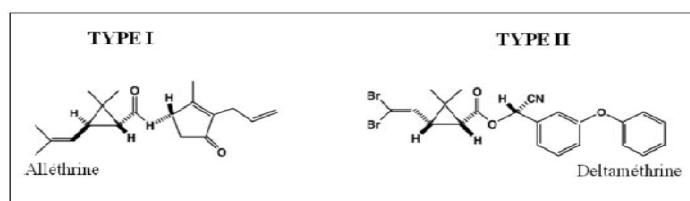
Le carbofuran, l'aldicarbe, le carbofuran ou encore le fénoxycarbe sont également des carbamates largement utilisés [9].



**Figure N ° 03:** Structure générale de carbamate

**d- Les pyréthrinoïdes de synthèse :**( deltaméthrine, perméthrine, esfenvalérate) ont une toxicité que les familles précédentes [3].

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 [8].



**Figure N ° 0 4 :** Les insecticides de la classe des pyréthrinoïdes.

### I 3- 3- Les herbicides :

Les herbicides sont des substances chargées de ralentir la croissance ou de détruire les plantes cibles, nommées adventices ou mauvaises herbes. Les plantes adventices sont considérées comme ennemis des cultures car elles entrent en compétition avec la culture elle-même pour la ressource organique et minérale du sol, l'eau, l'espace et la lumière. Cette définition inclue également les repousses de cultures (repousses de colza, pommes de terre, tournesols, ...) à l'endroit où elles sont indésirables, c'est à dire dans les cultures qui suivent. Les herbicides peuvent agir dans le sol au niveau des racines ou Directement sur feuilles [10].

Ce sont les plus utilisées dans le monde en tonnage et en surface ; ils permettent d'éliminer les mauvaises herbes des cultures [11].

Par souci de simplicité, les herbicides seront traités en fonction de leur mode d'application et de leur mode d'action [8].

### **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 les herbicides qui altèrent la synthèse d'acides aminés aromatiques 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 : Les bypyridilium et les diphenyl esters sont les deux principales familles d'herbicides qui altèrent la membrane Cellulaire.
- Les inhibiteurs de la photosynthèse : Les herbicides de la famille des triazines et des phénylurées agissent en interférant avec la photosynthèse [9].

### **a- Les herbicides appliqués au niveau du sol**

- les inhibiteurs de la division cellulaire : Certains herbicides de la famille des dinitroanilines (dinitrobenzenamines) comme la trifluraline, la prodiamine ou la pendiméthaline agit en inhibant les étapes de division cellulaire responsables de la séparation des chromosomes H et de la formation de la paroi cellulaire au niveau des racines de la plante. Les conséquences étant un nombre de racines trop faible pour pouvoir correctement assurer la nutrition de la plante.

D'autres herbicides de la famille des thiocarbamates (butylate, cycloate ...) et de la famille des amides substituées (alachlore, métolachlore...) agissent en interférant avec la division cellulaire des tissus méristématiques, ils agissent donc sur les plantules après absorption racinaire et empêchent l'évolution vers la plante adulte

- les destructeurs de pigments : Le clomazone ou la norflurazon agissent sur les plantes en détruisant la chlorophylle, ce qui empêche ensuite la plante de pouvoir réaliser correctement la photosynthèse [8].

#### **I- 4- 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 [12].

Les pesticides ont à l'origine un rôle sanitaire contre les organismes nuisibles. Cependant, les pesticides sont transportés dans tous les compartiments environnementaux (air, eaux et sols) induisant des effets immédiats ou à long terme néfastes pour l'homme ainsi que pour l'environnement [13].

#### **I 4-1- 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 [14].

#### **I 4-1-1- La Contamination des pesticides dans le sol :**

Le sol est un milieu hétérogène, complexe, variable dans ses constituants physiques, chimiques et biologique. Il est composé de trois phases : solide, liquide, gazeuse [1].

Le sol est un matériau à la fois minérale et organique. La partie minérale qui représente la fraction la plus importante environ : 80% à 95% en masse et constituée de minéraux primaires (issus de l'altération du substrat géologique sous l'action conjuguée de la température, de l'air et de l'eau) et des minéraux secondaires (produit d'altération) comme les argiles, les oxydes, les hydroxydes dont dépendent les propriétés physico-chimiques du sol [15].

Le sol joue un rôle fondamental dans le devenir des pesticides qui peuvent être appliqués [9].

Apporté dans le sol, un pesticide peut être retenu par les minéraux et la matière organique, transporté dans l'eau et dans l'air et transformé à des degrés divers jusqu'à sa complète dégradation. Ces processus contribuent au devenir des pesticides et sont le résultat

de plusieurs phénomènes qui interviennent simultanément ou successivement. A leur tour, ces phénomènes mettent en jeu plusieurs propriétés de façon plus ou moins complexes selon leur nombre et leurs interactions mutuelles. Le devenir d'un pesticide dans le sol dépend de quatre facteurs généraux : le climat, le sol, les propriétés moléculaires et les pratiques agricoles [5].

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 [9].

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

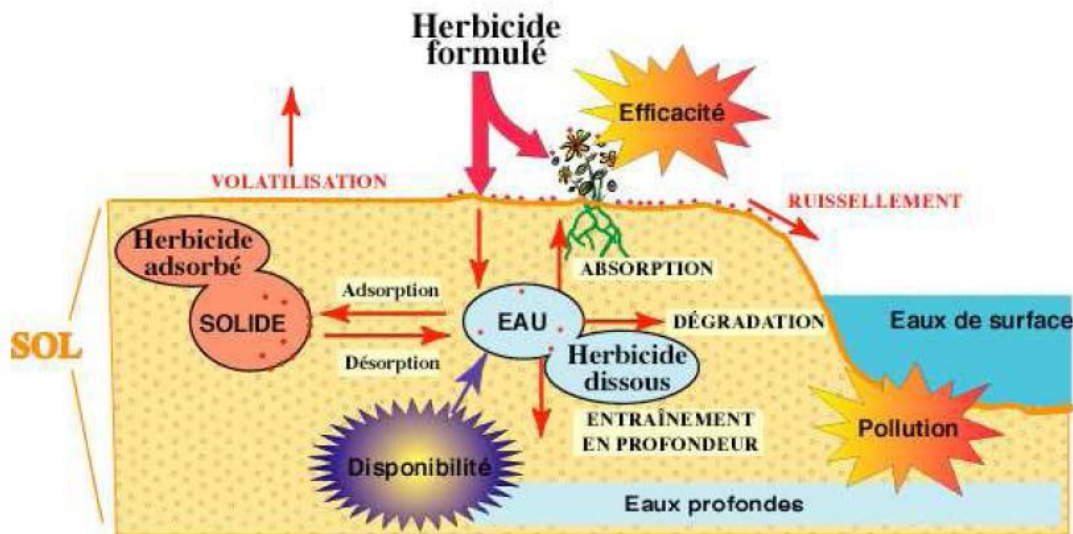


Figure N° 05 : Processus et transfert des pesticides dans les sols

#### I 4-1-2- La contamination des pesticides dans l'air :

L'air est un compartiment assez mal connu. Ce sont pourtant quelques  $15 \text{ m}^3$  qui transitent dans nos poumons chaque jour, sans oublier la respiration cutanée. Les mesures réalisées en France dans l'atmosphère depuis 2000 indiquent que les pesticides sont présents aussi bien dans les zones agricoles qu'en milieu urbain. Des biocides sont également détectés en zone urbaine.

En l'absence de réglementation relative à la présence de pesticides dans l'air que nous respirons, des limites indicatives de concentration sont définies. Ces valeurs limites sont établies à partir d'informations fournies par les tests toxicologiques des substances.

Il existe deux types de valeurs en France :

- La VLCT (valeur limite d'exposition à court terme): Ce sont les valeurs mesurées sur une durée maximale de 15 minutes. Elles sont établies pour prévenir les risques toxiques immédiats ou à court terme ;

- La VME (valeur limite moyenne d'exposition) : Mesurée ou estimée sur la durée moyenne d'un poste de travail de 8 heures, elles sont indicatives pour la protection des travailleurs à moyen ou long terme [2].

### **I 4-1-3- La Contamination des pesticides dans 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 [16].

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 [17].

Les eaux continentales sont les milieux pour lesquels les données sont les plus nombreuses, et font l'objet d'une compilation annuelle par l'IFEN. Ces données mettent en évidence une contamination quasi-généralisée des eaux de surface et des eaux souterraines par les pesticides, et la prépondérance des herbicides parmi les molécules les plus fréquemment détectées. Une contamination significative peut être générée par des pertes en pesticides très faible : une fuite de moins de 1/1000e de la masse d'herbicide épanchée sur une parcelle peut suffire par exemple pour contaminer l'eau qui s'en écoule au-dessus du seuil de potabilité.

Les données collectées ne permettent toutefois pas de quantifier avec précision les niveaux de contamination, ou de calculer l'exposition des organismes. Elles sont en effet souvent :

- très hétérogènes et difficilement comparables, car produites par des réseaux de suivi divers (ayant des objectifs différents: surveillance de l'eau potable, suivi de bassin versant...), avec des méthodes analytiques et des listes de molécules recherchées différentes... .
- non représentatives, car les prélèvements sont peu fréquents, et ne permettent notamment pas la détection des pics de pollution notamment dans le cadre de projets soutenus par l'ADEME, mais les données restent fragmentaires (campagnes de suivi ponctuelles dans le temps et dans l'espace) et les listes de molécules suivies sont limitées [2].

### **I-5- Les propriétés physico-chimiques :**

Les pesticides possèdent des caractères physico-chimiques qui favorisent ou non leur passage dans l'air [18].

De manière générale, la distribution d'une substance entre les différents compartiments de l'environnement : l'air, l'eau, les sols et les êtres vivants dépend du composé en lui-même (solubilité, pression de vapeur) mais également du milieu (température de l'air, de l'eau, structure du sol). Cette distribution permet de nous informer sur la mobilité de la substance dans l'hydrosphère, la pédosphère et l'atmosphère, afin de déterminer son impact sur l'environnement [19].

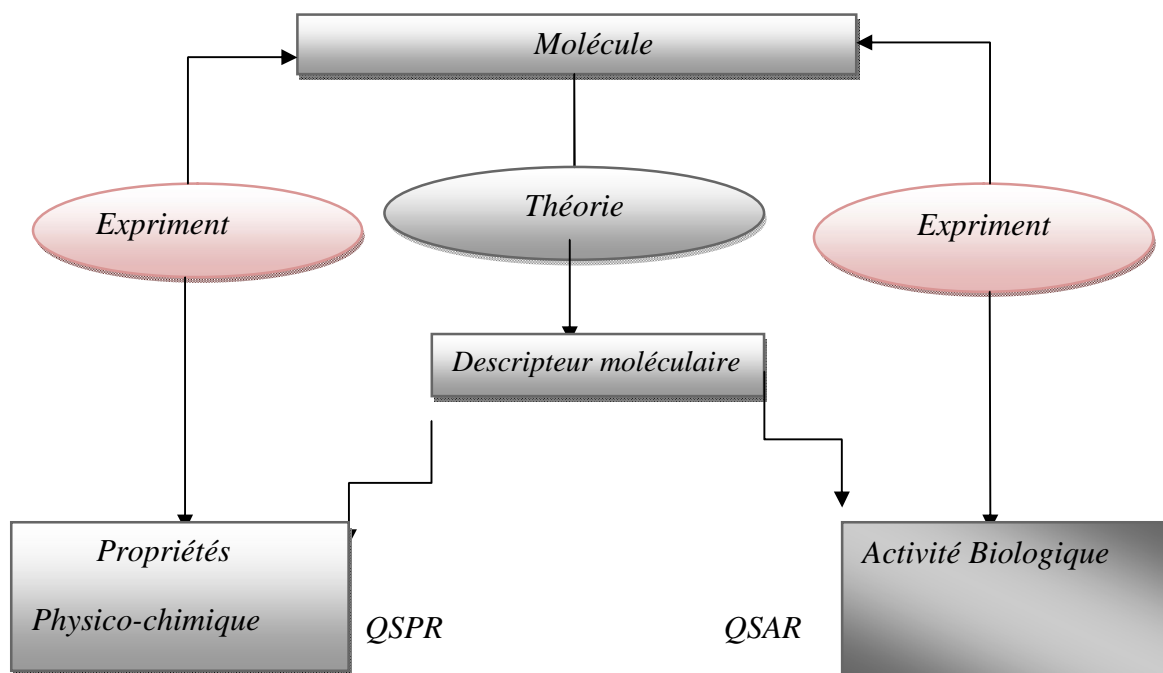
Les principaux facteurs physiques qui influencent leur transfert dans l'atmosphère sont la constante de Henry, la pression de vapeur et la demi-vie dans l'air [18].

Ces facteurs chimiques vont soit influencer la volatilisation des pesticides vers l'atmosphère, soit leur déposition. Ces paramètres nous permettent également de prédire leur devenir et leur répartition entre les différentes phases de l'atmosphère.

Tous les phénomènes de transfert des pesticides vont avoir une influence dépendante des propriétés physico-chimiques des composés. La pression de vapeur, la solubilité et le coefficient d'adsorption sont les caractéristiques chimiques qui vont le plus influencer sur la volatilisation [20].

## II-1 –Introduction :

La connaissance des propriétés et des activités est d'une importance capitale pour pouvoir classer et utiliser les composés chimiques. La caractérisation expérimentale complète est difficile, voire impossible, pour des raisons de temps, de coût, de dangerosité de certains essais ou d'éthique (limitations des essais sur les animaux). L'utilisation des méthodes alternatives à l'expérience est devenue plus qu'indispensable. Parmi ces méthodes, on trouve les méthodes de modélisation moléculaire qui permettent de justifier les données expérimentales disponibles et prédire les propriétés/activités pour des composés nouveaux ou des composés pour lesquels les données expérimentales ne sont pas disponibles. Parmi ces méthodes de modélisation les plus utilisées, on peut citer les méthodes QSPR (Quantitative Structure-Property Relationships) et QSAR (Quantitative Structure-Activity Relationships). Ces méthodes s'appuient sur le principe que les propriétés physico-chimiques et les activités biologiques des molécules dépendent fortement de leurs structures chimiques [21].



**Figure N°06:** QSPR / QSAR relatives structures moléculaires avec critères d'évaluation

## II- 2 -Historique

Il y a plus d'un siècle et demi, en 1863, Crois a observé que le point d'ébullition et le point de fusion des alcanes augmente avec le nombre d'atomes de carbone et la masse moléculaire. Il a observé également une diminution de la solubilité dans l'eau des alcools avec

l'augmentation du nombre d'atomes de carbone et la masse moléculaire, cela est considéré depuis comme la première formulation générale en RQSP [22].

Cinq ans après, En fait, les premiers travaux QSPR/QSAR remontent au 19<sup>ème</sup> siècle. En effet, dès 1868, Crum-Brown et Fraser [4] ont postulé l'existence de relations entre les activités physiologiques et les structures chimiques en reliant les variations de l'activité biologique à des modifications structurales. Cependant, à cette époque, les structures

A la fin du 19<sup>ème</sup> siècle, Meyer en 1899 et Overton en 1901, ont indépendamment observé « une relation linéaire entre l'activité des narcotiques et leur coefficient de partage

Six ans après, en 1907, Fühner et Neubauer ont montré pour une série de narcotiques homologues, que l'activité augmentait en fonction de la progression géométrique de la série de composés, ceci montrant l'importance de la contribution d'additivité de groupements fonctionnels pour l'activité biologique [22].

En 1962, Hansen [10] a montré l'existence d'une corrélation entre la toxicité des acides benzoïques substitués et les constantes électroniques «  $\sigma$  » des substituants [22].

L'année 1964 est considérée comme le début des méthodes RQSA modernes. Hansch et Fujita ont établi les premières corrélations entre les propriétés physico-chimiques (log P, pKa, paramètres stériques et électroniques) et l'activité biologique (activité enzymatique, pharmacologique), Ces méthodes seront appelées par la suite l'analyse de Hansch et l'analyse de Free Wilson) [23-24].

Sept ans plus tard, Hansch et Lien ont réalisé une étude RQSA sur différentes familles d'antifongiques : benzoquinones, sels d'alkylpyridinium, imidazoles et phénols. Ils ont observé que quels que soient la famille et le champignon utilisé, l'activité antifongique dépend du coefficient de partage Eau-Octanol, expérimental ou calculé [25].

Ces études ont été extrapolées aux techniques séparatives en corrélant les propriétés physico-chimiques des analytes avec les temps de rétention obtenus expérimentalement : c'est quantitative des relations structure temps de rétention noté RQSR [26].

Maintenant, des méthodes 3D comme l'étude CoMFA (Comparative Molecular Field Analysis) et CoMSIA (Comparative Molecular Similarity Indices Analysis) [27-28] permettent de traiter les relations structure-activité en trois dimensions, 3D-RQSA/RQSP.

Au cours de ces dernières décennies, l'utilisation des méthodes QSPR/QSAR n'a pas cessé de progresser. Elle est même devenue indispensable en chimie pharmaceutique, en toxicologie et pour la conception de médicaments [21].

### II- 3 - Définition :

Une relation QSPR/QSAR est un modèle ou une formule mathématique qui permet de relier, d'une manière quantitative, la structure d'une molécule à une propriété ou à une activité donnée. Les méthodes QSPR/QSAR sont de plus en plus utilisées, du fait de la croissance des moyens de calculs (machines, logiciels,...) [21].

Les méthodes QSAR/QSPR sont basées sur l'hypothèse que l'activité ou la propriété d'un composé chimique est liée à sa structure, plus précisément cette approche affirme que l'activité (ou la propriété) et la structure d'un composé chimique sont liées d'un certain algorithme mathématique, cela est basé sur le postulat de base « les composés chimiques similaires ont des activités similaires ». De plus, lorsque les paramètres moléculaires sont exprimés par des chiffres, on peut proposer une relation mathématique, ou relation quantitative structure activité/propriété, entre les deux.

Par définition, Une RQSA/RQSP est un modèle mathématique qui associe un ou plusieurs paramètres quantitatifs dérivés de la structure chimique, à une mesure quantitative d'une propriété ou d'une activité [22].

### II- 4 –Principe :

Le principe d'une étude RQSA/RQSP (Figure 07), consiste à trouver une relation mathématique reliant de manière quantitative une activité biologique, ou une propriété, mesurée pour une série de composés similaires dans les mêmes conditions expérimentales, avec des descripteurs moléculaires à l'aide des méthodes statistiques.

L'objectif de ces études est d'analyser les données structurales afin de détecter les facteurs déterminants pour l'activité ou la propriété étudiée. Pour ce faire, différents types de méthodes statistiques peuvent être employées (voir plus loin : les méthodes statistiques).

L'expression mathématique obtenue peut alors être utilisée comme moyen prédictif de l'activité/propriété étudiée pour de nouvelles molécules ou des molécules pour lesquelles les données expérimentales ne sont pas disponibles [22].

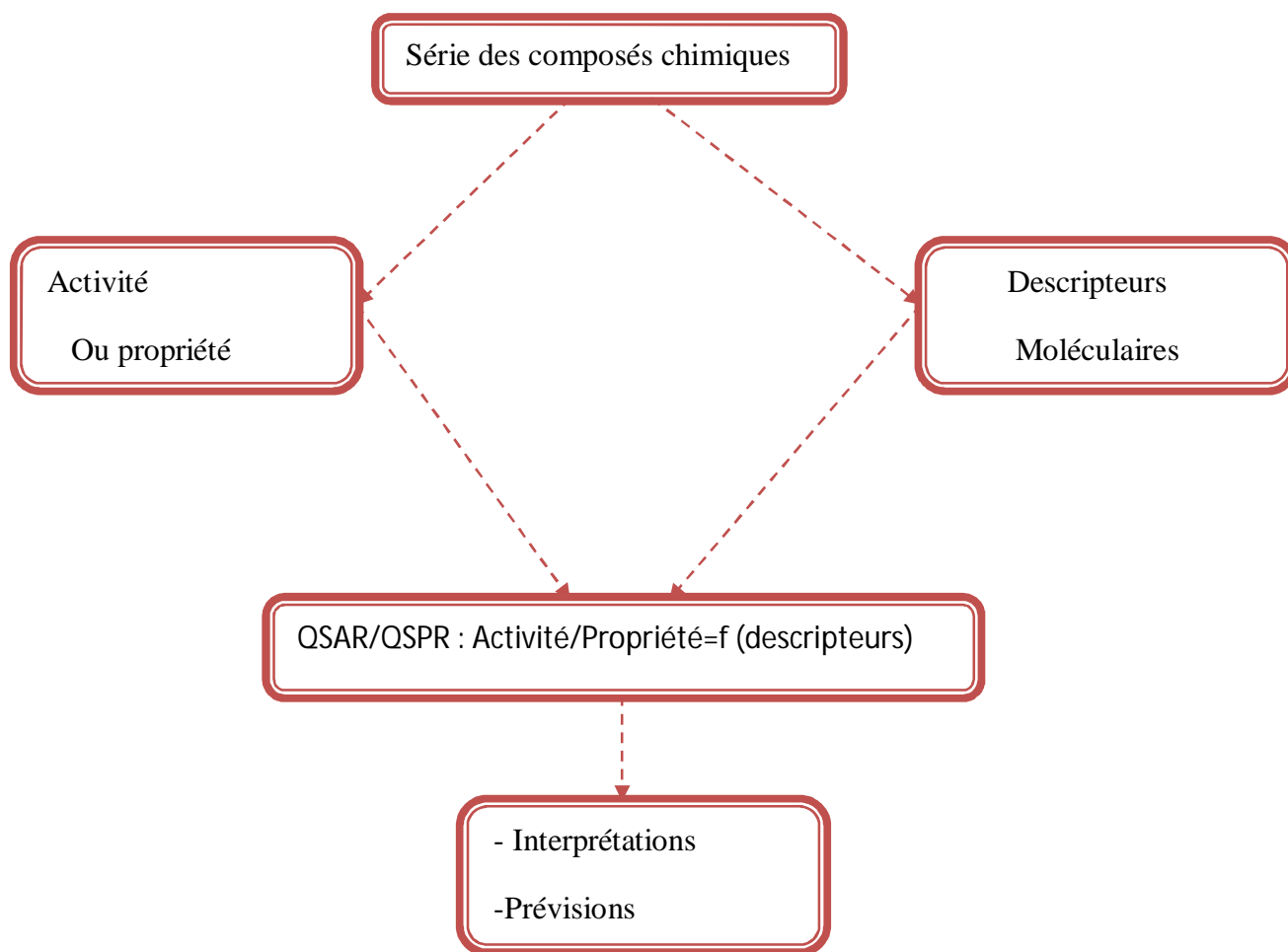
Ceci peut être traduit par l'équation suivante :

$$\text{Activité/Propriété} = f(\text{descripteurs moléculaires}) \dots\dots\dots(01)$$

Ou :

$$\text{Propriété/ Activité} = f(D1, D2, \dots Dn, \dots\dots\dots (02)$$

On a : D1, D2, ...Dn sont des descripteurs des structures moléculaires.



**Figure N° 07:**Modèle de l'étude de relation structure activité/propriété

L'objectif d'une telle méthode est d'analyser les données structurales afin de détecter les facteurs déterminants pour la propriété /activité mesurée. Pour ce faire, différents types d'outils statistiques peuvent être employés :

- Régressions linéaires simples et multiples [29].
- Régressions aux moindres carrées partielles (PLS) [30].
- Arbres de décision [31].
- Réseaux de neurones [32-33].
- Algorithmes génétiques [34].
- Vecteurs Machines [33].

Une fois cette relation est établie et validée, elle peut alors être employée pour la prédiction de la propriété /activité de nouvelles molécules, pour lesquelles les valeurs expérimentales ne sont pas disponibles. De tels modèles peuvent être également utilisés pour mieux comprendre les mécanismes et les modes d'action [21].

## II- 5- Méthodologie générale d'une étude QSPR/QSAR :

Le développement d'un modèle débute par la recherche du maximum possible des données expérimentales fiables. Ensuite, le développement d'une série de descripteurs qui caractérisent les structures moléculaires des composés de la base de données en vue de les relier à l'activité/propriété expérimentale étudiée. Une fois développé, le modèle doit être validé en termes de corrélation (sur le jeu de données d'entraînement). L'influence des composés du jeu d'entraînement sur le modèle (robustesse du modèle) est estimée par des méthodes de validation interne. Pour estimer le pouvoir prédictif du modèle, il est nécessaire de disposer de données expérimentales supplémentaires (jeu de données de validation externe) afin de déterminer la capacité du modèle à prédire ces valeurs. Enfin, pour tout modèle, il est important de savoir pour quel type de molécules il est utilisable ou non, c'est-à-dire connaître son domaine d'applicabilité.

Un modèle RQSA/RQSP relie, d'une manière qualitative ou quantitative, la structure des molécules à une activité ou propriété donnée. La stratégie de développement de tels modèles, en respectant les cinq règles mises en place par l'OCDE (Organisation de Coopération et de Développement Economique) pour la validation des modèles RQSA/RQSP (voir plus loin : les principes OECD de validité des modèles RQSA/RQSP) [24].

La méthodologie générale d'une étude QSAR/QSPR est la suivant : [35]

- a- Constituer une base de données à partir des mesures expérimentales fiables de la propriété ou de l'activité de chaque composé.
- b- Sélectionner les descripteurs en relation avec la propriété ou l'activité étudiée
- c- Diviser cette base de données, aléatoirement, en une série d'apprentissage (training set) qui contient généralement les 2/3 de la base de données et une série de test (test set) constituée par le 1/3 restant.
- d- Etablir des modèles mathématiques en utilisant la série d'apprentissage.
- e- Caractériser les modèles élaborés par leurs indices de validation internes et vérifier leur robustesse par un test de hasardisation (Randomisation) de la variable dépendante Y (réponse).
- f- Valider les modèles élaborés en utilisant la série de test et calculer leurs paramètres statistiques de validation externe.
- g- Elaborer le domaine d'applicabilité du modèle retenu.

h- Explorer et exploiter les modèles validés pour comprendre les mécanismes et les modes d'action.

## **II-6-Base de données :**

Le choix de la base de données expérimentale initiale est une étape critique pour le développement des modèles RQSA/RQSP. Généralement, les composés testés ont deux origines possibles (dans la plupart des cas sont issues de la littérature), soit des produits de synthèse ou bien des produits d'extraction à partir de plantes. Quelle que soit son origine, il arrive qu'un échantillon ne soit pas pur mais corresponde à un mélange racémique. Le résultat du test d'un tel échantillon pose problème : il est impossible de savoir quelle est la contribution de chaque énantiomère dans l'activité observée. Les structures dont la propriété étudiée est mesurée sur un mélange racémique ne peuvent pas être utilisées dans les études RQSA/RQSP.

Pour être de qualité, une base de données doit être composée de données expérimentales fiables, puisque les barres d'erreurs sur celles-ci se propageront dans le modèle final. Il est donc important de choisir des données présentant de faibles incertitudes afin de limiter les barres d'erreur expérimentales. D'autre part, l'homogénéité des données est fondamentale. Si l'on veut comparer l'activité/propriété d'une série de molécules, il faut s'assurer, si cela est possible, qu'elle est le résultat de leur interaction avec une seule et même cible et plus précisément avec le même site actif, et l'activité doit être mesurée par un seul et même test, avec des conditions expérimentales identiques pour chaque molécule.

En fin, la diversité des structures est un facteur important dans la qualité des modèles construits, elle définit l'espace chimique que l'analyse va couvrir [22].

## **II- 7-Descripteurs moléculaires :**

Un descripteur moléculaire est un paramètre (une valeur numérique) propre à une structure chimique donnée. Ces valeurs peuvent être obtenues expérimentalement ou calculées à partir de la structure de la molécule. Les descripteurs calculés, permettent d'effectuer des prédictions sans avoir à synthétiser les molécules, ce qui est l'un des objectifs de la modélisation moléculaire.

Les descripteurs moléculaires jouent un rôle fondamental dans les études de la relation quantitative structure activité/propriété. Ils sont utilisés en tant que variables indépendantes pour prédire une variable dépendante (activité ou propriété).

L'utilisation des descripteurs moléculaires dans le développement de modèles RQSA/RQSP n'est pas une tâche aisée. Tout d'abord, un très grand nombre de descripteurs moléculaires, de différentes complexités et de conceptions diverses a été introduit au cours des dernières années. Ensuite, pendant ce temps, aucune règle stricte n'a été établie pour la sélection de descripteurs adaptés parmi le grand nombre de descripteurs disponibles. Ce choix a souvent été basé sur l'intuition chimique des chercheurs, ou en se pliant à la tradition [22].

### **II-7-1 : Types de descripteurs :**

L'importance du nombre des descripteurs (plus de 6000 descripteurs répertoriés) pouvant décrire une molécule rend toute classification ou présentation de ces descripteurs non exhaustive. Dans ce qui suit, nous allons présenter que les descripteurs moléculaires les plus utilisés et ceux qui ont été utilisés dans l'ensemble de nos travaux, en commençant par les descripteurs les plus simples, qui nécessitent peu de connaissances sur la structure moléculaire, Nous verrons ensuite comment les progrès de la modélisation moléculaire ont permis d'accéder à la structure 3D de la molécule, et de calculer des descripteurs à partir de cette structure. Historiquement, deux grands schémas pour la classification des descripteurs moléculaires ont été établis : l'un en fonction de leur origine (constitutionnel, topologique, géométrique, quantique, thermodynamique...), et un autre sur leur dimensionnalité (1D, 2D, 3D ou 4D) [36].

### **II-8- Applications des méthodes QSPR/QSAR :**

Les applications des méthodes QSPR/QSAR sont très nombreuses, elles touchent tous les domaines où la structure chimique intervient, entre autre on peut citer:

- Propriétés physico-chimiques: point d'ébullition, Point de fusion, densité, indice de réfraction, température critique, viscosité, solubilité, pression de vapeur, tension superficielle, Coefficients de partition: eau/Octanol, air/eau, huile/air, lait/plasma [37].
- Autres propriétés/activités : [36]
  - Prédiction de la toxicité aquatique des composés chimiques vis-à-vis des espèces environnementales.
  - Toxicité des nanoparticules
  - Toxicité des pesticides et des colorants
  - Propriétés inhibitrices de corrosion

- Concentration micellaire critique
- Prédiction de plusieurs propriétés dangereuses telle que l'explosibilité et l'inflammabilité de certaines familles de molécules chimiques.
- Conception des médicaments et de nombreux autres produits tels que les agents tension- actifs, parfums, les colorants et les produits de la chimie fine.

## **II- 9 - 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 [38].

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 [39].

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).



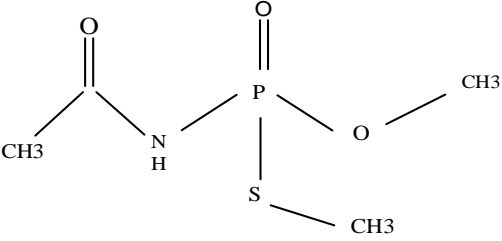
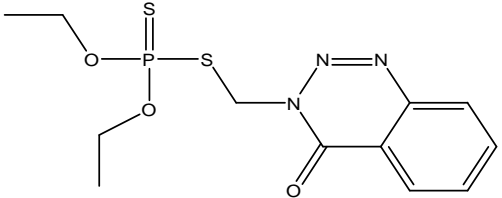
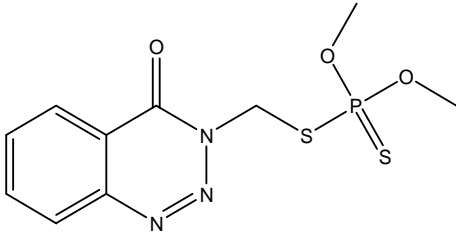
Application

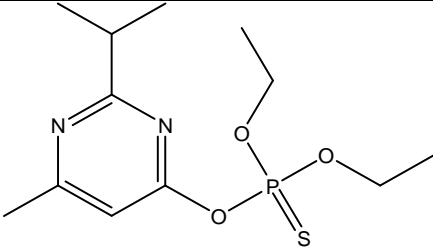
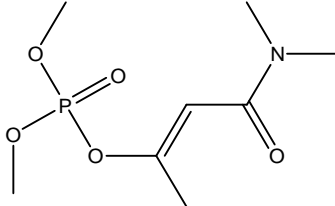
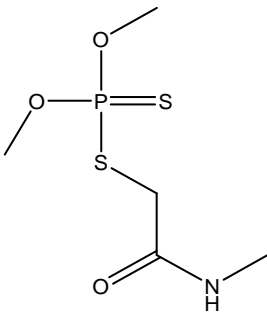
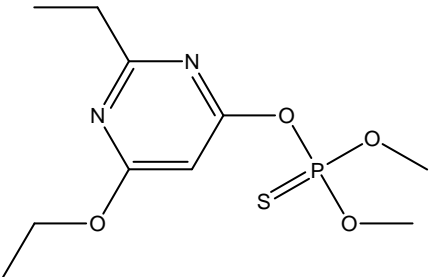
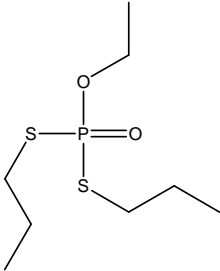
Dans cette partie nous avons modélisés la constante de Henry d'une série d'insecticides. L'ensemble des données est divisé en deux sous-ensembles, le premier contient 16 composés pour le calibrage et le second contient 6 composés pour la validation externe.

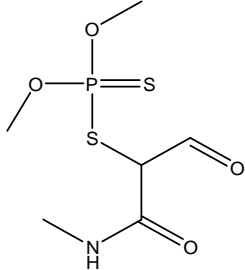
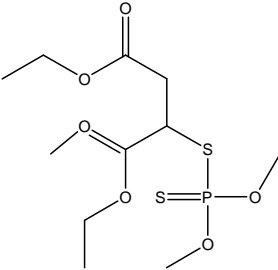
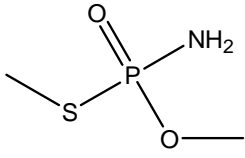
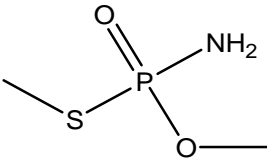
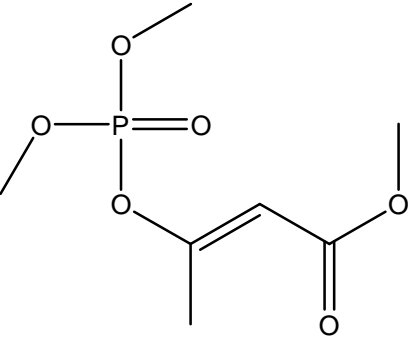
Le meilleur modèle obtenu il est validé par l'approche hybride algorithmique génétique/régression linéaire multiple (AG/MLR)

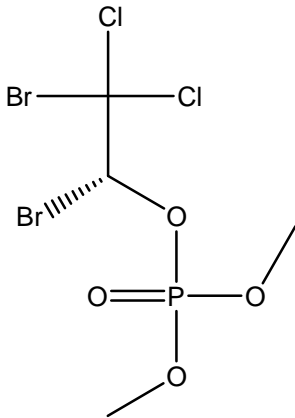
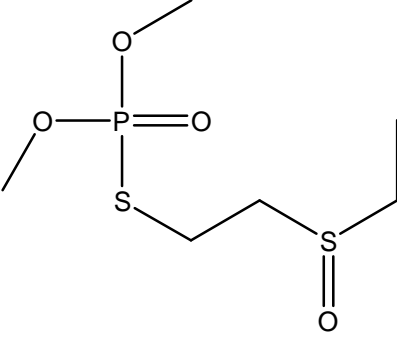
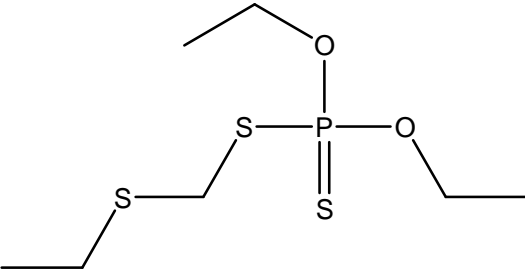
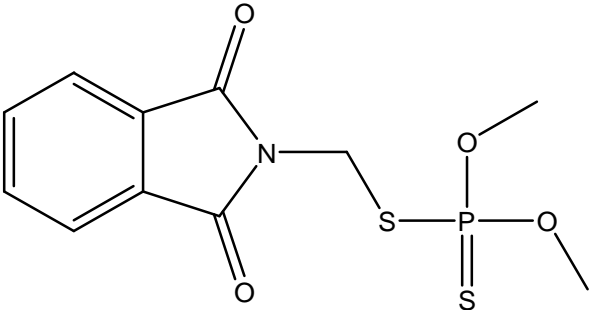
La liste des molécules étudiées dans notre travail présentée dans le Tableau I, cette base des données contient 22 composés, avec la propriété à étudier.

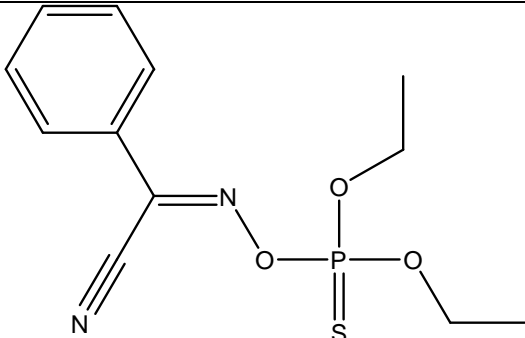
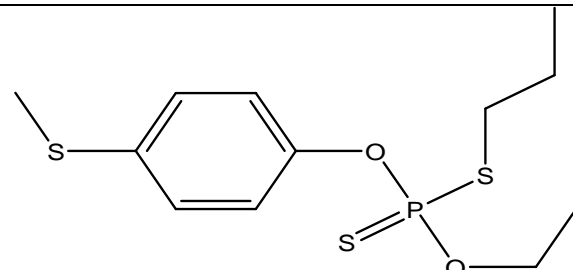
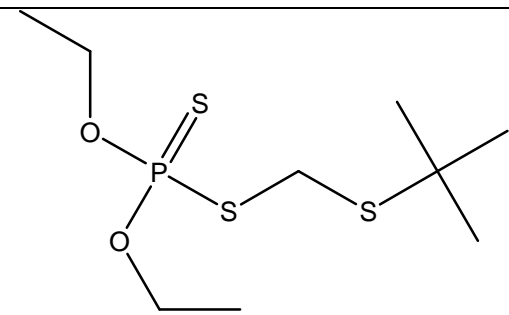
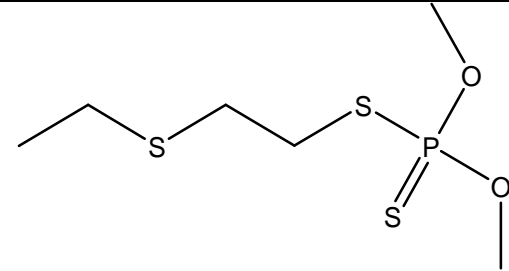
**Tableau I : les composés étudiés et leurs solubilités**

N°	Composé	Structure	Log S
1	Acéphate		0
2	azinephos-ethyl*		3.07
3	azinephos-methyl*		2.38

N°	Composé	Structure	Log S
4	Diazinon		3.2
5	Dicrotophos		0.2
6	Dimethoate		0.47
7	Etrimfos*		2.96
8	Ethoprophos		2.73

N°	Composé	Structure	Log S
9	Fomothion		0.04
10	Malathion		1.75
11	Methamidophos*		0.37
12	Methidathion*		0.37
13	Mevinphos		0.2

N°	Composé	Structure	Log S
14	Naled		2.57
15	oxydometon-methyl		0.38
16	Phorate		3.22
17	Phosmet		1.65

N°	Composé	Structure	Log S
19	phoxim		4.26
20	Sulprofos*		2.54
21	Terbufos		3.71
22	Thiometon		2.17

\* composés de validation externe

## I- Calcul du modèle :

### I- 1- Choix du modèle :

Les molécules sont dessinées par le logiciel :

**Chem Draw (Chem Draw ultra 7.0)** : le ChemDraw fournit des chimistes avec un ensemble d'outils riche, faciles à utiliser pour créer des publications prêt, dessins scientifiquement

significatifs de molécules et de réactions enregistré dans différents formats. ChemDraw est très pratique pour les réactions d'écriture à l'aide de produits chimiques. Vous devriez être en mesure de réaliser des structures dans différents formats disponibles [42].

Puis elles sont optimisées en utilisant le logiciel :

**Hyper Chem :** le Hyper Chem rassemble dans une même interface un ensemble d'outils dédiés à la modélisation moléculaire, qui est connu pour sa qualité, flexibilité, et facilité d'usage.

HyperChem est le logiciel qui vous permet de faire réellement de la modélisation : il possède plus de méthodes de calculs (mécanique moléculaire, semi empirique et ab- initio ..... ) pour que vous puissiez calculer plus de propriétés.

Hyperchem est utilisé dans cette étude pour construire et optimiser les molécules, chaque molécule est enregistrée comme un fichier nommé "Hin" après l'optimisation. Pour l'optimisation (stabilisation des structures des molécules par la minimisation de l'énergie potentielle) nous avons utilisés la méthode de mécanique moléculaire MM+ suivis par la méthode semi empirique AM1. On a 22 molécules pour l'étude, donc on obtient 22 fichiers Hin, en suite en calcul les descripteurs moléculaires pour chaque molécule à partir de ce fichier par le logiciel DRAGON [42].

**Le Logiciel DRAGON :** c'est une application pour le calcul des descripteurs moléculaires. Ces descripteurs peuvent être utilisés pour évaluer l'influence de la structure moléculaire ou les relations propriétés- structure, aussi pour l'analyse de symétrie et la projection des bases de données des molécules. DRAGON fournis 1664 descripteurs moléculaires qui sont divisés en 20 blocs logiques [42].

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 (AGs), Les AGs ont été conçus comme outil d'optimisation stochastique. Ils sont inspirés des mécanismes de la génétique et de l'évolution naturelle des êtres vivants. Ils cherchent la solution globale d'une manière aléatoire. Ces algorithmes sont les plus connus et utilisés des algorithmes évolutionnaires [41].

Dans la version MobyDigs, 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 6 composés, alors que sa calibration par 16[02].

## 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 II :

**Tableau II** : Descripteurs moléculaires intervenant dans la modélisation du (log P).

N°	Descripteur	Classe	Signification
01	IVDE	Information indeces	Mean information content on the vertex degree equality
02	H6p	Getaway descriptes	H autocorrelation of lag 6/ unweighted
03	R5u	Getaway descriptes	R autocorrelation of lag 5/ weighted

Les valeurs expérimentales de la constante de Henery et les descripteurs choisis par algorithme génétique sont présentées dans le tableau suivant:

**Tableau III** : Valeurs de log S, IVDE, H6p, R5u

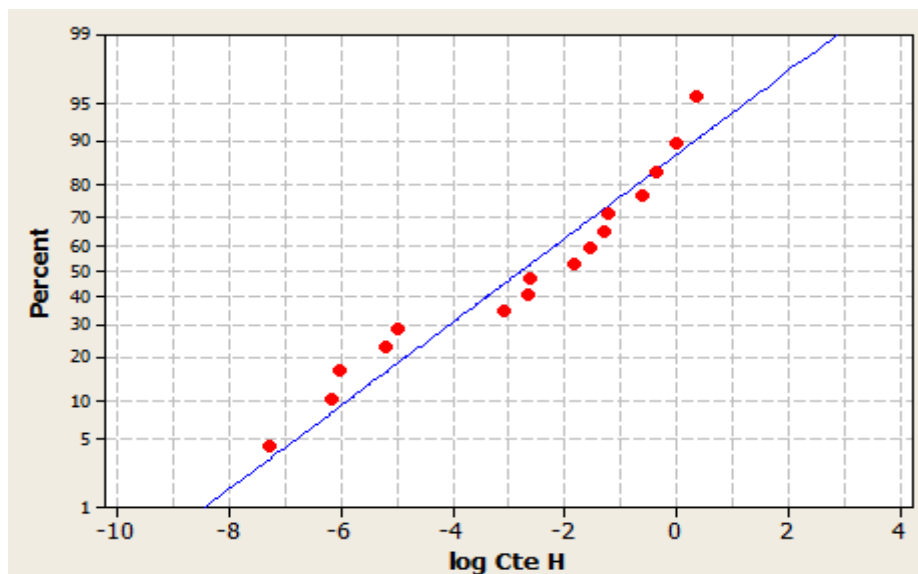
Nom ISO	log Cte H	IVDE	H6p	R5u
acéphate	-7.295849483	1.685	0.028	1.033
azinephos-ethyl	-1.655607726	1.582	0.126	0.856
azinephos-methyl	-2.690369833	1.657	0.147	0.748
diazinon	-1.215382707	1.747	0.155	0.675
dicrotophos	-2.655607726	1.746	0.212	0.939
dimethoate	-4.995678626	1.65	0.072	0.832
etrimfos	-1.202040356	1.676	0.081	0.606
ethoprophos	-1.823908741	1.239	0.046	0.685
fenitrothion	-0.623423043	1.727	0.165	0.9

Nom ISO	log Cte H	IVDE	H6p	R5u
methamidophos	-6.178486472	1.379	0.011	0.855
methidation	-3.209714836	1.379	0.011	0.855
mevinphos	-5.197226275	1.727	0.14	1.015
naled	-1.295849483	1.669	0.128	0.808
oxydometon-methyl	-6.028724151	1.614	0.05	0.916
phorate	0.004321374	1.239	0.144	0.84
phosmet	-3.083546051	1.763	0.108	0.775
phosphamidon	-6.181114585	1.747	0.224	1.502
phoxim	-0.378823718	1.457	0.118	0.773
sulprofos	-1.058488567	1.5	0.177	0.958
terbufos	0.346352974	1.429	0.171	0.824
thiometon	-1.548213564	1.281	0.135	0.734

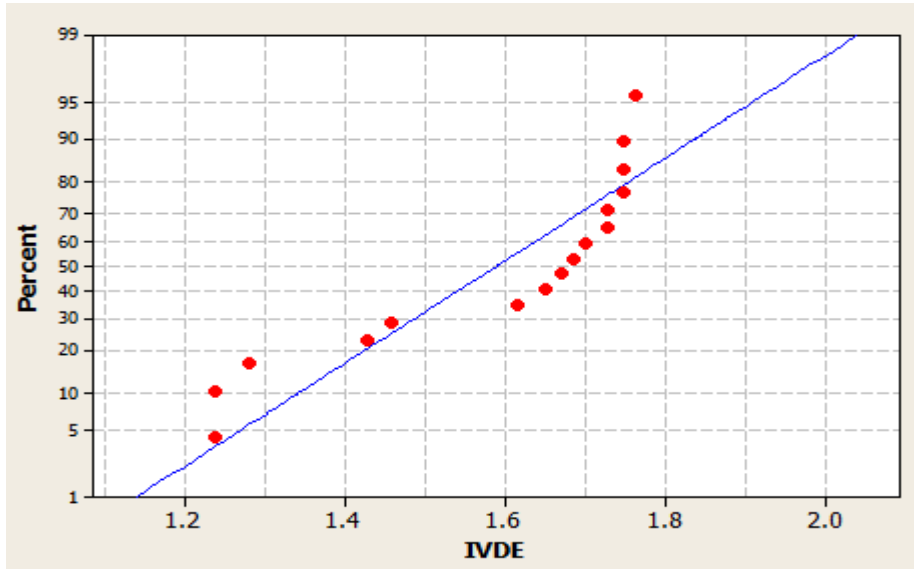
\*composés de validation externe

## II- Développement du modèle :

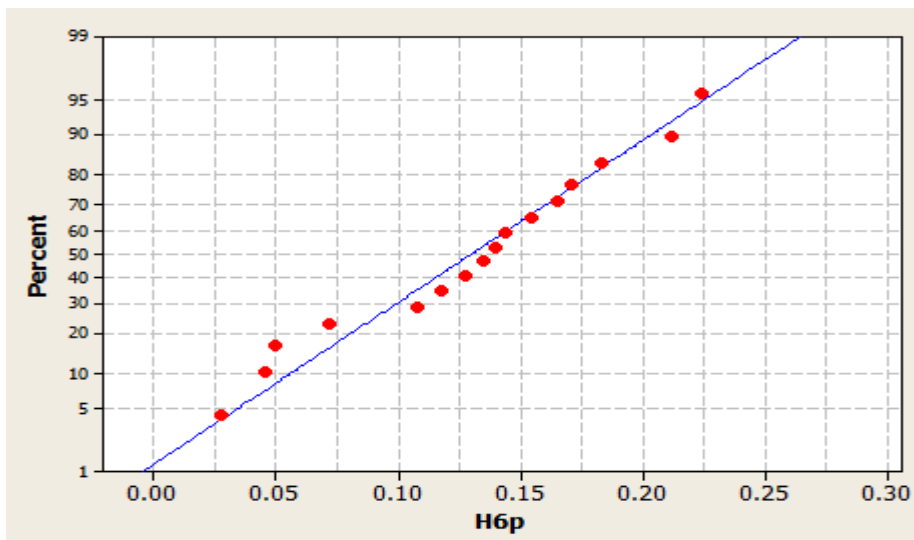
### II-1- Test de normalité



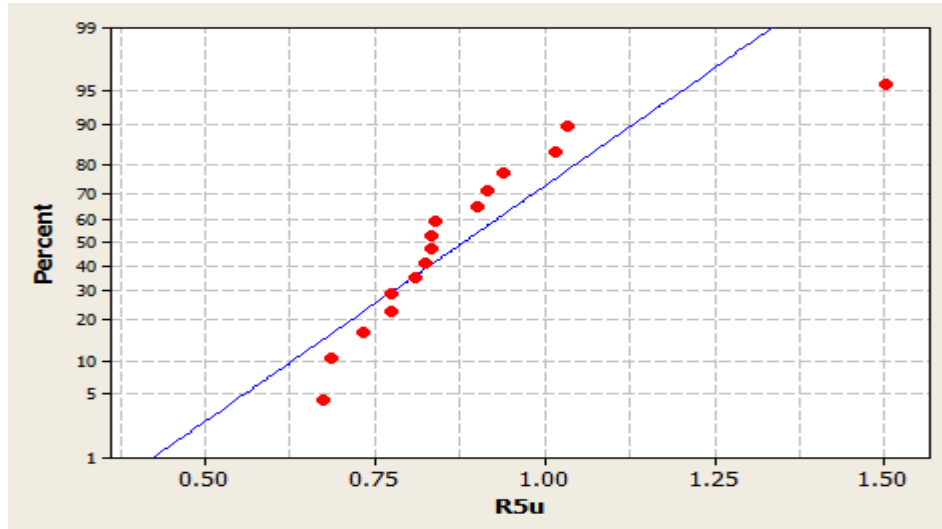
**Rj= 0.967**



**Rj = 0.904**



**Rj = 0.986**



$R_j = 0.868$

Comparant les valeurs de  $R_j$  des trois descripteurs choisis par algorithme génétique et le  $R_j$  de la variable dépendante avec la valeur critique on observe que la distribution des données suit une loi normale.

## II- 2- 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 corrélations:(log Cte H; IVDE; H6p; R5u)

	Log Cte H	IVDE	H6p
IVDE	-0.481 0.059		
H6p	0.336 0.204	0.254 0.343	
R5u	-0.619 0.010	0.408 0.117	0.346 0.189

D'après les valeurs des coefficients de corrélation et les valeurs de  $p$ , on remarque que tous les descripteurs sont non corrélés entre eux. On remarque aussi que la constante de Henry est corrélée avec l'IVDE et R5u par contre n'est pas corrélée avec le H6p que son élimination du modèle conduit à un mauvais résultat.

**II- 3- Equation et analyse de régression :**

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

**Log Cte H = 8.54 (± 2.378) - 4.59 (± 1.623) IVDE + 28.2 (±5.274) H6p- 8.74 (±-1.653) R5u (1)**

**Tableau V:** Analyse de régression

<b>Descripteur</b>	<b>Coef</b>	<b>SE Coef</b>	<b>T</b>	<b>P</b>	<b>VIF</b>
<b>Constant</b>	8.539	2.378	3.59	0.004	
<b>IVDE</b>	8.539	1.623	-2.83	0.015	1.220
<b>H6p</b>	28.219	5.274	5.35	0.000	1.156
<b>R5u</b>	-8.743	1.653	-5.29	0.000	1.297

Les valeurs de VIF présentées dans le tableau ci-dessous, montrent l'indépendance des trois descripteurs, valeur inférieur à 5.

Le descripteur H6p présent la plus importante valeur de T, ce qui confirme sa grande contribution dans le modèle, de même le descripteur R5u.

Le descripteur IVDE il a une contribution proche.

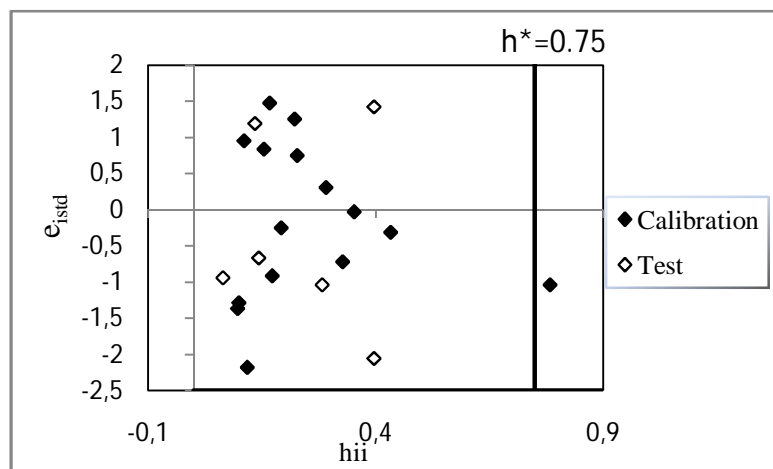
Les valeurs des paramètres statistiques sont présentes dans le tableau suivant :

**Tableau VII:** Les paramètres statistiques pour l'ensemble de calibration

<b>Les paramètres statistiques</b>	
F	20.46
R <sup>2</sup>	83.65
R <sup>2</sup> <sub>adj</sub>	79.56
Q <sup>2</sup> <sub>Loo</sub>	76.19
Q <sup>2</sup> <sub>ext</sub>	74.53
S	1.098
EQMP	1.147
EQMC	0.951
EQMP <sub>ext</sub>	1.186
Kx	33.77
Kxy	43.93

Les paramètres statistiques reproduit dans le tableau ci-dessus montrent une bonne explication de la variabilité log cte H, par les descripteurs choisis, de l'ordre de 83 pour le coefficient de détermination, et une robustesse du modèle due à la valeur de  $Q^2_{LOO}$ , alors que les valeurs de l'erreur quadratique moyenne de prédiction et de calcul sont petits et proches; en outre ce modèle est significatif avec une valeur du paramètre de Fisher: (F=20,46), Les statistiques montrent aussi que le modèle à une très grande capacité prédictive externe, ceci traduit par la grande valeur de  $Q^2_{ext}$  et la faible valeur de  $EQMP_{ext}$

## II- 4- Diagramme de Williams

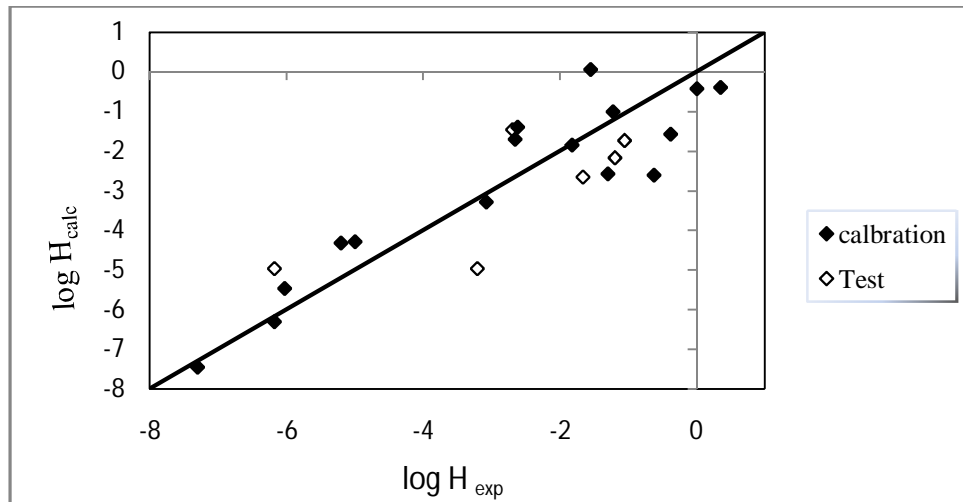


**Figure 08:** Diagramme de Williams

D'après le graphe on remarque que tous les erreurs standardisé sont comprise entre les limites  $\pm 2, 5$ , donc une absence des points aberrant, on remarque aussi une présence d'un point influents c'est-à-dire il a une valeur de levier supérieur à la valeur critique  $h^*=0,75$ .

## II- 5- 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 logarithme de la constante de Henry.



**Figure 09** : Qualité de l'ajustement.

D'après le graphe on remarque une certaine dispersion autour de la première bissectrice des deux ensembles de calibration et de validation. Ceci est confirmé par la valeur de  $Q^2_{LOO}$ , qu'elle n'est pas trop élevée (76,19).

## II-6- 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 10), dans cette figure on à présenter le graphe des coefficients statistiques  $Q^2$  et  $R^2$ . Il est clair que les statistiques obtenues pour les log cte H 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.

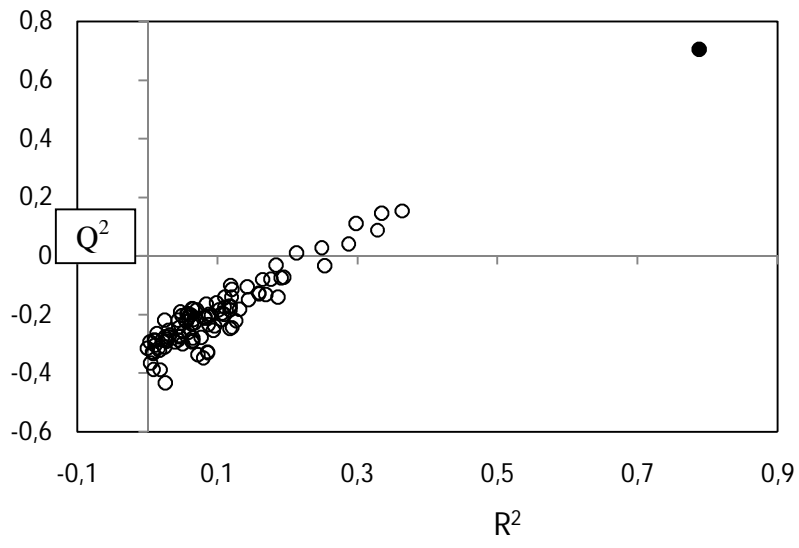
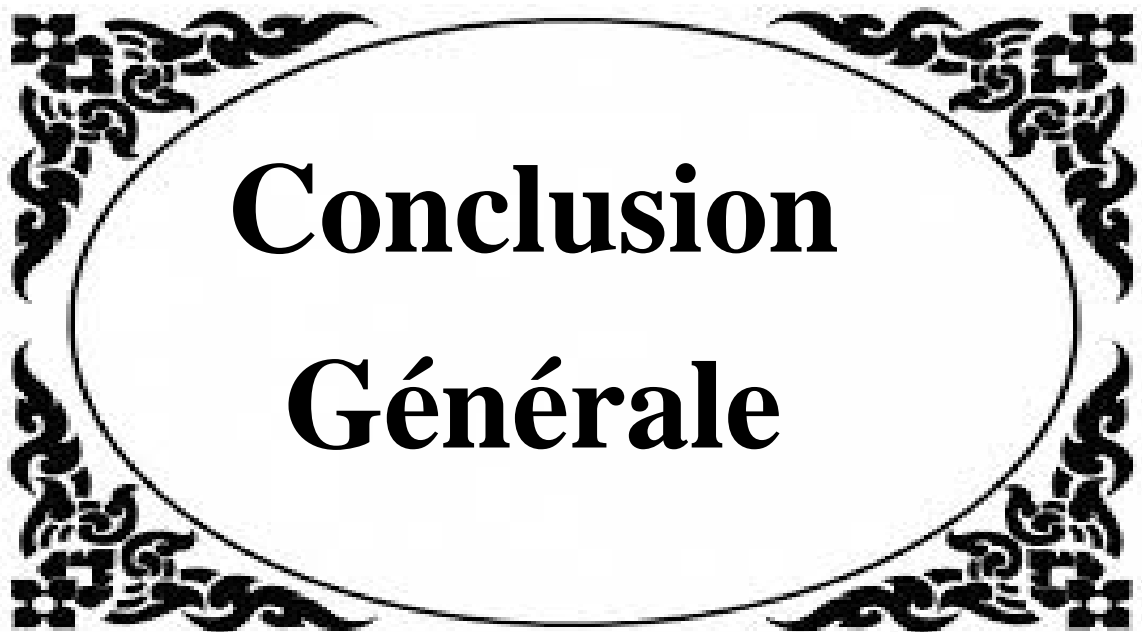


Figure N°10 : Test de randomisation



**Conclusion  
Générale**



### Conclusion générale

Dans notre travail nous avons utilisé la méthodologie QSPR pour prédire la constante de Henry d'une série de 22 insecticides, pour cela nous avons relié cette propriété à des descripteurs moléculaires théoriques.

Pour cette étude nous avons appliqué la méthode de régression linéaire multiple, les descripteurs intervenant dans le modèle sont choisis par algorithmes génétiques.

Le modèle choisi conduit aux bons résultats : qualité de l'ajustement, robustesse interne et externe, capacité prédictive... ce qui prouve que la relation entre le coefficient de Henry et les variables explicatives (IVDE, H6p, et R5u) elle est fondamentalement linéaire.

### Les Références

- [01] N. Mokhtari. Identification et dosage des pesticides dans l'agriculture et les problèmes d'environnement liés, (15 octobre 2012), mémoire pour l'obtention du diplôme de magister option chimie organique, l'université d'Oran, PP : 11-28.
- [02] H. Naima, M. Fayza. Contrôle de la pollution des eaux par les pesticides: modèle QSPR, (2017/2018), Mémoire de fin d'études Pour l'obtention du diplôme de Master, université : khenchela, PP : 14-18.
- [03] D. M. Hassenzahi, L.R. Berg, P.H. Raven. Environnement, (2008), Ed, 6 édition Américaine, PP : 582 – 584.
- [04] A. O. Mehmt. J. Mouchel. Pesticides : impact environnementaux, gestion et traitements, (2009), Ed, 5730, Paris, PP : 31- 32.
- [5] R. Calvet, E. Barriuso, C. Bedos, P. Benoit, M. Carnay, Y. Coquet. Les pesticides dans le sol (conséquences agronomique),(2005), Ed, France agricole, France, PP : 26- 159.
- [6] A. Berrah. Master -2- en toxicologie appliquée, (2011), Université de Tébessa Algérie.  
[https://www.memoireonline.com/11/12/6459/m\\_Etude-sur-les-pesticides.html](https://www.memoireonline.com/11/12/6459/m_Etude-sur-les-pesticides.html) Accès le 9/6/2019
- [7] A. Mamane. Effets sanitaire aigus de l'exposition aux pesticides en milieu rural - étude dans un pays du nord - Thèse présentée pour obtenue le grade de docteur, (2015), l'université de bordeaux, France, PP : 17- 20.
- [8] A. Periquet, M. Boisset, F. Casse, M. Catteau, J. Lecerf, C. Leguilie, J. Lavile, S. Barnat. Comité sécurité Alimentaire d'Aprifel : pesticides, risques et sécurité alimentaire, Op.cit., PP : 9 - 12.
- [9] M. Merghid, M. Debbache, I Foughali. Impacts des pesticides utilisés dans la plasticulture sur la santé humaine En Algérie - Etude de cas la wilaya de Constantine – en Toxicologie, (2017), Université des Frères M'entouré Constantine, PP : 22- 34.
- [10] B. Dorothée. L'impact des pesticides sur la santé humaine, (2011), université d'Nancyi, PP : 14 - 27.
- [11] M. Porta, E. Zumeta. Implementing the Stockholm treaty on persiste organic pollutants, Occupational and environmental médecine, (2002), PP : 59.
- [12] M. Aktar, S. Dwaipayan, CH. Ashim. Impact of pesticides use in agriculture: théier benefits and hazards, Toxicol Interdiscip, Slovak Toxicology Society setox, Inde, (2009), PP : 15.

- [13] J. Socorro. Etude de la réactivité hétérogène de pesticides adsorbés sur des particules modèles atmosphériques, cinétiques et produits de dégradation, Doctorat, (2015), Université d'Marseille, PP : 36.
- [14] E. Kegly Susan. Distrupting the balance, Ecological Impacts of pesticides In California Californians for pesticides reform, pesticides action Network (Group), California, (1999), PP : 19.
- [15] H. El Bakouri. Développement de la nouvelle technique de détermination du pesticides et contribution à la réduction de leur impact sur les eaux par utilisation des substances organique naturelle (S.O.N), Thèse de doctorat ,(2006), Université Abedelmalek Essadi, Maroc, PP : 16.
- [16] V. Feigenbrugel. Devenir Atmosphérique des pesticides, Distribution entre les déférentes phases de l'atmosphère et oxydation photochimique, (2005), L'universités d'Louis Pasteur de Strasbourg PP : 26 -27.
- [17] Ministère de l'Environnement, de l'Énergie et de la Mer, Commissariat général au Développement durable, Les impacts des pesticides. Cit. PP : 2.
- [18] A. Auvergne, Les pesticides dans l'air en France et en Auvergne : Etat des lieux (2006), version 1, PP : 6 <http://www.atmoauvergne.asso.fr> Accès le 11/6/2019.
- [19] H. Blanchoud. Répartition des pesticides dans les phases gazeuse, dissoute et particulaire du compartiment atmosphérique : influence sur le transport à longue distance, Université Pierre et Marie Curie, (2003), Université Paris-Sud, Ecole des Mines de Paris, PP : 8.
- [20] A. Sameut Bouhik. Modélisation de dépôt des pesticides sur des plans d'eau, (2009-2010), Université Paris-Est Marne-la-Vallée, PP : 11.
- [21] Mme. Errahouinée, KH. Bellifa. Etude des relations quantitatives structure–toxicité des composés chimiques à l'aide des descripteurs moléculaires. « Modélisation QSAR », Soutenu publiquement, le (08 /10 /2015), thèse : Doctorat en Chimie Physique Université Abou Bekr Belkaid, Tlemcen, PP : 08-10.
- [22] S. Chtita. Modélisation de molécules organiques hétérocycliques biologiquement actives par des méthodes QSAR/QSPR. Recherche de nouveaux médicaments, (le 08 Juillet 2017), thèse En vue de l'obtention de Doctorat en Chimie Spécialité : Chimie Physique et Modélisation, université Moulay Ismail, PP : 20-46.

- [23] C. Hansch, T. Fujita. P- $\sigma$ - $\pi$  Analysis, a Method for the Correlation of Biological Activity and Chemical Structure (1964), *Journal of the American Chemical Society*, 86(8), PP: 1616–1626.
- [24] S.M. Free, J.W. Wilson. A Mathematical Contribution to Structure-Activity Studies (1964), *Journal of Medicinal Chemistry*, 7(4), PP: 395–399.
- [25] C. Hansch, E.J. Lien. Structure activity relationships in anti fungal agents. A survey, (1971), *Journal of Medicinal Chemistry*, 14(8), PP: 653–670.
- [26] S.Y. Tham, S. Agatonovic-Kustrin. Application of the artificial neural network in quantitative structure gradient elution retention relationship of phenyl thiocarbamyl amino acids derivatives, (2002), *Journal of Pharmaceutical and Biomedical Analysis*, 28(3), PP: 581-590.
- [27] R.D. Cramer, D.E. Patterson, J.D. Bunce. Comparative molecular field analysis (CoMFA): Effect of shape on binding of steroids to carrier proteins, (1988) , *Journal of the American Chemical Society*, 110(18), PP: 5959-5967.
- [28] G. Klebe, U. Abraham, T. Mietzner. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity, (1994), *Journal of Medicinal Chemistry*, 37(24), PP: 4130-4146.
- [29] S. Ghasemi, Saaidpour, S.D. Brown. QSPR studies for estimation of acidity constants of some aromatic acids derive actives using multiple linear regression (MLR) analysis. *J. Mol. Struct. (Theochem)*, 805, (2007) PP: 27-32.
- [30] P. Geladi, B.R. Kowalski. Partial Least Squares Regression: a Tutorial, (1986), *Anal. Chim. Acta* 185, PP: 1-17.
- [31] A.J. Myles, R.N. Feudale, Y. Liu, N.A. Woody, S.D. Brown, An introduction to decision tree modeling,( 2004), *J. Chemom* , 18, PP:275-285.
- [32] A.F. Duprat, T. Huynh, G. Dreyfus. Toward a principle methodology for neural network design and performance evaluation in QSAR, application to the prediction of Log PJ, (1998), *Chem. Inf. Comput, Sci*, 38, PP: 586-594.
- [33] J. Gasteiger, J. Zupan. Neural Networks in Chemistry *Angew. Chem. Int. Ed. Engl.* (1993), 32, PP: 503-527.
- [34] R. Leardi. 2001. Genetic algorithms in chemo metrics and chemistry: are view J. *Chemometr*, 15, PP: 559-569.

- [35] A.Z. Dudek, T. Arodz, J. Gálvez. Computational methods in developing quantitative structure-activity relationships (QSARs), (2006), A review, *Combinatorial Chemistry and High Throughput Screening*, 9, PP : 213–228.
- [36] A.R. Katritzky, M.Kuanar, S.Slavov, C. D.Hall. Quantitative Correlation of Physical and Chemical Properties with Chemical Structure, (2010), Utility for Prediction, *Chem. Rev.*, 110, PP: 5714–5789.
- [37] K. Roy, S. Kar, R. Narayan Das. Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment, Chapter 6 – Selected Statistical Methods in QSAR, (2015), Academic Press, Boston, PP: 191–229.
- [38] B. Philippe. Pratique de la modélisation Statistique janvier, (2003), université Paul Sabatier Toulouse Version, PP : 9-21.
- [39] N. Zerari. Les algorithmes génétiques en maintenance, (2006), Université El Hadj Lakhdar Batna, PP : 19.
- [40] B. Hamad, S. K hala. Contribution à la Prédiction de Coefficient de partage Octanol /eau par la technique QSPR (Quantitative Structure Property Relationship), (2014), Université D’El –Oued, PP : 27 -28.