REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE YAHIA FARES DE MEDEA



Faculté de **T**echnologie

Mémoire de fin d'études de Master

Filière : Génie des procédés
Spécialité : Génie chimique

Modélisation de l'adsorption multi-composants par la méthode machine à vecteurs de support

---

Présenté par :

Mme ALOUANE Khadidja et Mlle SAIDOUN Meryem

Proposé et dirigé par :

Dr. LAIDI Maamar

Promotion 2021

Dédicace

je dédie d'abord ce modeste travail à mes parents,
pour leurs incessants encouragements, les moyens qu'ils mettent à ma disposition,
leurs extrêmes exigences, leur joie de vivre et leurs savoir-faire incommensurables. Je ferais
mon maximum pour ne pas les décevoir, à l'avenir.

Á mes frères: khaled et omar

Á mes sœurs : Dalila, Razika et karima

Á mon promoteur : Ms Maamar Laidi

Á mon binome : Alouane khadidja

A mes copines Ahlem, Kamila, Feryel, Achauak,Asma et Manel

Sans oublier à mes collègues de la promotion 2020/2021

À tous ceux qui me sont proches et qui me sont chers.



Dédicace

Je dédie ce travail à :

Les deux prunelles de mes yeux,

Mes parents

Moussa et Benissa

Hourai et zohra ,

À celle qui m'a été le soutien moral pendant ce projet… A mon mari ***DJAMEL HATCHANE qu'elle*** **trouve dans ce modeste travail mes gratitudes et reconnaissances ;**

À ma fille Açil, que dieu la protège ;

À mes frères

Rachid, Fathi, Mohamed

Et mes sœurs, surtout

Amina et Feziya

À toutes ma famille

Ma binôme meryem

Et à tous mes amis,

A tous ceux qui m'ont aidé, de près ou de loin, même avec un mot d'encouragement et de gentillesse.

Á Toutes la promotion de génie chimiques et génie des procédés

« Master »2020-2021.

KHADIDJA

© www.celestill.net

Remerciements

Au terme de ce Mémoire nous tenons à remercier tout naturellement en premier lieu ***ALLAH*** *qui nous a donné la force, le courage et la patience de bien mener ce travail.*

Ce travail a été réalisé sous la direction de ***M****onsieur* ***M. LAIDI****, Maître de Conférences de Yahia Farès de Médéa, qu'il trouve l'expression de notre profonde reconnaissance pour sa compétence, ses encouragements et ses précieux conseils durant le déroulement de ce travail.*

Nos remerciements au ***M****onsieur* ***M. HENTABLI*** **ET MLL M .ACHOUAK** pour *ses encouragements et ses conseils.*

Nos plus vifs remerciements vont également à ***M****essieurs* ***S.RAHAL*** *est* ***M****me* ***H KERMET SAID*** *membres jury.*

Nos sentiments vont également à toutes les personnes qui ont contribué de près ou de loin à la réalisation de ce travail.

Liste des Abréviation

BET Brunaue, Emmett, Teller.

EAM Erreur absolue moyenne.

EPM Erreur prédictive du modèle.

ER Erreur relative.

ERA Erreur relative absolue.

IAST Théorique de la solution adsorbée idéale.

pH Potentiel hydrogène.

R Coefficient de corrélation.

R^2 Coefficient de détermination.

RBF Radial basis fonction.

REQM Racine de l'erreur quadratique moyenne.

RMSE Racine de l'erreur quadratique moyenne.

RPD Ecart prédictif résiduel.

SEP Erreur standard de prédiction.

SVM Machine à vecteur de support.

Liste des Nomenclatures

A et B	Constantes de l'isotherme de Temkin.
A_s	Surface spécifique.
b	Biais d'un modèle linéaire .
b_A	Constante de Langmuir.
b_i,b_j	Paramètres de modèle de Langmuir généralisé .
b_T	Fonction de la chaleur d'adsorption.
C_0	Concentration initiale de l'adsorbat.
C_e	Concentration de l'adsorbat à équilibre.
C_{ei},C_{ej}	Concentration des espèces i et j à l'équilibre.
d	Degré d'un modèle linéaire .
F	Ensemble de fonction de x dans y.
H	Ensemble des hyperplans
ḱ	Constante liée à l'énergie d'adsorption.
K(xi,xj)	Fonction du noyau (kernel).
K_1	Constantes de vitesse d'adsorption pour le premier ordre.
K_e	Constante d'équilibre d'Elovich.
K_F	Constante associées respectivement à la capacité d'adsorption.
K_L	Constante de Langmuir.
K_T	Constante d'équilibre de Toth.
mT	L'exposant du modèle de Toth .
M_w	Masse molaire.

n	Constante déduite du modèle de Freundlich.
$\mathbf{q_e}$	Quantité du métal adsorbée par unité de masse de l'adsorbant .
$\mathbf{q_{max}}$	Quantité adsorbée maximale nécessaire pour la monocouche.
R	Constante universelle de gaz.
T	Température absolue.
$\mathbf{K_B}$	Constante de B.E.T.
ε	Potentiel de polanyi.
$\mathbf{\partial_{i,j}}$	Coefficient de compétition de l'espèce i en présence de l'espèce.
w	Vecteur de poids d'un modèle SVM.
δ	Probabilité de confiance sur le majorant de l'erreur de généralisation.
$\boldsymbol{\alpha_i, \beta_i}$	Multiplicateur de Lagrange.
σ	Largeur de la bande d'un noyau gaussien.
$\|.\|$	Norme euclidienne**.**

نمذجة الامتزاز متعدد المكونات بواسطة شعاع الدعم الآلي

الملخص :

في هذه الدراسة، تم تطوير نموذج شعاع الدعم الآلي (SVM) باستخدام قاعدة بينات للعديد من آنَضمة امتصاص المعادن الثقيلة. تم تطوير هذا النموذج لحساب سعة الامتزاز من خلال استخدام عدد قليل من المعلمات التي تميز الممتزات والممتازات كمصفوفة الإدخال. بناء على نضرية طريقة النمذجة غير الخطية SVM، تم وضع برنامج MATLAB لتدريب وتحسين معلمات طريقة مثل : {رقم الدعم المتجه والمعلمات الفائقة . في هذا العمل تم اختيار خمسة مداخل وهي {جهد الهيدروجين (pH) ، درجة حرارة الخليط (° C) ، السطح المحدد للمادة الممتزة ($m^2.g^{-1}$) ، الكتلة الجزيئية للأيونات المعدنية (mol-1) وتركيز توازن الأيونات في الممتزات} والمخرجات التي تمثل كمية خرج الأنظمة الثلاثة العديد من المعلمات ، مثل وظيفة النواة ، {linear, polyline RBF}. تم العثور على النموذج الأمثل بالدالة RBF و C .

تم التحقق من صحة آداء هذا النموذج على بيانات حقيقية وبا ستخدام بعض المعايير الإحصائية مثل{EAM {0.547،،0.478، 2.135}، REQM {1.13، 0.694 3.362}،

R^2 {0.9963،0.9963، 0.9963} اضهرت النتائج ان هذا النموذج قادر على التنبؤ بالكميات الثلاثة الممتصة بدقة عالية جدا.

النتائج مشجعة وتفتح آفاق بحثية جديدة. اكثر من مجرد دراسة حساسية توضح الاهتمام بكل مدخلات مقارنة بالمخرجات،فقد تم تصميم واجهة رسومية سهلة الاستخدام لتسهيل حساب سعات الامتصاص دون معرفة الضاهرة الفزيائية.

الكلمات المفتاحية : امتزاز متعدد المكونات،المعادن الثقيلة، شعاع الدعم الآلي ، نمذجة .

Modélisation d'Adsorption Multi-composants en utilisant l'approche Régression de vecteur de support

Résumé :

Dans cette étude, un modèle de type machine à vecteur de support (SVM) a été développé en employant une base de données regroupant plusieurs systèmes d'adsorption de métaux lourds. Ce modèle a été développé pour calculer la capacité d'adsorption en employant quelques paramètres caractérisant les adsorbats et l'adsorbant comme matrice d'entrée. En se basant sur la théorie de la méthode de la modélisation nono-linéaire SVM, un programme Matlab a été fait pour entrainer et optimiser les paramètres de la méthode SVM tels que {nombre de support vecteur et

les hyper paramètres}. Dans ce travail cinq entrés sont été sélectionnées qui sont {le potentiel hydrogène (pH), température du mélange (°C), la surface spécifique de l'adsorbant ($m^2.g^{-1}$), la masse moléculaire des ions métalliques ($g.mol^{-1}$) et la concentration à l'équilibre des ions dans l'adsorbants}et une sortie qui représente la quantité est sortie des trois systèmes plusieurs paramètres, ont été changés tels que les fonction kernel {(RBF,linear et polynomial), } .le modelé optimal a été trouve avec la fonction RBF, C , ε .

La performance de ce modèle été validée sur des données réelles et en employant quelques critères statistiques tels que REQM {1.136, 0.694, 3.362}, EAM {0.547,0.478, 2.135} et R^2 {0.9963, 0.9963, 0.9963}, les résultats montrent que ce modèle est capable de prédire les quantités adsorbées avec de très grande précision.

Les résultats sont encourageants et ouvrent de nouvelles perspectives de recherche. Plus d'une étude de sensibilité montrant l'intérêt de chaque entrée par rapport à la sortie, une interface graphique conviviale a été conçu pour faciliter le calcul des capacités d'adsorption sans connaitre le phénomène physique.

Mots-clés : adsorption multi-composants, métaux lourds, régression de vecteur de support, Modélisation

Multicomponent adsorption modeling using Support Vector Regression approach

Abstract:

In the work, a machine vector support (SVM) model was developed using a database grouping together several heavy metal adsorption systems. This model was developed to calculate the adsorption capacity by employing a few parameters characterizing the adsorbates and the adsorbent as the input matrix. Based on the theory of the SVM nonlinear modeling method, a MATLAB program was made to train and optimize the parameters of the SVM method such as (vector support number and hyperparameters). In this work five entries were selected which are {the hydrogen potential (pH), temperature of the mixture (° C), the specific surface of the adsorbent ($m^2.g-1$), the molecular mass of the metal ions (g.mol-1) and the equilibrium concentration of ions in the adsorbent} and an output which represents the amount output of the three systems several parameters, have been changed such as kernel

functions {(RBF, linear and polynomial), }. the optimal model was found with the function RBF, C, ε.

The performance of this model has been validated on real data and by using some statistical criteria such as RMSE {1.136, 0.694, 3.362}, MAE {0.547,0.478, 2.135} and R^2 {0.9963, 0.9963, 0.9963}, the results show that this model is able to predict the three adsorbed quantities with very high precision.

The results are encouraging and open up new research perspectives. More than a sensitivity study showing the interest of each input compared to the output, a user-friendly graphical interface has been designed to facilitate the calculation of the adsorption capacities without knowing the physical phenomenon.

Keywords: Multicomponent adsorption, heavy metals, Support vector Regression, molding.

Table des matières

Table des matières

Liste des figures

Liste des tableaux

Abréviations

Nomenclature et symbole

Introduction générale

Chapitre I : Généralité sur l'adsorption

I.1 Introduction 3

I.2 Définition de l'adsorption 3

I.3 Nature de l'adsorption 3

I.3.1 Adsorption physique (la physisorption) 4

I.3.2 Adsorption chimique (la chimisorption) 4

I.4 Description du mécanisme d'adsorption 5

I.4.1 Adsorption par liaison ionique ou échange d'ions 6

I.4.2 Adsorption par liaison hydrogène 6

I.4.3 Adsorption par les forces de Van der Waals 7

I.4.4 Rétention hydrophobe 7

I.5 Facteur influençant le phénomène d'adsorption 7

I.5.1 Surface spécifique 7

I.5.2 Température 7

I.5.3 PH 8

I.5.4 Nature de l'adsorbant 8

I.5.5 Nature de l'adsorbat 8

I.5.6 Polarité 8

I.6 Isotherme d'adsorption 8

I.7 Classification des isothermes d'adsorption 9

I.8 Modélisation des isothermes d'adsorption 11

I.8.1 Isotherme de Langmuir 11
I.8.2 Isotherme de Freundlich 12
1.8.3 Isotherme de Langmuir–Freundlich 12
I.8.4 Isotherme de Temkin 13
I.8.5. Isotherme BET (Brunaue, Emmett, Teller) 13
I.8.6 Isothermes de Dubinin-Radushkevish 14
1.8.7 Isotherme de Toth 15
1.8.8 Isotherme d'Elovich 15
I.9 Modélisation des isothermes d'adsorption du mélange binaire 16
I.9.1 Modèle de Langmuir généralisé 16
I.10 Adsorption statique multi-composant 17
I.10.1 Modèles mono-constituant étendus en modèles multi-constituants 17
I.10.1.1 Modèles de Freundlich étendu 17
I.10.1.2 Modèles de Langmuir étendu 18
I.10.1.3 Modèles étendus pour toute isotherme 18
I.10.2 Modèles d'adsorption en multi-constituants 18
I.10.2.1 Modèle de Langmuir 18
I.11 Cinétique d'adsorption 19
I.11.1 Modèle cinétique du pseudo premier ordre 19
I.11.2 Modèle de la cinétique du pseudo- second ordre 19
I .12 Généralités sur les adsorbants 20
I.12.1 Définition d'un adsorbant 20
I.12.2 Types d'adsorbants 20
I.12.2.1 Charbon actifs 20
I.12.2.2 Adsorbants minéraux 20
I.12.2.3 Les argiles activées 21
I.12.2.4 Les gels de silice 21

I.12.2.5 Les zéolithes 21

I.12.2.6 Les alumines activées 21

I.13 Application 21

II.14 Conclusion 22

Chapitre II : La machine vecteur de support

II.1 Introduction 23

II.2 Historique 23

II.3 Définition SVM 24

II.4 Principe de fonctionnement général de SVM 24

II.4.1 Notions de base 24

II.4.1.1 Hyperplan 24

II.4.1.2 Vecteurs de support 25

II.4.1.3 Marge 26

II.5 Linéarité et non-linéarité 27

II.5.1 Cas linéairement séparable 27

II.5.2 Cas non linéaire 28

II.6 Architecture de SVM 29

II.7 Apprentissage statistique 29

II.8 SVM binaires 30

II.8.1 SVM à marge dure 31

II.8.2 SVM à marge souple 34

II.9 Utilisation des noyaux 36

II .10 SVM pour la régression 40

II .11 Utilisation des noyaux 42

II.12 Applications 42

II.13 Avantages et inconvénients 43

II.13.1 Avantage……43
II.13.2 Inconvénients 43
II.14 Conclusion 44

Chapitre III : Modélisétion de l'adsorption multi-composaants par la méthode machine à vecteurs du support

III.1 Introduction 45
III.2 Modélisation par SVM 45
III.2.1 Méthodologie de travail 45
III.2.2 Collection et préparation de la base des données 46
III.3 Résultats de modélisation par SVM 48
III.3.1 Choix et optimisation des paramètres de SVM 48
III.3.2 Interprétation des résultats 50
III.3.3 Comparaison entre les résultats expérimentaux est résultats calcule 55
III.4 Analyse de sensibilité des modèles SVM 56
III.5 Interface graphique 58
III.6 Conclusion 59

Conclusion généralé.

Référénces bibliographiques .

Liste des tableaux

Chapitre I : Généralité sur l'adsorption.

Tableau I.1: Principales différences entre l'adsorption physique et chimique 5

Chapitre III : Modélisétion de l'adsorption multi-composaants par la méthode machine à vecteurs du support.

Tableau III.1 : Sources bibliographiques des données expérimentales. 46

Tableau III. 2 : Domaines des données expérimentales. 46

Tableau III.3 : Expressions mathématiques des fonctions du noyau. 48

Tableau III. 4 : Les paramètres optimisés du modèle SVM. 48

Tableau III. 5 : Performances des modèles SVM. 55

Liste des figures

Chapitre I : Généralité sur l'adsorption

Figure I.1 : Phénomène d'adsorption 4

Figure I.2 : Processus de transport lors de l'adsorption par un adsorbant poreux 6

Figure I.3 : Classes des isothermes d'après Gilles et Coll 10

Chapitre II : La machines vecteurs de support

Figure II.1 : Exemple d'un hyperplan sépare deux ensembles de points 25

Figure II.2 : Exemple de vecteurs de support 25

Figure II.3 : Exemple de marge maximal (hyperplan valide) 26

Figure II.4 : **a)** Hyperplan avec faible marge, **b)** Meilleur hyperplan séparateur 27

Figure II.5 : Exemple de classification d'un nouvel élément 27

Figure II.6 : **a)** linéairement séparable et ,**b)** cas non linéairement séparable 28

Figure II.7 : Exemple de changement de l'espace de données 28

Figure II.8 : Architecture de SVM 29

Figure III.9 : SVM binaire 31

Figure II.10 : SVM binaire à marge souple 34

Figure II.11 : Mal-adaptation de l'hyperplan aux problèmes réels de classification 37

Figure II.12 : Représentation idéale de la fonction de décision 38

Figure II.13 : Transformation d'espace 38

Figure II.14 : Hyper-tube modélisant la fonction de régression 41

Figure II .15 : Utilisation des noyaux pour la résolution de la régression non linéaire 42

Chapitre III : Modélisétion de l'adsorption multi-composaants par la méthode machine à vecteurs du support

Figure III.1 : Méthodologie de travail 45

Figure III.1 : Organigramme des étapes pour la création du modèle SVM 49

Figure III.3 : Diagramme de dispersion de la capacité d'adsorption calculée en Fonction de la capacité calorifique expérimentale par le modèle SVM : « RBF » 51

Figure III.4 : Erreur relative (%) de q_{e1} en fonction de C_e 52

Figure III.5 : Erreur relative (%) de q_{e2} en fonction de C_e 52

Figure III.6 : Erreur relative (%) de q_{e3} en fonction de C_e 53

Figure III.7 : Diagrammes de distribution d'ER [%] du modèle SVM de la q_{e1} 53

Figure III.8 : Diagrammes de distribution d'ER [%] du modèle SVM de la q_{e2} 54

Figure III.9 : Diagrammes de distribution d'ER [%] du modèle SVM de la q_{e3} 54

Figure III.10 : Comparaison statistique entre la quantité adsorbée (q_{e1}, q_{e2}, q_{e3}) 55

Figure III.11 : (a) et (b) Comparaison entre q_e^{cal} par SVM et q_e^{exp} des trois adsorbats (Pb^{+2},Hg^{+2}, Zn^{+2}) à différente température **(25 et 30°C)** en fonction de la concentration C_e. 56

Figure III.12 : (a) et (b) Comparaison entre q_e^{cal} par SVM et q_e^{exp} des trois adsorbats (Pb^{+2}+Cu^{+2}+Cd^{+2}) à différent pH **(4.1 et 6.45)** en fonction de la Concentration C_e. 56

Figure III.13: Résultats de l'analyse de sensibilité réalisés sur les modèles SVM. 57

Figure III.14 : Interface graphique développée pour le calcul de la quantité adsorbée. 58

Introduction générale

Introduction générale

L'industrialisation est vitale pour le développement durable, la gestion des ressources en eau est essentielle pour protéger l'environnement. Cependant, l'augmentation de la population mondiale conduit à une croissance plus industrielle et à des eaux usées de mauvaise qualité contenant des polluants nocifs pour la santé humaine et l'environnement , La pollution des eaux par certains produits chimiques d'origine industrielle (Métaux lourds, colorants, produits pharmaceutique, hydrocarbures, matière organique et inorganiques) provoque la dégradation de l'environnement et suscite à l'heure actuelle un défi particulier à l'échelle internationale (Kulkarni ., 2010).

Les eaux usées générées dans les procédés industriels nécessitent l'utilisation de méthodes et procédés de prétraitement pour réduire et / ou éliminer tous les types de polluants, avant rejet de l'effluent dans la nature (San, 2003).

Les eaux usées traversent les usines de traitement pour en éliminer les polluants. La plupart des eaux usées contiennent de nombreuses substances qui pourraient être classées en polluants organiques et inorganiques (Farag & Mohamed, 2004).

La majorité des polluants présents dans les effluents industriels sont difficiles à analyser et ne peuvent pas être traités par des méthodes conventionnelles. Cela est dû à des débits volumiques illimités et des concentrations plus élevées, leur présence pourrait entraîner des dommages environnementaux importants et des risques pour la santé. Plusieurs études et recherches sont en cours sur l'élimination de ces polluants environnementaux à l'aide de diverses méthodes physiques et chimiques. Ces méthodes comprennent la précipitation chimique, l'oxydation, l'échange d'ions, la coagulation, floculation, l'extraction par solvant, la séparation par membrane et l'adsorption (Carmen & Daniela, 2010).

L'une de ces techniques est l'adsorption multi-composant qui est le procédé le plus utilisée pour la séparation de l'eau (des colorants, les métaux lourds et les produits pharmaceutiques). L'adsorption est le procédé le plus utilisé dans les applications environnementales des industries chimiques, en particulier pour la séparation et la purification des effluents, en raison de son rendement élevé, de son faible coût et de sa facilité d'utilisation (Khandegar & Saroha, 2013)

L'adsorption séquentielle en mode statique des systèmes adsorbant(s)/ adsorbat(s) est très complexe et suscite aujourd'hui un intérêt considérable dans de nombreux secteurs industriels. Le processus séquentiel a une particularité que les espèces chimiques

s'adsorbent en même temps avec un degré de compétition différent dépendant de plusieurs paramètres thermo physicochimiques et morphologiques, il est à l'origine d'une multitude d'applications, notamment dans le domaine agro-alimentaire, dans le secteur pharmaceutique, ou encore dans le domaine chimique et pétrochimique.

La technologie étudiée dans ces derniers temps « l'intelligences artificiels ». Les chercheurs développent des méthodes et des logiciels pour avoir une meilleure capacité de traitement et modélisation les connaissances qui permettent de diminuer le coût de ces recherches qui est souvent étonnant, aussi bien un gain du temps.

Parmi les modèles qui se dérive de l'intelligence artificielle est « machine à vecteur de support ».

Dans notre travail de modélisation, nous avons collectés des données à partir de la littérature sur l'adsorption multi-composants des métaux lourds (tertiaires) afin de mettre une base de données pour essayer de construire un modèle fiable.

L'objectif principal de notre travail est le modélisation eu se basant de sur la minimisation de l'erreur afin d'obtenir des résultats confirmant la robustesse des modèles pour eu se basait cela nous appliquons les techniques les plus récentes qui sont les Support Vecteur Machines (SVM), Ces techniques sont sélectionnées pour être appliquées dans nos différentes expérimentations, ce choix est justifié par la simplicité de leur algorithme d'apprentissage.

Notre travail de Master est structuré en trois principaux chapitres en plus de l'introduction générale.

- Le premier Chapitre présente la théorie d'adsorption statique englobant plus de détails sur les phénomènes d'adsorption, les facteurs d'influence, caractéristique de la molécule, les types d'adsorption, les isothermes et la modélisation des isothermes.
- Dans le deuxième chapitre, nous parlerons brièvement sur le support vecteur machines.
- Dans le troisième chapitre, nous allons présenter les résultats de modélisation de l'adsorption multi-composant des métaux lourds (tertiaires) par la méthode

  Support Vecteur Machines

Enfin, le mémoire se termine par une conclusion générale en résumant l'essentiel de notre travail avec des perspectives qui peuvent être envisagées dans le futur.

Chapitre I : généralité sur l'adsorption multi-composants

I.1 Introduction

La contamination de l'eau par des métaux lourds toxiques par le rejet d'eaux usées industrielles est un problème environnemental mondial. De nombreux procédés chimiques tels que la précipitation chimique, l'électrofloatation, l'ion l'échange et l'osmose inverse ont été utilisés pour l'élimination des métaux lourds. Cependant, l'adsorption est la plus approprié pour résoudre ces problèmes (Ajmal , 2003). L'adsorption est l'une des techniques les plus fiables pour l'enlèvement des substances toxiques. Elle est très utilisée pour la séparation et la purification des gaz et des liquides dans différents domaines à cause de sa facilité d'emploi, son efficacité et son faible cout. En 1881, le terme d'adsorption a été proposé la première fois par Kayser afin de différencier entre la condensation et l'adsorption du gaz

Dans ce chapitre nous allons donner quelques généralités sur l'adsorption pour mieux présenter l'importance de ce processus, expliquer sa cinétique, ses caractéristiques et les éléments influant l'adsorption (Meriem, 2017).

I.2 Définition de l'adsorption

L'adsorption est un phénomène de surface spontané par lequel des molécules de gaz ou de liquides se fixent sur les surfaces des solides selon divers processus plus ou moins intenses. Le terme surface correspond à la totalité de la surface du solide, surface géométrique pour un solide non poreux, à laquelle s'ajoute pour un solide poreux, la surface interne des pores, accessible aux molécules du fluide. Elle est traduite par une augmentation de la densité du fluide à l'interface des deux phases. Elle peut donc aller de quelques molécules sur la surface, puis une monocouche et jusqu'à plusieurs couches formant une véritable phase liquide, surtout dans les micropores figure (I.1).

On appelle « adsorbat » la molécule qui s'adsorbe et « adsorbant » le solide sur lequel s'adsorbe la molécule. Le phénomène inverse par lequel les molécules se détachent est la désorption (Boulkrah, 2008) .

I.3 Nature de l'adsorption

L'adsorption dépend de l'existence d'un champ de force à la surface d'un solide, qui réduit l'énergie potentielle d'une molécule adsorbée en dessous de celle de la phase fluide ambiante. Il est utile de distinguer deux grandes classes d'adsorption (adsorption physique et chimique) en fonction de la nature du forces de surface (Ruthven, 2008) .

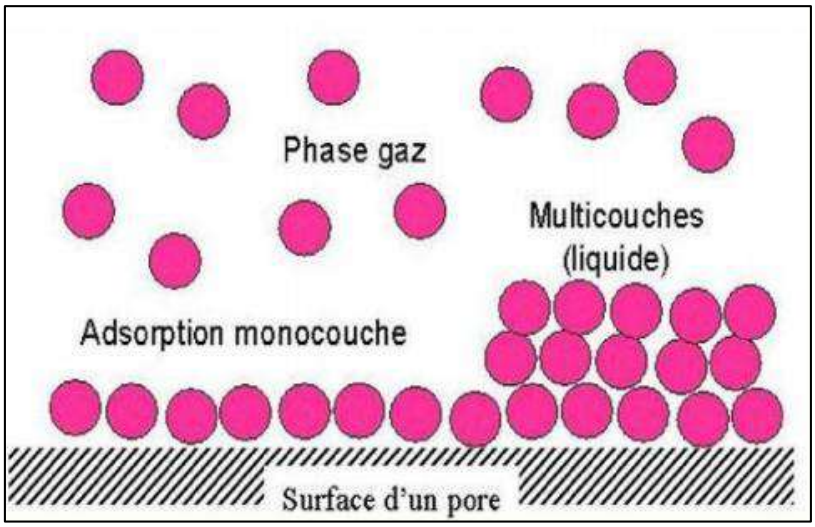


Figure I.1 phénomène d'adsorption

I.3.1 Adsorption physique (la physisorption)

L'adsorption physique met en jeu de très faibles interactions entre entités moléculaires comme les forces d'attraction de van der Waals et des forces dues aux interactions électrostatiques de polarisation. Elle est réversible et peu spécifique. L'adsorption physique est rapide et généralement limitée par les phénomènes de diffusion. La force des interactions mises en jeu peut être estimée par l'énergie d'adsorption qui est comprise entre 5 et 40 kJ.mol-1 et considérée comme faible : la désorption peut donc être totale (Mustapha, 2014).

I.3.2 Adsorption chimique (la chimisorption)

Dite chimisorption est un phénomène irréversible due à une liaison chimique forte de type covalente entre les atomes superficiels du solide et les molécule sorbées, ce type d'adsorption met en jeu des énergies d'attractions élevées, qui conduisent à des chaleurs d'adsorption élevées, approchant souvent aux énergies de liaisons chimiques covalentes de l'ordre de 20 à 80 K cal/ mol. D'après WEBER l'adsorption chimique ne concerne qu'une fraction très petite de la surface intra particulaire totale du charbon (Nabil, 2007).

Tableau I.1: Principales différences entre l'adsorption physique et chimique (Nabil, 2007)

Propriétés	**Adsorption physique**	**Adsorption chimique**
Adsorbant	Tous les solides	Quelques solides
Adsorbat	Tous les gaz en dessous de la température critique	Quelques gaz chimiquement réactifs
Température de processus	Inférieur à la température	Elevée
Nature de liaison	Physique (Van Der Waals)	Chimique
Energie d'adsorption	5 à 10 Kcal/mol	20 à 100 Kcal/mol
La désorption	Plus ou moins parfaite	Difficile
Energie d'activation	Non appréciable	Peut être mise en jeu
La cinétique	Très rapide	Lent
Etat de surface	Formation de multicouches	Conduit tout ou plus à une monocouche

I.4 Description du mécanisme d'adsorption

Le processus d'adsorption est l'un des transferts de matière qui s'effectue entre le fluide et l'adsorbant. Il se produit principalement en plusieurs étapes. Au cours de l'adsorption d'une espèce sur un solide, le transfert de matière a lieu de la phase fluide vers les sites actifs de l'adsorbant en passant par les étapes suivantes (Nandi , 2009) :

1. Diffusion de l'adsorbat de la phase liquide externe vers celle située au voisinage de la surface de l'adsorbant.
2. Diffusion extra granulaire de la matière (transfert du soluté à travers le film liquide vers la surface des grains).
3. Transfert intra granulaire de la matière (transfert de la matière dans la structure poreuse de la surface extérieure des graines vers les sites actifs).
4. Réaction d'adsorption au contact des sites actifs, une fois adsorbée, la molécule est considérée comme immobile (Mohammed, 2013) .

Pour mieux qualifier et quantifier la rétention, il convient de s'intéresser aux phénomènes se produisant à l'échelle moléculaire, c'est-à-dire aux mécanismes d'adsorption. Les liaisons composés/adsorbant sont de deux types :

- Liaisons de fortes énergies (> 80 kJ. Mol-1) : liaisons ioniques et échanges de ligands ;
- Liaisons de faibles énergies (< 80 kJ. mol-1) : interactions dipôle-dipôle, liaisons hydrogène, interactions hydrophobes (Bouziane, 2007) .

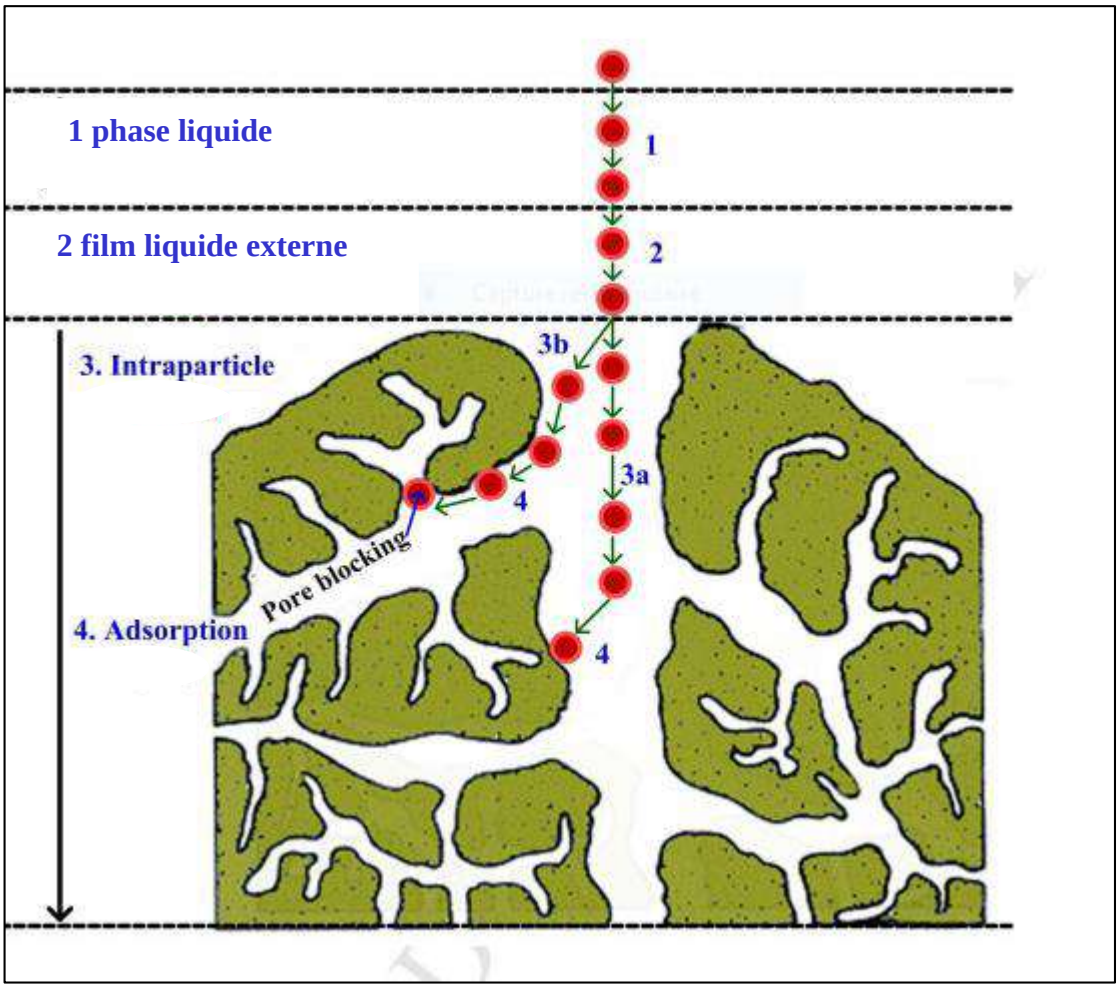


Figure I.2 Processus de transport lors de l'adsorption par un adsorbant poreux (Tran., 2017).

Sur la base de ces liaisons, quatre mécanismes principaux peuvent être distingués :

I.4.1 Adsorption par liaison ionique ou échange d'ions

Ce mécanisme ne concerne que les produits sous forme cationique ou ceux qui peuvent le devenir, comme les bases faibles. Il s'agit d'une interaction entre la molécule et les groupements ionisés ou facilement ionisables de l'adsorbant, tels que les groupements phénoliques et carboxyliques présents dans la structure des charbons. Ces liaisons sont le plus souvent peu réactives (Bouziane, 2007).

I.4.2 Adsorption par liaison hydrogène

Certains supports peuvent présenter à leur surface de nombreuses fonctions (oxygénées, hydroxyles…) qui peuvent interagir par liaisons hydrogène avec les groupes complémentaires des molécules adsorbat, bien qu'il y ait une forte compétition avec les molécules d'eau pour ces sites d'adsorption (Bouziane, 2007) .

I.4.3 Adsorption par les forces de Van der Waals

Les forces de Van der Waals sont des attractions dipolaires de faible intensité qui agissent à courte distance. Leur action s'additionne à d'autres types d'interactions plus fortes. Dans le cas de molécules non-ioniques et apolaires, leur participation à la rétention n'est pas négligeable. Si la molécule est assez volumineuse et qu'elle peut s'adapter à la surface de l'adsorbant, ces forces sont alors additives et leur rôle peut devenir important (Bouziane, 2007).

I.4.4 Rétention hydrophobe

Ce type de mécanisme concerne en particulier les composés non-ioniques et apolaires. Deux approches peuvent être distinguées :

- Certains auteurs proposent un mécanisme d'adsorption hydrophobe indépendant du pH. Les molécules des produits phytosanitaires apolaires interagissent avec les sites hydrophobes du support comme les chaînes aliphatiques, les fractions lipidiques, les dérivés de la lignine (riches en carbone avec peu de groupes fonctionnels). Ces sites sont relativement accessibles car il n'y a quasiment aucune compétition avec les molécules d'eau.
- D'autres auteurs préfèrent ne pas parler d'adsorption mais plutôt d'une partition. Ils assimilent la surface de l'adsorbant à un solvant liquide non-miscible à l'eau (Bouziane, 2007).

I.5 Facteur influençant le phénomène d'adsorption

L'équilibre d'adsorption entre un adsorbant et un adsorbat, dépend de nombreux facteurs. Les principaux facteurs influençant l'adsorption sont .

I.5.1 Surface spécifique

L'ampleur de l'adsorption est généralement considérée comme proportionnelle à la surface spécifique. La surface spécifique est la proportion de la surface totale disponible pour l'adsorption. Les adsorbants plus finement divisés et plus poreux devraient donner plus d'adsorption par unité de poids d'adsorbant. La surface peut être qualifiée d'externe lorsqu'il s'agit de renflements ou de cavités de largeur supérieure à la profondeur ou interne lorsqu'il s'agit de pores et de cavités de profondeur supérieure à la largeur (Libra & Saupe, 2011).

I.5.2 Température

L'adsorption implique des relations spécifiques entre les propriétés du charbon actif et le soluté. Par conséquent, les effets quantitatifs de la température ne sont pas les mêmes avec tous les carbones et solutés. L'ampleur de l'adsorption devrait augmenter avec la diminution température car les réactions d'adsorption sont exothermiques. Cependant, a augmenté la

température augmente également la vitesse de diffusion du soluté à travers le liquide vers le sites d'adsorption, ce qui conduit finalement à une augmentation de l'adsorption (Libra & Saupe, 2011).

I.5.3 PH

Le pH est un facteur important dans toute étude d'adsorption du fait qu'il peut influencer à la fois la structure de l'adsorbant et de l'adsorbat ainsi que le mécanisme d'adsorption. Ce facteur dépend de l'origine des eaux à traiter et du procédé de leurs traitements (coagulation, floculation, oxydation…). Donc, il est judicieux de connaître l'efficacité de l'adsorption à différents PH (Zerfa , 2015) .

I.5.4 Nature de l'adsorbant

On peut classer les adsorbants en : Adsorbants apolaires, ayant plus d'affinité pour les substances non polaires (charbon, graphite, talc, etc.). Adsorbats polaires, qui adsorbe les électrolytes, dans ce cas l'adsorption est sélective. Parmi ces adsorbants polaires, on cite les gels de silice et les différentes argiles (Khelifi, 2018).

I.5.5 Nature de l'adsorbat

Pour qu'il y ait une bonne adsorption il faut qu'il y ait d'abord une affinité entre le solide et le soluté (Khelifi , 2018) .

I.5.6 Polarité

En règle générale, les solides polaires, adsorbent préférentiellement d'autres corps polaires. Par contre les solides non polaires, adsorbent préférentiellement des substances non polaires et l'affinité pour le substrat croît avec la masse moléculaire de l'adsorbat (Khelifi, 2018).

I.6 Isotherme d'adsorption

L'adsorption du soluté est en général limitée à la monocouche. En effet, les interactions soluté-solide sont assez fortes pour concurrencer avec succès les interactions solvant-solide dans la monocouche, mais ce n'est plus le cas dans les couches suivantes. Toutefois l'adsorption du soluté en multicouche a été observée dans certains cas, en se basant sur la forme des isothermes établies et sur le fait que la surface spécifique évaluée à partir de l'hypothèse d'une monocouche est trop faible. Une telle isotherme est une courbe qui représente la relation entre la quantité adsorbée par unité de masse de solide et la concentration de la phase fluide. On obtient une telle courbe à partir des résultats d'essais effectués à une température constante.

Pour ce faire, on introduit des quantités connues de solide dans des volumes de solution à traiter, et après un temps de contact donné, on mesure la concentration résiduelle de la solution (Souheila, 2012)

On peut calculer la quantité adsorbée à l'aide de l'équation suivante :

$$q_e = \frac{(C_0 - C_e).v}{m} \qquad (I.1)$$

Avec :

q_e : quantité du métal adsorbée par unité de masse de l'adsorbant en ($mg.g^{-1}$)

C_0 : concentration initiale du soluté adsorbat en ($mg.l^{-1}$)

C_e : concentration d'acide à l'équilibre dans la solution en ($mg.l^{-1}$)

V : volume de la solution (l)

m : masse de l'adsorbant (g)

I.7 Classification des isothermes d'adsorption

Expérimentalement, on distingue quatre classes principales nommées : S (Sigmoïde), L (Langmuir), H (Haute affinité) et C (partition Constante).

La figure (I.3) illustre la forme de chaque type d'isotherme (Souheila, 2012) .

Cette classification tient compte entre autres d'un certain nombre d'hypothèses :

- Le solvant s'adsorbe sur les mêmes sites que le soluté. Ceci implique l'existence d'une compétition d'adsorption entre le solvant et le soluté.
- Le nombre de sites susceptibles d'accueillir les molécules de soluté à la surface du solide diminue quand la quantité adsorbée augmente.
- L'orientation des molécules à la surface. On peut citer le cas où les molécules sont adsorbées verticalement ou horizontalement sur la surface.
- Enfin, les interactions attractives ou répulsives entre les molécules adsorbées se manifestent d'une façon notable dans le phénomène d'adsorption

S : Sigmoïde

L : Langmuir

H : Haute affinité

C : Partition Constante

Après cette description nous abordons l'interprétation des différentes classes d'isothermes. Nous commençons par la plus fréquente : l'isotherme de Langmuir.

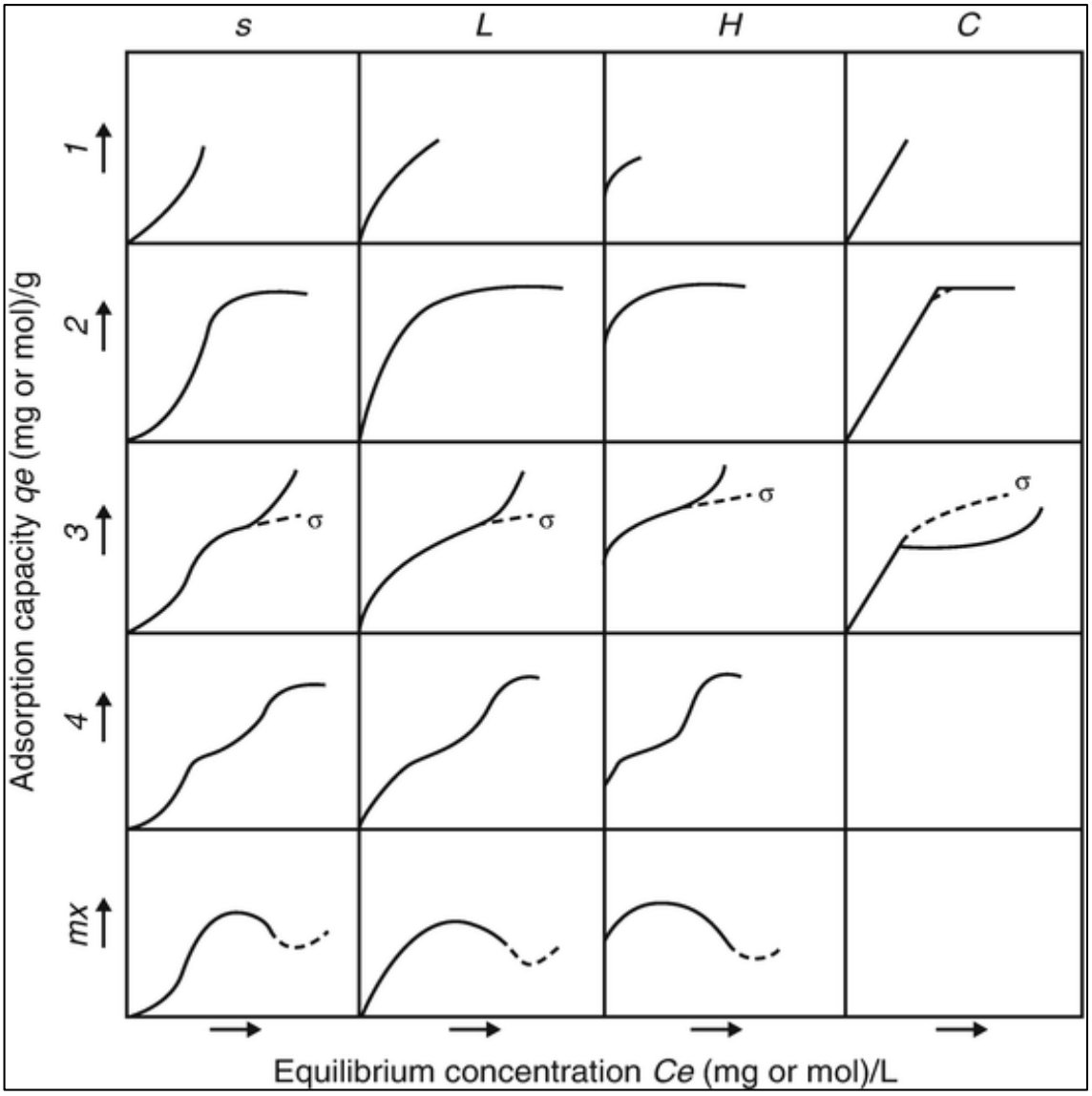


Figure I.3: Classes des isothermes d'après Gilles et Coll (Van Der Stap & Klaasse, 2016).

a) Classe L

Les isothermes de classe L présentent, aux faibles concentrations de la solution, une concavité tournée vers le bas qui traduit une diminution des sites libres au fur et à mesure de la progression de l'adsorption. Ce phénomène se produit lorsque les forces d'attraction entre les molécules adsorbées sont faibles. Elle est souvent observée quand les molécules sont adsorbées à plat, ce qui minimise leur attraction latérale. Elle peut également apparaître quand les

molécules sont adsorbées verticalement et lorsque la compétition d'adsorption entre le solvant et le soluté est faible. Dans ce cas, l'adsorption des molécules isolées est assez forte pour rendre négligeable les interactions latérales (Khoualdi & Medjani, 2016) .

b) Classe S

Les isothermes de cette classe présentent, à faible concentration, une concavité tournée vers le haut. Les molécules adsorbées favorisent l'adsorption ultérieure d'autres molécules (adsorption coopérative), ceci est dû aux molécules qui s'attirent par des forces de Van Der Waals, et se regroupent en îlots dans lesquels elles se tassent les unes contre les autres. Ce comportement est favorisé, d'une part, quand les molécules de soluté sont adsorbées verticalement comme c'est le cas des molécules possédant un seul groupe fonctionnel et d'autre part, quand les molécules se trouvent en compétition d'adsorption forte avec le solvant (Khoualdi & Medjani, 2016) .

a) Classe H

La partie initiale de l'isotherme est presque verticale, la quantité adsorbée apparaît importante à concentration quasiment nulle du soluté dans la solution. Ce phénomène se produit lorsque les interactions entre les molécules adsorbées et la surface du solide sont très fortes. L'isotherme de classe H est aussi observée lors de l'adsorption de micelles ou de polymères formées à partir des molécules de soluté (Khoualdi & Medjani, 2016).

b) Classe C

Les isothermes de cette classe se caractérisent par une partition constante entre la solution et le substrat jusqu'à un palier. La linéarité montre que le nombre de sites libres reste constant au cours de l'adsorption. Ceci signifie que les sites sont crées au cours de l'adsorption. Ce qui implique que les isothermes de cette classe sont obtenues quand les molécules de soluté sont capables de modifier la texture du substrat en ouvrant des pores qui n'avaient pas été ouverts préalablement par le solvant (Khoualdi & Medjani, 2016) .

I.8 Modélisation des isothermes d'adsorption

L'étude de l'adsorption est générée par plusieurs lois proposées, et de plusieurs modèles sont développés pour caractériser les processus impliqués. L'approche la plus adaptée a pour but de mesurer les isothermes d'adsorption qui présentent le partage à l'équilibre d'un soluté entre le solvant et la phase adsorbant. Parmi les modèles les plus connus, on peut citer :

I.8.1 Isotherme de Langmuir

Langmuir définit l'équilibre d'adsorption comme un processus dynamique entre les molécules arrivant à la surface et celles quittant la surface. Langmuir considère que les sites d'adsorption sont équivalents, que la surface est homogène et qu'il n'y pas d'interactions latérales entre les molécules adsorbées (Meriem ., 2017) .

Ce modèle est décrit par l'équation (I.2) exprimée comme suit :

$$q_e = q_{max} \cdot \frac{K_L . C_e}{(1+K_L C_e)} \qquad (I.2)$$

Avec :

q_e : la quantité adsorbée à l'équilibre ($mg.g^{-1}$) .

q_{max} :la quantité adsorbée maximale nécessaire pour la monocouche ($mg.g^{-1}$) .

K_L:la constante d'équilibre d'adsorption dite constante de Langmuir et elle est relative à l'énergie d'adsorption. Une valeur élevée de KL signifie une forte adsorption ($L.mg^{-1}$ou $L.mol^{-1}$)

C_e : est la concentration à l'équilibre ($mg. L^{-1}$).

I.8.2 Isotherme de Freundlich

Sont les modèles les plus simples et les plus rencontrés d'après (Van Bemmelen, 1888 ; Freundlich, 1909). Cette isotherme s'applique surtout dans le cas de l'adsorption multicouche et il traite la relation entre la quantité adsorbée et la concentration résiduelle du soluté.

La relation est donnée par l'équation ci-dessous :

$$q_e = K_F . C_e^{1/n} \qquad (I.3)$$

Le modèle de Freundlich est une équation empirique basée sur la sorption sur une surface hétérogène ou des sites supportant des surfaces variées affinités (Ahmad & Rahman, 2011) .

La forme logarithmique bien connue de Freundlich est donnée par l'équation suivante :

$$Ln\ q_e = Ln\ K_F + (\frac{1}{n}) . Ln\ C_e \qquad (I .4)$$

1.8.3 Isotherme de Langmuir–Freundlich

C'est une équation basée sur les isothermes de Langmuir et Freundlich. Elle décrit bien les surfaces hétérogènes. L'équation peut être écrite comme (Bouaghi, 2008) :

$$q_e = \frac{q_{max}(K_L.C_e).n}{[1+K_L.C_e]} \qquad (I.5)$$

Avec :

q_e: capacité à l'équilibre(mg. g^{-1}) .

q_{max}: capacité maximale d'adsorption(mg. g^{-1}) .

K_L: constante de Langmuir.

C_e: concentration à l'équilibre(mg. L^{-1}).

n : constante déduite du modèle de Freundlich.

I.8.4 Isotherme de Temkin

L'isotherme de Temkin est basée sur l'hypothèse que la chaleur d'adsorption diminue linéairement avec l'augmentation du taux de recouvrement de la surface solide. L'adsorption est caractérisée par une répartition uniforme des énergies de liaison, jusqu'à un maximum énergie de liaison (Ahmad & Rahman, 2011).

Le modèle de Temkin est établi comme (Ahmad & Rahman, 2011) :

$$q_e = (\frac{RT}{b_T}) \cdot \text{Ln}\,(A.C_e) = B.\text{Ln}A + B.\text{Ln}C_e \qquad \text{(I.6)}$$

$$\text{ou} \qquad B = (\frac{RT}{b_t})$$

Avec :

b_T : fonction de la chaleur d'adsorption **.**

A et B : constantes de l'isotherme de emkin (J. mol^{-1}**).**

R : constante des gaz parfaits (8.344 J. K^{-1} mol^{-1}).

C_e : concentration à l'équilibre (mg. L^{-1}).

T : température absolue (k).

I.8.5 Isotherme BET (Brunaue, Emmett, Teller)

On 1938 Brunauer, Emmett & Teller ont proposé une théorie de l'adsorption physique reposant sur les hypothèses suivantes (Richard, 2013) :

- Les interactions gaz-gaz et gaz-solide sont prises en considération et chaque molécule adsorbée dans une couche est un site d'adsorption pour une molécule de la couche suivante.

- La vitesse d'évaporation des molécules adsorbées dans une couche est égale à la vitesse de condensation sur la couche du dessous ;
- La chaleur d'adsorption dans les couches (à l'exception de la première) est égale à la chaleur de liquéfaction du gaz ;
- A la saturation on peut considérer que le nombre de couches adsorbées est infini.

L'isotherme de BET est représentée par l'équation suivante (Souad, 2015) :

$$\frac{q}{q_{max}} = \frac{K_B(\frac{C}{C_0})}{\left(1-\frac{C}{C_0}\right)[1+(K_B-1)\frac{C}{C_0}]} \quad (I.8)$$

Avec :

C: concentration au temps t (mg. L^{-1}).

C_0: Concentration initiale(mg. L^{-1}).

q_e: Capacité adsorbée au temps t (mg. g^{-1}).

q_{max}: Capacité de rétention mono moléculaire (mg. g^{-1}) .

K_B : Constante de B.E.T.

La linéarisation et la représentation graphique de l'équation en $C/q_e(C_0-C)$ en fonction de C/C_0 permettent de déterminer le K_B et q_{max} .

$$\frac{C}{q_e(C_0-C)} = \frac{1}{q_{max}\,K_B} + \frac{K_B-1}{q_{max}\,K_B}\left(\frac{C}{C_0}\right) \quad (I.9)$$

I.8.6 Isothermes de Dubinin-Radushkevish

Cette isotherme suppose une surface hétérogène et est exprimée comme suit (Bouaghi, 2008) :

$$q_e = q_m . e^{-\acute{K}.\varepsilon^2} \quad (I.10)$$

Avec :

ε : le potentiel de Polanyi ε = (R.T).ln (1 + 1/Ce) .

q_e: la quantité d'ions métallique adsorbés par poids spécifique (mg .g^{-1}),

q_m :la capacité d'adsorption (mg .g^{-1})

C_e : la concentration en équilibre des ions métalliques en solution ($mg.L^{-1}$)

ǩ : une constante liée à l'énergie d'adsorption ($mol^2 .k^{-1}J^{-2}$)

R : la constante universelle de gaz ($kJ .K^{-1} .mol^{-1}$)

T : température (K)

L'isotherme de D–R peut être exprimée par sa forme linéaire comme suit :

$$\ln (q_e)= \ln(q_m)- K' . \varepsilon^2 \quad (I.11)$$

K' est calculé à partir de la pente du tracé de ln q_e en fonction de ε^2, et l'énergie moyenne d'adsorption E ($kJ.mol^{-1}$) peut être obtenue à partir des valeurs de K' en employant équation suivante :

$$E = (-2K')^{-1/2} \quad (I.12)$$

Les constantes d'isotherme de Langmuir n'expliquent pas les propriétés du processus d'adsorption physique ou chimique. Cependant, l'énergie moyenne d'adsorption (E) calculée à partir de l'isotherme de D–R fournit des informations importantes au sujet de ces propriétés

- $E< 8KJ\ mol^{-1}$: la physisorption domine le mécanisme de sorption.

- E est entr 8 et 16 $KJ\ mol^{-1}$, l'échange ionique est le facteur dominant.

-$E>16KJ\ mol^{-1}$: la sorption est dominée par diffusion intra particule.

1.8.7 Isotherme de Toth

Toth a modifié l'isotherme de Langmuir pour minimiser l'erreur entre les données expérimentales de l'équilibre d'adsorption et les valeurs prédites. L'application de cette équation est plus convenable pour l'adsorption en multicouches, similairement à l'isotherme de BET qui est un type spécifique de l'isotherme de Langmuir (Bouaghi, 2008) .

L'équation peut être écrite comme :

$$q_e = \frac{C_e . q_{mt}}{(\frac{1}{K_T}+C_e.mT)^{1/mT}} \quad (I.13)$$

Avec :

K_T : la constante d'équilibre de toth.

C_e : la concentration à l'équilibre ($mg.l^{-1}$).

q_{mT} : la capacité maximun de toth ($mg.g^{-1}$).

mT : l'exposant du modèle de toth .

1.8.8 Isotherme d'Elovich

L'équation est basée sur le principe cinétique qui suppose que le nombre des sites d'adsorption augmente exponentiellement avec l'adsorption, ce qui implique une adsorption à multicouches (Bouaghi, 2008) .

Il est décrit par la relation :

$$\frac{q_e}{q_m} = K_e . C_e . \exp^{\frac{q_e}{q_m}} \qquad (I.14)$$

Avec :

K_e :la constante d'équilibre d'Elovich ($l.mg^{-1}$)

q_m : la capacité maximum d'Elovich ($mg.g^{-1}$)

Si l'adsorption est décrite par l'équation d'Elovich, la constante d'équilibre et la capacité maximum d'Elovich peuvent être calculées à partir du tracé de ln (q_e /C_e) en fonction du q_e.

I.9 Modélisation des isothermes d'adsorption du mélange binaire

De nombreux modèles sont proposés pour représenter l'adsorption multi-constituante à l'équilibre, notamment les modèles de Langmuir généralisé, de Langmuir généralisé modifié et le modèle théorique de la solution adsorbée idéale (I.A.S.T). Chacun de ces modèles a été établi à partir d'hypothèses effectuées sur les mécanismes d'adsorption. Ainsi, compte tenu de la forme des isothermes obtenues en mono constituant et en mélange binaire (Abdellah&Hasna , 2017).

I.9.1 Modèle de Langmuir généralisé

Le modèle de Langmuir généralisé repose sur les mêmes hypothèses que le modèle de Langmuir à un seul constituant : une molécule par site, surface homogène, pas d'interactions latérales. Au cours d'une adsorption multi-constituants, ce modèle simple permet de représenter les isothermes de chaque espèce à partir des coefficients de Langmuir calculés pour des produits purs. Le modèle de Langmuir généralisé est représenté sous la forme suivante (Abdellah&Hasna., 2017) :

$$q_{e,i} = \frac{q_{max,i} . b_i . C_{e,i}}{1 + \sum_j . b_j . C_{e,j}} \qquad (I.15)$$

Avec :

$\mathbf{q_{max,\,i}}$: quantité d'adsorption maximale pour l'espèce i en produit pur ($mg.g^{-1}$).

$\mathbf{b_i}$: constante de Langmuir de l'espèce i en produit pur ($l.mg^{-1}$).

$\mathbf{q_{e,i}}$: quantité de l'espèce i adsorbée à l'équilibre en compétition ($mg.g^{-1}$).

$\mathbf{C_{e,i}}$: quantité de l'espèce i en solution à l'équilibre en compétition ($mg.l^{-1}$).

Pour des mélanges binaires, différents auteurs ont repris l'équation de Langmuir généralisée sous la forme suivante :

$$q_{e,A} = q_{max,A} \frac{b_A.C_{e,A}}{1+b_A.C_{e,A}.b_B.C_{e,B}} \quad (I.16)$$

$$q_{e,B} = q_{max,B} \frac{b_B.C_{e,B}}{1+b_A.C_{e,A}+b_B.C_{e,B}} \quad (I.17)$$

Avec :

$\mathbf{q_{max,A}}$: Quantité d'adsorption maximale pour l'espèce A en produit pur ($mg.g^{-1}$).

$\mathbf{b_A}$: Constante de Langmuir de l'espèce A en produit pur ($l.mg^{-1}$).

$\mathbf{q_{eA}}$: Quantité de l'espèce A adsorbée à l'équilibre en compétition ($mg.g^{-1}$).

$\mathbf{C_{Ea}}$ **:** Quantité de l'espèce i en solution à l'équilibre en compétition ($mg.l^{-1}$).

$\mathbf{q_{max,B}}$: Quantité d'adsorption maximale pour l'espèce i en produit pur ($mg.g^{-1}$).

$\mathbf{b_B}$ **:** Constante de Langmuir de l'espèce B en produit pur ($l.mg^{-1}$).

$\mathbf{q_{eB}}$: Quantité de l'espèce B adsorbée à l'équilibre en compétition ($mg.g^{-1}$).

$\mathbf{C_{eB}}$: Quantité de l'espèce B en solution à l'équilibre en compétition ($mg.\,l^{-1}$)

I.10 Adsorption statique multi-composant

L'air à traiter est souvent un mélange dont plusieurs des constituants sont susceptibles d'être adsorber. La capacité d'adsorption d'un polluant peut, de ce fait, être affectée par des phénomènes d'adsorption compétitive. C'est le cas notamment de l'adsorption de vapeur de solvant en présence d'humidité. La sélectivité de l'adsorbant est alors déterminante. Elle représente l'affinité relative entre l'adsorbant et chacun des composés adsorbables. Par exemple, au sein d'une même famille de molécule, la capacité augmente avec la masse moléculaire. De grosse molécules pourront ainsi venir déplacer de plus petit (élution préférentielle). La cinétique d'adsorption et la capacité de différents adsorbants ont été très étudiées. Dans la plupart des cas, les études ont été limitées aux faibles températures (273-400K) et aux faibles pressions partielles (0-1bar). La plupart des données publiées

correspondent à des isothermes de corps purs. On dispose de très peu de données d'adsorption de mélanges multi-composants qui correspondent plus à la réalité industriel (Etudes, 2017) .

I.10.1 Modèles mono-constituant étendus en modèles multi-constituants

I.10.1.1 Modèles de Freundlich étendu

Les différents modèles existants dit de « Freundlich étendu » sont tous empiriques et sont généralement de la forme suivante (Boulkrah, 2008):

$$q_{e,i} = K^0_{F,i} . C_{e,i} . \left(\sum_{j=1}^{N} \partial_{i,j} . C_{e,j}\right)^{n^0_{i-1}} \quad \text{(I.18)}$$

Avec :

$\boldsymbol{\partial_{i,j}}$: Coefficient de compétition de l'espèce i en présence de l'espèce j.

$\mathbf{K^0_{F,i}}$ et $\mathbf{n^0_i}$: Coefficient de l'isotherme (mono-constituant) de Freundlich ($q_e = K_F . C_e^n$) de l'espèce.

I.10.1.2 Modèles de Langmuir étendu

Bien que l'isotherme mono-constituant de Langmuir soit largement utilisé, sa linéarisation donne souvent des résultats médiocres que certains auteurs expliquent par l'existence de plusieurs types d'adsorption (Boulkrah, 2008).

La forme étendue du modèle de Langmuir pour la compétition entre N espèces est la suivante :

$$q_{e,i} = q^0_{m,i} \left[\frac{K^0_{L,i} . C_{e,i}}{1 + \sum_{j=1}^{N} K^0_{L,j} . C_{e,j}}\right] \quad \text{(I.19)}$$

Avec :

$\mathbf{q^0_{m,i}}$ **et** $\mathbf{K^0_{L,i}}$**:** Paramètres de l'isotherme mono-constituant de Langmuir

I.10.1.3 Modèles étendus pour toute isotherme

D'autres modèles sont cités dans la littérature, par exemple par Hinz (2001) qui propose une équation qui pourrait décrire chaque type d'isotherme, ou encore sous des formes intégrant l'isotherme expérimental q_e vs C_e. A titre d'exemple, on peut citer le modèle de Langmuir-Freundlich étendu (Boulkrah, 2008) .

$$q_{e,i} = q^0_{m,i} . \frac{K^0_{L,j} . C_{ei}^{ni}}{1 + \sum_{j=1}^{N} K^0_{L,j} . C_{ej}^{nj}} \quad \text{(I.20)}$$

I.10.2 Modèles d'adsorption en multi-constituants

Dans les cas de mélanges binaires ou multi-constituants, la mesure expérimentale des isothermes d'adsorption est généralement très délicate en raison du nombre de variables mises en jeu. C'est pourquoi, la prédiction des équilibres binaires ou multi- constituants à partir des équilibre des corps purs s'est beaucoup développée (Ettlili, 2018). On trouve par exemple une équation de Langmuir dérivée de l'approche cinétique et une extension de la théorie du potentiel pour décrire l'adsorption de plusieurs composés (Clausse, 2006).

I.10.2.1 Modèle de Langmuir

Ce modèle s'écrit de la manier suivante :

$$n_i = n_{i,sat} \frac{b_i.P_i}{1+\sum_j b_j.p_j} \qquad (I.21)$$

Cependant les hypothèses permettant d'aboutir à cette relation étant les mêmes que pour un constituant seul (pas d'interactions entre les molécules adsorbées, etc.) il arrive souvent que les équilibres d'adsorption ne soient pas décrits correctement par cette équation. On obtient alors les équations étendues de Langmuir-Freundlich et de Toth qui permettent de prédire avec une exactitude accrue les équilibres d'adsorption (Clausse, 2006).

I.11 Cinétique d'adsorption

Dans un système d'adsorption, la vitesse de réaction aboutissant à l'équilibre thermodynamique entre l'adsorbant en phase liquide et l'adsorbant fixé sur le solide.

La définition de la cinétique d'adsorption est le taux d'approche de l'équilibre, l'équilibre d'adsorption n'apparait pas instantanément car le taux d'adsorption est habituellement limité par le transport de masse suivant mécanismes et dépond à la fois des propriétés de l'adsorbant et de l'adsorbant.

Le but de cette étude est de déterminer les quantités fixées de l'adsorbat choisi depuis sa mise en contact jusqu'à l'atteinte de l'équilibre thermodynamique. Tandis que les prélèvements des échantillons sont réalisés avec des intervalles de temps variables (Mohamed&Ismail., 2017).

I.11.1 Modèle cinétique du pseudo premier ordre

Il a été supposé dans ce modèle que la vitesse d'adsorption à l'instant t est proportionnelle à la différence entre la quantité adsorbée à l'équilibre q_e et la quantité q_t adsorbée à cet instant et que l'adsorption est réversible. La constante de vitesse d'adsorption du premier ordre est déduit à partir du modèle établi par Lagergreen (Aarfane , 2014).

La loi de vitesse s'écrit :

$$\frac{dq_t}{dt} = K_1(q_e - q_t) \qquad (I.22)$$

Avec :

q_e: quantité d'adsorbat à l'équilibre, par gramme d'adsorbant (mg .g^{-1})

t : le temps de contact (min)

$\boldsymbol{K_1}$: Constantes de vitesse d'adsorption pour le premier ordre(min^{-1})

L'intégration de l'équation (I.22) donne :

$$\log(q_e - q_t) = \log(q_e) - \frac{K_1}{2{,}303}.t \qquad (I.23)$$

I.11.2 Modèle de la cinétique du pseudo- second ordre

L'équation du pseudo-second ordre est souvent utilisée avec succès pour décrire la cinétique de la réaction de fixation des polluants sur l'adsorbant. Le modèle du pseudo-second ordre permet de caractériser la cinétique d'adsorption en prenant en compte à la fois le cas d'une fixation rapide des solutés sur les sites les plus réactifs et celui d'une fixation lente sur les sites d'énergie faible (Aarfane , 2014) :

$$\frac{dq_t}{dt} = K_2(q_e - q_t)^2 \qquad (I.24)$$

L'intégration de l'équation (eq) donne :

$$\frac{t}{q_t} = \frac{1}{K_2.q_e^2} + \frac{1}{q_e}.t \qquad (I.25)$$

Avec :

K_2: Constantes de vitesse d'adsorption pour le pseudo second ordre (g.mol^{-1}.min^{-1}).

q_e: Quantité d'adsorbat à l'équilibre par gramme d'adsorbant (mg.g^{-1}).

La quantité adsorbée à l'équilibre q_e et la constante du pseudo-second ordre K_2 peuvent être déterminées expérimentalement à partir de la pente et de l'ordonnée à l'origine de t/q_t en fonction de t.

I .12 Généralités sur les adsorbants

I.12.1 Définition d'un adsorbant

Les adsorbants sont des solides microporeux présentant des surfaces par unité de masse importantes (de 100 m^2 /g et jusqu'à ou plus de 1000 m^2 /g) donc ils sont caractérisés par leurs propriétés extérieures telles que leur surface spécifique et leur porosité.

I.12.2 Types d'adsorbants

I.12.2.1 Charbon actifs

Les charbons actifs sont préparés par pyrolyse d'une matière contenant du carbone, charbon ou matériau végétal, pour conduire à un charbon de bois qui est ensuite oxydé par la vapeur d'eau dans des conditions contrôlées pour créer une structure microporeuse. Il existe plusieurs centaines de qualités de charbons actifs, suivant le précurseur et les conditions de traitement. On peut aussi trouver des charbons actifs dits « chimiques », car activés à chaud en présence d'agents chimiques déshydratants, acide phosphorique ou chlorure de zinc. Les charbons actifs sont des adsorbants organophiles amorphes. Leur structure n'est donc pas régulière, contrairement à un cristal. Cette structure amorphe se traduit par une répartition continue de taille de pores dont l'étalement (l'écart entre les plus petites et les plus grandes valeurs) peut atteindre plusieurs ordres de grandeur (Krou, 2010) .

I.12.2.2 Adsorbants minéraux

Les adsorbants minéraux peuvent exister à l'état naturel ou synthétisés.

I.12.2.3 Les argiles activées

Les argiles activées sont des alumino-silicates de formule brute proche des zéolithes, mais de structure cristalline différente. Ce sont des produits naturels, utilisés surtout pour le séchage (Krou, 2010).

I.12.2.4 Les gels de silice

Les gels de silice sont préparés à partir de $Si(OH)_4$ en phase aqueuse, obtenu par acidification d'un silicate de sodium, ou bien à partir d'un sel de silice (suspension dans un liquide, tel que l'eau, de microparticules (20 à 100 nm), appelées micelles, stables car trop petites pour décanter, ou bien par hydrolyse d'un alcoxy-silane. La solution fluide obtenue ne tarde pas à polymériser, ce qui conduit à un gel qui conserve sa structure lâche après rinçage et séchage. Les groupements Si-OH conduisent à des liaisons hydrogène. Il existe deux types de gels de silice : les microporeux, assez hydrophiles, et les macroporeux, versatiles, qui diffèrent par la taille des pores comme le nom l'indique (Krou, 2010).

I.12.2.5 Les zéolithes

Les zéolithes sont des alumino-silicates cristallisés microporeux de formule globale (AlO_2M, $nSiO_2$) où M représente le plus souvent un métal alcalin ou alcalino-terreux et $n \geq 1$. Il existe plus de 100 espèces de zéolithes, différant par la valeur de n et la structure

cristallographique. La présence de cations dans les micropores génère des champs électriques de l'ordre de 10^{10} V.m^{-1}, ce qui fait de ces corps de puissants adsorbants polaires.

I.12.2.6 Les alumines activées

Les alumines activées sont obtenues par thermolyse flash du trihydroxyde d'aluminium $Al(OH)_3$ qui conduit à un produit de composition approximative Al_2O_3, 0,5 H_2O, possédant une structure poreuse résultant du départ de molécules d'eau. La surface des pores est couverte de groupements Al-OH, et l'adsorption se fait préférentiellement par liaison hydrogène. Les alumines activées sont des adsorbants amorphes, moyennement polaires et hydrophiles (Krou, 2010).

I.13 Application

Les nombreuses applications techniques de l'adsorption résultent de trois caractéristiques qui la différencient des autres procédés de séparation, à savoir (Sedira, 2013) :

- La rétention de très petites particules, comme par exemple les colloïdes.
- La rétention des composants à très faible concentration, par exemples des impuretés ou des molécules et ions métalliques qui confèrent au produit couleurs, odeurs, ou saveurs désagréables, voire une toxicité.
- La sélectivité de l'adsorbant par apport à certains constituants du mélange. Parmi

Les applications, on cite :

- Le séchage, purification et désodorisation des gaz.
- Le raffinage des produits pétroliers.
- La catalyse de contact.
- La récupération des solvants volatils et de l'alcool dans le processus de fermentation.
- La décoloration des liquides.
- La chromatographie gazeuse. La déshumidification et la désodorisation de l'aire.

II.14 Conclusion

Au cours de ce chapitre, il nous a été possible de définir l'adsorption comme étant essentiellement une réaction de surface et de différencier entre la physisorption et la chimisorption. Nous avons présenté aussi les différents types d'isotherme d'adsorption et exposé les modèles les plus utilisés pour illustrer ces isothermes. Enfin, nous avons cité quelques matériaux adsorbants provenant de différentes origines (Sedira, 2013) .

Chapitre II :
la machine vecteurs de support

II.1 Introduction

Les machines à vecteurs de support, ou SVM (Support Vector Machines), sont des méthodes relativement récentes de résolution des problèmes de régression est calssification. Les SVM sont introduit par Vladimir Vapnik au début des années 90,(Guyon , 1993) elles reposent sur une théorie mathématique solide à l'inverse des méthodes de réseaux de neurones. Cet outil devient largement utilisé dans de nombreux types d'applications et s'avère un concurrent sérieux des algorithmes les plus performants (agrégation de modèles). L'introduction de noyaux, spécifiquement adaptés à une problématique donnée, lui confère une grande flexibilité pour s'adapter à des situations très diverses (reconnaissance de formes, de séquences génomiques, de caractères, détection de spams, diagnostics...). À noter que, sur le plan algorithmique, ces algorithmes sont plus pénalisés par le nombre d'observations, c'est-à-dire le nombre de vecteurs supports potentiels, que par le nombre de variables. Néanmoins, des versions performantes des algorithmes permettent de prendre en compte des bases de données volumineuses dans des temps de calcul acceptables. Elle a pu, en quelques années depuis sa proposition, conquérir la plupart des autres systèmes d'apprentissage dans une grande variété de domaines d'application (Chapel & Deffuant, 2010).

II.2 Historique

Les séparateurs à vastes marges reposent sur deux idées clés : la notion de marge maximale et la notion de fonction noyau. Ces deux notions existaient depuis plusieurs années avant qu'elles ne soient mises en commun pour construire les SVM.

L'idée des hyperplans à marge maximale a été explorée dès 1963 par Vladimir Vapnik et A. Lerner, et en 1973 par Richard Duda et Peter Hart dans leur livre Pattern Classification. Les fondations théoriques des SVM ont été explorées par Vapnik et ses collègues dans les années 70 avec le développement de la Théorie de Vapnik-Chervonenkis, et par Valiant.

L'idée des fonctions noyaux n'est pas non plus nouvelle : le théorème de Mercer date de 1909, et l'utilité des fonctions noyaux dans le contexte de l'apprentissage artificiel a été montrée dès 1964 par Aizermann, Bravermann et Rozoener.

Ce n'est toutefois qu'en 1992 que ces idées seront bien comprises et rassemblées par Boser, Guyon et Vapnik dans un article, qui est l'article fondateur des séparateurs à vaste marge. L'idée des variables ressorts, qui permet de résoudre certaines limitations pratiques importantes, ne sera introduite qu'en 1995. À partir de cette date, qui correspond à la publication du livre de Vapnik, les SVM gagnent en popularité et sont utilisés dans de nombreuses applications .Un

brevet américain sur les SVM est déposé en 1997 par les inventeurs originaux (Hasan & Boris, 2006).

II.3 Définition SVM

. Les SVM sont un ensemble de techniques d'apprentissage supervisé de nouvelle génération basés sur les progrès récents de la théorie de l'apprentissage statistique, ils sont basés sur le principe de la maximisation de la marge (séparation des classes . Principalement conçues pour résoudre ces problèmes (Hasan & Boris, 2006).

- Des problèmes de discrimination, permettant de décider à quelle classe appartient un échantillon (Hasan & Boris, 2006).

- Cependant, ils peuvent aussi résoudre des problèmes de régression visant à prédire la valeur numérique d'une variable (Scholkopf et Smola, 2001).C'est-à-dire de prédiction d'une variable continue en fonction d'autres variables, Comme c'est le cas par exemple dans de la prédiction de consommation électrique en fonction de la période de l'année, de la température, etc (Hasan & Boris, 2006).

La résolution de ces deux problèmes se base sur la construction d'une fonction *h* qui, à un vecteur d'entrée *x*, fait correspondre une sortie *y* (*y* = *h*(*x*)) (Hasan & Boris, 2006).

Les SVM constituent la forme la plus connue parmi les méthodes à noyaux. Ce sont des class ifieurs qui reposent sur deux idées clés :

- ✓ La notion de marge maximale.
- ✓ La notion de Fonction noyau.

II.4 Principe de fonctionnement général de SVM

II.4.1 Notions de base

II.4.1.1 Hyperplan

On appelle hyperplan séparateur un hyperplan qui sépare les deux classes, en particulier il sépare leurs points d'apprentissage. Comme il n'est en générale pas possible d'en trouver un, on se contentera donc de chercher un hyperplan discriminant qui est une approximation au sens d'un critère à fixer (maximiser la distance entre ces deux classes) c).

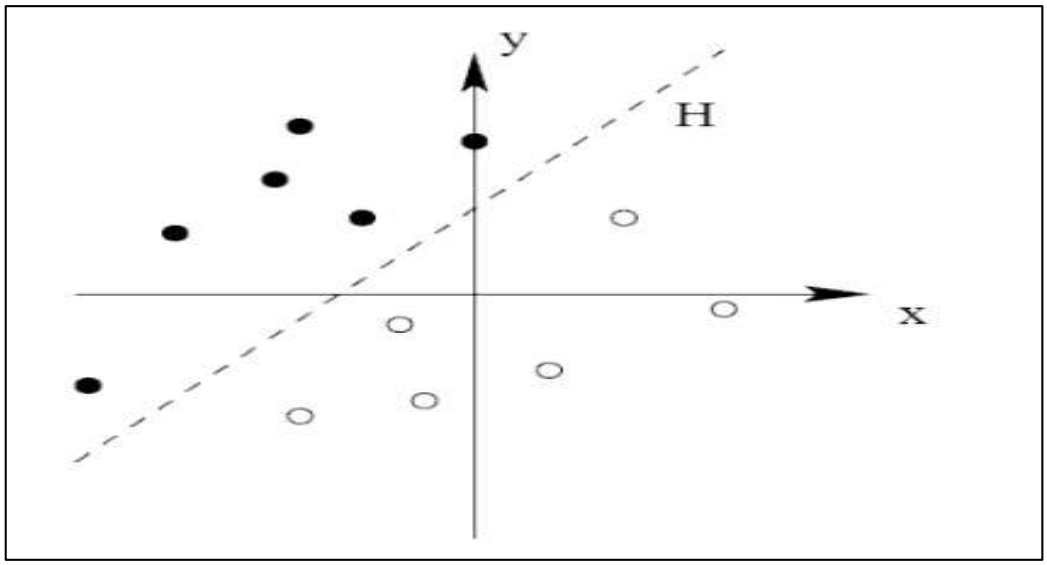


Figure II.1: Exemple d'un hyperplan sépare deux ensembles de points (Hasna&Bouris,2006).

II.4.1.2 Vecteurs de support

Ce sont les points de la frontière entre les deux classes des données parmi l'ensemble total d'apprentissage, ces points sont appelés vecteurs de support

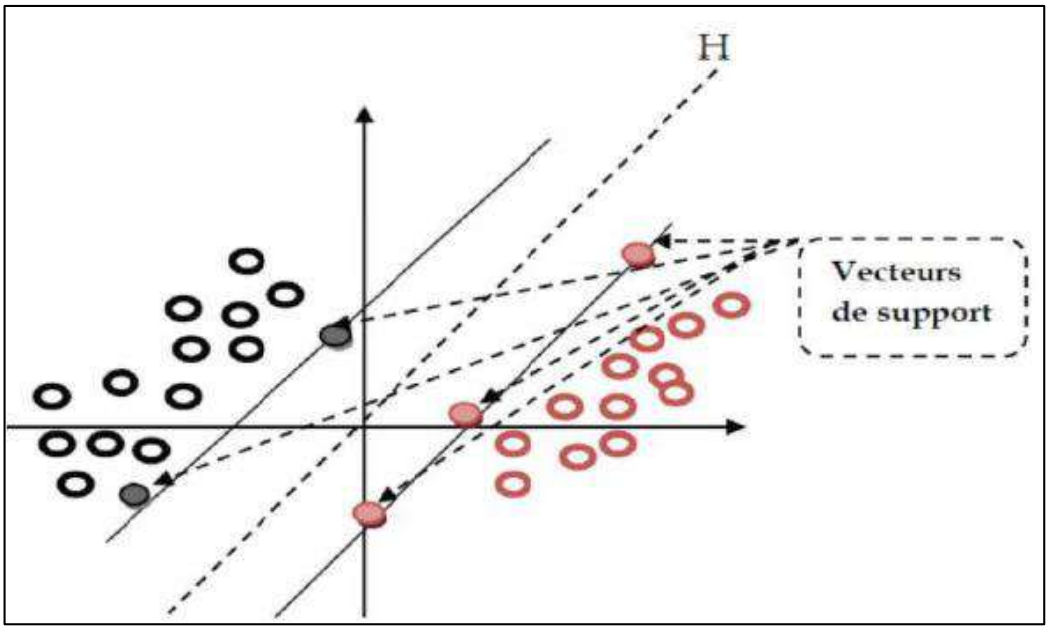


Figure II.2: Exemple de vecteurs de support (Hasna&Bouris,2006)

Il est évident qu'il existe une multitude d'hyperplan valide mais la propriété remarquable des SVM est que cet hyperplan doit être optimal. Nous allons donc en plus chercher parmi les hyperplans valides, celui qui passe ≪ au milieu ≫ des points des deux classes d'exemples. Intuitivement, cela revient à chercher l'hyperplan le ≪ plus sur ≫. En effet,

supposons qu'un exemple n'ait pas été décrit parfaitement, une petite variation ne modifiera pas sa classification si sa distance a l'hyperplan est grande. Formellement, cela revient à chercher un hyperplan dont la distance minimale aux exemples d'apprentissage est maximale (Hasna,2006).

II.4.1.3 Marge

Il existe une infinité d'hyperplans capable de séparer parfaitement les deux classes d'exemples. Le principe des SVM est de choisir celui qui va maximiser la distance minimale entre l'hyperplan et les exemples d'apprentissage (i.e. la distance entre l'hyperplan et les vecteurs de support), cette distance est appelée la marge(Boris, 2006)

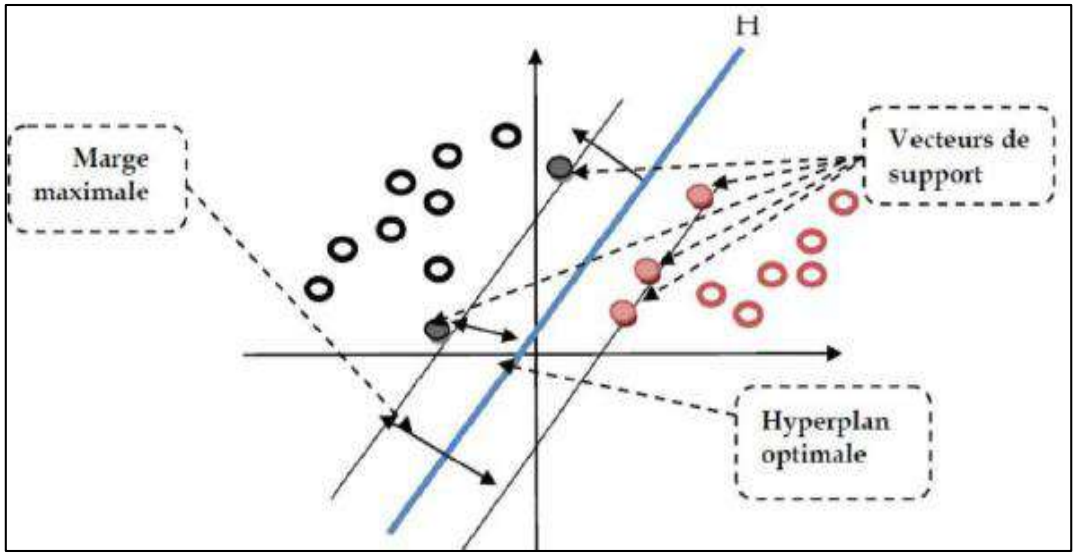


Figure II.3: Exemple de marge maximal (hyperplan valide) (Boris, 2006).

Intuitivement, le fait d'avoir une marge plus large procure plus de sécurité lorsque l'on classe un nouvel exemple. De plus, si l'on trouve le classificateur qui se comporte le mieux vis-à-vis des données d'apprentissage, il est clair qu'il sera aussi celui qui permettra au mieux de classer les nouveaux exemples. Dans le schéma qui suit, la partie droite nous montre qu'avec un hyperplan optimal, un nouvel exemple reste bien classé alors qu'il tombe dans la marge. On constate sur la partie gauche qu'avec une plus petite marge, l'exemple se voit mal classé

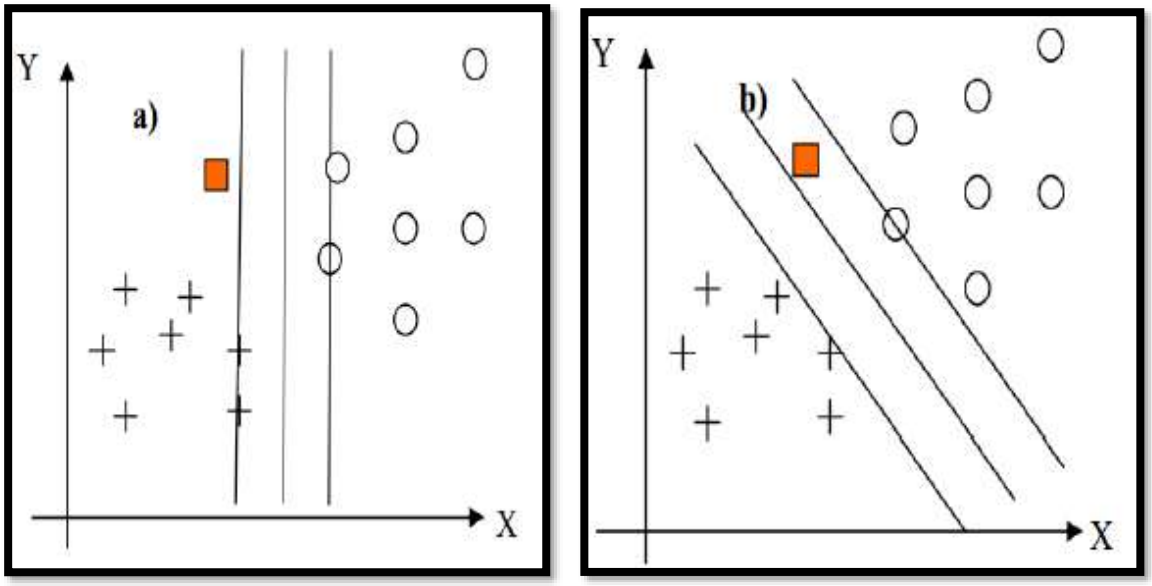


Figure II.4: a) Hyperplan avec faible marge, **b)** Meilleur hyperplan séparateur (Algerienne , 2010)

En général, la classification d'un nouvel exemple inconnu est donnée par sa position par rapport à l'hyperplan optimal. Dans le schéma suivant, le nouvel élément sera classé dans la catégorie des « + ».

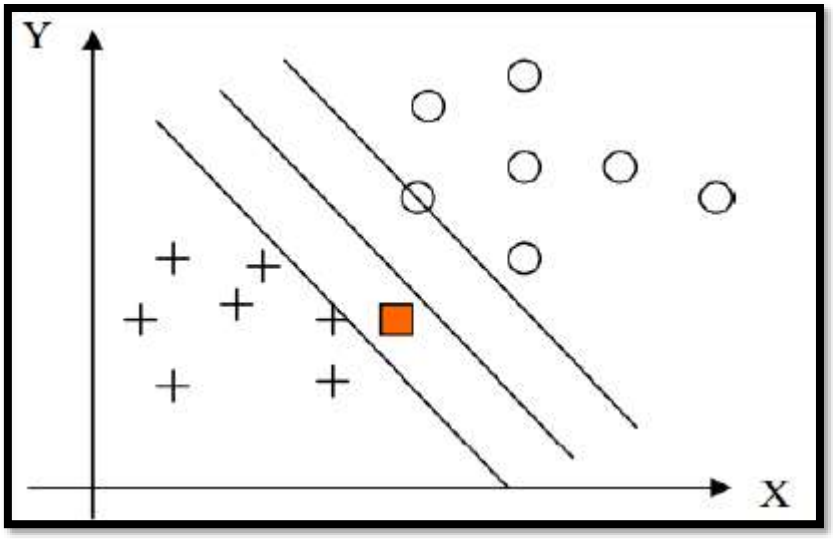
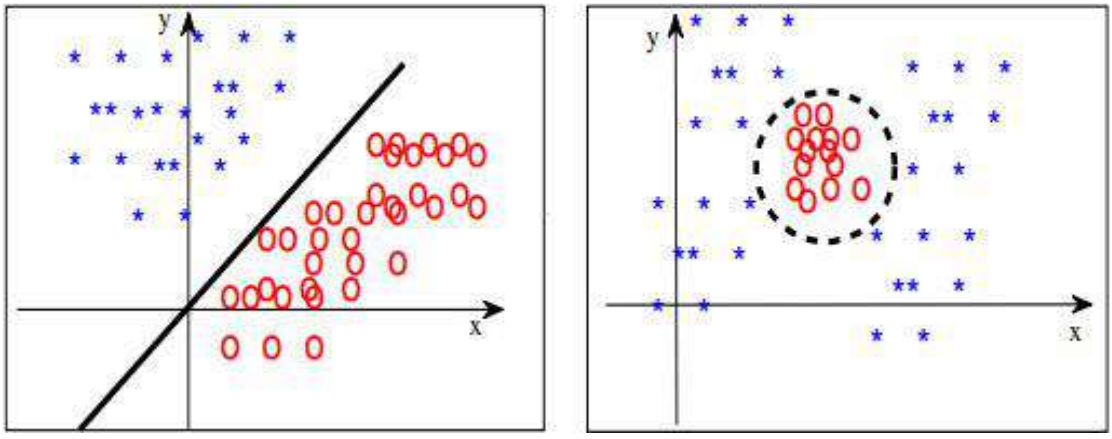


Figure II.5: Exemple de classification d'un nouvel élément.

II.7 Linéarité et non-linéarité

II.7.1 Cas linéairement séparable

Parmi les modèles des SVM, on constate les cas linéairement séparables et les cas non linéairement séparable (Algerienne, 2010).



(a) **(b)**

Figure II.6: (a) linéairement séparable et **(b)** cas non linéairement séparable (Algerienne, 2010).

II.7.1 Cas linéairement séparable

Les cas linéairement séparables sont les plus simples de SVM car ils permettent de trouver facilement le classificateur linéaire (Algerienne , 2010).

II.7.2 Cas non linéaire

Pour surmonter les inconvénients des cas non linéairement séparable, l'idée des SVM est de changer l'espace des données. La transformation non linéaire des données peut permettre une séparation linéaire des exemples dans un nouvel espace. On va donc avoir un changement de dimension. Cette nouvelle dimension est appelé « espace de re-description ». En effet, intuitivement, plus la dimension de l'espace de re-description est grande, plus la probabilité de pouvoir trouver un hyperplan séparateur entre les exemples est élevée. Ceci est illustré par le schéma suivant (Algerienne, 2010) :

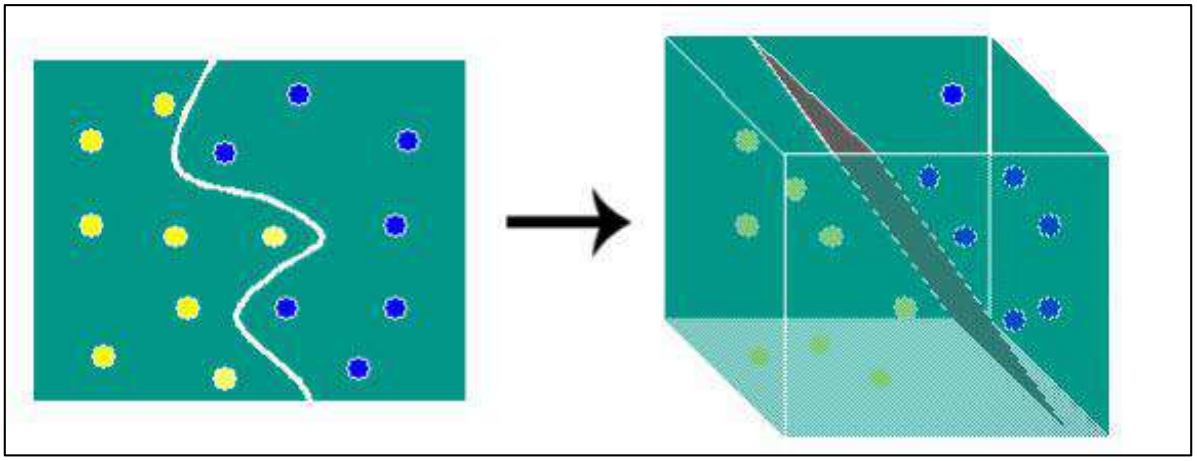


Figure II.7: exemple de changement de l'espace de données .

On a donc une transformation d'un problème de séparation non linéaire dans l'espace de représentation en un problème de séparation linéaire dans un espace de re-description de plus grande dimension. Cette transformation non linéaire est réalisée via une fonction noyau.

II.8. Architecture de SVM

L'architecture du modèle SVM n'est pas déterminée a priori. Les vecteurs d'entrée supportant la structure du modèle sont sélectionnés via un processus d'apprentissage du modèle illustré à (la figure II.11) (Qu ., 2020) .

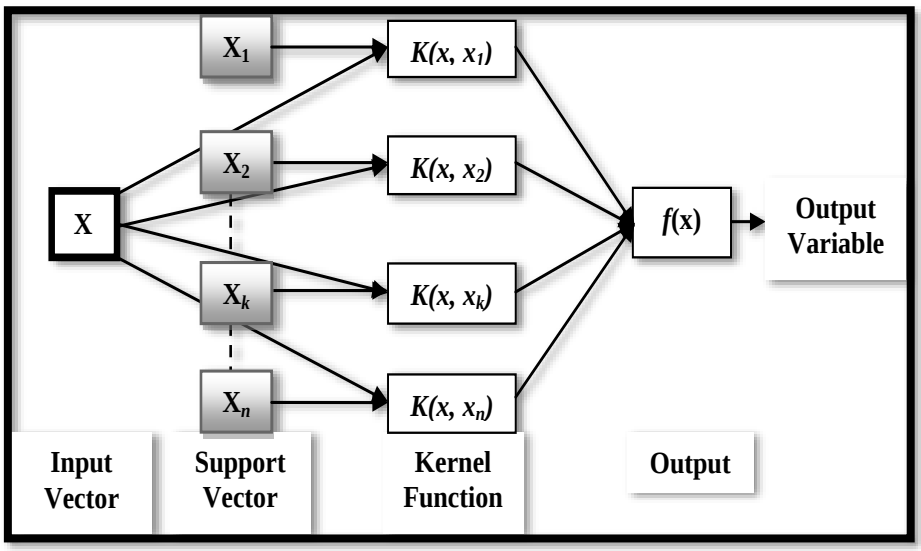


Figure II.8: Architecture de SVM**.**

II.9 Apprentissage statistique

La théorie d'apprentissage statistique étudie les propriétés mathématiques des machines d'apprentissage. Ces propriétés représentent les propriétés de la classe de fonctions ou modèles que peut implémenter la machine. L'apprentissage statistique utilise un nombre limité d'entrées (appelées exemples) d'un système avec les valeurs de leurs sorties pour apprendre une fonction qui décrit la relation fonctionnelle existante, mais non connue, entre les entrées et les sorties du système.

Effectuer une classification consiste à déterminer une règle de décision capable, à partir d'observations externes, d'assigner un objet à une classe parmi plusieurs. Le cas le plus simple consiste à discriminer deux classes. D'une manière plus formelle, dans le cas d'une classification binaire cette étiquette est soit +1 ou -1. L'ensemble des exemples et leurs

étiquettes correspondantes est appelés ensemble d'apprentissage. Une machine d'apprentissage est dite efficace quand elle apprend de l'ensemble d'entrainement une fonction qui minimise les erreurs de classification sur l'ensemble lui-même. Soit D = {(x1, y1),...(x_n, y_n)}, l'ensemble des exemples d'entrainement ; avec $X_i \in \Re^m$ et $Y_i = \pm 1$ et soit $f(x)$ fonction apprise par la machine d'apprentissage. On appelle risque empirique $\Re_{emp}[f]$ le taux d'erreurs effectuées par la fonction f sur l'ensemble d'entrainement D.

$$\Re_{emp} = \frac{1}{N}\sum_{i=1}^{n} L(y_i, f(x_i))$$

Avec (II .1)

$$L = \begin{cases} 1 \; si \; y = f(x_i) \\ 0 \; sinon \end{cases}$$

Puisque l'ensemble D ne représente qu'une simple partie de tout l'espace d'exemples, la fonction apprise f, qui minimise le risque empirique, peut se comporter mal avec les nouveaux exemples non vus à l'entrainement. C'est un phénomène très connu en apprentissage automatique appelé le sur-apprentissage ou apprentissage par cœur. Pour garantir que f, prenne en charge même les exemples non jamais vus, il faut contrôler sa capacité de généralisation mesurée souvent sur un autre ensemble d'exemples appelé ensemble de test réservé uniquement pour tester la machine apprise. La fonction f recherchée doit donc minimiser les erreurs de classification sur les deux ensembles : d'entrainement et de test.

Trouver la fonction optimale f revient toujours à un problème d'optimisation, ce qui explique la forte relation entre l'apprentissage et l'optimisation. Avant de rechercher la fonction f, il faut définir son type puis rechercher ses paramètres.

Pour les machines à vecteurs support, la fonction recherchée est linéaire. Par conséquent, SVM est un système d'apprentissage qui utilise un espace d'hypothèses de fonctions linéaires dans un espace de caractéristiques de grande dimension. Cette stratégie d'apprentissage proposée par Vapnik et ses co-auteurs (Vapnik, 2013)est une méthode très efficace. De plus, vous devez en savoir plus (Vapnik, 2013).

Depuis que la proposition a été faite, il a été en mesure de conquérir la plupart des autres systèmes d'apprentissage dans un large éventail d'applications.

II.10 SVM binaires

Le cas le plus simple est celui où les données d'entrainement viennent uniquement de deux classes différentes (+1 ou -1), on parle alors de classification binaire. L'idée des SVM est

de 8rechercher un hyperplan (droite dans le cas de deux dimensions) qui sépare le mieux ces deux classes (figure II.3) (Mohan & Singh, 2002).

.

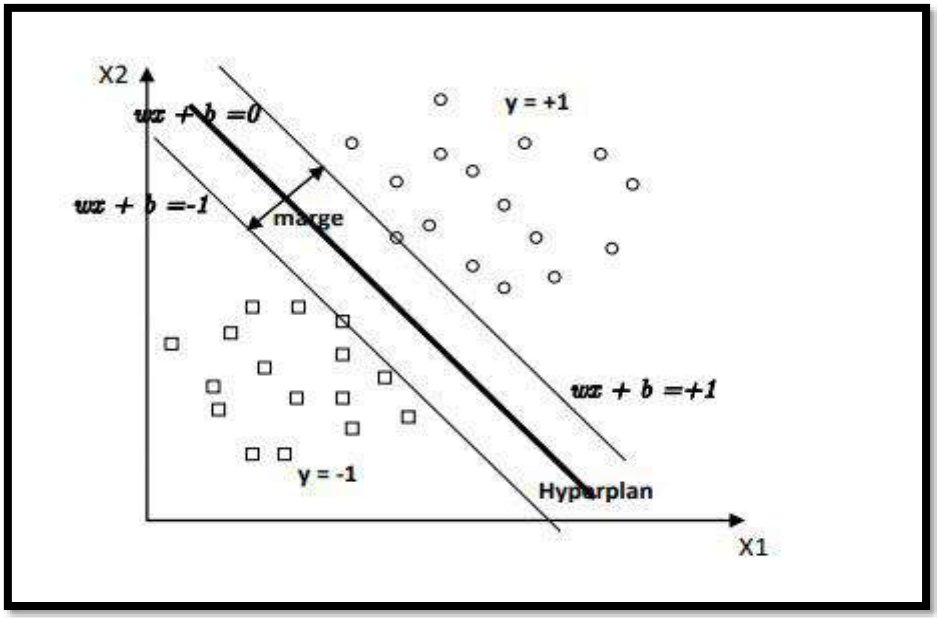


Figure II.9: SVM binaire.

Si un tel hyperplan existe, c'est-à-dire si les données sont linéairement séparables, on parle d'une machine à vecteur support à marge dure (Hard margin)

II.10.1 SVM à marge dure

L'hyperplan séparateur est représenté par l'équation (II.2) suivante (Mohan & Singh, 2002) :

$$H(x) = W^T x + b \qquad \text{(II.2)}$$

Où w est un vecteur de m dimensions et b est un terme. La fonction de décision, pour un exemple x, peut être exprimée comme suit :

$$\begin{cases} \text{Clase} = 1 \ \ \text{Si } H(x) > 0 \\ \text{Clase} = -1 \ \ \text{Si } H(x) < 0 \end{cases} \qquad \text{(II.3)}$$

Puisque les deux classes sont linéairement séparables, il n'existe aucun exemple qui se situe sur l'hyperplan, c-à-d qui satisfait H(x) = 0 . Il convient alors d'utiliser la fonction de décisions suivante :

$$\begin{cases} \text{Classe} = 1 \ \ \text{Si } \ H(x) > 1 \\ \text{Classe} = -1 \ \ \text{Si } (H) < -1 \end{cases} \qquad \text{(II.4)}$$

Les valeurs +1 et -1 à droite des inégalités peuvent être des constantes quelconques +a et -a, mais en divisant les deux parties des inégalités par a, on trouve les inégalités précédentes qui sont équivalentes à l'équation (II.5) :

$$y_i(W^T x_i + b) \geq 1, i = 1 \ldots . n \qquad \text{(II.5)}$$

L'hyperplan $w^T x + b = 0$ représente un hyperplan séparateur des deux classes, et la distance entre cet hyperplan et l'exemple le plus proche s'appelle la marge (figure II.3). La région qui se trouve entre les deux hyperplans $w^T x + b = -1$ et $w^T x + b = +1$ est appelée la région de généralisation de la machine d'apprentissage. Plus cette région est importante, plus est la capacité de généralisation de la machine. La maximisation de cette région est l'objectif de la phase d'entrainement qui consiste, pour la méthode SVM, à rechercher l'hyperplan qui maximise la région de généralisation c-à-d la marge. Un tel hyperplan est appelé "hyperplan de séparation optimale". En supposant que les données d'apprentissage ne contiennent pas des données bruitées (mal-étiquetées) et que les données de test suivent la même probabilité que celle des données d'entraînement, l'hyperplan de marge maximale va certainement maximiser la capacité de généralisation de la machine d'apprentissage.

La détermination de l'hyperplan optimal passe par la détermination de la distance euclidienne minimale entre l'hyperplan et l'exemple le plus proche des deux classes. Puisque le vecteur w est orthogonal sur l'hyperplan séparateur, la droite parallèle à w et reliant un exemple x à l'hyperplan est donnée par la formule (II.6) :

$$\frac{\mathrm{aw}}{\|\mathrm{w}\|} + \mathrm{x} = 0 \qquad \text{(II.6)}$$

Où a représente la distance entre x et l'hyperplan. La résolution de (II.6), donne (II.7) :

$$a = -\frac{W^T x_i + b}{\|w\|} \qquad \text{(II.7)}$$

La distance de tout exemple de l'hyperplan doit être supérieure ou égale à la marge δ :

$$\frac{y_i(W^T x_i + b)}{\|w\|} \geq \delta \qquad \text{(II.8)}$$

Si une paire (w, b) est une solution alors (aw, ab) est une solution aussi où a est un scalaire. On impose alors la contrainte (II.9) suivante :

$$\|w\|\delta \geq 1 \qquad \text{(II.9)}$$

Pour trouver l'hyperplan séparateur qui maximise la marge, on doit déterminer, à partir des deux dernières inégalités, le vecteur w qui possède la norme euclidienne minimale et qui vérifie la contrainte de l'équation (II.6), de bonne classification des exemples d'entrainement. L'hyperplan séparateur optimal peut être obtenu en résolvant le problème de l'équation (II.10) :

$$\begin{cases} \text{Minimiser} & \frac{1}{2}\|W\|^2 \\ \text{Sous contraints} & \\ y_i(W^T x_i + b) \geq 1 \forall i = 1 \dots n \end{cases} \quad \text{(II.10)}$$

Remarquons que nous pouvons obtenir le même hyperplan même en supprimant toutes les données qui vérifient l'inégalité de la contrainte. Les données qui vérifient l'égalité de la contrainte s'appellent les vecteurs supports, et ce sont ces données seules qui contribuent à la détermination de l'hyperplan. Dans la figure (II.1), les données qui se trouvent sur les deux droites +1 et −1 représentent les vecteurs supports. Le problème de l'équation (II.10) est un problème de programmation quadratique avec contraintes linéaires. Dans ce problème, les variables sont w et b, c-à-d que le nombre de variables est égal à m + 1. Généralement, le nombre de variables est important ce qui ne permet pas d'utiliser les techniques classiques de programmation quadratique. Dans ce cas le problème (II.10) est convertit en un problème dual équivalent sans contraintes de l'équation (II.10) qui introduit les multiplicateurs de Lagrange :

$$\mathbb{Q}(w, b, \alpha) = \frac{1}{2} w^T w - \sum_{i=1}^{n} \alpha_i \{y_i (w^T x_i + b) - 1\} \quad \text{(II.11)}$$

Où les α_i sont les multiplicateurs non négatifs de Lagrange. L'optimum de la fonction objective Q peut être obtenu en la minimisant par rapport à w et b et en la maximisant par rapport aux α_i . À l'optimum de la fonction objective, ses dérivées par rapports aux variables w et b s'annulent ainsi que le produit des α_i aux contraintes (équation II.11) :

$$\begin{cases} \frac{\partial \mathbb{Q}(w,b,\alpha)}{\partial w} = 0 & \text{(a)} \\ \frac{\partial \mathbb{Q}(w,b,\alpha)}{\partial b} = 0 & \text{(b)} \\ \alpha_i \{y_i (w^T x_i + b) - 1 = 0 & \text{(c)} \\ \alpha_i \geq 0 & \text{(d)} \end{cases} \quad \text{(II.12)}$$

De (II.12) on déduit :

$$\begin{cases} w = \sum_{i=1}^{n} \alpha_i y_i x_i \\ \sum_{i=1}^{n} \alpha_i y_i \end{cases} \quad \text{(II.13)}$$

En remplaçant dans (II.11), on obtient le problème dual à maximiser suivant :

$$\begin{cases} \text{Maximiser } \mathbb{Q}(\alpha) = \sum_{i=1}^{n} \alpha_i - \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j x_i^T x_j \\ \text{Sous contraintes} \\ \sum_{i=1}^{n} \alpha_i y_i \\ \alpha_i \geq 0 \end{cases} \quad \text{(II.14)}$$

Si le problème de classification est linéairement séparable, une solution optimale pour les α_i existe. Les exemples ayant des $\alpha_i \neq 0$ représentent les vecteurs supports appartenant aux deux classes. La fonction de décision est donnée par :

$$H(x) = \sum_s \alpha_i y_i x^T x_i + b \quad \text{(II.15)}$$

Où S représente l'ensemble des vecteurs supports. b peut être calculé à partir de n'importe quel vecteur support par l'équation (II.16) :

$$b = y_i - w^T x_i \quad \text{(II.16)}$$

II.10.2 SVM à marge souple

En réalité, un hyperplan séparateur n'existe pas toujours, et même s'il existe, il ne représente pas généralement la meilleure solution pour la classification. En plus une erreur d'étiquetage dans les données d'entrainement (un exemple étiqueté $+1$ au lieu de -1 par exemple) affectera crucialement l'hyperplan.

Dans le cas où les données ne sont pas linéairement séparables, ou contiennent du bruit (outliers : données mal étiquetées), et il y a nécessité de les relaxer un peu. Ceci peut être fait en admettant une certaine erreur de classification des données (figure II.10) ce qui est appelé "SVM à marge souple (Soft Margin)".

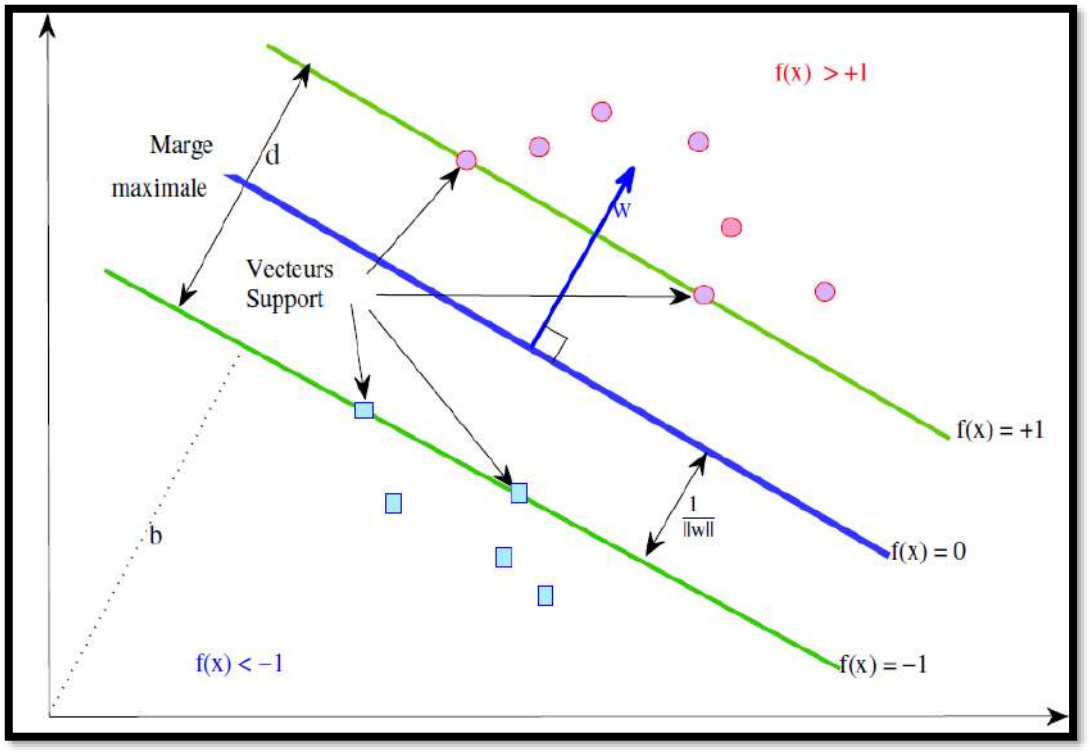


Figure II.10. : SVM binaire à marge souple

L'hyperplan séparateur optimal peut être obtenu en résolvant le problème de l'équation (II.17) :

$$\begin{cases} \text{Minimiser}\frac{1}{2}\|\mathrm{w}\|^2 \\ \text{sous contraintes} \\ \mathrm{y_i}(\mathrm{w^T x_i} + \mathrm{b}) \geq 1 \; \forall_\mathrm{i} = 1 \ldots . \mathrm{n} \end{cases} \quad \text{(II.17)}$$

On introduit alors sur les contraintes des variables ξ_i dites de relaxation pour obtenir la contrainte de l'équation (II.18) :

$$y_i(w^T x_i + b) \geq 1 - \xi_i, \; i = 1 \ldots . n \quad \text{(II.18)}$$

Grâce aux variables de relaxation non négatives i, un hyperplan séparateur existera toujours.

Si $\xi_i < 1, x_i$ ne respecte pas la marge mais reste bien classé, sinon x_i est mal classé par l'hyperplan. Dans ce cas, au lieu de rechercher uniquement un hyperplan séparateur qui maximise la marge, on recherche un hyperplan qui minimise aussi la somme des erreurs permises c-à-d minimisé $Q(w) = \sum_{i=1}^{n} \xi_i$

En combinant avec l'équation (II.18), on obtient le problème primal de l'équation (II.19) suivant :

$$\begin{cases} \text{Minimiser}\frac{1}{2}\|\mathrm{w}\|^2 + \mathrm{C}\sum_{i=1}^{n} \xi_i \\ \text{sous contraintes} \\ y_i(w^T x_i + b) \geq 1 - \xi_i, \; i = 1 \ldots . n \\ \xi_i \geq 0 \end{cases} \quad \text{(II.19)}$$

Où C est un paramètre positif libre (mais fixe) qui représente une balance entre les deux termes de la fonction objective (la marge et les erreurs permises) c-à-d entre la maximisation de la marge et la minimisation de l'erreur de classification. On obtient le problème dual de l'équation (II.20) où on introduit les multiplicateurs de Lagrange $\alpha_i$et β_i :

$$Q(w, b, \alpha, \xi, \beta) = \frac{1}{2} w^T w + \mathrm{C}\sum_{i=1}^{n} \xi_i - \sum_{i=1}^{n} \alpha_i y_i (w^T x_i + b) - 1 + \xi_i - \sum_{i=1}^{n} \beta_i \xi_i \quad \text{(II.20)}$$

A la solution optimale, les dérivées par rapport aux variables w, b, α, βs'annulent ainsi que le produit des contraintes aux multiplicateurs. Les conditions suivantes sont alors vérifiées:

$$\begin{cases} \frac{\partial Q(W,b,\xi,\alpha,\beta)}{\partial W} = 0 & \text{(a)} \\ \frac{\partial Q(W,b,\xi,\alpha,\beta)}{\partial b} = 0 & \text{(b)} \\ \alpha_i\{y_i(w^T x_i + b) - 1 + \xi_i\} = 0 & \text{(c)} \\ \beta_i \xi_i = 0 & \text{(d)} \\ \alpha_i \geq 0; \; \beta_i \geq 0; \; \xi_i \geq 0 & \text{(e)} \end{cases} \quad \text{(II.21)}$$

On déduit :

$$\begin{cases} w = \sum_{i=1}^{n} \alpha_i y_i x_i \\ \sum_{i=1}^{n} \alpha_i y_i = 0 \\ \alpha_i + \beta_i = 0 \end{cases} \quad \text{(II.22)}$$

En remplaçant l'équation (II.20) dans l'équation (II.22), on obtient le problème dual (II.23) suivant :

$$\begin{cases} \text{Maximiser } Q(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j x_i^T x_j \\ \text{sous contraintes} \\ \sum_{i=1}^{n} \alpha_i y_i = 0 \\ 0 \leq \alpha_i \leq C \end{cases} \quad \text{(II.23)}$$

La seule différence avec la SVM à marge dure est que les α_i ne peuvent pas dépasser C ils peuvent être dans l'un des trois cas suivants :

1. $\alpha_i = 0 \Longrightarrow \beta_i = C \Longrightarrow \xi_i = 0 : x_i$ est bien classé ;
2. $0 < \alpha_i < C \Longrightarrow \beta_i > 0 \Longrightarrow \xi_i = 0 \Longrightarrow y_i(w^T x_i + b) = 1 : x_i$ est un vecteur support et est appelé dans ce cas vecteur support non borné (unbounded) ;
3. $\alpha_i = C \Longrightarrow \beta_i = 0 \Longrightarrow \xi_i \geq 0 \Longrightarrow y_i(w^T x_i + b) = 1 - \xi_i : x_i$ est un vecteur support appelé dans ce cas vecteur support borné (bounded). Si $0 \leq \xi_i < 1, x_i$est bien classé, sinon x_i est mal classé. Ces conditions sur les α_ii sont appelées les conditions de Karush Kuhn-Tucker (KKT) (Karush, 1939; Kuhn, 1951), elles sont très utilisées par les algorithmes d'optimisation pour rechercher les α_i optimums et par conséquent l'hyperplan optimal. La fonction de décision est alors calculée de la même manière que dans le cas des SVM à marge dure mais uniquement à base des vecteurs supports non bornés par :

$$H(x) = \sum_{i \in U} \alpha_i y_i x_i^T x + b \quad \text{(II.24)}$$

Pour les vecteurs supports non bornés, nous avons :

$$b = y_i - w^T x_i \quad \text{(II.25)}$$

Pour garantir une bonne précision, on prend la moyenne de b pour tous les vecteurs supports non bornés :

$$b = \frac{1}{|U|} \sum_{i \in U} y_i - w^T x_i \quad \text{(II.26)}$$

II.11 Utilisation des noyaux

Le fait d'admettre la mal-classification de certains exemples, ne peut pas toujours donner une bonne généralisation pour un hyperplan même si ce dernier est optimisé (figure II.11).

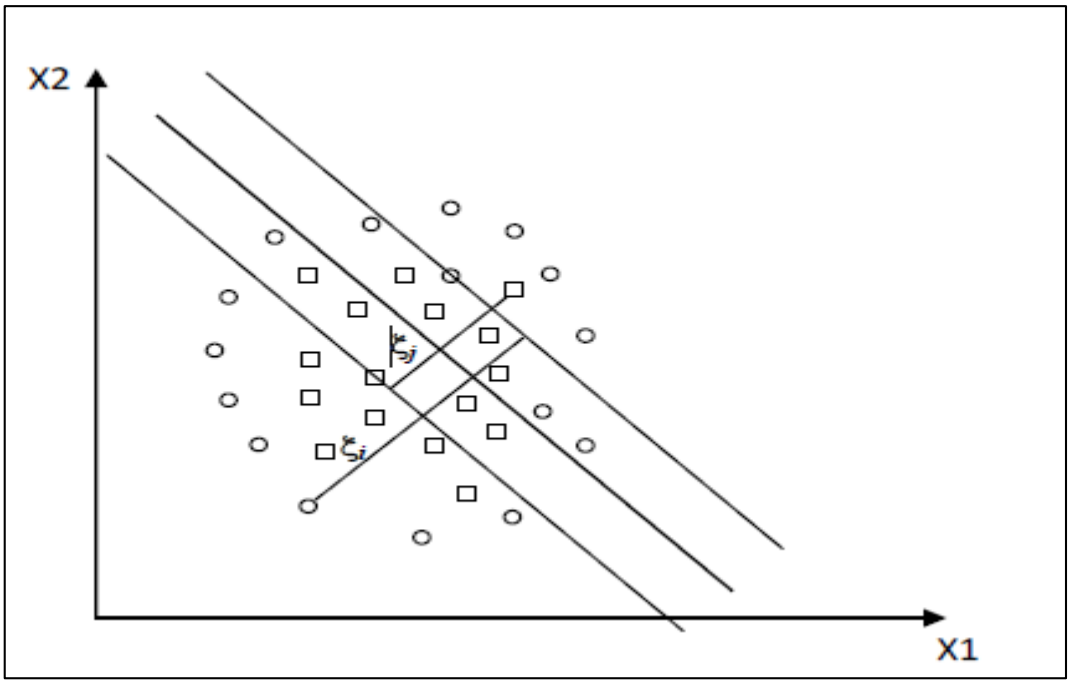


Figure II.11 : Mal-adaptation de l'hyperplan aux problèmes réels de classification

Plutôt qu'une droite, la représentation idéale de la fonction de décision serait une représentation qui colle le mieux aux données d'entrainement (figure II.11).

La détermination d'une telle fonction non linéaire est très difficile voire impossible. Pour cela les données sont amenées dans un espace où cette fonction devient linéaire (figure II.12), cette astuce permet de garder les mêmes modèles de problèmes d'optimisation vus dans les sections précédentes, utilisant les SVM basées essentiellement sur le principe de séparation linéaire.

Cette transformation d'espace est réalisée souvent à l'aide d'une fonction $F = \{\emptyset(x)|x \in X\}$ appelé Mapping function" et le nouvel espace est appelé espace de caractéristiques "Features space".

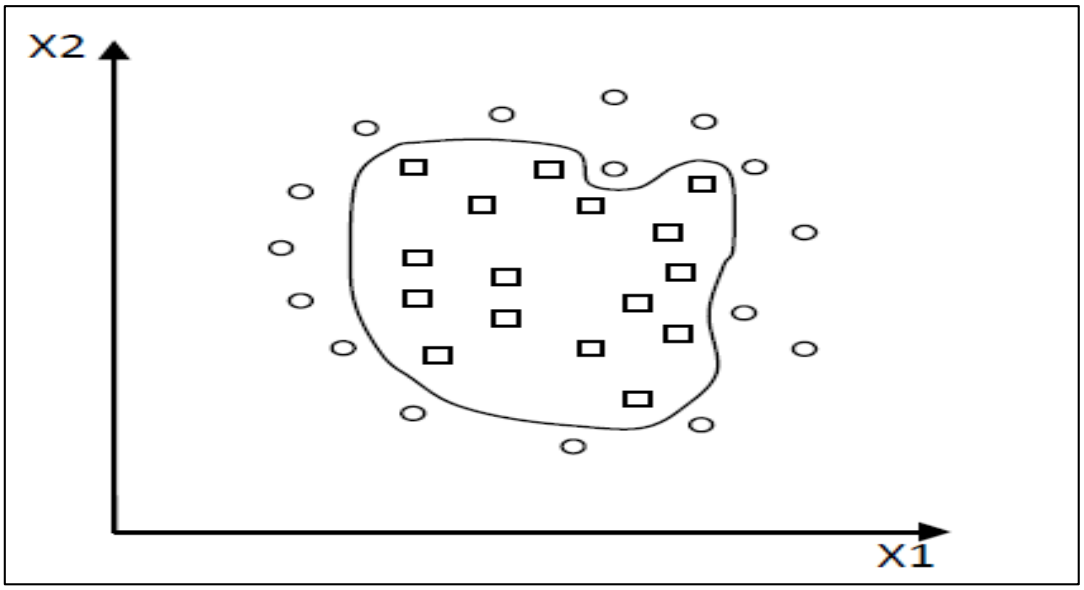


Figure II.12 : Représentation idéale de la fonction de décision

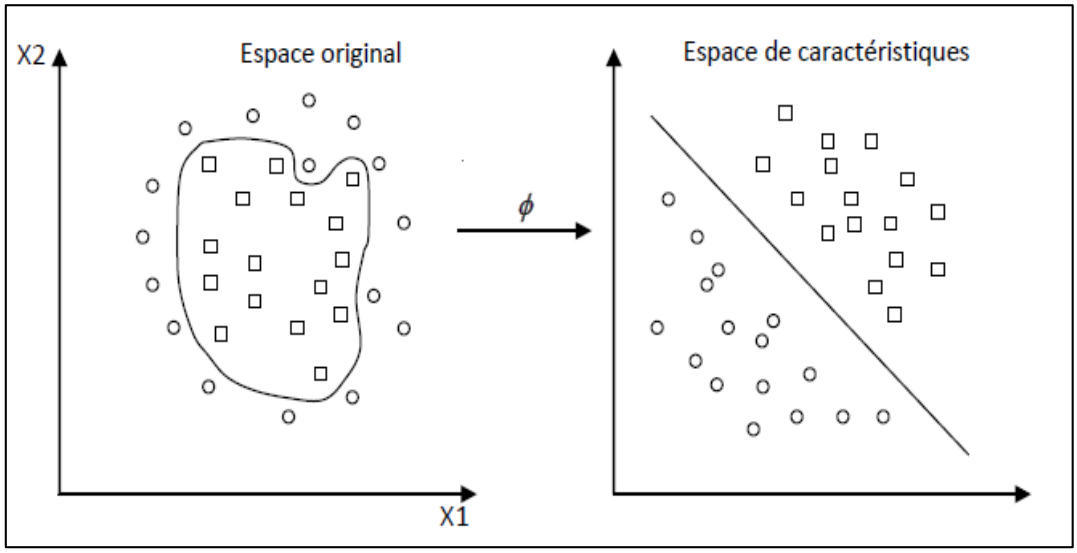


Figure II.13 : Transformation d'espace

Dans ce nouvel espace de caractéristiques, la fonction objective à optimiser de l'équation (II.27) devient :

$$\mathbb{Q}(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j \langle \phi(x_i), \phi(x_j) \rangle \qquad \text{(II.27)}$$

Où $\langle \phi(x_i)\phi(x_j) \rangle$ est le produit scalaire des deux images des vecteurs xi et xj dans le nouvel espace et dont le résultat est un scalaire. Dans le calcul de l'optimum de la fonction (II.26), on utilise une astuce appelée "Noyau" ("Kernel"), au lieu de calculer $\phi(x_i), \phi(x_j)$ et leur produit scalaire, on calcule plutôt une fonction $K(x_i,x_j)$ qui représente à la fois les deux transformations (qui peuvent être inconnues) et leur produit scalaire. Cette fonction permet de

surmonter le problème de détermination de la transformation Ø et permet d'apprendre des relations non linéaires par des machines linéaires. La fonction $K(x_i,x_j)$ peut être vue comme une matrice G[n,n] dite de Gram (Kivinen et al., 2002) qui représente les distances entre tous les exemples (Kivinen et al., 2002) :

$$\begin{bmatrix} \mathbf{K(x_1, x_1)} & \cdots & \mathbf{K(x_1, x_n)} \\ \vdots & \ddots & \vdots \\ \mathbf{K(x_n, x_n)} & \cdots & \mathbf{K(x_n, x_n)} \end{bmatrix}$$

Pour qu'une matrice G (fonction K) soit un noyau, il faut qu'elle respecte les conditions de Mercer (Milenova et al., 2005), (Equations, 1991), c-à-d qu'elle doit être semi-définie positive (symétrique et n'a pas de valeurs propres négatives). La construction de tels noyaux est largement étudiée dans (Herbrich, 2001)(Soman et al., 2009), néanmoins, il existe certains noyaux qui sont très utilisés et qui sont considérés comme standards (II.12) Une fois le noyau choisi, la fonction objective (II.28) peut être calculée comme suit :

$$\mathbb{Q}(\boldsymbol{\alpha}) = \sum_{\mathbf{i=1}}^{\mathbf{n}} \boldsymbol{\alpha_i} - \frac{\mathbf{1}}{\mathbf{2}} \sum_{\mathbf{i=1}}^{\mathbf{n}} \sum_{\mathbf{j=1}}^{\mathbf{n}} \boldsymbol{\alpha_i \alpha_j y_i y_j K(x_i, x_j)} \qquad \text{(II.28)}$$

Et la fonction de décision devient :

$$\mathbf{H(x)} = \sum_{\mathbf{i \in S}} \boldsymbol{\alpha_i y_i K(x_i, x)} + \mathbf{b} \qquad \text{(II.29)}$$

Ou S représente l'ensemble des vecteur supports.

Des exemples des fonction noyau :

a) **Noyau polynomial** : Le noyau polynomial élève le produit scalaire à une puissance naturelle d :

$$K = (x_i, x_j) = (x_i . x_j + 1)^d \qquad \text{(II.30)}$$

b) **Noyau linéaire :** Si les données sont linéairement séparables, on n'a pas besoin de changer d'espace, et le produit scalaire suffit pour définir la fonction de décision :

$$K = (x_i, x_j) = x_i^T x_j \qquad \text{(II.31)}$$

c) **Noyau RBF :** Les noyaux RBF (Radial Basis fonctions) sont des noyaux qui peuvent être écrits sous la forme :$K(x_i, x_j) f\left(d(x_i, x_j)\right)$ où d est une métrique sur X et f est une fonction dans $\Re$.

Un exemple des noyaux RBF est le noyau Gaussien (II.32) :

$$K = (x_i, x_j) = e^{\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)} \qquad \text{(II.32)}$$

On général le choix d'un Noyau correspond à (Schuhmacher et al., 1996) :

- Choisir une mesure de similarité pour les données.
- Choisir une représentation linéaire aux données.
- Choisir un espace de fonction pour l'apprentissage

II .12 SVM pour la régression

Dans leur origine, les SVM ont été développées pour des problèmes de classification. Cependant, leur nature leur permet de résoudre également des problèmes de régression. La régression est un cas particulier de classification où les classes des exemples ne sont pas dénombrables c-à-d continues. Le problème consiste à trouver, en utilisant D = $\{(x_1, y_1), \dots, (x_n, y_n)\}$, une fonction $\hat{f}: \Re^m \to \Re$ qui rapproche le plus possible des y_i, en d'autre terme qui minimise la différence entre les $\hat{f}(x_i)$ et les yi.Souvent, $\hat{f}$ est considérée comme fonction linéaire $\hat{f}$=$\langle w, x\rangle + b$,

Avec :

w : est un vecteur

b : est un scalaire.

Le problème revient donc à trouver un hyperplan caractérisé par w et b qui minimise l'écart global entre $\hat{f}$ et les y_i (équation II .33)

$$(w^*, b^*)\text{argmin}_{w,b} \sum_{i=1}^{n} (y_i - \langle w, x_i\rangle - b)^2 \qquad \text{(II.33)}$$

Pour résoudre ce problème, les SVM utilisent une astuce semblable à celle utilisée en classification (Kadirvelu et al., 2008) (Su et al., 2015) (Hamel, 2009) On propose de modéliser la fonction de régression par un hyperplan qui se situe au centre d'un hyper-tube de largeur contenant tous les exemples d'entrainement (figure II.14)

Plusieurs hyper-tubes, de largeur 2ε contenant tous les exemples d'entrainement, peuvent Exister. L'hyper-tube optimal est celui qui minimise la distance entre les exemples d'entrainement et ses frontières, autrement dit, qui maximise la distance des exemples de l'hyperplan du centre (figure II .14).

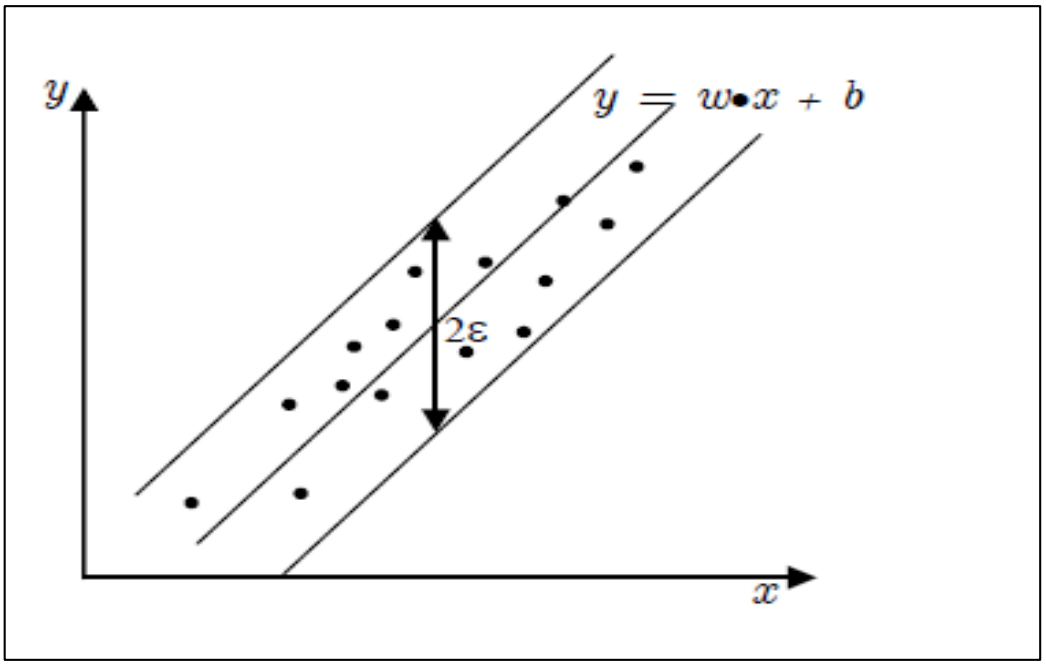


Figure II.14: Hyper-tube modélisant la fonction de régression (Hamel, 2009).

La détermination de l'hyper-tube optimal est semblable à la détermination de l'hyperplan optimal de marge maximale dans le cas de classification. On doit donc rechercher un hyper tube de marge maximale avec tous les exemples d'entrainement à l'intérieur. Par une analyse similaire à celle du problème de classification, la solution du problème de régression est réduite à la résolution du problème dual d'optimisation quadratique de l'équation (II.34) :

$$\begin{cases} \textbf{Maximiser}_{\alpha,\acute{\alpha}} & -\frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i-\acute{\alpha}_i)(\alpha_j-\acute{\alpha}_j)\langle x_i,x_j\rangle \\ & -\epsilon\sum_{i=1}^{n}(\alpha_i+\acute{\alpha}_i)+\sum_{i=1}^{n}y_i(\alpha_i,\acute{\alpha}_i) \\ \textbf{saus contrainte} & \sum_{i=1}^{n}(\alpha_i-\acute{\alpha}_i)=0 \\ & 0\le\alpha_i,\acute{\alpha}_i\le C \end{cases} \quad \text{(II.34)}$$

Où le α_i et les $\acute{\alpha}_i$ sont les coefficients des exemples respectivement au dessus et au dessous de l'hyperplan et C est un paramètre pour leur pénalisation. La fonction de sortie $\hat{f}(x)$ peut être donnée par l'équation :

$$\hat{\mathbf{f}}(\mathbf{x})=\sum_{i=1}^{n}(\alpha_i-\acute{\alpha}_i)(x_i,x)+b \quad \text{(II.35)}$$

Où b est calculé à partir d'un exemple dont 0 < _i < C (vecteur support) par l'équation (II.36) :

$$b=y_i-\langle w,x_i\rangle-\epsilon \quad \text{(II.36)}$$

II .13 Utilisation des noyaux

Parmi les motivations les plus solides du développement des machines à vecteur support pour la régression, est leur extension simple aux cas non linéaires, grâce à l'utilisation des noyaux. En effet, d'une manière similaire au cas de classification, on fait une transformation d'espace Ø pour se trouver toujours face à une régression linéaire. La transformation d'espace inverse $Ø^{-1}$, permet de retourner à l'espace d'origine après la résolution dans le nouvel espace (figure II.15) (Djeffal, 2012).

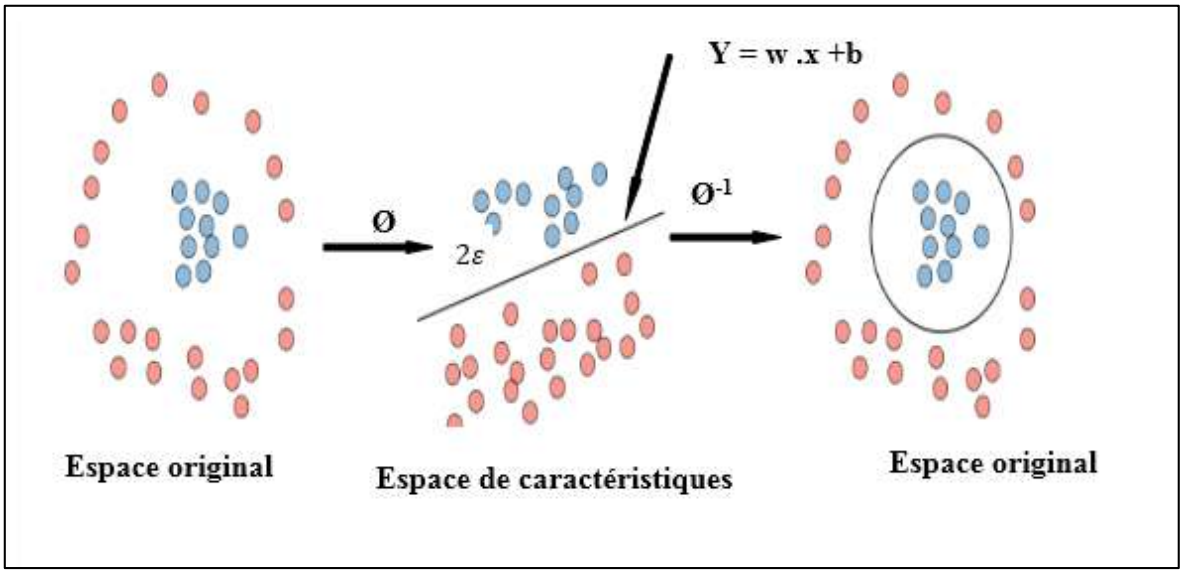


Figure II .15 : utilisation des noyaux pour la résolution de la régression non linéaire.

II.14. Applications

SVM est une méthode de classification qui montre de bonnes performances dans la Résolution de problèmes varies. Cette méthode a montré son efficacité dans de nombreux Domaines d'applications tels que le traitement d'image, la catégorisation de textes ou le Diagnostiques médicales et ce même sur des ensembles de données de très grandes dimensions (Hasan & Boris, 2006).

Réalisation d'un programme d'apprentissage par SVM se ramené a résoudre un problème d'optimisation impliquant un système de résolution dans un espace de dimension Conséquente. L'utilisation de ces programmes revient surtout a sélectionner une bonne famille de fonctions noyau et a régler les paramètres de ces fonctions . Ces choix sont le plus souvent faits par une technique de validation croisée, dans laquelle on estime la performance du système en la mesurant sur des exemples n'ayant pas été utilises en cours d'apprentissage. L'idée est de chercher les paramètres permettant d'obtenir la performance maximale. Si la mise en œuvre

d'un algorithme de SVM est en général peu couteuse en temps, il faut cependant compter que la recherche des meilleurs paramètres peut requérir des phases de test assez longues (Hasan & Boris, 2006) .

Les SVM fournissent des performances de pointe dans les applications du monde réel telles que la catégorisation des textes, la reconnaissance des caractères manuscrits, la classification des images, l'analyse des bio séquences, etc., et sont maintenant définis comme l'un des outils standard pour l'apprentissage et la transmission des machines exploitation minière(Howie, 1984).

Dans ce chapitre, nous avons introduit les machines à vecteur support, leur fondement théorique et leurs différentes variantes sans être très exhaustifs. Nous avons donné une vision générale et le fondement mathématique des SVM. Nous avons exposé les cas linéairement séparables et les cas non linéairement séparables qui nécessitent l'utilisation de fonctions noyaux (kernel) pour changer d'espace. Les SVM représentent aujourd'hui l'une des méthodes les plus utilisées grâce à leur pouvoir de généralisation.

Ils sont fondés rigoureusement et simplement (Howie, 1984).

II.15 Avantages et inconvénients

II.15.1 Avantages

- Très efficace en dimension élevée.
- Un nombre de paramètres faible à régler.
- Une grande vitesse d'apprentissage.
- Ils sont aussi efficaces dans le cas où la dimension de l'espace est plus grande que le nombre d'échantillon d'apprentissage.
- Traitement des problèmes linéaires ou non linéaires selon la fonction du noyau.
- Un nombre restreint d'échantillons suffit à la détermination des vecteurs supports permettant la discrimination entre les classes. En conséquence, ces algorithmes demandent moins de mémoire (Lajevardi & Lech, 2008) .

II.15.2 Inconvénients

- Si le nombre d'attributs est beaucoup plus grand que le nombre d'échantillons, les performances seront moins bonnes.

- Comme il s'agit de méthodes de discrimination entre les classes, elles ne fournissent pas des estimations de probabilités (Lajevardi & Lech, 2008).

II.16 Conclusion

Dans ce chapitre, nous avons introduit les machines à vecteur support, leur fondement théorique et leurs différentes variantes sans être très exhaustifs. Nous avons donné une vision générale et le fondement mathématique des SVM. Nous avons exposé les cas linéairement séparables et les cas non linéairement séparables qui nécessitent l'utilisation de fonctions noyaux (kernel) pour changer d'espace. Les SVM représentent aujourd'hui l'une des méthodes les plus utilisées grâce à leur pouvoir de généralisation. Ils sont fondés rigoureusement et simplement.

Chapitre III : Modilisation de L’adsorption multi-composant par la méthode machine à vecteur de support

III.1 Introduction

Les méthodes de traitement des eaux sont nombreuses et variées. Mais l'une des techniques les plus efficaces reste l'adsorption de polluants par des matériaux poreux. Le charbon actif est le plus utilisé dans l'adsorption des matières organiques et des métaux dissouts dans l'eau ainsi que pour l'élimination des odeurs des eaux. Actuellement, beaucoup de travaux de recherche qui ont été publiés sur l'adsorption des métaux lourds par [le charbon actif] (Mohan 2002). Ces travaux montrent que celui-ci est très efficace. L'adsorption demeure l'une des techniques d'élimination des polluants organiques et métalliques la plus utilisée pour son efficacité, la facilité de sa mise en œuvre et son coût d'investissement accessible. Cette méthode nécessite le choix d'un adsorbant qui présente de bonnes caractéristiques (surface spécifique élevée, disponibilité, prix de revient bas et son impact sur l'environnement...). Pour bien comprendre et contrôler le phénomène d'adsorption multi-composants qui est considéré comme compétitive et complexe, plusieurs méthodes de modélisation ont été proposé tels que la méthode dite machine à support vecteur qui présente quelques avantages en comparaison avec d'autres techniques d'intelligence artificiel notamment elle nécessite une petite base de données, ne présente pas le problème de minimum local. L'objectif principal de cette contribution est de tester la faisabilité de modélisation par la méthode support à vecteur machine à prédire la quantité adsorbée par l'adsorbant de chaque polluant (métaux lourds) en fonction de quelques paramètres indépendant caractérisant le phénomène.

III.2 Modélisation par SVM

III.2.1 Méthodologie de travail

La présente étude est effectuée selon la procédure de travail illustrée dans la Figure III.1

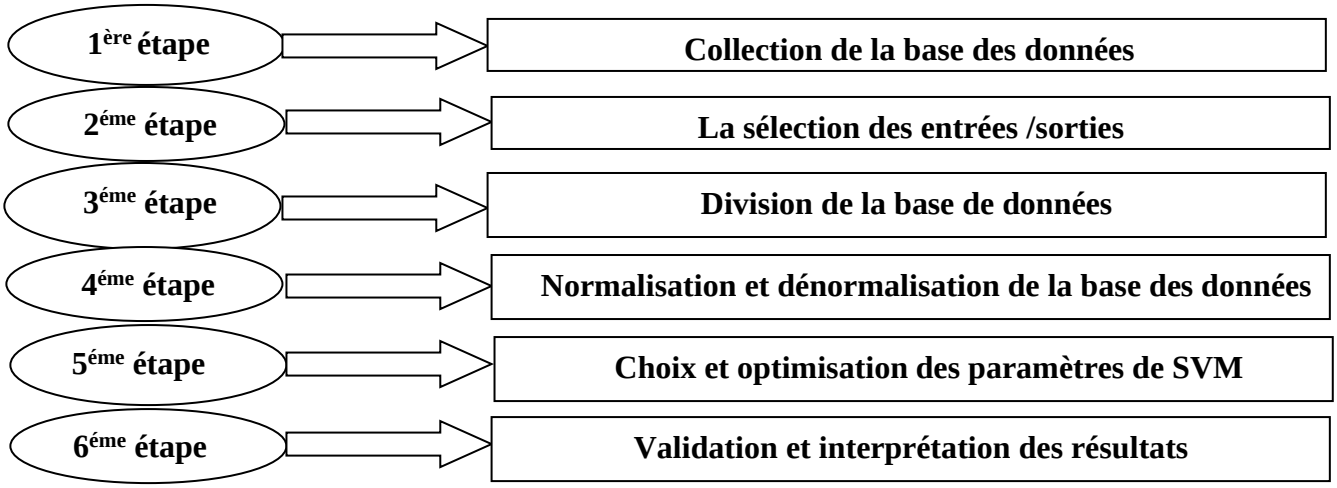


Figure III.1 : Méthodologie de travail

III.2.2 Collection et préparation de la base des données

Notre objectif principal est de proposer un modèle basé sur un nombre réduit d'entrées facilement accessibles. Il convient aussi de signaler que lors d'une recherche bibliographique, l'efficacité d'élimination {removal efficiency, R(%)} et la capacité d'adsorption {adsorption capacity, q_e ($mg.g^{-1}$)} sont importants et significatifs pour décrire le processus d'adsorption, mais ces deux paramètres ne sont pas toujours disponible dans la littérature, raison pour laquelle, plusieurs méthodes ont été proposé tels que {modèles empiriques notamment Langmuir et Freundlich, méthodes d'intelligences artificiels et d'autres}.

Dans cette étude la capacité d'adsorption est sectionnée comme paramètre de sortie en raison de sa disponibilité dans les articles choisis afin de construire notre base de sonnées. Cinq variables d'entrée sont retenues et choisies par rapport à deux critères importants {leurs influences sur le phénomène physique et leurs communes disponibilités dans les articles choisis}, ces paramètres sont {le potentiel hydrogène (pH), température du mélange (°C), la surface spécifique de l'adsorbant ($m^2.g^{-1}$), la masse moléculaire des ions métalliques ($g.mol^{-1}$) et la concentration à l'équilibre des ions dans l'adsorbants}.

Tableau III.1 Sources bibliographiques des données expérimentales

Références bibliographiques	Adsorbats	Adsorbants	Points
(Mohan & Singh, 2002)	Zn^{+2}+Cu^{+2}+Cd^{+2}	Charbon actif	12*3
(Qin et al., 2006)	Pb^{+2}+Cu^{+2}+Cd^{+2}	Tourbe du Heilongjiang	10*3
		Tourbe danoise	10*3
(Kadirvelu et al., 2008)	Pb^{+2}+Hg^{+2}+Cd^{+2}	Charbon actif	15*3
(Roy & Bhattacharya, 2013)	Cd^{+2}+Ni^{+2}+Co^{+2}	Les nanotubes	14*3
(Mendoza-castillo et al., 2019)	$Pb+^{2}$,Hg^{+2}+Zn^{+2}	Charbon actif	39*3
(Qu et al., 2020)	Pb^{+2}+Cd^{+2}+Ni^{+2}	Cellulose	24*3
		Nombre total des points : 372	

Le tableau (III.1) montre les sources bibliographiques de chaque système dans la base de données respectivement. Comme le modèle SVM ne peut pas extrapoler, nous avons représenté dans le tableau (III.2) les domaines de variation de chaque paramètre sélectionné dans la base de données

Tableau III.2 Domaines des données expérimentales.

Analyses statistiques des paramètres	Entrées					Sortie
	M_w ($g.mol^{-1}$)	As ($m^2.g^{-1}$)	pH	T (°C)	C_e ($mg.\ g^{-1}$)	q_e ($mg.\ g^{-1}$)
Maximum	207.2	960	8.46	105	869.24	203,627
Minimum	58.693	3.795	4.1	20	0,099734	0,416987

Pour la construction de notre modèle, l'ensemble des données d'apprentissage est constitué de 75% des données choisies aléatoirement par l'utilisation de la fonction

dividerand de MATLAB 2018. Les 25% restantes sont utilisées comme ensemble de test pour SVM. Cette division a pour but d'éviter le sur-apprentissage.

La fiabilité et la précision des modèles, développés à partir des méthodes d'intelligence computationnelle, reposent sur diverses quantités statistiques calculés pour les ensembles d'apprentissage et de test :

- Erreur absolue moyenne, {EAM}
- Erreur prédictive du modèle (%), {EPM}
- Racine de l'erreur quadratique moyenne {REQM}
- Erreur standard de prédiction (%) {ESP}
- Écart type de la valeur résiduelle {RPD}
- Coefficient de détermination (R^2)
- Coefficient de corrélation (R)
- Erreur relative absolue (%), {ERA}

Les formules mathématiques de ces critères statistiques sont données comme suit [(Kavand., 2017)- (Baghban.,2018)] :

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

$$EPM(\%) = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_{i,exp}-y_{i,cal}}{y_{i,exp}}\right| \quad \text{(III.2)}$$

$$REQM = \sqrt{\frac{\sum_{i=1}^{n}\left(Y_{i,exp}-Y_{i,cal}\right)^2}{n}} \quad \text{(III .3)}$$

$$ESP(\%) = \frac{RMSE}{Y_e} \times 100 \quad \text{(III.4)}$$

$$RPD = \frac{SD}{RMSE} \quad \text{(III.5)}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_{i_{exp}}-y_{i_{cal}}\right)^2}{\sum_{i=1}^{N}\left(y_{i_{exp}}-\bar{y}_{i_{exp}}\right)^2} \quad \text{(III.6)}$$

$$R = \frac{\sum_{K=1}^{N}\left(y_{i_{exp}}-\bar{y}_{i_{exp}}\right)\left(y_{i_{cal}}-\bar{y}_{i_{cal}}\right)}{\sqrt{\sum_{K=1}^{N}\left(y_{i_{exp}}-\bar{y}_{i_{exp}}\right)^2 \sum_{K=1}^{N}\left(y_{i_{cal}}-\bar{y}_{i_{cal}}\right)^2}} \quad \text{(III.7)}$$

$$ERA(\%) = \left|\frac{q_{exp}-q_{cal}}{q_{exp}}\right| \quad \text{(III.8)}$$

Avec :

n= Nombre de l'ensemble de données (apprentissage ou test).

$\mathbf{Y_{i,exp}}$: Valeurs expérimentales de la propriété étudiée pour ensembles d'apprentissage et test

$\mathbf{Y_{i,cal}}$: Valeurs calculées de la propriété étudiée pour les ensembles d'apprentissage et de test.

$\mathbf{Y_e}$: Valeur moyenne des données expérimentales

SD : L'écart type des données

$\bar{\mathbf{y}}_{\mathbf{i_{exp}}}$: La moyenne des éléments de vecteur

III.3 Résultats de modélisation par SVM

Pour SVM, il est souvent utile de normaliser les entrées et les sorties, de manière à ce qu'elles se situent dans une plage spécifiée, ce qui peut permettre d'éviter les débordements numériques et d'accélérer la convergence de l'algorithme.

Les valeurs numériques ont été normalisées afin d'améliorer l'optimisation. La normalisation se fait par ligne.

- Normalisation des entrées : $\mathrm{X_n} = X^{0.013}$
- Normalisation des sorties : $\mathrm{Y_n} = 3 \times log10(Y)$

Pour dé normalisation des réponses, nous utilisons la fonction suivante :

$$\mathrm{Y} = 10^{\frac{\mathrm{Y_n}}{3}}$$

III.3.1 Choix et optimisation des paramètres de SVM

Pour SVM nous avons choisi le modèle SVM régression de MATLAB, l'optimisation de ses paramètres notamment la fonction noyaux (sigmoid, RBF et polynomial), les valeurs optimales de la constante C (box Constraint), epsilon et le paramètre de la fonction noyaux (Kernel Scale) du modèle. Dans cette étude nous avons choisir la méthode de Cross-validation « holdout » que nous avons fixé à 0.25 qui signifie de 75% des données sont utilisées pour l'apprentissage et 25% pour la validation. Les expressions mathématiques des fonctions du noyau testé sont présentées dans le (tableau III.3), et les paramètres optimisés du modèles SVM sont présentés dans le (tableau III.4).

Tableau III.3 Expressions mathématiques des fonctions du noyau

RBF	$K(x, x_i) = exp\left(\frac{-\|\|x - x_i\|\|^2}{2\sigma^2}\right).b$
Polynomial	$\mathrm{K}(x, x_i) = (\alpha x + \mathrm{c})^{\mathrm{d}}$
Linear	$\mathrm{K}(x, x_i) = \mathrm{y}x^T + \mathrm{c}$
Sigmoid	$\mathrm{K}(x, x_i) = \frac{1}{\alpha\sqrt{2\pi}} \exp\left(\frac{-\|\|x - x_i\|\|^2}{2\delta^2}\right)$

Tableau III. 4 Les paramètres optimisés du modèle SVM.

C	ε	σ	Quantité de Vecteur de Support	Fonction kernel
280	0.034	1	180	RBF

L'équation du modèle SVM peur être écrite sous la forme suivante :

$$q_e(x, \alpha_i, \alpha_i^*) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) \ . exp\left[\left(\frac{-\|x - x_i\|^2}{2\sigma^2}\right)\right] . b \qquad \text{(III.8)}$$

Avec :

$\alpha_i - \alpha_i^*$ et b sont déterminé au cours de la phase d'apprentissage.

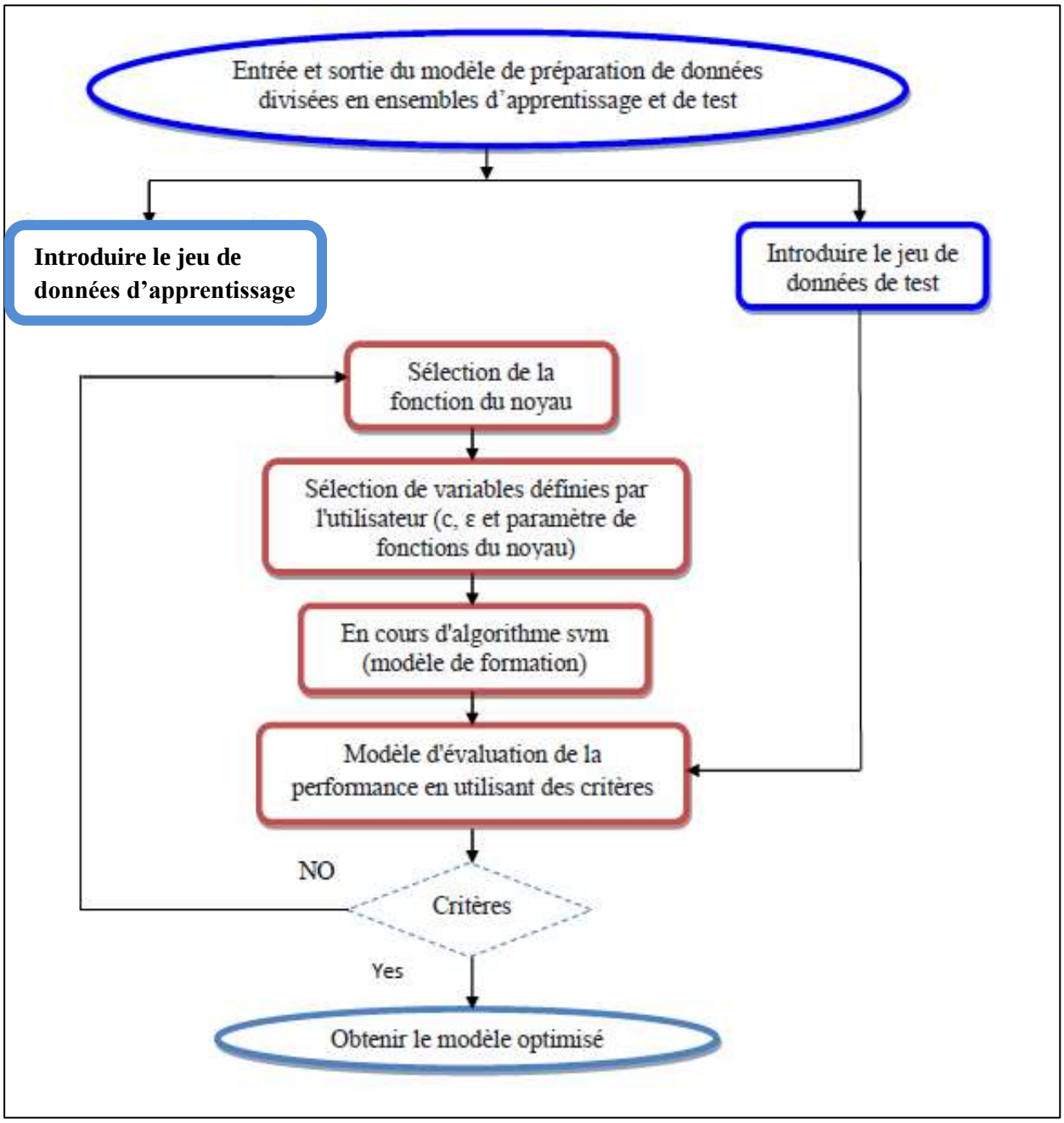
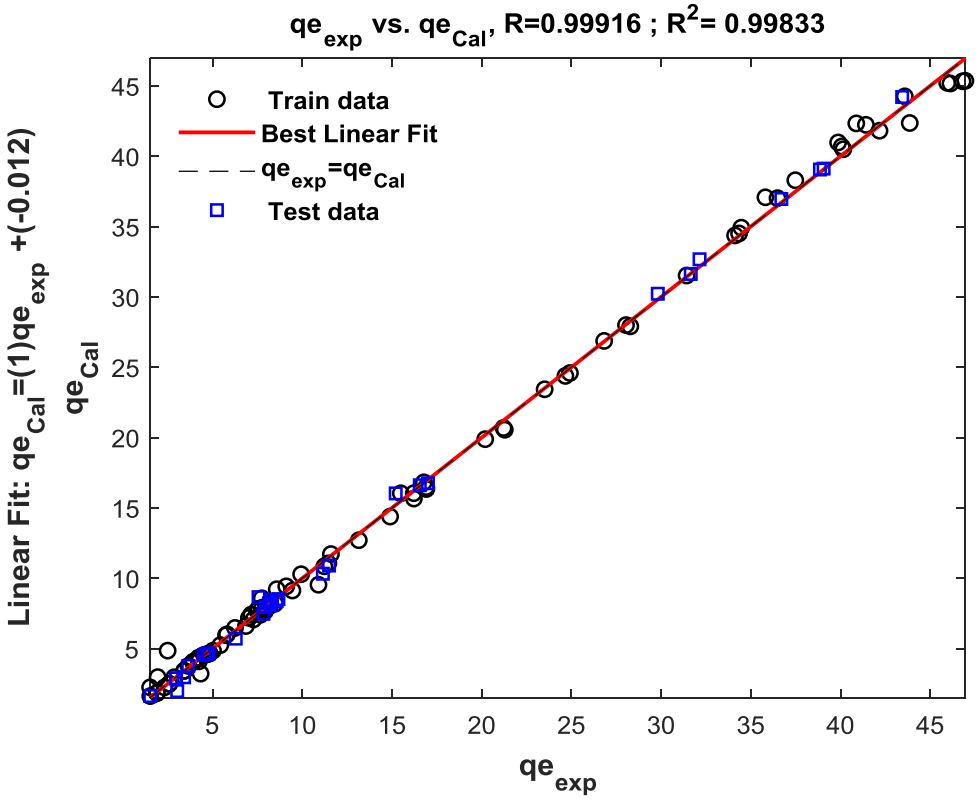


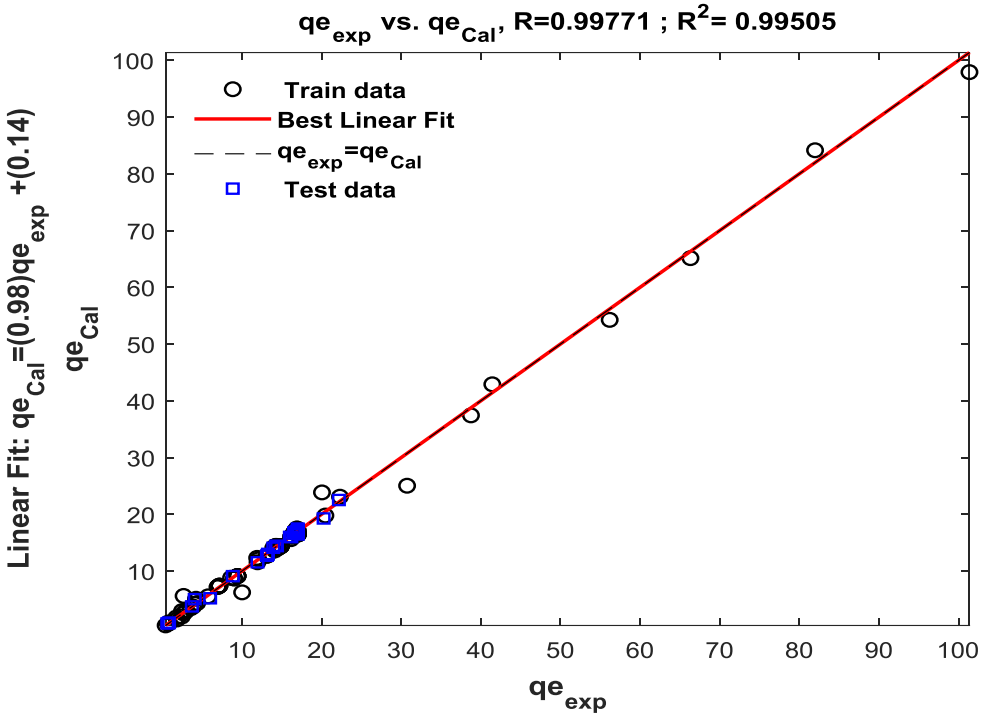
Figure III.2 : Organigramme des étapes pour la création du modèle SVM

III.3.2 Interprétation des résultats

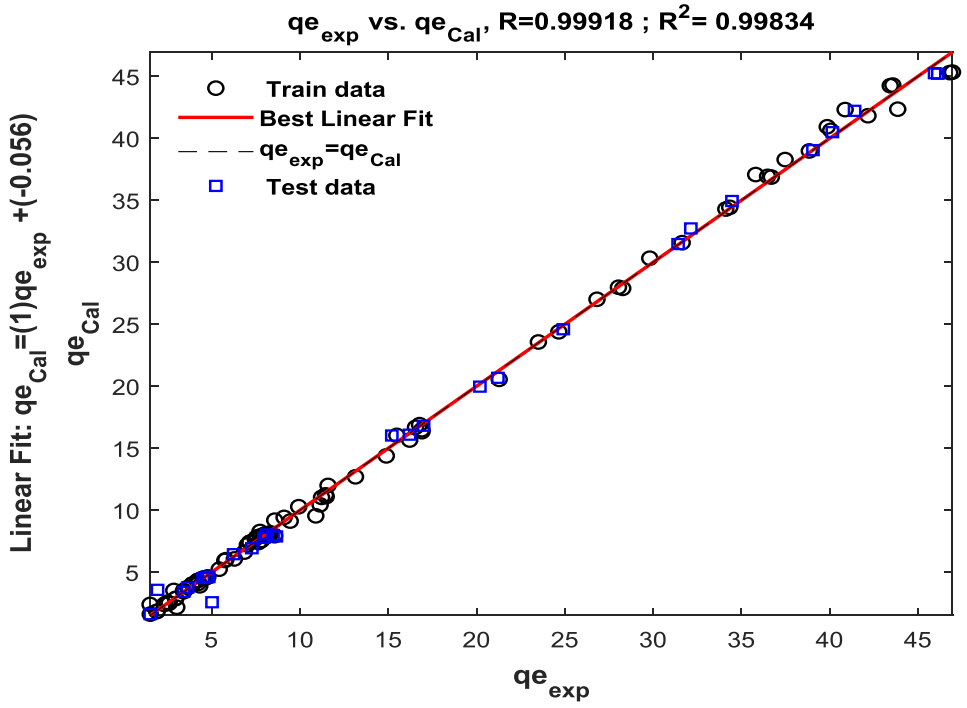
Les figures (a,b,c) montrent les diagrammes de dispersion de la capacité d'adsorption calculée en fonction de la capacité d'adsorption expérimentale. Ces diagrammes montrent clairement la concordance entre les valeurs calculées et les valeurs expérimentales à travers une dispersion serrée des points sur la première bissectrice. Cependant on peut remarquer l'existence d'un certain nombre de points éloignés qui peuvent être attribué à l'existence de points aberrants et la qualité de généralisation du modèle



(a)



(b)



(c)

Figure III.3 : Diagramme de dispersion de la capacité d'adsorption calculée en fonction de la capacité d'adsorption expérimentale par le modèle SVM « RBF » : (a) premier polluant (q_{e1}), (b) deuxième polluant (q_{e2}), (c) troisième polluant (q_{e3}).

Afin de visualiser quantitativement la distribution des erreurs relatives nous avons adoptés une deuxième comparaison en termes d'analyse statistiques à travers le calcul de l'erreur relative en fonction de la concentration à l'équilibre montrant la distribution de la valeur absolue de cette erreur en %. Les figures (4 ,5 ,6) représentent l'erreur relative en fonction de la concentration à l'équilibre pour chaque ion métallique respectivement. Pour les trois cinétiques la majorité des points sont situés autour de la ligne zéro. Quantitativement l'ER pour les quelques points aberrant ne dépasse pas 15% pour la topologie optima**.**

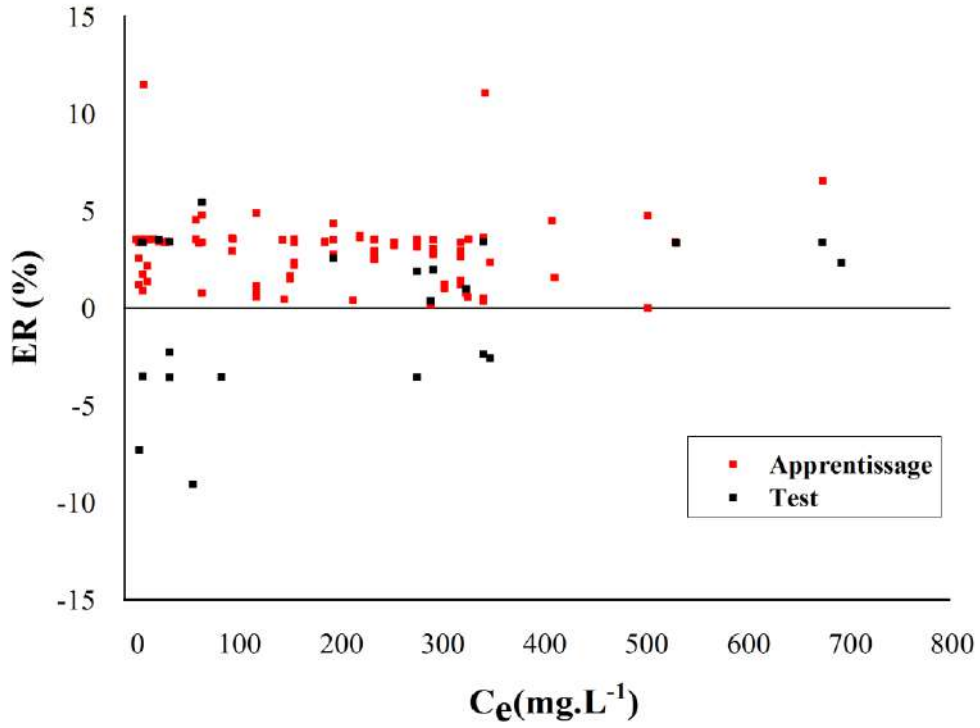


Figure III.4 : Erreur relative (%) de la 1ère quantité adsorbée en fonction de C_e

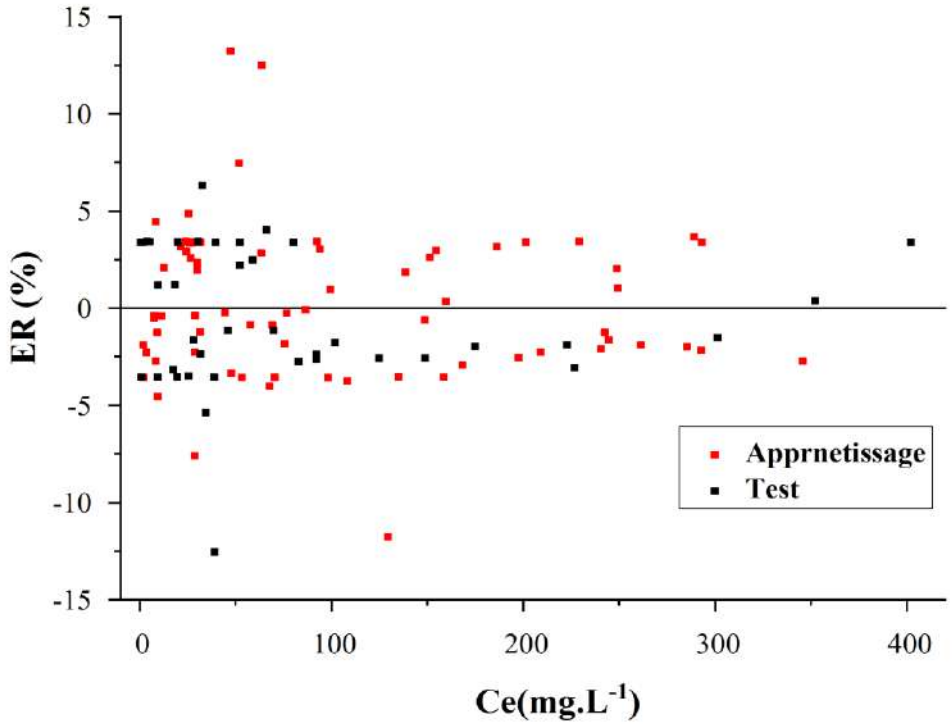


Figure III.5 : Erreur relative (%) de la 2ème quantité adsorbée en fonction de C_e

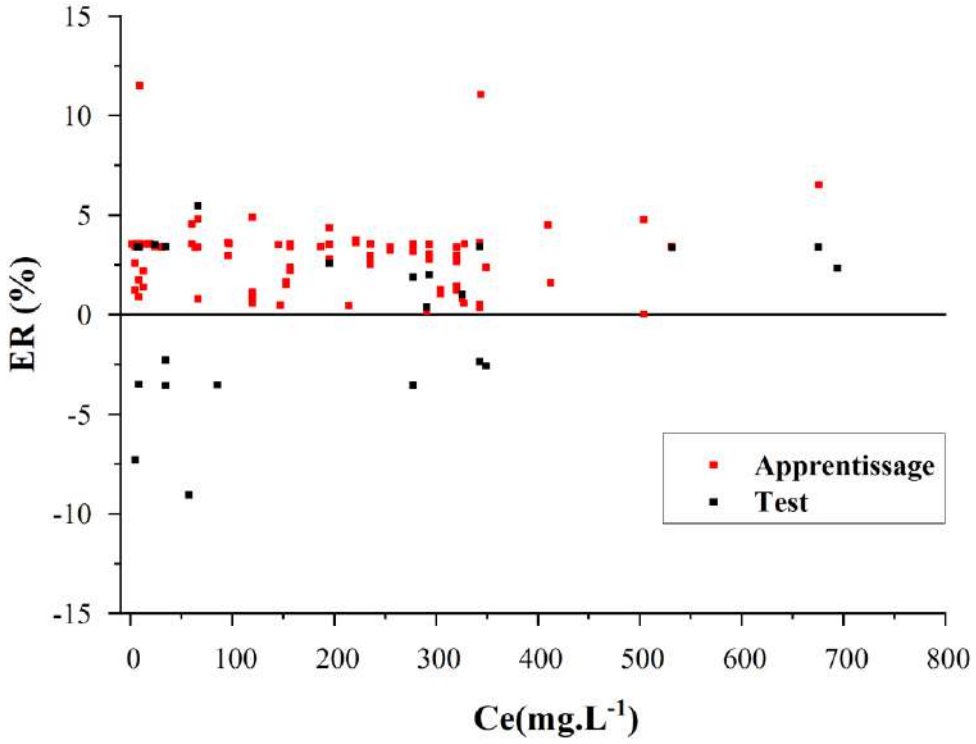


Figure III.6 : Erreur relative (%) de la 3ème quantité adsorbée en fonction de C_e

Ces remarques relatives à l'ER sont mieux illustrées sur les diagrammes de distribution des figures (III.7 a III.9)

ERA [%]

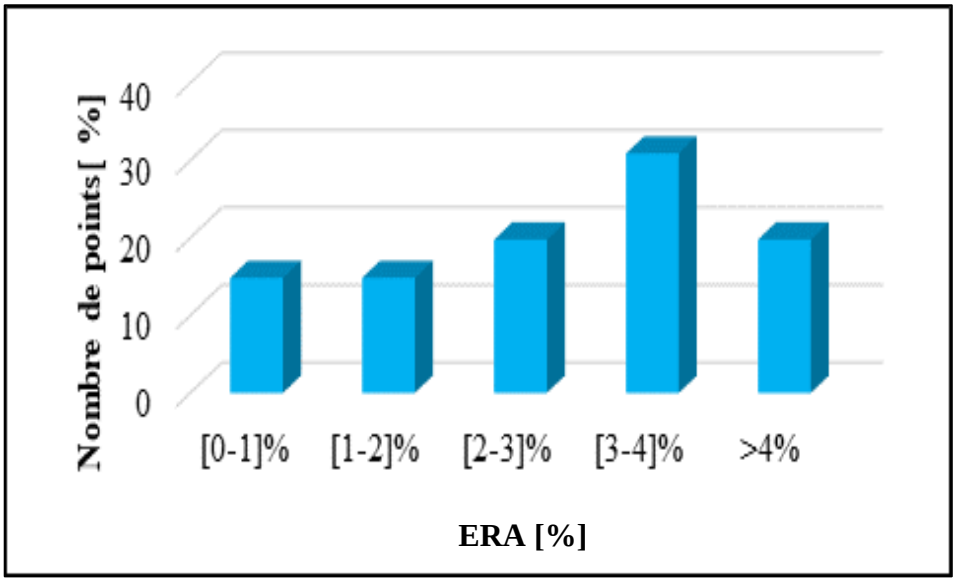


Figure III.7: Diagrammes de distribution d'ER [%] du modèle SVM de La quantités adsorbées 1

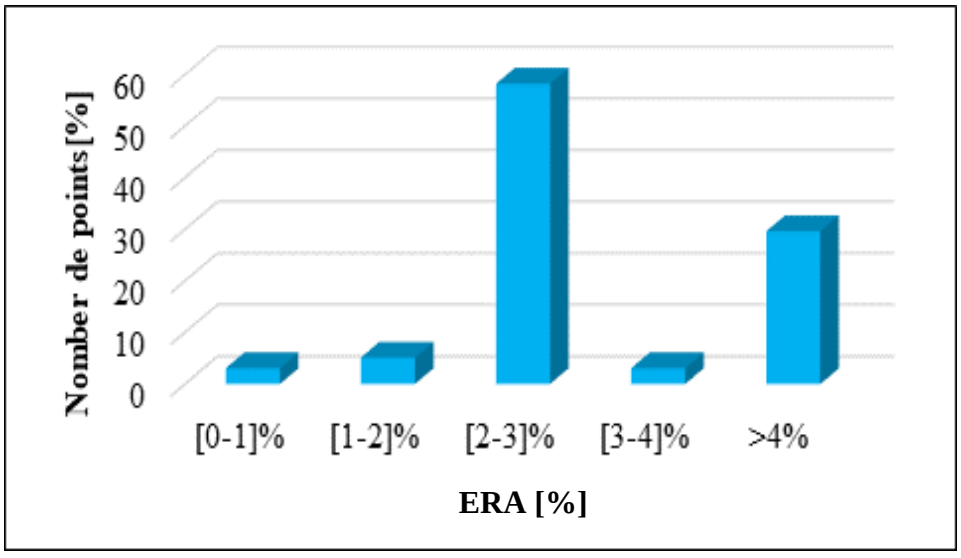


Figure III.8 : Diagrammes de distribution d'ER [%] du modèle SVM de La quantités adsorbées 2

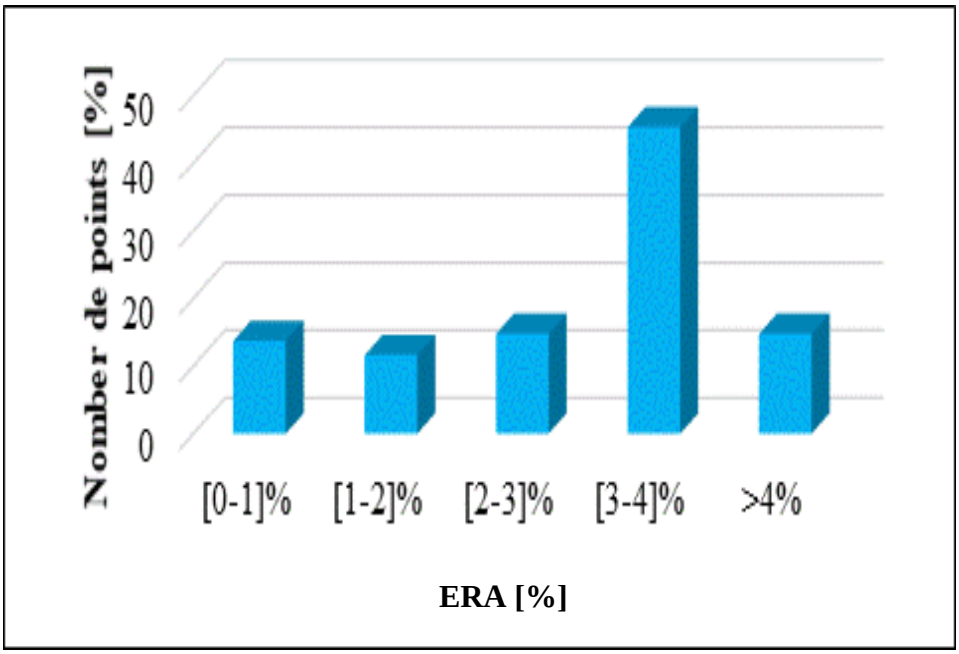


Figure III.9 : Diagrammes de distribution d'ER [%] du modèle SVM de La quantités adsorbées 3.

Le calcul de la fonction d'erreur est nécessaire pour évaluer les courbes l'adsorption statistique aux celles obtenues expérimentalement. Par conséquent, à l'aide de MATLAB, Le tableau III. Montre l'erreur absolue moyenne (MAE), le modèle erreur prédictive (MPE), erreur quadratique moyenne (RMSE), erreur standard de prédiction (SEP), et écart prédictif résiduel (RPD) du quantités adsorbées (q_e1, qe2, qe3) sur la phase d'apprentissage la phase de test, et la phase totale.

Tableau III.5 Performances des modèles SVM.

L'erreur / q_e	MAE	MPE (%)	RMSE	SEP(%)	RPD
q_e1	0.54776149	4.4350559	1.13694365	7.8130479	12.1574916
q_e2	0.47890223	6.22211801	0.69407949	4.50765948	19.9508236
q_e3	2.13516987	9.3623792	3.36208489	7.37799326	15.021334

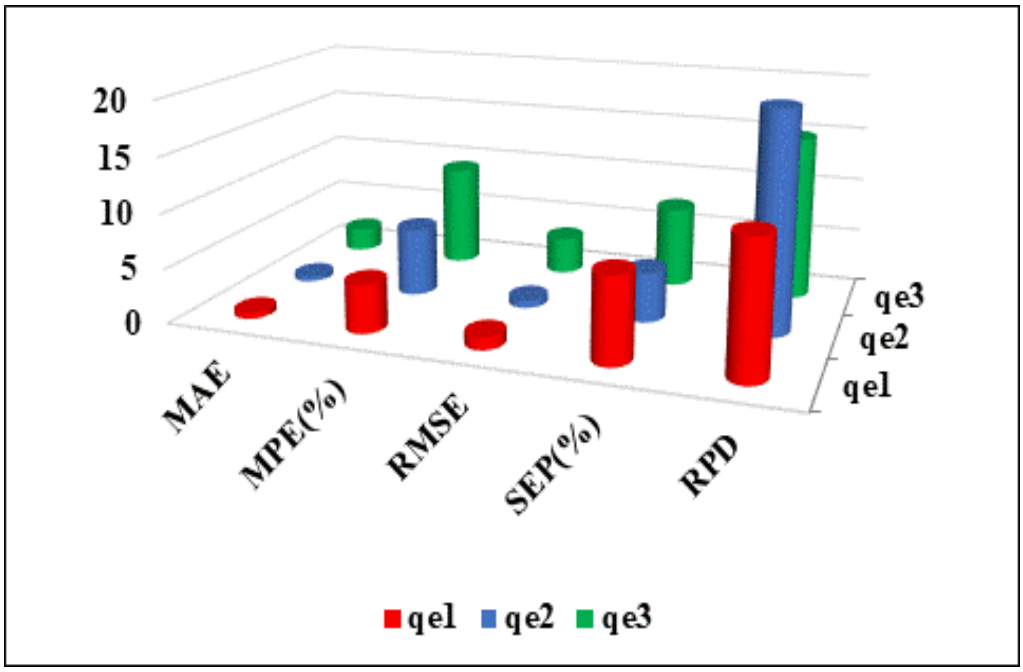
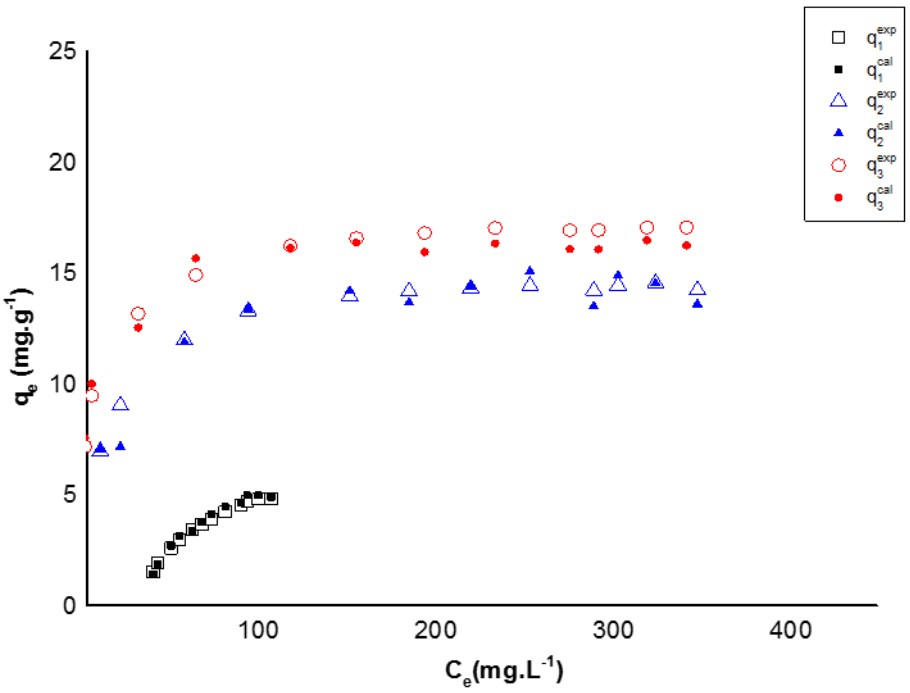


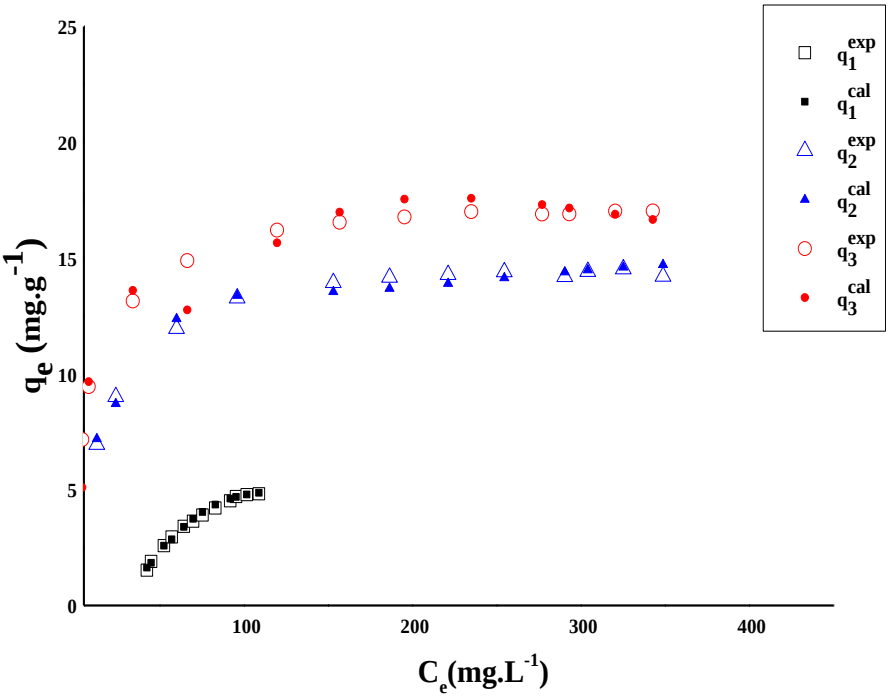
Figure III.10 : Comparaison statistique entre la quantité adsorbée (q_{e1}, q_{e2}, q_{e3})

III.3.3 Comparaison entre les résultats expérimentaux est prédits

Les figures (11,12) représentent la quantité adsorbée (q_e^{exp}) en fonction de la concentration à l'équilibre (C_e) sous l'effet du pH et la température, elles montrent bien que les points calculés par le modèle SVM sont juxtaposés sur celles de l'expérimentales, ce qui confirme une autre fois la performance du modèle SVM optimisé.

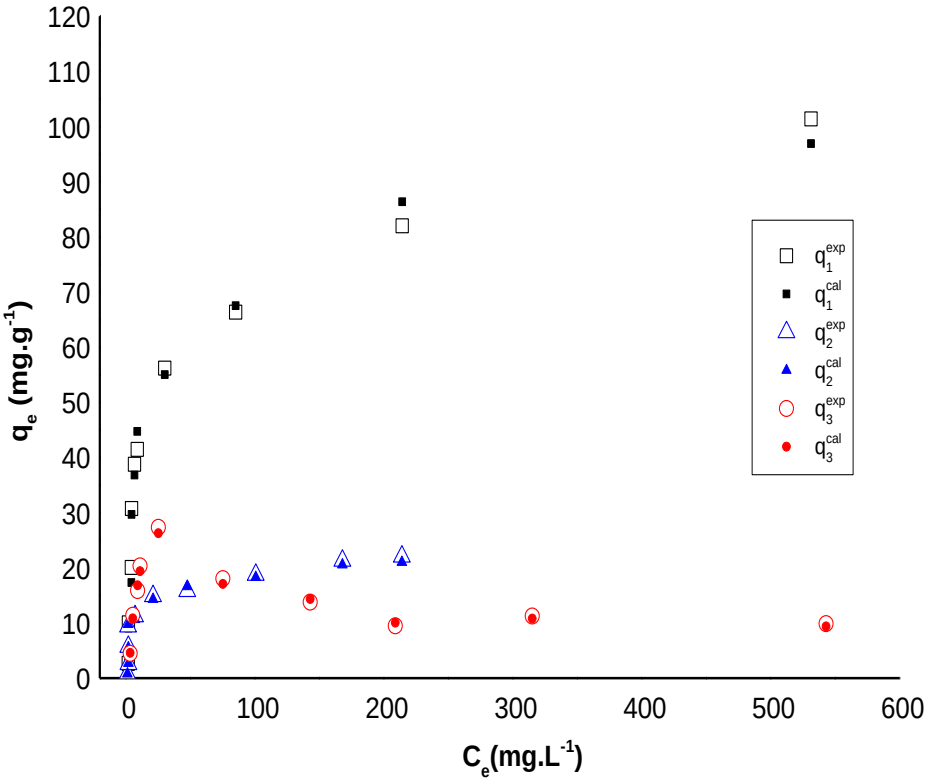


(a)

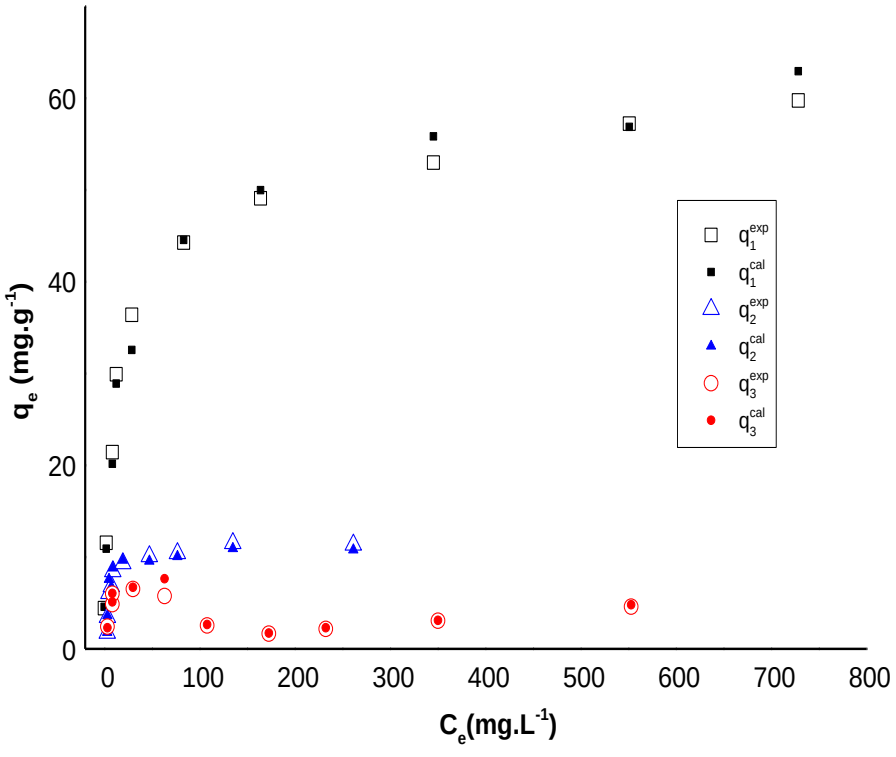


(b)

Figure III.11 : (a)et **(b)** Comparaison entre q_e^{cal} par SVM et q_e^{exp} des trois adsorbats (Pb^{+2}, Hg^{+2}, Zn^{+2}) à différente température **(25 et 30°C)** en fonction de la concentration C_e.



(a)



(b)

Figure III.12 : (a)et **(b)** Comparaison entre q_e^{cal} par SVM et q_e^{exp} des trois adsorbats (Pb^{+2}+Cu^{+2}+Cd^{+2}) à différent pH **(4.1 et 6.45)** en fonction de la concentration C_e.

III.4 Analyse de sensibilité des modèles SVM

Une analyse de sensibilité a été réalisée pour évaluer l'effet des paramètres d'entrée sur la quantité adsorbée. L'entrée la plus pertinente peut être reconnue par le facteur de pertinence r, compris entre -1 et +1. Le facteur de pertinence (r) a été appliqué pour évaluer le degré d'influence de chaque paramètre sur la quantité adsorbée, les valeurs ont été calculées à l'aide de la formule suivante (Baghban et al., 2018) :

$$r_k = \frac{\sum_{i=1}^{n}(X_{k,i}-\overline{X_k})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(X_{k,i}-\overline{X_k})^2 \sum_{i=1}^{n}(y_i-\bar{y})^2}} \text{ , k=1……5} \qquad \text{(III.9)}$$

Avec :

$\mathbf{X_{ki}}$: est l'ième valeur de l'entrée k

$\mathbf{Y_i}$: est l'ième valeur de sortie

$\overline{\mathbf{X_k}}$: est la valeur moyenne de l'entrée k

$\bar{\mathbf{y}}$: est la valeur moyenne du paramétré de sortie

La valeur absolue de r_k a une relation directe sur la sortie ,comme on peut le voir sur la figure (III.13) :

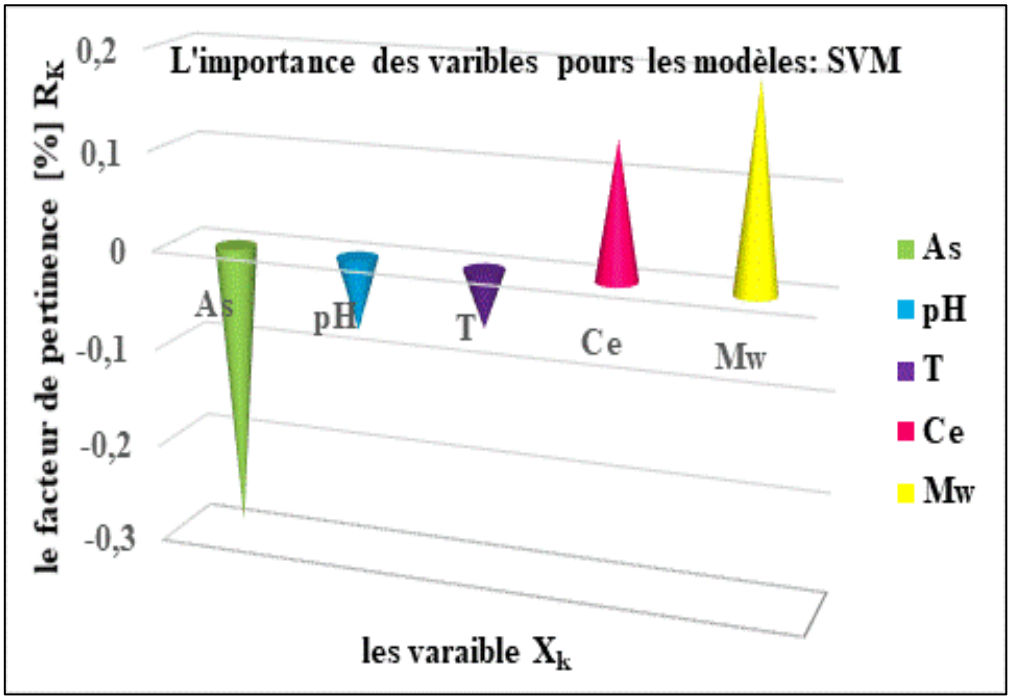


Figure III.13 : Résultats de l'analyse de sensibilité réalise sur les modèles SVM

III.5 Interface graphique

Après avoir construit le modèle SVM, c'est-à-dire déterminer ses poids synaptiques et ses biais, nous avons développé un programme informatique en langage MatLab®. Cette démarche permet à l'utilisateur d'avoir tous les inputs nécessaires à l'exécution de ce modèle pour calculer la sortie. Cette démarche a été faite dans le but d'apporter plus de souplesse à l'utilisation de modèle SVM élaboré. Le programme va donc permettre de faire un calcul simple et rapide du qe.

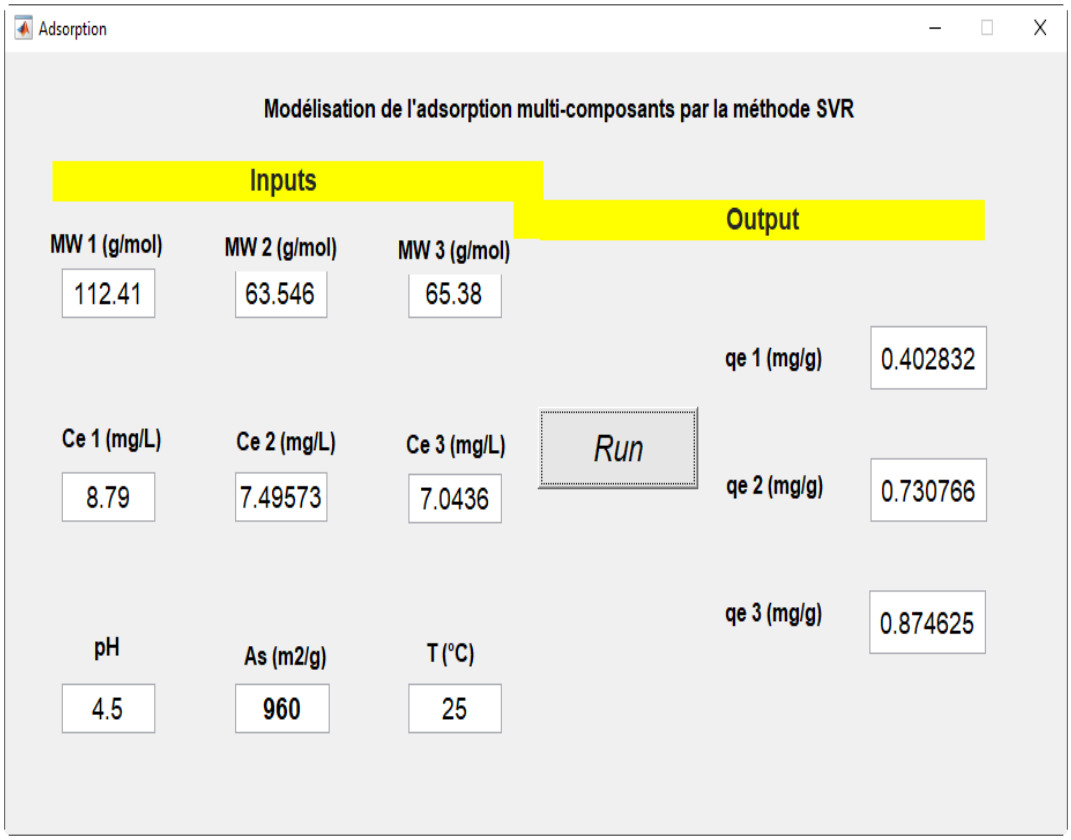


Figure III.14 : Interface graphique développée pour le calcul de la quantité adsorbée

III.6 Conclusion

Cette contribution avait comme objectif de modéliser le processus d'adsorption en employant les techniques d'intelligences artificiels. Les résultats sont présentés comme suit :
Un seul modèle de type support vecteur machine (SVM) a été développé en employant une base de données regroupant plusieurs systèmes d'adsorption de métaux lourds. Ce modèle a été développé pour calculer la capacité d'adsorption en employant quelques paramètres caractérisant les adsorbats et l'adsorbant comme matrice d'entrée. En se basant sur la théorie de la méthode de la modélisation nono-linéaire SVM, un programme Matlab a été fait pour entrainer et optimiser les paramètres de la méthode SVM tes que {nombre de support vecteur

et les hyperparamètres}. La performance de ce modèle été validée sur des données réelles et en employant quelques critères statistiques tels que **REQM {1.13, 0.694, 3.362}, EAM {0.547,0.478, 2.135}** et **R^2 {0.9963, 0.9963, 0.9963},** les résultats montrent que ce modèle est capable de prédire les trois quantités adsorbées avec de très grande précision. Les résultats sont encourageants et ouvrent de nouvelles perspectives de recherche. Plus d'une étude de sensibilité montrant l'intérêt de chaque entrée par rapport à la sortie, une interface graphique conviviale a été conçu pour faciliter le calcul des capacités d'adsorption sans connaitre le phénomène physique.

Conclusion générale

Conclusion générale

Notre travail a pour objectif modélisation de phénomène de l'adsorption multi-composants de différentes adsorbats (métaux lourds) sur différents adsorbants par la méthode machines à vecteur de support et qui font partie de nouvelles techniques sous l'égide de l'intelligence computationnelle ou des science cognitive .Les données expérimentales sont extraites de courbes expérimentales récemment publiées dans des articles scientifiques et analysées de manière statistique à l'aide du logiciel Digitizer 4.1.

Les entrées choisies pour le modèle machine à vecteur de support qui sont usuelles et disponible dans bibliographie, les rendent encore plus pratique dans leur utilisation et par conséquent plus compétitives. Par contre la recherche de la machine à vecteur de support optimal nécessite la confection d'un programme en langage MATLAB comportant plusieurs boucles afin de tenir compte des paramètres constituant la topologie et l'architecture du (SVR).

La modélisation des quantités adsorbées pour un mélange tertiaire sur divers adsorbants. Nous avons utilisé un model (machine à vecteur de support (SVR)) où suivi les fonctions de noyaux (RBF, linear et polynomial) et les paramètres de modèle (la constante C, potentiel de polanyi et largeur de la bande d'un nouyan goussian) avec une différence dans la division de la base de données pour (test, apprentissage), nous avons obtenu un REMS = [2.1451 ,1.9043 ,2.0880], R^2 = [0.9963,0.9963,0.9963],
RPD = [12.1574 ,19.9508 ,15.0213], EAM [0.547,0.478, 2.135].

Les résultats obtenus ont montré qu'un modèle présentent de bonnes aptitudes à prédire les quantités adsorbées bien qu'une minime préférence va pour le modèle SVR.

Dans le futur, on envisage :

- Modéliser l'adsorption multi-composants avec d'autres techniques l'intelligences artificiels tels que RNA, ANFIS et autres.
- Employer autre fonction de minimisation tels que algorithmes d'optimisation de libellule et autre.

Références bibliographiques

Arfane, A., Salhi, A., Krati, M., Tahiri, S., Monkade, M. (2014). « Etude cinétique et thermodynamique de l ' adsorption des colorants red195 et bleu de méthylène en milieu aqueux sur les cendres volantes et les mâchefers ».Journal of Materials and Environnemental Science, 5(6), 1927-1939.

Ahmad, M. A., Rahman, N. K. (2011). « Equilibrium, kinetics and thermodynamic of remazol brilliant orange 3R dye adsorption on coffee husk-based activated carbon ». Chemical Engineering, 170(1),154-161.

Ajmal, M., Ali Khan Rao, R., Anwar, S., Ahmad, J., Ahmad, R. (2003). « Adsorption studies on rice husk: removal and recovery of Cd(II) from wastewater ». Bioresource Technology, 86(2), 147-149.

Abdallah, I. (2017). « Modélisation neuronale et empirique de l'adsorption d'un système multi-composants ».Mémoire de Master .Université Yahia Fares de Medea.

Al Mardini, F. (2008). « Etude de l'adsorption du pesticide Bromacil sur charbon actif en poudre en milieu aqueux Effet compétiteur des matières organiques naturelles ».Thése de Doctorat, Université de Frères Mentouri Constantine.

Bouaghi,O. (2008). « Etude expérimentale et identification du processus de rétention des cations métalliques par des matériaux naturels ». Thése de Doctorat, Université de Ferhat Abbas Setif-1.

Boulkrah,H. (2008). « Etude comparative de l'adsorption des ions plomb sur différents adsorbants ». Mémoiere de Magister. Université de Mouloud Mammeri de Tizi-Ouzou.

Bouziane,N. (2007). « Elimination du 2-mercaptobenzothiazole par voie photochimique et par adsorption sue la bentonite et le charbon actif en poudre ». Mémoire de Magister. Université de Kasdi Merbah Ouargla.

Boris, F. (2006). « SVM : Machines à vecteurs de support ou séparateurs à vastes marges ». Egyptien Informatiques Journal, 13(3), 177-183.

Boser, B. Vapnik, V. (1992). « Training algorithm margin for optimal classifiers ». Perception. 144‑152.

Baghban, A., Kardani, M. N., Mohammadi, A. H. (2018). « Improved estimation of Cetane number of fatty acid methyl esters (FAMEs) based biodiesels using TLBO-NN and PSO-NN models ». Fuel, 232, 620‑631.

Carmen, Z., Daniela, S. (2010). « Textile organic dyes – characteristics , polluting effects

and separation / elimination procedures from industrial effluents – a critical overview. textile organic dye». Rijeka: Intechopen,55‑86.

Clausse, M. (2006). « Etude d'un procédé d'adsorption (température swing adsorption) à Chauffage et refroidissement indirects marc clause to cite this version (température swing adsorption) ».Thèse de doctorat, Université de Abou Bakr Belkaid-Tlemcen.

Chapel, L.Deffuant, G. (2010). « Utiliser des machines à vecteurs de support pour approcher les contours d'une fonction valeur dans des problèmes d'atteinte de cible ». Thèse de Doctorat, Université de Sciences et Technologies de Lille France.

Djeffal, A. (2012). « Utilisation des méthodes support vector machine (SVM) dans l'analyse des bases de données ».Thése de Doctorat, Université Mohamed Khider – Biskra.

Ettlili, N. (2018). « Procédé d ' adsorption avec régénération électrothermique pour la capture du dioxyde de carbone en post-combustion » .Thése de doctorat, Université Institut National Polytechnique de Lorraine Français .

Esmaeelza,I. (1991). « Operator theory: advances and applications. integral equations and operator theory». Basel,14(1), 153‑155.

Guyon, I. Boser, B.& Vapnik, V. (1993). « Automatic capacity tuning of very large VC-dimension classifiers. advances in neural information processing systems ».In Advances In Neural Information Processing Systems, 147‑155.

Hamel, L. (2009). « Knowledge discovery with support vector machines, inn knowledge discovery with support vector machines ». John Wiley Sons ,150-152.

Hasan, M., Boris, F. (2006). « SVM : Machines à vecteurs de support ou séparateurs à vastes marges ». Survey,16(1),112-116

Howie, J. (1984). « Cohomology of one-relator products of locally indicable groups ». Journal of the London Mathematica Society, 2(3), 419‑430.

Khelifi, O. (2018). « Etude de l'adsorption du nickel et du cuivre sur un charbon actif préparé à partir des boues de station d'épuration ». Thèse de Doctorat. Université 8 Mai 1945 Guelma .

Guesmia,S. (2015). « L'Adsorption des colorants (Bleu De methylene et chrome Iii) sur La bentonite ». Mémoire de Master . Université Mohamed Khider – Biskra

Khoualdi, A., Medjani, A. (2016). « Adsorption du bleu de méthylène par le charbon actif et l'argile ». Memoire de Master. Université Larbi Ben M'hidi Oum El Bouaghi.

Krou, N. J. (2010). « Etude expérimentale et modélisation d'un procédé séquentiel AD-OX d'élimination de polluants organiques ». Thése de Doctorat.Université de Abderrahmane Mira Bejaia.

Kadirvelu, K., Goel, J., Rajagopal, C. (2008). « Sorption of lead , mercury and cadmium ions in multi-component system using carbon aerogel as adsorbent ». Journal of Hazardous Materials, 153(1-2), 502‑507.

Khandegar, V., Saroha, A. K. (2013). « Electrocoagulation for the treatment of textile industry effluent - Areview ». Journal of Environnemental Management, 128, 949‑963.

Kulkarni, V. R., Mujawar, S., Apte, S. (2010). « Ash unction implementation using a rtificial n eural n etwork. soft computing », Modeling Earth Systems and Environment, 1(1), 1‑8.

Kivinen, J., Smola, A. J., Williamson, R. C. (2002). « Online learning with kernels. advances in neural information processing systems ». In Proceedings of the International Conference on Artificial Neural Networks, 100(10), 1‑12.

Kuhn, H. W. (1951). « Tucker,"nonlinear programming,". proc. and berkeley symposium on mathematical statistics and probability ». Chemical Engineering,10, 481‑492.

Kadirvelu, K., Goel, J., Rajagopal, C. (2008). « Sorption of lead , mercury and cadmium ions in multi-component system using carbon aerogel as adsorbent », Watre Research, 153, 502‑507.

Kavand, M., Fakoor, E., Mahzoon, S. (2017). « An improved film – pore – surface diffusion model in the fixed-bed column adsorption for heavy metal ions : single and multi-component systems », Process Safety and Environmental Protection, 113, 330‑342.

Kidwell, A. (2015). « Estimation of subsurface temperature anomaly in the indian ocean during recent global surface warming hiatus from satellite measurements: A support vector machine approach ». Remote Sensing of Environment, 160(2), 63‑71.

Lajevardi, S. M., Lech, M. (2008). « Facial expression recognition from image sequences using optimized feature selection ». International Conference Image and Vision Computing New Zealand ,26(1),1-6.

Meriem,S., (2017). « Contribution a la moélisation de ladsorption dynamique : étude compactive » . Mémoire de Magister, Université Yahia Fares de Medèa.

Mustapha, D. (2014). « Argile de maghnia: purification et adsorption de polluants ». Thèse de Doctorat, Université de Ahmed Ben Bbila d'Oran.

Milenova, B. L., Yarmus, J. S., Campos, M. M. (2005). « SVM in oracle database 10g: removing the barriers to widespread adoption of support vector machines». Proceedings, 3, 1152‑1163.

Mohan, D., Singh, K. P. (2002). « Single- and multi-component adsorption of cadmium and zinc using activated carbon derived from bagasse an agricultural waste ». Water research, 36(9), 2304-2318.

Mohamed, R., Farag, A. (2004). « Regression using support vector machines: basic foundations ».Sci Total Environ,22(2),87-94.

Nandi, B.Goswami, A. Purkait, M. (2009). « Adsorption characteristics of brilliant green dye onkaolin ». Journal of Hazardous Materials, 161(1), 387‑395.

Ruthven, D. (2008). « Fundamentals of adsorption equilibrium and kinetics in microporous solids » . In Adsorption and Diffusion, 1‑43.

Sekirifa, M. (2013). « Étude des propriétés adsorbantes des charbons activés issus des noyaux de dattes, application au traitement d'efflue nt aqueux ». Thèse de Doctorat, Université de Annaba-Badji Mokhtar.

Sedira, N. (2013). « Etude de l'adsorption des métaux lourds sur un charbon actif issu de noyaux de dattes ». Mémoire de Magister, Université de Annaba-Badji Mokhtar.

Souheila, B. (2012). « Eliminations des polluants spécifiques par adsorption sur charbon actif et argile traitée et non traitée ». Memoire de Magister,Université Mohamed Chérif Messaadia –Souk-Ahras .

Schuhmacher, M., Granero, S., Bellés, M., Llobet, J. M., Domingo, J. L. (1996). « Levels of metals in soils and vegetation in the vicinity of a municipal solid waste incinerator». Toxicological Environmental Chemistry, 56(1‑4), 119‑132.

Tran, H. N., You, S. J., Hosseini-Bandegharaei, A., Chao, H. P. (2017). « Mistakes and inconsistencies regarding adsorption of contaminants from aqueous solutions: acritical review». Water Research, 120, 88‑116.

Todeschini, R., Ballabio, D., Grisoni, F. (2016). « Beware of unreliable Q2 A comparative study of regression metrics for predictivity assessment of QSAR models ». Journal of Chemical Information and Modeling, 56(10), 1905‑1913.

Van Der Stap, J., & Klaasse, S. (2016). « Labwaarden: acute nierinsuffi ciëntie ». Nursing 22(6), 36‑38.

Vapnik, V. (2013). « The nature of statistical learning theory ». Springer Science Business Media,20(1),40-42.

Zerfa abdnour, k. (2015). « Etude expérimentale de l'adsorption d'un colorant sur des déchets agricoles (écorces d'orange et de tiges des dattes) », Mémoire de Master ,Universite L'arbi Ben M'hidi Oum El Bouaghi.

Zhang, Y., Gao, X., Smith, K., Inial, G., Liu, S., Conil, B., Pan, B. (2019). « Integrating water quality and operation into prediction of water production in drinking water treatment plants by genetic algorithm enhanced artificial neural network».Water Research, 164, 114888.