République Algérienne Démocratique Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITE YAHIA FARES DE MEDEA
Faculté de la Technologie
Département de Génie de Procédés et Environnement

MEMOIRE DE MASTER

Présenté par :

OULD MEDJEBER Manel

ZEGUAI Hamza

En vue de L'obtention du Diplôme de master en Génie des Procédés

Option : Génie des Procèdes de l'environnements

Dirigé par :

- Dr. LAIDI Maamar
- Pr. HANINI Salah
- Mr. Moussaoui Mohamed

Thème

Modélisation par réseaux de neurones artificiels et machines à vecteurs supports :

Contribution à l'étude de l'influence des descripteurs sur la solubilité de quelques solutés solides dans le CO_2 supercritiques

Soutenue le 15 mai 2018 Devant l'honorable jury composé de :

Président : Mr.BITTAM Said (M.A.A,Université de Médéa)
Examinateur : Mr.MESELEM Yamin (M.A.B,Université de Médéa)

Promotion 2017-2018

Remerciements

Au terme de ce Mémoire nous tenons à remercier tout naturellement en premier lieu ***Dieu le tout Puissant*** *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 Monsieur ***M. LAIDI****, Maître de Conférences au Pôle Universitaire Dr. Yahia Farés 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. Si ce travail, a pu faire l'objet de communications nationales et international c'est bien grâce à ses directives et son soutien.*

Un hommage particulier au Co-promoteur de cette étude ***Mr. S. HANINI*** *qu'il trouve ainsi l'expression de notre profonde reconnaissance pour sa compétence, ses encouragements et ses précieux conseils durant le déroulement de ce travail. Si ce travail, a pu faire l'objet de communications nationales et international c'est bien grâce à ses directives et son soutien. Nos remerciements vont aussi à Monsieur* ***C. SI MOUSSA*** *Maître de Conférences au Pôle Universitaire Dr. Yahia Farés de Médéa, pour leur aide et encouragements.*

Nous profitons de l'occasion de la présentation de ce travail pour exprimer Mes vifs remerciements à Monsieur **M.MOUSSAOUI** *notre Co-promoteur.*

Nous remercions le Président de Jury *Monsieur* **S.BITTAM***, qui a bien voulu assurer la Présidence de notre jury ainsi que Monsieur* **Y.MESELLEM** *et Monsieur* **M.HENTABLI***, pour avoir accepté de juger ce modeste travail.*

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

OULD MEDJEBER MANEL
ZEGUAI HAMZA

Dédicaces

Au-delà des personnes, des lieux et des temps, je dédie ce travail avec ma profonde conviction, à tous ceux qui ont toujours cru et sans doute à la Science. Cette lumière qui éclaire les esprits et leurs permet de transcender les limites installées par les sociétés et les cultures dans leurs périodes de dégénérescence.

À ceux qui savent concrétiser leurs idées et savent réussir.

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

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

Je dédie ce travail à mes chers parents :

L'épaule solide l'œil attentif compréhensive et la personne la plus digne de mon estime et de mon respect. A l'homme de ma vie, mon exemple éternel, mon soutien moral et source de joie et de bonheur, celui qui s'est toujours sacrifié pour me voir réussir, que dieu te garde dans son vaste paradis, aucun dédicace ne serait exprimer mes sentiments ; ***Mon père.***

À la lumière de mes jours, la source de mes efforts, la flamme de mon cœur, ma vie et mon bonheur, qui a œuvré pour ma réussite, de par son amour, son soutien, tous les sacrifices consentis et ses précieux conseils, pour toute son assistance et sa présence dans ma vie, reçois à travers ce travail aussi modeste soit-il, l'expression de mes sentiments et de mon éternelle gratitude; ***Ma mère****.*

À mes frères et à mes sœurs.

À mon binôme Hamza.

À mes collègues et amis de l'Université Dr. Yahia Fares de Médéa

À tous les enseignants de l'Université Dr. Yahia Fares de Médéa et ma promotion (2017-2018).

À tous qui m'ont encouragé et soutenu dans mes moments les plus difficiles.

Manel

Dédicaces

Je dédie ce mémoire à :

***Mon père**, qui peut être fier et trouver ici le résultat de longues années de sacrifices et de privations pour m'aider à avancer dans la vie. Puisse Dieu faire en sorte que ce travail porte son fruit ; Merci pour les valeurs nobles, l'éducation et le soutient permanent venu de toi.*

***Ma mère**, allahyrhmha qui a œuvré pour ma réussite, de par son amour, son soutien, tous les sacrifices consentis et ses précieux conseils, pour toute son assistance et sa présence dans ma vie, reçois à travers ce travail aussi modeste soit-il, l'expression de mes sentiments et de mon éternelle gratitude.*

***Mes frères** et **ma sœur** et **ma femme** qui n'ont cessé d'être pour moi des exemples de persévérance, de courage et de générosité.*

*Mon binôme **Manel.***

*Mes **enseignants** à l'**UYFM** qui doivent voir dans ce travail la fierté d'un savoir bien acquis.*

*À mes **collègues** et **amis** à l'université Dr. Yahia Farès et ma **promotion** génie des procédés **PCGE** (2017/2018).*

À tous qui m'ont encouragé et soutenu dans mes moments les plus difficiles.

Hamza

قابلية ذوبان بعض لانظمة المذابة الصلبة في نمذجة ثاني أكسيد الكربون فوق الحرج باستخدام نماذج RNA و SVM

ملخص :

ان الـ غرض من هذه الدراسة هو تقييم إمكانيات النماذج القائمة على الذكاء لاصطناعي لنمذجة ذوبان بعض لانظمة المذابة الصلبة في ثاني أكسيد الكربون فوق الحرج. نقترح دراستين خوارزميتين جديدتين للنمذجة، وتتمثل في النمذجة العصبية لاصطناعية RNA و شعاع الدعم لالي SVM ، للحصول على ذوبان 25 نظامًا على نحو الدقيق نستخدم توليفات مختلفة من المعلومات التشغيلية الشائعة ونوعين من الواصفات في قاعدة بيانات محكمة تم جمعها من المقالات العلمية ، والتي تحتوي على انظمة مذابة مختلفة. تم برمجة هذه الطرق باستخدام برنامج Matlab وحصلنا على عمارة عصبية بتصميم بنية لخلايا عصبية اصطناعية متكونة من {10 ،16، 1} في المداخل المموهة و الطبقة المخفية و طبقة الخروج على التوالي.

تم تحسين هذا لأخير و اختباره لتنبؤ بذوبان لانظمة مع {0.99692 , 0.0472} من حيث R^2 و MSE(test) (%)على التوالي . مقارنتا مع نموذج SVM ، تظهر خوارزمية RNA دقة أعلى.

لتطور نتائج لاختبار مع {0.910806 , 0.174} من حيث R^2 و(%) MSE(test) على التوالي. بحيث تؤكد النتائج أن نموذج RNA كان أكثر دقة واستقرارًا وسرعة من نموذج SVM.

من خلال النظر في دقة التنبؤ واستقرار النماذج والكفاءة الحسابية، يوصى بشدة استخدام نموذج RNA لتقدير قابلية الذوبان في انظمة المذابة الصلبة المحددة في ثاني أكسيد الكربون فوق الحرج.

الكلمات الرئيسية: الذوبان، المذابة الصلبة، السوائل فوق الحرجة، الشبكات العصبية لاصطناعية، شعاع الدعم لالي.

La modélisation de la solubilité de quelque soluté solide dans le dioxyde de carbone supercritique avec la méthode de RNA et SVM

Résumé :

L'objectif principal de cette étude est d'évaluer le potentiel des modèles basés sur l'intelligence artificielle pour modéliser la solubilité de certains solutés solides dans le dioxyde de carbone supercritique en utilisant. Cette étude propose deux nouveaux algorithmes de modélisation, à savoir les réseaux neuronaux artificiels (RNA) et Support Vecteur Machine (SVM), pour une prédiction précise de la solubilité de 25 systèmes utilisant différentes combinaisons de paramètres opérationnels couramment disponibles et deux types de descripteurs. Dans ce travail,la base de données a été recueillie à partir de documents scientifiques, qui contiennent la solubilité de divers systèmes.

Ces approches ont été programmées en utilisant le logiciel Matlab. Le RNA conçu a été obtenu avec l'architecture de neurones {10, 16, 1} dans l'entrée, la couche cachée et la couche de sortie respectivement. Ce dernier a été optimisé et testé pour prédire la solubilité avec {0.0472 , 0.99692} en termes de MSE(test) (%) et R^2 respectivement. En comparant avec le modèle ANN, l'algorithme SVM, cette comparaison montre une plus grande précision pour la phase de test avec {0.174 , 0.910806} en termes de MSE(test) (%) et R^2 respectivement. Les résultats confirment que le modèle RNA était plus précis, stable et rapide que la méthode SVM.

En considérant la précision de la prédiction, la stabilité du modèle et l'efficacité de calcul, le modèle SVM est fortement recommandé pour estimer la solubilité du soluté solide sélectionné dans le dioxyde de carbone supercritique.

Les mots de clés : solubilité, soluté solide, fluides supercritique, réseaux neuronaux artificiels, support vecteurs machines.

Solubility of some solid solute in supercritical carbon dioxide modeling using ANN and SVM approach

Abstract:

The main objective of this study is to evaluate the potential of artificial intelligence-based models to model the solubility of some solid solute in supercritical carbon dioxide using. This study proposes two novel modeling algorithms, i.e. Artificial Neural Networks (ANN) and Support Vector Machine (SVM), for accurate prediction of the solubility of 25 systems using different combinations of commonly available operational parameters and two type of descriptors. In this work database has been collected from scientific papers, which contains the solubility of various systems.

These approaches were programmed using Matlab software. The designed ANN was obtained with the architecture of {10, 16, 1} neurons in the input, hidden and the output layer respectively.

This latter was optimized and tested to predict the solubility with {0.0472 , 0.99692} in terms of MSE(test) (%) and R^2 respectively. Comparing with the ANN model, the SVM algorithm show higher accuracy for test phase with {0.174 , 0.910806} in terms of MSE(test) (%) and R^2 respectively. Results confirm that the SVM model was more accurate, stable and rapid than the ANN method.

By considering the prediction accuracy, model stability and computational efficiency, the SVM model is highly recommended to estimate the solubility of the selected solid solute in supercritical carbon dioxide.

Keywords: Solubility, solid solute, supercritical fluids, artificial neural networks, support vector machine.

Table de matières

Table de matières……...……I

Nomenclature……V

Abréviation ……...VII

Liste des figures……..……...….VIII

Liste des tableaux……..……....X

Chapitre I : La solubilité des solutés solides dans le co_2 supercritique.

Introduction générale **1**

I.1. Introduction **5**

I.2 Historique **5**

I.3 Principe **6**

I.4 Etat supercritique **6**

I.4.1 Propriétés des fluides supercritiques **7**

I.5 Le CO_2 supercritique **8**

I.5.1 Le pouvoir solvant à géométrie variable **8**

I.5.2 Polarité du co_2 supercritique **9**

I.6 Nature d'équilibre **10**

I.7 Condition thermodynamique des équilibres **10**

I.7.1 Degré de liberté **11**

I.7.2 Règle de phase **11**

I.7.3 Condition d'équilibre solide – fluide **11**

I.8 Fugacité 12

I.9 La solubilité des solutés dans les fluides 14

I.10 Les modelés de prédiction de la solubilité 14

I.10.1 Les modèles semi empirique 15

I.10.1.1 Modèle de chrastile 15

I.10.1.2 Modèle de del valle et aguilera 16

I.10.1.3 Modèle d'adachi et lu 16

I.10.1.4 Modèle de bartle

17

I.10.1.5 Modèle de mendez-santiago et teja 17

I.10.1.6 Modèle de kumar et johnston 18

I.10.2 Les équations d'état 18

I.10.2.1 Equation d'état du gaz parfait – loi des états correspondants 19

I.10.2.2 Equations d'état cubique 20

I.10.2.3 Equations d'état non cubique (basée sur la théorie de perturbation) 26

I.10.3 Les réseaux de neurones artificiels 29

I.11 Règles de mélange 30

I.12 Analyse numérique et mathématique 30

I.13 Conclusion 31

Chapitre II : Le réseau de neurone artificiel.

II.1. Introduction 32

II.2 Historique sur les réseaux de neurones 32

II.3 Les réseaux de neurones **33**

II.3.1 Principe du neurone artificiel **33**

II.3.2 Définition **35**

II.3.3 Architecture des réseaux de neurones **36**

II.3.3.1 Les réseaux de neurones non bouclés **36**

II.3.3.2 Les réseaux de neurones bouclés **37**

II.3.4. Apprentissage des réseaux de neurones **39**

II.3.4.1. Type d'apprentissage **41**

II.3.4.2. Algorithme d'apprentissage **42**

II.4. Conception d'un réseau de neurones **44**

II.4.1. Détermination des entrées/sorties du réseau de neurones **44**

II.4.2. Choix et préparation des échantillons **44**

II.4.3. Elaboration de la structure du réseau **45**

II.4.4. Apprentissage **46**

II.4.5. Validation et tests **46**

II.5. Conclusion **47**

Chapitre III : Contribution à l'étude de l'influence des descripteurs sur la solubilité de quelques solutés solides dans le CO_2 supercritiques.

III.1. Introduction **48**

CONTRIBUTION 1 : Modélisation par réseaux de neurones artificiels.

III.2 La modélisation neuronale **49**

III.2.1 Méthodologie de travail **49**

III.2.1.1 Sélection et collecte des données thermodynamiques 51

III.2.1.2 Calcul des descripteurs et prétraitement de la base de données 52

III.2.1.3 Séparation de la base de données 57

III.2.1.4 Normalisation des données 58

III.2.1.5 Sélection des paramètres de réseau de neurones 59

III.2.1.6 Validation et généralisation 60

III.2.1.7 Exploitation des résultats 60

III.2.1.8 Conclusion 66

CONTRIBUTION 2 : Modélisation par support vecteurs machine.

III.3 La modélisation par support vecteurs machine 67

III.3.1 Le principe des SVM 68

III.3.2 Exploitation des résultats (SVM) 71

III.3.3 Analyse du modèle de régression de support vecteur machine 72

III.3.4 Comparaison entre RNA et SVM 73

III.3.5 Conclusion 73

Conclusion générale 74

Nomenclature

A	Energie libre de Helmholtz, (J).
A_0, A_1, A_2	Paramètres du modèle de Chrastile.
A, B, A', B'	Paramètres du modèle de Menèz-Santiago-Teja.
a	Paramètre d'énergie de l'équation d'état.
b	Paramètre de volume de l'équation d'état.
am, bm	Constantes des mélanges.
b_1, b_2, b_3	Paramètres du modèle de Kumar-Johnston.
b_i	Biais du neurone.
B'', C''	Coefficients du Viriel.
C_1, C_2, C_3, K_1, K_2	Paramètres du modèle de Bartle.
C	Le nombre de constituants
d_m	Diamètre de sphère dure, (A°).
E	Facteur du poiyting.
e_0, e_1, e_2	Paramètres du modèle d'Adachi et lu.
F	Paramètre de l'équation Patel-Teja-Valderrama.
f	Fugacité
${}_g E$	Energie d'excès, (J).
$g_n^{hs}(\eta)$	Fonction de distribution radiale moyenne.
I	Paramètres du modèle de Del valle et aguilera.
I_n , C_n,	Constants relatif à l'équation PC-SAFT.
K	Conctante de Boltzman, (J/K).
k_{ij}, l_{ij}	Paramètres d'interactions binaires ajustables.
{n ; (m_l) ; k}	Nombre de neurones dans la couche d'entrée, (n).
	Nombre de neurones dans la 1ère couche cachée, (m).
	Nombre de neurones dans la 2ème couche cachée, (l).
	Nombre de neurones dans la couche de sortie, (k).
m	Paramètre de l'équation Patel-Teja-Valderrama.
N	Nombre d'Avogadro. .
P_c	Pression critique, (bar).
P^{sub}	Pression de sublimation, (bar).
R	Coefficient de corrélation.
S	Solubilité.
T_c	Température critique, (K).
T_r	Température réduite, (K).
μ_k / kT	Contribution de la perturbation dans l'équation PC-SAFT.
v	Variance.
v_i^S	Volume molaire, (cm^3/mole).
w	Facteur acentrique**.**
$w_{k,j}$	Poids synaptiques du neurone entre le neurone, (k) et le neurone, (j).
y_2	Fraction molaire du soluté solide.
Z	Facteur de compressibilité.

Caractères grecs

α , β	Paramètres du modèle d'Adachi et lu.
$\alpha(Tr)$	Fonction dans le paramètre attractive d'une équation d'état**.**
$\beta(T_r, w)$	Fonction dans le paramètre répulsif d'une équation d'état.
δ	Paramètre de solubilité d'Hildebrand.
ε / kT	Paramètre d'énergie pour corps purs.
□	le nombre de phase
ϕ	Coefficient de fugacité.
η	Fraction de garnissage.
π	Pression interne, (bar).
Ω_b , Ω_c	Paramètre de l'équation Patel-Teja-Valderrama.
ρ_1	Densité du solvant, (kg/m^3).
ρ_{fs}	Masse volumique du fluide à l'état SC, (kg/m^3).
ρ_{liq}	Masse volumique du fluide à l'état liquide, (kg/m^3).
μ	Potentiel chimique, (J).
σ_{ii}	Diamètre du segment, (A°).

Indices et exposants

assoc	association.
att	attractif.
C	Critique.
hc	hard sphere.
fs	Fluide supercritique
i	Composé (i)
j	Composé (j)
pert	perturbation.
ref	Référence.
rep	Répulsif.
RES	Résidue
RS	Real solution.
seg	Segment
sub	Sublimation

Abréviation	
1D	une Dimension
2D	deux Dimensions
3D	trois Dimensions
AINS	Anti-Inflammatoire Non Stéroïdienne
ARP	Algorithme de RétroPropagation
ANN	Artificial Neural Networks
AS	Adachi-Sugi.
CFC	Chloro Fluoro Carbones.
CO_2 SC	Dioxyde de Carbone Supercritique
COV	Composés Organiques Volatiles.
CPA	Cubic Plus Association.
EFSC	Extraction par Fluide Super Critique.
MAARD	Moyenne Absolute Average Relative Deviation $\left(MAARD(\%) = \frac{100}{n} \sum_{i}^{n} \frac{\left\| y_i^{\exp} - y_i^{cal} \right\|}{y_i^{\exp}} \right)$
MSE	l'Erreur Quadratique Moyenne
MLP	Multi Layer Perceptron
MLR	Regression Linéaire Multiple
MPR	Modefied- Peng-Robinson
Nc	Neurones cachés
NC	Nombre de Composés.
OF	Objective Function.
PACT	Perturbed Association Chain Theory.
PC	Point Critique.
PC-SAFT	Perturbed Chain Statistical Associated Fluid Théory.
PHCT	Perturbed Hard Chain Theory.
PR	Peng-Robinson
PTV	Patel-Teja-Valderrama
QSAR	Quantitative Struture-Activity Relationship
RBF	Radial Basis Function
RLM	Régression Linéaire Multiple
RK	Redlich et Kwong
RNA	Réseau de Neurones Artificiels.
RNS	Structure du Réseau de Neurones
RNSO	Structure du Réseau de Neurones Optimisé
RNSO	Réseau de Neurone Statique Optimisé.
RSD	Déviation Standard Relative
SAFT	Statistical Associated Fluid Théory.
SAR	Relation Activité Structure
SC	Supercritique
SC	Super-Critical.
SMILE	Simplified Molecular Input Line Entry Specification
SRK	Soave-Redlich-Kwong.
SVM	Machine à Vecteurs Supports
vdW1	Ven-Der-Waal

Liste des figures

Chapitre I : Modèles de prédiction de la solubilité des solutés solides dans les fluides supercritiques.	
Figure I.1 : Représentation bidimensionnelle (Pression, Température) du diagramme de phases d'un corps pur, [Ziemons ,2006].	6
Figure I.2 : Solubilité du naphtalène dans le dioxyde de carbone, [MCHugh et Krukonis, 1986].	7
Figure I.3: Paramètres de solubilité de différents liquides et fluides supercritiques [De Castro, 1994].	10
Figure I.4 : Classification des différents modèles de prédiction de la solubilité des Solides dans les fluides supercritiques.	15
Figure I.5 : Les étapes de formation de modèle moléculaires à partir de théorie associative SAFT, [Müller et Gubbins, 2001].	27
Figure I.6 : Le modèle moléculaire dans un mélange de deux substances selon l'équation d'état PC-SAFT. [Tumakak et al, 2005].	28
Figure I.7 : Classification des différents types des équations d'état selon les théories développées à partir des interactions entre les molécules, [Valderrama, 2003].	29

Chapitre II : Le réseaux de neurones.	
Figure II.1 : Neurone artificiel.	34
Figure II.2 : Différents types de fonction de transfert pour le neurone artificiel.	34
Figure II.3 : Réseau de neurones à n entrées, une couche de Nc neurones cachés et N0 neurones de sortie.	37
Figure II.4 : Réseau de neurone bouclé.	38
Figure II.5 : Erreur moyenne sur la base d'apprentissage en fonction du nombre d'itérations.	40
Figure II.6: Organigramme de conception d'un réseau de neurones.	47

Chapitre III : Modélisation de la solubilité de quelques solutés solides dans le CO_2 supercritique par RNA et SVM.	
Figure III.1 : Méthodologie de calcul de la solubilité par les réseaux de neurones.	50
Figure III.2 : Capteur d'écran du site intitulé pubchem.	54
Figure III.3 : Capteur d'écran du site intitulé PaDel.	55
Figure III.4 : Régression linéaire de l'architecture de 1 à 10 → 11 calculée par le RNA optimisé pour les phases d'apprentissage, test et validation.	64
Figure III.5 : Régression linéaire de l'architecture de 4 à 10 → 11 calculée par le RNA optimisé pour les phases d'apprentissage, test et validation.	64
Figure III.6 : Régression linéaire de l'architecture de 1 à 3+7 à 10 → 11 calculée par le RNA optimisé pour les phases d'apprentissage, test et validation.	65
Figure III.7 : Réseau de neurones artificiels.	65
Figure III.8: Architecture d'une machine de régression construite par le SV algorithme.	71
Figure III.9 : le régression l'linéaire de support vecteur machine.	72

Liste des tableaux

Chapitre I : Modèles de prédiction de la solubilité des solutés solides dans les fluides supercritiques.	
Tableau I.1 : Propriétés physiques pour les gaz, le CO_2 supercritique et les liquides, [De Castro, 1994].	8
Tableau I.2 : Paramètres critiques des fluides à l'état supercritique [De Castro, 1994].	8
Tableau I.3 : Les modèles sélectionnés pour la fonction de terme attractif $\alpha(Tr)$ dans l'équation d'état.	26
Tableau I.4 : Le nombre des équations, des variables et le processus de calcul par les différents modèles d'estimation, [Gani et al, 2006].	30

Chapitre III : Modélisation de la solubilité de quelques solutés solides dans le CO_2 supercritique par RNA et SVM.	
Tableau III.1 : Source et domaine des données utilisées pour l'apprentissage et la validation du modèle de RNA pour le système de 25 solutés solides qui sont solubles dans le CO_2 supercritiques.	51
Tableau III.2 : Les composés et ses formules SMILE.	54
Tableau III.3 : Nombre de descripteurs calculés.	55
Tableau III.4: Les descripteurs et leurs types et leurs descriptions (www.scbdd.com).	57
Tableau III.5 : Les différentes entrées de la matrice.	57
Tableau III.6 : Les sept différents systèmes de réseau de neurones.	58
Tableau III.7 : Structure du réseau de neurones optimisé (RNSO).	60
Tableau III.8 : Architecture des RNA utilisés.	61
Tableau III.9 : les poids et les biais**.**	63
Tableau 10 : Paramètres du modèle SVM pour la solubilité de quelques solutés solides dans le CO_2 supercritique.	72
Tableau 11 : Évaluation de SVM model par rapport au modèle MLR pour les données de test.	73

Introduction générale

Avant qu'un médicament soit disponible à ses récepteurs, il doit être dissoute dans le liquide biologique entourant la solubilité des récepteurs et par conséquent, est un sujet important en sciences pharmaceutiques. Il y a un problème de solubilité avec près de 40 % des candidats médicaments, et toute tentative de prédire la solubilité est très important dans le processus de découverte de médicaments d'enquêtes. La plus ancienne règle pour la prédiction de la solubilité est "comme se dissout comme". Cette règle est appliquée à toutes équations de solubilité, (Jouyban, 2010). La solubilité est définie comme la quantité maximale d'un médicament dissous dans un volume donné d'un solvant/solution.

La technologie d'extraction de fluide supercritique (SFET) a attiré beaucoup d'attention dans la séparation et la purification des substances naturelles. En particulier, le CO_2 supercritique ($scCO_2$) offre des caractéristiques intéressantes en tant que solvant non toxique, non inflammable, à faible coût et facile à traiter, avec des propriétés critiques modérées, (Zabihi *et al.*, 2010; Zhu *et al.*, 2015). Afin de développer un système efficace d'extraction ou de purification supercritique, il est nécessaire d'avoir une idée approfondie du comportement de la phase et des propriétés d'équilibre de la substance cible dans les milieux supercritiques, (Ismadji *et al.*, 2011).

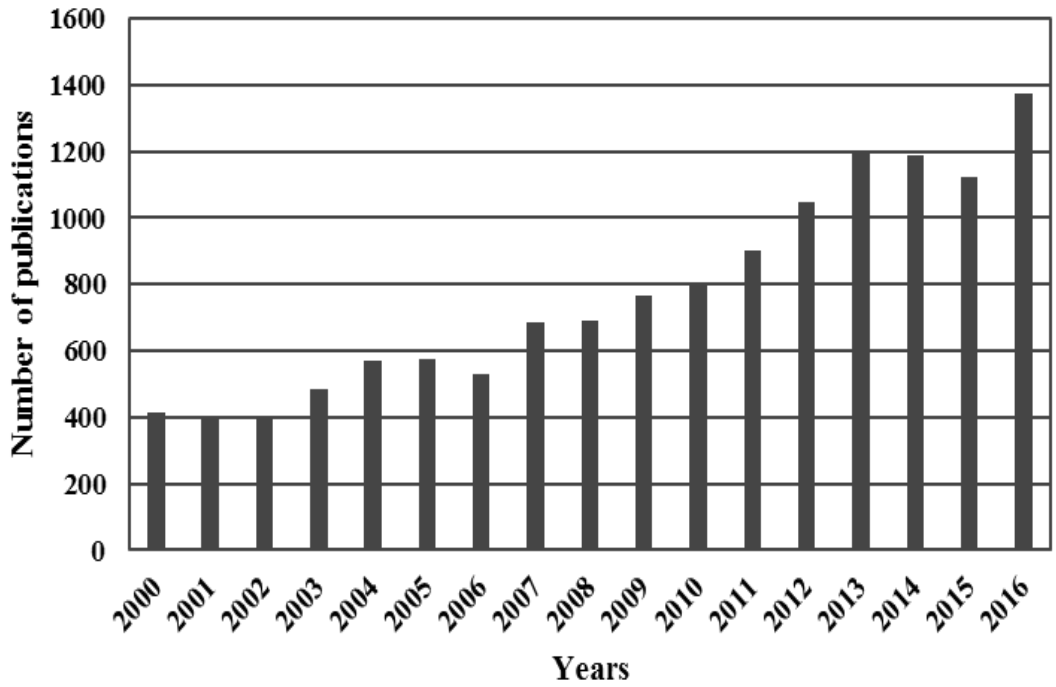


Figure 1. Histogramme des publications scientifiques par ans de la solubilité dans les fluides supercritique, d'après le site web sciencedirect (*ScienceDirect.coms.*, 2017)

Le principe des méthodes QSAR (*Quantitative Structure Activity Relationship*) consiste comme leurs noms l'indiquent à mettre en place une relation mathématique à l'aide de méthodes d'analyse de données, reliant des propriétés moléculaires microscopiques appelées descripteurs, à un effet expérimental (activité biologique, toxicité, affinité pour un récepteur), pour une série de composés chimiques similaires. Le point de départ de telles méthodes se construit sur la définition des descripteurs moléculaires empiriques ou théoriques. Ces dernières prennent en compte des informations sur la structure et les caractéristiques physico-chimiques des molécules, comme dans ce cas présent, les perturbateurs endocriniens.

Les descripteurs théoriques se déclinent eux-mêmes en sous-classes de descripteurs : les constitutionnels, basés sur la composition chimique de la molécule, les topologiques, obtenus à partir de la structure bidimensionnelle (table de connectivité des atomes de la molécule), les géométriques, évalués à partir de la structure tridimensionnelle et les quantiques, issus de la structure électronique de la molécule. Les détails concernant les descripteurs quantiques sont disponibles dans l'ouvrage de Karelson, (Karelson, 2000).

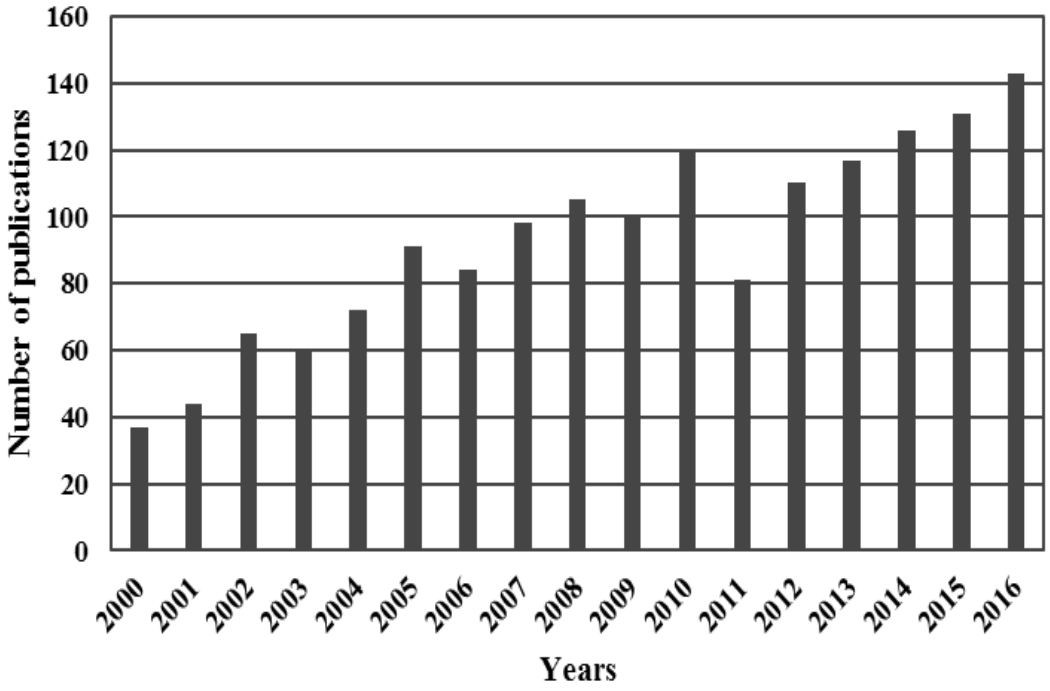


Figure 2. Histogramme des publications scientifiques par ans de l'utilisation de la méthode QSAR/QSPR, d'après le site sciencedirect (*ScienceDirect.com.*, 2017)

Enfin, le lien entre les descripteurs et la base de données est déterminé grâce à des outils d'analyse comme les régressions linéaires multiple (MLR), les régressions aux moindres carrés partiels (PLS), les arbres de décisions, les réseaux de neurones, et les algorithmes génétiques.

L'apprentissage connexionniste est une discipline scientifique qui recouvre plusieurs aspects d'études mathématiques, statistiques et algorithmiques. Les systèmes d'apprentissage connexionnistes (ou réseaux de neurones artificiels [RNA] sont des systèmes numériques permettant la modélisation de processus généraux par l'établissement de modèles fonctionnels. Ceux-ci sont identifiés à partir des entrées-sorties du processus par des algorithmes dits "d'apprentissage" qui s'apparentent à des techniques d'estimation statistiques. Nés en informatique dans le domaine de l'intelligence artificielle, ils ont connu depuis le début des années 80 un développement intensif dû au succès rencontré dans une très large gamme d'applications. Les réseaux de neurones artificiels offrent une panoplie de techniques adaptatives pour de nombreux problèmes génériques : la classification, le classement, la modélisation, la prévision. Parmi les domaines où ils ont été utilisés, on peut citer le traitement du signal, de l'image, de la parole, le contrôle, la prévision, l'optimisation combinatoire, le diagnostic, la fouille de données, ...etc. Il s'agit donc d'un outil général pouvant être utilisé dans des cadres différents avec une grande souplesse. Les modèles RNAs apparaissent comme une alternative ou plus souvent comme un complément aux techniques classiques. Les recherches dans le domaine sont extrêmement actives actuellement. Plusieurs laboratoires tant industriels ou universitaires participent à cette recherche en Europe, États-Unis et au Japon. Différentes voies ont déjà été explorées et les réseaux de neurones artificiels sont devenus un outil incorporé dans différents logiciels commercialisés pour les statistiques, la fouille de données, le traitement du signal et l'identification de systèmes.

L'objectif du présent mémoire est de développer deux méthodologies pour prédire la solubilité des solutés solides dans le dioxyde de carbone supercritique, la première c'est de développer un modèle basé sur la théorie des réseaux de neurones artificiels ou on a utilisé les propriétés thermodynamique et les descripteurs moléculaires comme des entrées et la deuxième c'est de développer un modèle basé sur la théorie de support vecteur machine.

Dans ce contexte notre travail est organisé en trois chapitres suivis d'une synthèse bibliographique aussi exhaustive que possible:

Dans le premier chapitre, après une introduction générale, nous rappelons les notations de la solubilité, les fluides supercritiques et les modèles de prédiction de la solubilité.

Dans le deuxième chapitre, nous parlons brièvement sur le réseau de neurones artificiel, les étapes de l'utilisation d'un réseau neuronale.

Dans le troisième chapitre, nous attaquons le but de ce travail qui est la modélisation, et nous avons divisé ce chapitre en 2 contributions. La première contribution vise la modélisation neuronale, et la deuxième contribution vise la modélisation par support vecteur machine pour prédire la solubilité des solutés solides dans les fluides supercritiques.

Pour chaque système, une comparaison, d'une part entre les valeurs des modèles introduits dans ce travail, de réseau de neurones Artificiel (RNA) et les données expérimentales est réalisée en termes d'erreurs.

Enfin une conclusion générale qui rassemble les résultats obtenus dans cette étude et des perspectives qui peuvent être envisagées pour le futur.

Chapitre I :
Modèles de prédiction de la solubilité des solutés solides dans les fluides supercritiques.

I.1 Introduction

La technologie des fluides supercritiques a attiré l'attention des scientifiques et des chercheurs. En 20 dernières années où ainsi, les applications de la technologie supercritique ont été principalement en extraction et chromatographie. Expérimental étendu et des investigations théoriques ont été visées vers une compréhension des propriétés des systèmes des fluides supercritiques, en particulier interactions intermoléculaires (soluté–solvant, solvant – soluté, et soluté – soluté) dans ces derniers. Cependant, le progrès significatif a été accompli dans l'utilisation de fluides et mélanges supercritiques comme médias de réaction pour des synthèses chimiques et préparations de polymère et en tant que systèmes dissolvants alternatifs pour le traitement de matériaux ainsi que pour de nombreux des substances à intérêt pharmaceutique.

L'industrie chimique a connu un développement étonnant dans ces dernières années notamment dans ces procédés de point de vue technologique et ça se voie dans la qualité d'équipements mis en œuvre ainsi que l'utilisation de l'outil informatique dans les différentes étapes de l'étude (simulation, modélisation, …).

L'industrie pharmaceutique ne sorte pas de cette règle et on peut la définir comme un ensemble des procédés qui visent à transformer les matières premières aux produits finis c'est-à-dire des médicaments qui doivent être de bonnes qualité, efficaces et compétents lorsqu'on les met sur marché, ceci signifié qu'il faut avoir un bon procédé de production c'est à dire maîtriser et optimiser tous les paramètres qu'ils peuvent être trouvés, y compris la solubilité des soluté solides dans les fluides plus précisément les fluides supercritiques qui possèdent beaucoup d'avantages qu'on a déjà détaillé dans le chapitre précédent, le présent chapitre va détailler les différentes méthodes de prédiction de ce paramètre**,** [Abdallah El Hadj, 2009].

I.2 Historique

L'existence d'un état supercritique fut découverte par *Cagniard* en 1822. Il constata la disparition de la frontière gaz -liquide de certaines substances chauffées en milieu fermé.

En 1879, le pouvoir solvant des fluides supercritiques fut décrit par *Hannay* et Hogarth. Ils observèrent la solubilité élevée de l'iodure de potassium dans l'éthanol supercritique (T_c = 243°C et P_c = 63 atm) et montrèrent que la réduction de la pression provoquait la précipitation du sel à partir du fluide homogène [Hannay et al, 1879].

Ils étudièrent ensuite le pouvoir solvant des fluides supercritiques avec d'autres sels (chlorure de cobalt (II), chlorure ferrique et le bromure de potassium). Une des applications les plus

connues au niveau industriel de l'extraction par fluide supercritique (EFSC) fut la décaféinisassions de grains de café en 1971.

En 1989, l'agence américaine de protection de l'environnement lança un programme de développement des méthodes EFSC en vue de diminuer de 95% la consommation de solvants organiques chlorés. La technologie utilisant les fluides supercritiques suscite actuellement beaucoup d'intérêt dans différents domaines : industriel, pharmaceutique et agroalimentaire.

I.3 Principe

Le diagramme d'un corps pur présente les différents états de la matière (solide, liquide et gazeux) en fonction de la pression et de la température (Figure I-1). Les domaines d'existence de ces trois états sont délimités par les courbes d'équilibre (solide-gazeux, solide-liquide et liquide-gazeux). Par définition, l'état supercritique est caractérisé par la disparition de l'interface séparant l'état liquide et l'état gazeux au-dessus d'une certaine température et d'une certaine pression. Si seulement l'un de ces deux paramètres satisfait à cette définition, on parle de l'état surcritique.

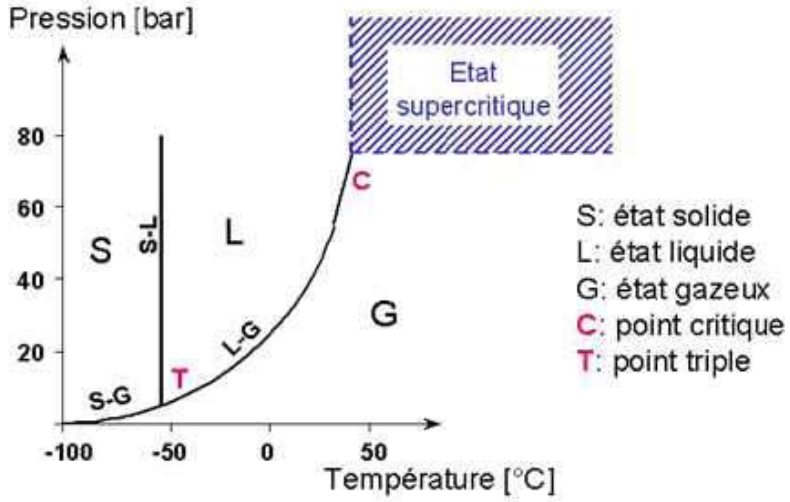


Figure I.1 : Représentation bidimensionnelle (Pression, Température) du diagramme de phases d'un corps pur, [Ziemons ,2006].

I.4 Etat supercritique

Les corps purs peuvent se trouver soit à l'état solide, soit à l'état liquide, soit à l'état gazeux ; dans le diagramme température- pression (figure I.1), les régions correspondant à ces trois états sont séparées par les courbes de changement d'état concourantes au point triple. Il se trouve que la courbe de vaporisation (liquide -gaz) présente un point d'arrêt dit point

critique (PC) correspondant à un couple de pression -température (Pc, Tc) assez facile à atteindre pour la plupart des composés organiques simples. Au-delà de ce point critique (PC), un seul état existe : le fluide est dit supercritique et présente des propriétés très particulières sur lesquelles vont reposer les procédés de mise en œuvre. En effet, les fluides supercritiques et les liquides « surcritiques » sont généralement utilisés dans un domaine de pression allant de 0,9 à 1,2 T_c, où ils ont une masse volumique voisine de celle des liquides, une viscosité à peine supérieure à celle des gaz, une diffusivité comprise entre celle des liquides et celle des gaz comme le montre le tableau I.1. De plus, on peut facilement faire varier très fortement leur masse volumique et leur pouvoir solvant.

Ainsi on peut citer, à titre d'exemple, la solubilité du naphtalène dans le CO_2, qui varie considérablement avec la masse volumique du fluide, c'est-à-dire, à température constante, avec sa pression comme présenté sur la figure I.2.

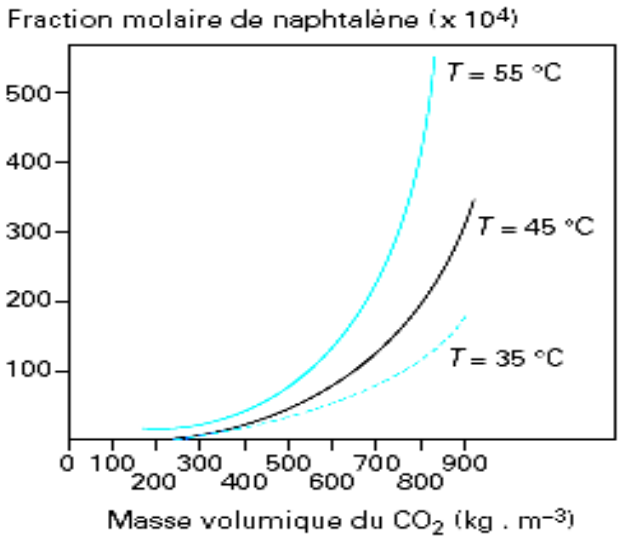


Figure I.2. Solubilité du naphtalène dans le dioxyde de carbone, [MCHugh et Krukonis, 1986].

I.4.1 Propriétés des fluides supercritiques

Les propriétés des fluides supercritiques sont intermédiaires entre celles des liquides et celles des gaz. Le fluide supercritique présente une viscosité faible et une diffusivité élevée comme les gaz et une masse volumique se rapprochant de celle d'un liquide. Un ordre de grandeur de la masse volumique, de la viscosité et de la diffusivité pour les gaz, les fluides supercritiques et les liquides est repris dans le tableau I-1.

Tableau I.1 : Propriétés physiques pour les gaz, le CO_2 supercritique et les liquides, [De Castro, 1994].

Propriétés	Gaz (30°C et 1 atm)	CO_2 SC (T_c et P_c)	Liquide (30°C et 1 atm)
Masse volumique, ρ (g/cm3)	10^{-3}	0.47	0,6 –1,6
Diffusivité, D (cm2/s)	10^{-1}	10^{-3}	$5 .10^{-6}$
Viscosité, η (g/cm s)	10^{-4}	10^{-4}	10^{-2}

Comme la montre le tableau I-2, les paramètres critiques (pression et température) varient fortement suivant le choix du fluide supercritique.

Tableau I.2 : Paramètres critiques des fluides à l'état supercritique [De Castro, 1994].

Constantes critiques			
Fluide	Pression (atm)	Température (°C)	MV (g/cm^3)
Dioxyde de carbone	73,8	31.1	0,468
Ethane	48,8	32.2	0.203
Ethylène	50,4	9.3	0.200
Propane	42,5	96.7	0.220
Propylène	46,2	91.9	0.230
Benzène	48,9	289	0.302
Toluène	41,1	318.6	0.290
Chlorotrifluorométhane	39,2	28.9	0.580
Trichlorofluorométhane	44,1	196.6	0.554
Protoxyde d'azote	71,0	36.5	0.457
Ammoniac	112.8	132.5	0.240
Eau	220.5	374.2	0.272

I.5 Le CO_2 supercritique

Le dioxyde de carbone présente de multiples avantages qui en font de loin le fluide le plus utilisé. Il est peu coûteux, peu toxique, ininflammable et chimiquement inerte. Sa température critique et pression critique facilement accessibles sont compatibles avec les composés thermiquement instables. En outre, son état gazeux à pression atmosphérique permet de le séparer spontanément du produit final et d'éviter ainsi la présence de solvants résiduels.

I.5.1 Le pouvoir solvant à géométrie variable

La solubilité d'une substance dans le CO_2 supercritique dépend de la masse volumique du fluide. Compte tenu de la compressibilité du SC-CO_2, la température et la pression de travail influencent la masse volumique et par conséquent la solubilité des composés. Ce

changement de masse volumique du fluide supercritique est à l'origine du pouvoir solvant "*à géométrie variable*" [Knutson et al, 1996].

Dans des conditions isothermes, il y a une relation proportionnelle entre la pression et la masse volumique de fluide supercritique c'est à dire une augmentation de la pression provoque une augmentation de la masse volumique (Figure I.3), et donc une augmentation du pouvoir solvant. L'augmentation de la solubilité d'un soluté ne sera effective qu'au-dessus d'une valeur précise de la pression qui a comme nom la "pression seuil".

Celle-ci est définie comme la pression à partir de laquelle le soluté commence à se solubiliser dans le fluide supercritique. Le tableau I-3 donne la pression seuil de certains stéroïdes diversement substitués [Stahl, 1983].

I.5.2 Polarité du CO_2 supercritique

La polarité du CO_2 supercritique est estimée par le paramètre de solubilité d'Hildebrand. Ce paramètre, calculé à partir de l'équation de Giddings, est fonction de la pression critique du fluide et du rapport des masses volumiques du fluide à l'état supercritique et à l'état liquide [Giddings, 1968].

$$\delta = \frac{1.25\sqrt{P_c}\cdot\rho_{fs}}{\rho_{liq}} \qquad \text{(I-1)}$$

Où *Pc* est la pression critique du fluide, ρ_{fs} la masse volumique du fluide à l'état supercritique et ρ_{liq} la masse volumique à l'état liquide. La polarité du CO_2 supercritique varie entre celle d'un alcane et celle de la pyridine (Figure I-3). Par conséquent, le pouvoir solvant est limité aux composés apolaires ou peu polaires. Il est toutefois possible de remédier à cet inconvénient en ajoutant un Co-solvant (modificateur ou entraîneur) en faibles quantités (5 à 15%) qui va conférer au mélange les propriétés idoines pour la molécule visée.

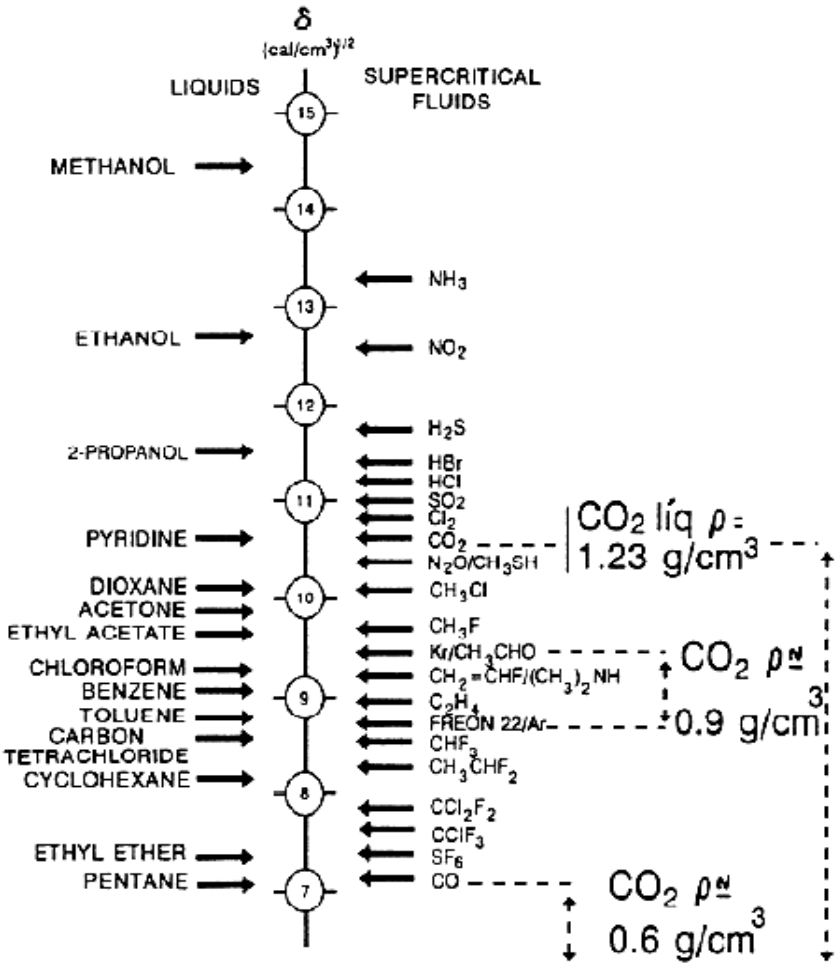


Figure I-3: Paramètres de solubilité de différents liquides et fluides supercritiques [De Castro, 1994].

I.6 Nature d'équilibre

L'équilibre est une condition statique où les propriétés macroscopiques ne subissent aucune variation avec le temps. Ceci implique un équilibre de tous les potentiels ou gradients, qui pourraient causer leur variation.

Cependant, à l'échelle microscopique, les conditions ne sont pas statiques, les molécules qui forment une phase à un instant donné ne sont plus les même à un autre instant, les molécules ayant des vitesses élevées et qui sont prés de l'interface arrivent à vaincre les tensions superficielles et passent à l'autre phase. Cependant le taux de passage dans les deux directions est le même.

I.7 Condition thermodynamique des équilibres

La plus grande partie du domaine de l'équilibre des phases est régie par deux relations théoriques fondamentales : La première est la règle de phase, qui définit l'allure générale du

diagramme des phases, la seconde est la formule de Clapeyron qui détermine la pente des courbes de ce diagramme.

I.7.1 Degré de liberté

Il est nécessaire, pour définir complètement un système, d'indiquer les valeurs numériques d'un certain nombre de variables, que l'on choisit par exemple parmi la pression, la température, l'énergie interne, l'entropie et les concentrations des différentes phases.
Parmi les variables, on distingue les variables extensives (qui dépendent de la masse) et les variables intensives dites variables d'état.

Un équilibre entre phase ne dépendant pas des quantités réelles des diverses phases en présence, on ne tiendra compte que des variables intensives, comme la température, la pression et la concentration pour le décrire.

On peut faire varier un certain nombre de ces variables de manière indépendante, mais les valeurs des autres sont déterminées par les valeurs choisies et par les conditions thermodynamiques d'équilibre. Alors, le degré de liberté ou variance du système est défini par le nombre de variables d'état qu'on peut faire varier de manière indépendante sans modifier le nombre de phase, [Hadj-Kali, 2004].

I.7.2 Règle de phase

La règle de phase proposée par Gibbs fournit une relation générale entre le nombre **v** de degré de liberté d'un système, le nombre de phase φ et le nombre de constituants **C**. Elle exprime qu'on a toujours pour un système non réactionnel :

$$v = C + 2 - \varphi \qquad \text{(I-2)}$$

Par exemple, pour un système binaire (C = 2) en équilibre solide- fluide (φ = 2), la règle de phase indique que la variance est égale à : v =2. Ceci signifie que l'on peut faire varier de manière indépendante deux variables d'état sans modifier le nombre de phases. Ces variables sont en général choisies parmi la température, la pression et la composition.

I.7.3 Condition d'équilibre solide – fluide

Pour chaque constituant *i* du mélange, la condition d'équilibre thermodynamique entre phase est donnée par l'égalité des potentiels chimiques des phases qui se réduit, par définition de la fugacité, à l'égalité des fugacités dans chaque phase pour chaque constituant. Dans le cas d'un équilibre solide-fluide, [Prausnitz et al, 1999].

$$\mu_i^s = \mu_i^f \tag{I-3}$$

Le problème fondamental est de relier ces fugacités aux compositions de chacune des phases. Dans ce qui suit, nous négligeons les effets dus aux forces de surface, gravité, champs électriques ou magnétiques et membranes semi-perméables. Alors, la fugacité d'un constituant dans un mélange ne dépend que de la température, de la pression et de la composition de ce mélange. On exprimera les compositions de solide dans la phase fluide (dans notre cas : un fluide supercritique) par la fraction molaire, y_i. Si la pression est faible et le mélange idéal (peu d'interactions entre les molécules présentes dans le mélange).

I.8 Fugacité

Cette propriété, qui signifie la tendance d'un système à occuper un plus grand volume, ou bien la tendance à la dispersion a été introduite par *LEWIS* pour tenir compte l'écart entre le comportement du gaz parfait et celui du gaz réel.

Le potentiel chimique d'un gaz idéal (ou parfait) est donné par :

$$\mu^{gp}(T,p) = \mu^{\circ}(T) + RT\ln\frac{P}{P^{\circ}} \tag{I-4}$$

Pour simplifier la notation la pression standard (P^0 = 1 bar) sera omise dorénavant (il faut travailler en bar, donc) et les dépendances avec T ou P aussi. Dans le cas d'un mélange de gaz parfaits, pour chaque constituant :

$$\mu_i^{gp} = \mu_i^0 + RT\ln y_i p \tag{I-5}$$

Où $y_i p = p_i$ est la pression partielle du constituant i.

Un fluide réel (gaz ou liquide) n'obéit pas à cette relation linéaire entre le potentiel chimique et le logarithme de la pression, mais il serait convenable de garder une forme équivalente pour les équations, en remplaçant la pression par une fugacité, soit une sorte de pression efficace. Pour une substance pure, la fugacité f est définie par :

$$\mu(T,p) \equiv \mu^{*}(T) + RT\ln\frac{f}{p^{\circ}} \tag{I-6}$$

Où μ^* est le potentiel chimique quand la fugacité vaut 1 bar.

Bien sûr, la fugacité est fonction de la température et de la pression, car elle prend en compte la non idéalité du fluide. Le coefficient de fugacité est le rapport entre la fugacité et la pression, donc une grandeur sans dimension :

$$\phi \equiv \frac{f}{p} \tag{I-7}$$

Pour un gaz idéal, $f = p$ et $\phi = 1$. Jusqu'ici, un fluide réel à l'état pur a été considéré, et la non idéalité introduite sous forme de fugacité concerne les écarts par rapport au gaz idéal sous effet de la pression et de la température.

Dans la situation plus générale d'un mélange de fluides réels, la composition affecte aussi les écarts à l'idéalité. La fugacité de chaque constituant i est définie par, [Ashour et al, 2000] :

$$\mu_i = \mu_i^* + RT \ln f_i \tag{I-8}$$

Comme les phases en équilibre sont obligatoirement à la même température, la condition d'équilibre de phases, équation (II-2), correspond à une égalité de fugacités :

$f_i^s = f_i^{SF}$ i= 1, NC, avec NC : nombre des composés

Ou f_i^s est la fugacité du composé *i* dans le solide(s) et f_i^{SF} est la fugacité de *(i)* dans la phase supercritique. Cette fugacité peut être calculée par la relation suivante :

$$f_i^{SF} = y_i \phi_i^{SF} p \tag{I-9}$$

Où ϕ_i^{SF} est l coefficient de fugacité de composé *(i)* dans la phase supercritique, et *P* est la pression du système.

Le coefficient de fugacité peut être calculé par la relation suivante :

$$RT \ln \phi_i = \int_V^\infty \left[\left(\frac{\partial P}{\partial n_i} \right)_{T,V,nj \neq n_i} - \frac{RT}{V} \right] dV - \ln Z \tag{I-10}$$

Où *R* est la constante des gaz, *V* est volume total du système, n_i et n_j sont les nombres de moles des composés *i* et *j*, et *Z* c'est facteur de compressibilité.

La fugacité du composé i comme une phase solide pure est :

$$f_i^s = P_i^{sub} \phi_i^{sub} \exp\left[\int_{P^{sub}}^{P} \left(\frac{v_i^s}{RT} \right) dP \right] \tag{I-11}$$

Où P^{sub} est la pression de sublimation de solide pure à la température de système, v_i^s est le volume molaire de solide pure, ϕ_i^{sub} est le coefficient de fugacité du solide pure.

Le mot fugacité est lié à la tendance du composant de *"fuir"* la phase en question. Ainsi dans l'exemple de l'eau liquide en équilibre avec sa vapeur, si la fugacité de l'eau liquide est égale à celle de la vapeur, les tendances de fuite sont identiques et le système se trouve en équilibre. Si la fugacité de la vapeur est supérieure à celle de l'eau liquide, alors la vapeur condense ; dans le cas opposé, l'eau s'évapore.

I.9 La solubilité des solutés dans les fluides

La solubilité du soluté solide n'est que la fraction molaire ou la fraction massique de ce solide dans la phase supercritique, et elle peut être calculée à partir de la relation thermodynamique suivant :

$$y_i = \left(\frac{P^{sub}}{P} \right) \tag{I-12}$$

Où y_i est la fraction molaire du solide dans la phase supercritique et c'est la solubilité idéale. La solubilité réelle est définie comme suit :

$$y_i = \left(\frac{P^{sub}}{P} \right) E \tag{I-13}$$

Avec E est le facteur de poiyting est donné par la relation suivante :

$$E = \frac{\phi_i^{sub} \exp\left(\frac{v_i^S}{RT} (P - P^{sub} \right)}{\phi_i^{FS}} \tag{I-14}$$

Ce facteur est introduit pour apporter une correction due à la compression du fluide, car la pression d'équilibre du mélange est en général différente de la pression de saturation qu'aurait le composant solide pur à la même température, [Prausnitz et al, 1986].

I.10 Les modèles de prédiction de la solubilité

La solubilité est une information très importante et cruciale pour le design des procédés pharmaceutiques, surtout ceux qui comportent des procédures de séparation (extraction, chromatographie, fractionnement...) et de traitement des solides (micronisation, micro encapsulation…). La modélisation des données expérimentales de la solubilité dans la phase supercritique peut aussi apporter une meilleure compréhension de phénomènes de dissolution

qui peut être utilisée dans la prédiction de solubilité à des pressions et températures intéressantes après la mesure d'un nombre minimum de données expérimentales qui peut accélérer le développement des procédés des fluides supercritiques.

Lors de la modélisation on devrait maintenir dans l'esprit que la solubilité des systèmes peut être catégorisée en trois groupe, le premier un seul type du soluté dans un fluide supercritique le deuxième un mélange de solutés dans un fluide supercritique et le dernier un soluté dans un mélanges de fluides supercritiques.

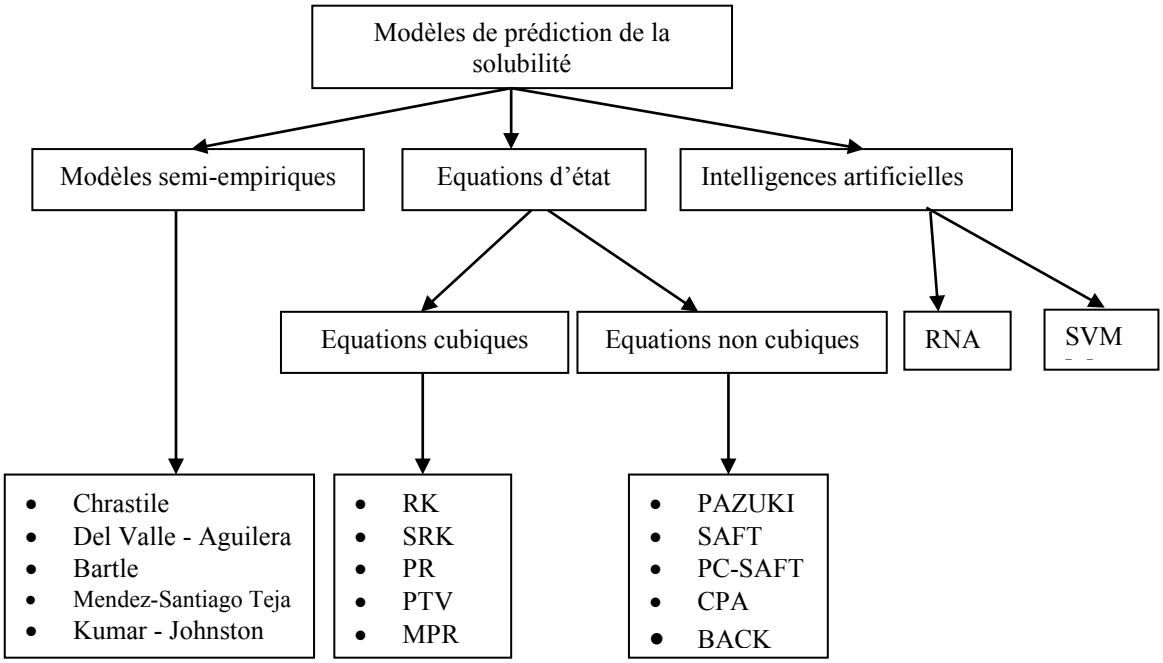


Figure I.4 : Classification des différents modèles de prédiction de la solubilité des solides dans les fluides supercritiques.

I.10.1 Les modèles semi empirique

La solubilité des solutés solides dans les phases supercritiques est une propriété thermo-physique très importante qui nécessite d'être modélisée comme un premier pas pour le développement des technologies supercritiques. La corrélation de solubilité peut être conventionnellement mise en œuvre en utilisant la densité de fluide supercritique, d'où beaucoup de modèles semi empiriques appelé autrement les modèles de densité existent, on cite à titre d'exemple :

I.10.1.1 Modèle de Chrastile

Chrastile (1982) a proposé un modèle pour la modélisation de solubilité des substances solides dans les phases supercritiques, ce dernier il relie la solubilité de soluté à la densité de fluide supercritique avec la supposition que l'association de soluté *(S)* avec N molécules de

fluide (C) donne un complexe solvaté (SCN) qui est en équilibre avec le gaz. L'expression du modèle peut être donnée, [Chrastile, 1982] :

$$S = \rho^{A_2} \exp(\frac{A_1}{T} + A_0) \quad \text{(I-15)}$$

Où, A_2 est le nombre d'association qui représente le nombre moyen des molécules de CO_2 dans le complexe solvate, le paramètre a dépend de température de vaporisation et les enthalpies de solvatation de soluté ($A_1 = \frac{\Delta H}{T}$), tandis que A_0 dépend de poids moléculaire de soluté (Ms) et ceux du fluide supercritique (Mscf), la transformation logarithmique de l'équation donne :

$$LnS = A_0 + \frac{A_1}{T} + A_2 Ln\rho \quad \text{(I-16)}$$

Où les paramètres A_0, A_1 et A_2 sont des constants qui peuvent être déterminés à partir des données expérimentales, [Chrastile, 1982].

I.10.1.2 Modèle de Del Valle et Aguilera

Del valle et Aguilera (1988) ont trouvé que la variation de *I*, défini comme suit $I = A_0 + \frac{A_1}{T}$, en fonction de $(1/T)$ est linéaire et la pente de cette droite est équivalente a A_1 de l'équation de Chrastile (II-16), et pour récompenser le changement de l'enthalpie en fonction de température ils ont proposé la modification de l'équation de Chrastile :

$$S = \rho^{A_2} \exp(\frac{A_3}{T^2} + \frac{A_1}{T} + A_0) \quad \text{(I-17)}$$

I.10.1.3 Modèle d'Adachi et Lu

Adachi et Lu (1983) ont modifié l'équation de Chrastile pour une bonne modélisation de la solubilité des triglycerides. Notons que les deux modèles de Chrastile et Del Vall ont supposé que le nombre d'association *k* est constant et dépend de la densité et de la température. Ces deux chercheurs ont donné ce nombre sous forme d'un polynôme de densité de second ordre :

$$k = e_0 + e_1\rho_1 + e_2\rho_1^2 \quad \text{(I-18)}$$

D'où, l'équation d'Adachi et Lu est de la forme :

$$C_2 = \rho_1^{e_0 + e_1\rho_1 + e_2\rho_1^2} \exp(\alpha + \frac{\beta}{T}) \quad \text{(I-19)}$$

I.10.1.4 Modèle de Bartle

Un autre modèle empirique proposé, [Bartle et al, 1991] qui décrive la relation entre la solubilité et la densité, notant qu'il y a une relation peut être existe entre le facteur de poiyting *(E)* et la densité de solvant :

$$\ln E = K_0 + K_1 \rho \tag{I-20}$$

Où E définie comme $E = y_2 P / P_2^S$ avec P_2^S est la pression de sublimation de soluté, les paramètres K_0, K_1 peuvent être obtenus par la régression des données de solubilité expérimentales. Cette équation permet la corrélation des solubilités des solides à chaque isotherme. Cependant, cette application est strictement limitée aux systèmes où la pression de sublimation est disponible. Pour cette raison Bartle et al ont réécris cette expression en introduisant une pression de référence P_{ref}= 1bar, et une densité de référence $\rho_{ref} = 700 kg/m^3$, connaissant que $P_2^S = A + \frac{B}{T}$ l'équation (I-20) devient :

$$\ln \frac{y_2 P}{P_{ref}} = C_0 + C_1(\rho - \rho_{ref}) + \frac{C_2}{T} \tag{I-21}$$

Les paramètres C_1, C_2 peuvent être obtenus par la régression des données expérimentales de y_2. A différentes valeurs de température et pression le paramètre C_0 peut être utilisé pour estimer l'enthalpie de vaporisation, [Bartle et al, 1991].

I.10.1.5 Modèle de Mendez-Santiago et Teja

C'est le modèle de densité le plus utilisé [Santiago et Teja, 1999], cette méthode nécessite la pression de sublimation de solide. Ils ont proposé un modèle basé sur la théorie des solutions diluées dans lequel le coefficient de poiyting est donné en fonction de densité de solvant. L'équation du modèle :

$$T \ln E = T \ln \frac{y_2 P}{P_2^{sub}} = A + B\rho_1 \tag{I-22}$$

Ou T est la température absolue, E facteur de poiyting, y_2 est la fraction molaire de soluté, P est la pression totale du système, P_2^{sub} est la pression de sublimation du soluté, ρ_1 est la densité de fluide supercritique, A et B sont des constantes de régression. Le facteur E représente le rapport entre la solubilité actuelle (réelle) et idéale, cette dernière est calculée selon la loi des gaz parfait (P/P^{sub}), [prausnitz et al. 1999]. Avantage de ce modèle c'est que les données expérimentales de solubilité pour différentes isothermes peuvent être représentés par

une seule droite ou la régression donne des constantes *A* et *B* indépendant de la température. Dans le plus part des cas la pression de sublimation n'est pas disponible, c'est pour cette raison que Mendez et Teja ont proposé une modification de leur expression pour l'utilisation dans le cas où P^{sub} est inconnue en remplaçant ce dernier par les constantes de l'équation d'Antoine :

$$T \ln y_2 P = A' + B' \rho_1 + C'^T \qquad \text{(I-23)}$$

Avec A', B' et C' sont des paramètres de régression, [Santiago et Teja., 1999].

I.10.1.6 Modèle de Kumar et Johnston

Kumar et Johnston ont observé qu'il y a une relation linéaire entre le logarithme de solubilité de soluté solide et la densité de solvant l'expression de modèle [Kumar et Johnston, 1988] est la suivante :

$$Lny_2 = b_0 + b_1 + \rho \frac{b_2}{T} \qquad \text{(I-24)}$$

Où les paramètres b_n (n=1,3) sont des paramètres obtenus par la corrélation des données expérimentales. Cependant, le paramètre b_2 relatif à l'enthalpie totale $\Delta H_{total} = -b_2 R$, [Kumar et Johnston, 1988].

I.10.2 Les équations d'état

De par leur origine historique et par la prédominance historique de la pétrochimie en Génie Chimique, les équations d'état jouent un rôle central dans la conception des unités de Génie des Procédés. Les articles foisonnent dans la littérature et les revues sont nombreuses. Sans intention d'exhaustivité, cette partie présente le développement d'équations d'état qui ont le plus marqué la discipline du Génie des Procédés.

L'utilisation d'une équation d'état pour reproduire les propriétés volumétriques et thermodynamiques des corps purs et des mélanges de corps purs fait l'objet de recherche par les thermodynamiciens depuis plus d'un siècle. En effet, depuis que Van der Waals en 1873 a proposé sa très célèbre équation d'état cubique, un nombre très important de publications a vu le jour. Avec le développement de l'informatique, l'utilisation d'expressions analytiques de plus en plus performantes permet la prédiction des propriétés thermodynamiques avec une grande précision permettant d'améliorer la conception des procédés de Génie Chimique tels que la distillation, l'absorption, …etc. Néanmoins, ce n'est que pour les gaz qu'on a mis ces relations sous une forme assez précise, non seulement en partant des données expérimentales

concernant les paramètres *P*, *V*, *T*, mais aussi grâce à des considérations théoriques sur la constitution des atomes et des molécules et des interactions qui restent faibles dans le cas des gaz, [Hadj-Kali, 2004].

En général, une équation d'état est développée initialement pour les substances pures et elle s'écrit sous la forme : $f(P, V, X, T) = 0$. Où X représente le vecteur composition (fraction molaire) ; et $X = 1$ pour un Corps pur.

Toute équation d'état doit vérifier le critère suivant aux coordonnées critiques (c'est-à-dire pour $T = T_c$, $P = P_c$ et $V = V_c$):

$$\left(\frac{\partial P}{\partial V}\right)_{T=Tc} = \left(\frac{\partial^2 P}{\partial V^2}\right)_{T=Tc} = 0 \qquad \text{(I-25)}$$

Ceci est dû au fait que l'isotherme critique possède une pente horizontale et un point d'inflexion au point critique.

I.10.2.1 Equation d'état du gaz parfait – Loi des états correspondants

La première équation d'état reliant les propriétés macroscopiques d'un système est la loi des gaz parfait découlant de la loi de Boyle-Mariotte, [Hadj-Kali, 2004] :

$$P\,v = RT \qquad \text{(I-26)}$$

Avec v représentant le volume molaire du gaz.

Nous avons aussi vu que la thermodynamique statistique permet d'établir cette loi en séparant les contributions cinétiques et potentielles de l'énergie du système et en faisant la simple hypothèse que l'énergie d'interaction entre les molécules est nulle.

Pour un gaz réel, on peut décrire commodément les écarts par rapport à l'état parfait en introduisant un facteur de compressibilité Z :

$$P\,v = ZRT \qquad \text{(I-27)}$$

Van der Waals a constaté que le facteur de compressibilité pour des fluides simples avec des molécules sphériques comme les gaz rares et le méthane ne dépend que des paramètres réduits *Tr* et *Pr* définis de la manière suivante :

$$T_r = \frac{T}{T_c} \qquad \text{Et} \qquad P_r = \frac{P}{P_c} \qquad \text{(I-28)}$$

Où T_c et P_c sont la température critique et la pression critique respectivement.

Alors, toutes les substances ayant les mêmes coordonnées réduites (*Tr, Pr*) auront forcement le même facteur de compressibilité Z : c'est la loi des états correspondants.

Cependant, pour des molécules non sphériques, on a constaté qu'il y a une déviation par rapport à la loi des états correspondants, d'où la nécessité d'introduire un troisième paramètre qui tient compte de la non sphéricité de la molécule. Plusieurs types de paramètres ont été proposés dont le facteur de compressibilité critique Z_c par Meissner et Seferian et le paramètre α_c proposé par Reidel, mais celui qui a connu le plus grand succès est le facteur acentrique de Pitzer ω défini par :

$$\omega = -\left(\log P_r^{sat}\right)_{T_C=0.7} = 1 \qquad \text{Avec} \quad P_r^{sat} = \frac{P^{sat}}{P_C} \tag{I-29}$$

Le facteur de compressibilité s'exprime alors par la relation suivante :

$$Z = Z^{(0)}\left(T_r, P_r\right) + \omega Z^{(1)}\left(T_r, P_r\right) \tag{I-30}$$

Le premier terme de cette égalité ($Z^{(0)}$) représente le facteur de compressibilité pour une molécule sphérique, alors que $Z^{(1)}$ exprime la déviation par rapport à la sphéricité idéale de la molécule. Pitzer et ses collaborateurs ont réparti dans un tableau les valeurs de $Z^{(0)}$ et $Z^{(1)}$ en fonction des coordonnées réduites T_r et P_r, alors qu'Edmister propose les mêmes valeurs sous forme de diagrammes, [Prausnitz, 1999]

II.10.2.2 Equations d'état cubique

L'équation des gaz parfaits n'est utilisable que dans un domaine de pression restreint (celui de basses pressions), pour cette raison cette équation a été modifiée sous la forme proposée par Hirn, [Hadj-Kali, 2004] :

$$\left(P + \pi\right)\left(v - b\right) = RT \tag{I-31}$$

Où b est appelé *covolume* représente la part du volume molaire qui n'est pas réellement disponible à cause de la présence d'autres molécules, et π est un terme correctif de pression dû à l'attraction mutuelle des molécules, appelé pression interne.

La première équation d'état à fournir une description quantitative des deux phases vapeur et liquide et une prédiction de la transition entre phases est celle de Van der Waals en 1873 qui a établi théoriquement que la pression interne est inversement proportionnelle au carré du volume molaire, d'où son équation d'état :

$$\left(P + \frac{a}{v^2}\right)\left(v - b\right) = RT \tag{I-32}$$

Ou bien : $$P = \frac{RT}{\left(v - b\right)} - \frac{a}{v^2} \tag{I-33}$$

Dans cette dernière équation, le premier terme associé au covolume b caractérise la répulsion et le second associé à la pression interne l'attraction. Si cette formule ne représente

pas parfaitement les isothermes d'un fluide, elle ne conduit pas à des absurdités physiques, d'où sa validité et son application pour des calculs approximatifs.

Pour la détermination des paramètres de cette équation (et ceux des autres équations d'état du même type), il existe deux alternatives :

1. la première consiste à choisir des paramètres qui permettent de reproduire des données expérimentales, généralement la pression de vapeur et la densité liquide ou vapeur. Avec l'équation de Van der Waals dont les coefficients ne dépendent pas de la température, une seule température suffit, pour les autres équations, par contre, où ces paramètres seront fonction de la température, il faudra régresser ces paramètres sur un intervalle de température.
2. la deuxième procédure, qui reste la plus recommandée, utilise la condition du point critique, donnée par l'équation (II-25), pour obtenir les valeurs de ces paramètres aux points critiques a_c et b_c ; et ensuite applique une correction qui dépendra de la température (c'est-à-dire $a(T)=a_c \times \alpha(T)$), égale à l'unité au point critique et ajusté pour permettre de mieux reproduire tout l'intervalle de température.

Hormis son caractère historique, l'équation de Van der Waals n'est pas quantitativement précise. Par exemple, elle prédit que le facteur de compressibilité critique :

$$Z_C = \frac{P_c v_c}{RT_c} \tag{I-34}$$

Z_C est égal à 3/8 = 0,375 pour tous les fluides ; alors que sa valeur pour différents hydrocarbures varie de 0,24 à 0,29, et cet intervalle est plus grand si d'autres fluides non hydrocarbures sont considérés. De même, les pressions de saturation prédites sont incorrectes. A cause de ces insuffisances, de nombreux chercheurs ont tenté de modifier la forme de l'équation de Van der Waals afin d'améliorer sa précision.

La deuxième alternative, [Hadj-Kali, 2004] est la plus recommandée et utilisée pour la prédiction de solubilité des solutés solides dans les fluides supercritiques.

Après cette petite revue sur les origines des équations d'état et les méthodes d'utilisation on passe maintenant à la description des étapes de calcul de la solubilité des solides à l'aide d'une équation d'état cubique qui a pour but de prédire le coefficient de fugacité du solide dans l'autre phase (supercritique) nécessaire dans ce processus.

A. L'équation d'état de Redlich et Kwong (RK)

Dans le but de mieux représenter les propriétés volumétriques des fluides, Redlich et Kwong ont proposé une modification empirique du terme d'attraction de Van der Waals en

introduisant également une variation du paramètre ***(a)*** avec la température. Ceci permet d'améliorer le calcul de propriétés de corps purs, [Redlich et Kwong, 1949].

$$P = \frac{RT}{(v-b)} - \frac{\alpha \times a}{v(v+b)} \tag{I-35}$$

(Avec $\alpha = T^{-1/2}$ pour l'équation de Redlich-Kwong)

Exprimée en fonction du facteur acentrique, l'équation de Redlich-Kwong a une forme cubique en posant :

$A = \frac{aP}{(RT)^2}$ Et $B = \frac{bP}{RT}$:

$$Z^3 - Z^2 + \left(A - B - B^2\right)Z - AB = 0 \tag{I-36}$$

La représentation de cette équation sur un diagramme de corps pur illustre bien les trois racines de cette équation sur l'isotherme d'équilibre liquide–vapeur : deux racines correspondent aux points de bulle et de rosée, tandis que la troisième située dans la région liquide – vapeur n'a pas de signification physique, [Hadj-Kali, 2004].

Comme le montre également la Figure 5. L'équation de Redlich-Kwong ne s'applique bien qu'à la phase vapeur.

B. L'équation d'état de Soave Redlich et Kwong (SRK)

Il a été remarqué que les équations d'état cubique à deux paramètres ne peuvent représenter avec une précision satisfaisante à la fois le comportement volumétrique et les tensions de vapeur, et qu'il convient de choisir une priorité. C'est ce qu'a compris Soave en 1972 qui a appliqué l'équation d'état de Redlich-Kwong au calcul des tensions de vapeur des hydrocarbures en modifiant l'expression du paramètre α devenu fonction de la température et du facteur acentrique de Pitzer ω, [Soave, 1972] :

$$P = \frac{RT}{(v-b)} - \frac{a(T)}{v(v+b)} \tag{I-37}$$

Où les constants des corps pures *a(T)* et *b* pour cette équation, [Soave, 1972]:

$$a(T) = 0.42747 \frac{R^2 T_c^2}{P_C} \alpha(T_r, w) \tag{I-38}$$

$$b = 0.08664 \frac{RT_c}{P_c} \tag{I-39}$$

Avec :

$$\alpha(Tr, w) = \left[1 + \left(0.480 + 1.574\omega - 0.176\omega^2\right) \times \left(1 - T_r^{0.5}\right)\right]^2 \tag{I-40}$$

Et sous forme cubique :

$$Z^3 - Z^2 + \left(A - B - B^2 \right) Z - AB = 0 \tag{I-41}$$

L'expression du coefficient de fugacité est donnée en logarithme :

$$\ln \phi_i = \frac{b_i}{b_m}(Z-1) - \ln(Z-B) - \frac{A}{B}\left(2\frac{\sqrt{a_i}}{\sqrt{a_m}} - \frac{b_i}{b_m}\right)\ln\left(1+\frac{B}{Z}\right) \tag{I-42}$$

L'équation de Soave-Redlich-Kwong conserve la même valeur du facteur de compressibilité critique (Zc = 0,333) qui reste donc trop élevée par rapport aux valeurs expérimentales constatées.

C. L'équation d'état de Peng et Robinson (PR)

En 1976, Peng et Robinson apportent en plus de la modification du paramètre α, une nouvelle expression pour le terme d'attraction. Le progrès est sensible pour la tension de vapeur et surtout pour le calcul des masses volumiques en phase liquide. En outre, le facteur de compressibilité critique est mieux estimé (Zc = 0,307), [Peng et Robinson, 1976].

$$P = \frac{RT}{(v-b)} - \frac{\alpha \times a}{v(v+b) + b(v-b)} \tag{I-43}$$

Avec : $\alpha = \left[1 + \left(0.37464 + 1.5422\omega - 0.26992\omega^2 \right) \times \left(1 - T_r^{0.5} \right) \right]^2$ (I-44)

Et en termes de facteur de compressibilité :

$$Z^3 - (1-B)Z^2 + \left(A - 2B - 3B^2 \right) Z - \left(AB - B^2 - B^3 \right) = 0 \tag{I-45}$$

Où les constants des corps pures a et b pour cette équation :

$$a = 0.45724 \frac{R^2 T_c^2}{P_c} \alpha(T_r, w) \tag{I-46}$$

$$b = 0.07780 \frac{RT_c}{P_c} \tag{I-47}$$

Pour l'équation de PR, l'expression de coefficient de fugacité est donnée par :

$$\ln \phi_i = \frac{b_i}{b_m}(Z-1) - \ln(Z-B) - \frac{A}{2\sqrt{2}B}\left(\frac{2\sum_{j=1}^{n} y_j a_{ij}}{a_m} - \frac{b_i}{b_m} \right) \ln\left(\frac{Z + (1+\sqrt{2})B}{Z + (1-\sqrt{2})B} \right) \tag{I-48}$$

Les équations de Peng-Robinson (PR) et de Soave-Redlich-Kwong (SRK) sont largement utilisées dans l'industrie, particulièrement pour le raffinage et la simulation des réservoirs. Leurs avantages résident dans le fait qu'elles suivent un processus de calcul successive pour

le calcul de coefficient de fugacité mais pour ce faire elles nécessitent les coordonnées critiques et le facteur acentrique qui sont généralement indisponible pour les nouvelles substances, un temps de simulation relativement court et surtout qu'elles conduisent à une bonne estimation [Valderrama, 2003].

D. L'équation d'état de Patel-Teja-Valderrama (PTV)

Abbott (1979), dans une étude comparative des équations d'état cubique, a conclu que la forme la plus adéquate doit faire intervenir des paramètres supplémentaires. De nouvelles équations ont été développées, toutes en convergeant vers les équations SRK ou PR lorsque des valeurs particulières sont attribuées à ces paramètres additionnels.

L'une de ces équations qui contiennent des paramètres autres que les deux paramètres *a* et *b* est l'équation de Patel-Teja-Valderrama, [Valderrama, 1990] et son expression est donnée ci-dessous : $P = \frac{RT}{(v-b)} - \frac{a(T)}{v(v+b)+c(v-b)}$ (I-49)

Où les constantes des corps purs *a*, *b* et *c* pour cette équation :

$$a = \Omega_a(Z_c)\frac{R^2T_c^2}{P_c}\alpha(T_r, w, Z_c) \quad \text{(I-50)}$$

Avec : $\Omega_a(Z_c) = 0.66121 - 0.761057.Z_c$ (I-51)

$$b = \Omega_b(Z_c)\frac{RT_c}{P_c} \quad \text{(I-52)}$$

Où : $\Omega_b(Z_c) = 0.02207 + .0.20868.Z_c$ (I-53)

$$c = \Omega_c(Z_c)\frac{RT_c}{P_c} \quad \text{(I-54)}$$

Avec : $\Omega_b(Z_c) = 0.57765 - 1.87080.Z_c$ (I-55)

L'expression de la fonction dépendante de la température $\alpha(T_r, w, Z_c)$ pour cette équation, [Valderrama, 1990]:

$$\alpha(T_r, w, Z_c) = \left[1 + F \times \left(1 - T_r^{0.5}\right)\right]^2 \quad \text{(I-56)}$$

Et $F = \left(0.46283 + 3.58230\omega Z_c - 8.19417(\omega Z_c)^2\right)$ (I-57)

E. L'équation de Peng et Robinson modifiée (MPR)

La fonction alpha *(α)* de l'équation d'état de Peng et Robinson a été modifiée [Gasem et al, 2001] pour donner l'expression suivante :

$$\alpha(Tr) = \exp\left[(2 + 0.836Tr) \times \left(1 - T_r^m\right)\right]^2 \quad \text{(I-58)}$$

Et $m = \left(0.134 + 0.508\omega - 0.0467\omega^2\right)$ (I-59)

Tous les autres paramètres utilisés sont ceux de l'équation PR (*a*, *b*…).

Les équations citées sont les équations d'état cubiques les plus utilisées dans la prédiction de la solubilité des solutés solides dans les phases supercritiques et d'autres paramètres, il existe d'autres équations mais elles ne sont pas assez robustes pour donner des résultats comparables a celles données par les équations suscitées.

- **Les modifications de $\alpha(Tr)$**

Les axes de recherche**s** sur les équations d'état cubiques après les contributions de Peng et Robinson et celles de Soave suivis trois voies principales, [Valderrama, 2003] :

(i) Modification de fonction $\alpha(Tr)$ dans les équation**s** PR et SRK.

(ii) Modification du volume dépend de terme attractif.

(iii) L'utilisation de troisième paramètre relatif au corps purs *(c)*.

Ces modifications de la fonction dépendante $\alpha(Tr)$ dans le terme attractif des deux équations (SRK et PR) ont été principalement proposées pour améliorer les capacités de prédiction des différentes propriétés des corps pures et des mélanges. Le tableau II.1 montre certaines expressions proposées dans la littérature.

Tableau I.3 : Les modèles sélectionnés pour la fonction de terme attractif $\alpha(Tr)$ dans l'équation d'état.

Expression de $\alpha(Tr)$	Référence
$1/\sqrt{Tr}$	Redlich et Kwong, 1949
$1+m(Tr-1)^2(1-\sqrt{Tr})^2$	Usdin et McAuliffe, 1976
$1+m(1-Tr)(m+n/Tr)^2$	Soave, 1979
$\exp[C(1-Tr^m)]$	Heyen,1983
$\left[1+m(1-\sqrt{Tr})-p(1-Tr)(0.7-Tr)\right]^2$	Raimondi, 1980
$\exp[C(1-Tr^d)]$	Ishikawa et al, 1980
$\left[1+m(1-\sqrt{Tr})-n\sqrt{1-Tr/0.7}\right]^2$	Boston and Mathias, 1980
$m1 = m_2/Tr + m_3/Tr^2$	Mathias, 1983
$\left[1+m(1-\sqrt{Tr})-m_2(1-Tr)^2+m_3(1-\sqrt{Tr})^3\right]$	Mathias et Copman, 1983
$[1+m(1-\sqrt{Tr})]^2/Tr$	Bazua, 1983
$m_1+m_2/Tr+m_3/T_r^2+m_4/T_r^3$	Adachi et Lu, 1984
$1+m_1(Tr-1)+m_2(\sqrt{Tr}-1)$	Gibbons Laughton, 1984
$[1+m(1-T_r^n)]^2$	Kabadl et Dannaer, 1985
$\left[1+m(1-\sqrt{Tr})-n(1-Tr)(0.7-Tr)\right]^2$	Stryjek et Vera, 1986
$\left[1+m(1-\sqrt{Tr})\right]^2+n(Tr-0.6)^2$	Adachi et Sugie, 1986
$\left[1+m(1-\sqrt{Tr})-p(1-Tr)(q-Tr)\right]^2$	Du et Guo, 1988
$\left\{\left[1+n\exp(-k)\right)\right]/[1+\exp(-kT_r^{2m})]\right\}T_r^{2m-1}\exp[(1-T_r^{2m})^k]$	Twu, 1988
$\exp[m(1-Tc)+n(1-\sqrt{Tr})]^2$	Melhem et al, 1989

I.10.2.3 Equations d'état non cubique (basée sur la théorie de perturbation)

Les équations de type SAFT (Statistical Associated Fluid Théory) ou, basées sur les théories de perturbation connaissent un développement important depuis 15 ans [Müller et Gubbins, 2001]. Elles évaluent l'énergie libre de *Helmoltz* comme la somme de plusieurs contributions formalisant une représentation des molécules et de leur interaction : contribution de monomère, contribution de liaison de ces monomères pour former une molécule ; contribution d'association intermoléculaire entre sites d'interaction. Ce formalisme est proche de celui imaginé en mécanique moléculaire et l'un des atouts de l'approche (SAFT) est le fait que les prédictions d'une telle équation peuvent être confrontées à des simulations moléculaires utilisant le même formalisme, [Hadj-Kali, 2004]. Cette gamme contient des équations d'état combinées avec la théorie associatives telle que l'équation d'état de CPA (Cubic Plus Association) et celle de Pasuki et des équations d'état complexes et non cubiques à fondement théorique basé sur thermodynamique statistique telles que les équations d'état SAFT (Statistical Association Fluid Thoery), PC-SAFT (Perturbed Chain Statistical Association Fluid Thoery), VR-SAFT.

Dans cette partie on va donner une revue sur ces équations, ses origines et ses bases théoriques.

L'équation d'état SAFT [Chapmen et al ,1989] est considérée comme une méthode qui combine entre la théorie de perturbation thermodynamique de Wertheim, [Wertheim, 1984] pour l'association des fluides avec l'idée de formuler une équation d'état. Dés le début, SAFT n'est pas une équation d'état rigide mais plutôt une méthode qui permet l'incorporation l'effet de l'association dans une théorie donnée. L'équation d'état d'un fluide associatif est une perturbation [Müller et Gubbins, 2001] donnée en termes de l'énergie molaire de Helmholtz résiduelle définie comme la différence entre l'énergie molaire de Helmholtz totale et celle d'un gaz idéal à température et densité constantes :

$$a^r(T,\rho) = a(T,\rho) - a^{gazidéal}(T,\rho) \qquad \text{(I-60)}$$

SAFT suppose qu'il y a trois types de contributions au potentiel intermoléculaire :

La contribution des segments individuels, la contribution des chaînes qui peuvent se former à partir des segments, la contribution due à la possibilité que certains segments forment des sites d'associations avec d'autres molécules. La figure II.1 montre les étapes de formations supposées par SAFT.

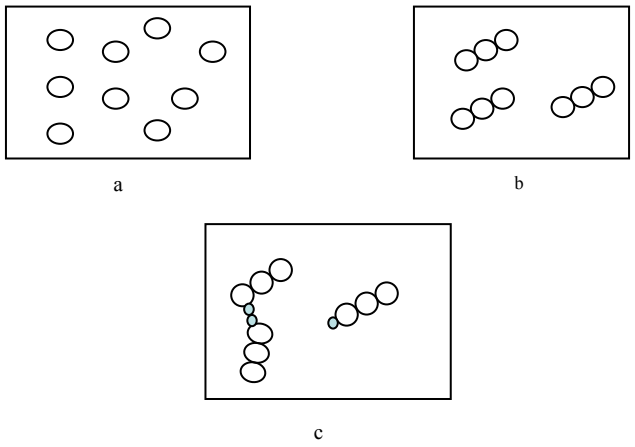


Figure I.5 : Les étapes de formation de modèle moléculaires à partir de théorie associative SAFT, [Müller et Gubbins, 2001].

L'énergie molaire de Helmholtz résiduelle est définie comme la somme de ces contributions :

$$a^r = a^{seg} + a^{chain} + a^{assoc} \qquad \text{(I-61)}$$

A. Equation d'état de PAZUKI

Cette nouvelle équation basée sur la théorie de perturbation [Pasuki et al, 2006] possède deux termes le terme répulsif et attractif selon la thermodynamique statistique [MacQuarrie, 1976], elle peut être écrite en termes de facteur de compressibilité :

$$Z = Z^{hs} + Z^{pert} \tag{I-62}$$

La forme cubique de cette équation est la suivante :

$$Z^3 - Z^2 + (A - B^2 - 2B)Z - (AB + B^2) = 0 \tag{I-63}$$

Avec les fonctions *alpha* et *bêta* pour le calcul des constantes des mélanges :

$$\alpha(Tr, w) = [1 + m_1(1 - Tr) + m_2(1 - Tr)^2 + m_3(1 - Tr)^3]^2 \tag{I-64}$$

$$\beta(Tr, w) = [1 + n(1 - Tr)] \tag{I-65}$$

B. Equation d'état PC-SAFT

L'équation d'état PC-SAFT [Sadowski et Gross, 2000] a été développée à partir de la théorie de perturbation comme celles citées auparavant où elle considère les molécules comme des chaînes linéaires composées des segments sphérique dures (hard chain) qui peuvent générer des sites d'associations, d'où le principe est le même que l'équation d'état SAFT. L'expression générale de cette équation est donnée par la relation suivante :

$$\frac{A}{NkT} = \frac{A^{hc}}{NkT} + \frac{A^{pert}}{NkT} \tag{I-66}$$

Où A est l'énergie libre de Helmholtz, N est le nombre total des molécules, k est la constante de Boltzmann, T est la température absolue. Dans cette terminologie le premier terme (hc) est relatif aux chaînes des segments durs dus aux interactions répulsives et la partie de perturbation signifie la contribution due aux interactions attractives des molécules. L'expression de cette équation peut être écrite en termes de facteur de compressibilité :

$$Z = Z^{hc} + Z^{pert} \tag{I-67}$$

Par exemple dans un mélange de deux substances c'est-à-dire selon PC-SAFT un contact de deux chaînes linéaires composées des sphères où les interactions intermoléculaires sont régit par trois types d'interactions : les interactions dispersives, celles attractives entre les chaînes et des forces d'associations entre les sites sur les deux chaînes.

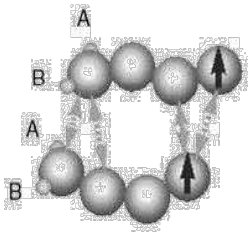


Figure I.6 : Le modèle moléculaire dans un mélange de deux substances selon l'équation d'état PC-SAFT. [Tumakak et al, 2005].

I.10.3 Les réseaux de neurones artificiels

Les réseaux de neurones sont des systèmes de traitement de l'information dans la structure s'inspire du celle du système nerveux. Leurs deux grands domaines d'application sont d'une part la mobilisation biologique, et d'autre part, la réalisation de machines destinées à effectuer des tâches auxquels les ordinateurs et les outils traditionnels semblent mois bien adapté que êtres vivants, telles que des tâches perceptives et motrices.

Les applications de réseau de neurones formels à la connaissance de formes et à la classification ont pris alors une place importante au sein des réalisations industrielles.

Dans le domaine du traitement du signal et de l'automatique et l'industrie chimique (thermodynamique), les années 90 ont été marquées par des avances significatives tant du point de vue fondamental que du point de vue des applications industrielles, signe incontestable de la maturation de ces outils, [Touzet, 1992]. Vue ainsi les capacités de ces réseaux à prédire la non linéarité ils peuvent représenter un outil intéressant pour prédire la solubilité des solutés solides dans les phases supercritiques.

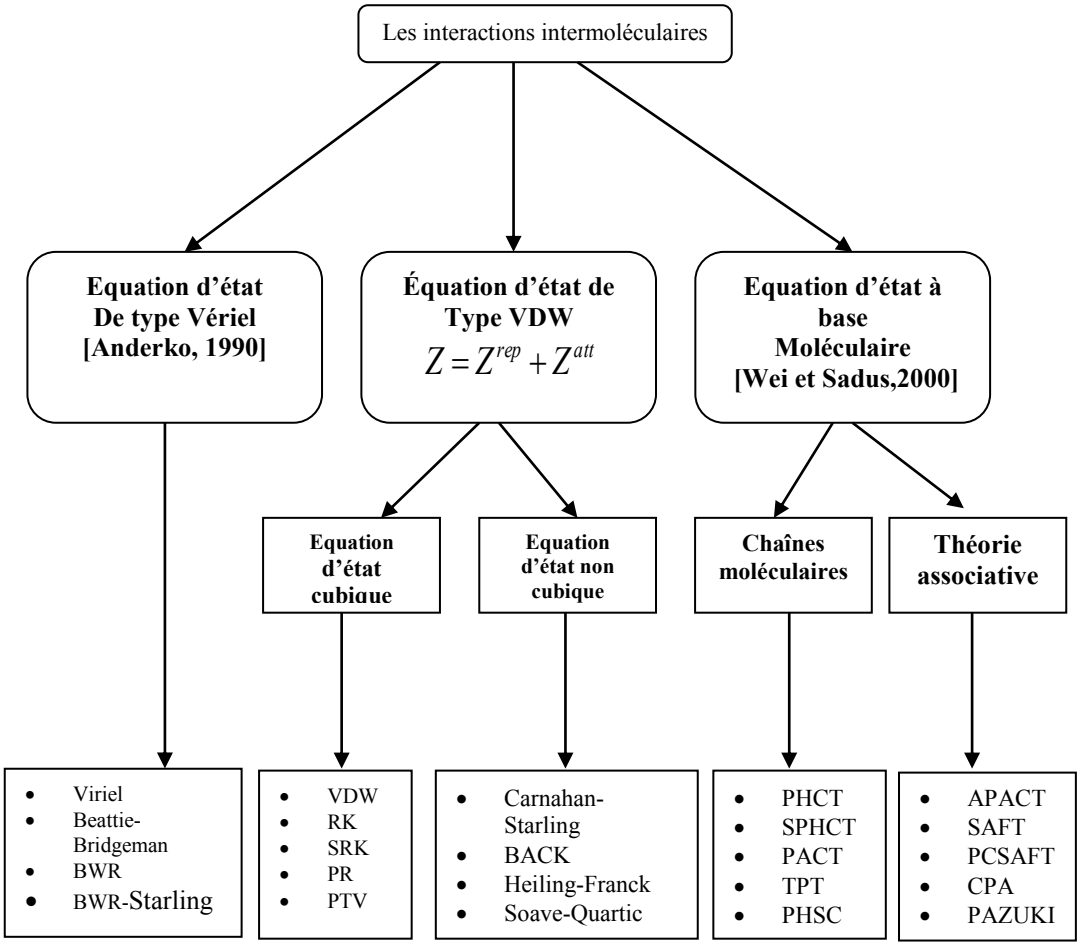


Figure I.7 : Classification des différents types des équations d'état selon les théories développées à partir des interactions entre les molécules, [Valderrama, 2003].

I.11 Règles de mélange

Comme nous l'avons signalé au début de ce chapitre, les équations d'état ont été développées initialement pour des fluides purs. Puis, leurs utilisations sont prolongées pour les mélanges. Cette extension nécessite l'introduction de règles de mélange permettant le calcul des paramètres de mélange en fonction de ceux des corps purs. Cette introduction fait intervenir à travers ces règles la régression ou l'ajustement des paramètres d'interactions binaires (dans le cas d'un mélange binaire) à partir des données expérimentales dans le but d'améliorer la prédiction obtenue par ces équations d'état. Dans ce qui suit on va détailler les différentes règles de mélanges les plus utilisées dans les processus de calcul.

I.12 Analyse numérique et mathématique

Les modèles qu'ils existent, pour estimer les différentes propriétés de corps purs et de mélanges, sont très importants pour les analyses de contrôle, de design et de synthèse de procédés chimiques et ceux de production, dans d'autres cas ils sont utilisés dans un rôle de service. Ce qui concerne l'estimation de la solubilité, le tableau ci-dessous donne le nombre d'équations, de variables et le processus de calcul par les différentes équations d'état, pour avoir une idée sur les difficultés rencontrées lors de prédiction de propriétés thermodynamiques d'un modèle à l'autre, où *NC* est le nombre de composés qu'ils forment le mélange.

Tableau I.4 : Le nombre des équations, des variables et le processus de calcul par les différents modèles d'estimation, [Gani et al, 2006]

Equation d'état	Nombre d'équations	Processus de calcul	Nombre de variables	Cas binaire NC=2	
				NE	NV
PR	5+8NC+2NC*NC	Successif	10+12NC+3NC*NC	29	46
SRK	5+8NC+2NC*NC	Successif	10+12NC+3NC*NC	29	46
Pazuki	5+11NC+2NC*NC	Successif	10+16NC+4NC*NC	35	58
SAFT	35+15NC+5NC*NC+2NCA*NS+ 4NCA*NCA+ NCA*NCA*NC+ NCA*NC*NS+NC^3	Successif et itératif	4NCA*NCA+6NC*NC+19NC+37+ 19NCNCA*NCA*NC+ NCA*NC*NS+NC^3	133	279
PC-SAFT	NC*(3NC+14) +Nz*(NC+1)	Successif et itératif	22+ NC*(3NC+18) + Nz*(NC+1) + J_a(NC+4) + J_b(NC+4) + NC*NC	180	170
CPA	11+9NC+NCA+3NCNA+ 2NC*NC+NS*NCA+3NCA*NC	Successif et itératif	17+6NCA+3NCNA+14NC+3NC*NC +NCA+3NCA*NCA	38	64

I.13 Conclusion

La connaissance de la solubilité des solutés solides dans les phases supercritiques est importante pour le design des procédés de séparation. De plus la disponibilité des résultats expérimentaux de solubilité des nouveaux composés intéressants de point de vue pharmaceutique nous offre l'opportunité de modéliser et voir les capacités des modèles existants à prédire cette propriété dans des conditions opératoires bien déterminées.

Les modèles de détermination de la solubilité allant de nature semi empirique jusqu'aux modèles purement théoriques notamment les équations d'état cubiques, non cubiques et celles basées sur la théorie de perturbation qui sont malgré la complexité de processus de calcul et les déviations qui peuvent être enregistrées restent les meilleurs modèles de fait qu'elles sont construit sur une base théorique où la majorité des paramètres à optimiser ont un sens physique. Ce qui concerne les mélanges l'utilisation de ces modèles est assurée par les règles des mélanges.

La tache sera plus compliquée pour les nouvelles substances récemment publiées dont les propriétés nécessaires pour la prédiction sont inconnues, et c'est l'objectif du troisième chapitre où on va estimer les solubilités de ces composés dans le dioxyde de carbone supercritique par les modèles des équations d'état en faisant une comparaison en terme d'erreur relative moyenne avec les données expérimentales issue de la littérature**,** [Abdallah El Hadj, 2009].

CHAPITRE 2
LES RESEAUX DE NEURONES

II.1 Introduction

Les réseaux de neurones connaissent depuis quelques années un succès croissant dans divers domaines des sciences de l'ingénieur, celui du génie des procédés n'échappe pas à cette règle. Dans ce chapitre, nous nous intéressons à l'application des réseaux de neurones aux procédés chimiques. Dans un premier temps, nous rappellerons les définitions et notations de base relatives aux réseaux de neurones. Nous poursuivrons en exposant les types et la méthodologie d'apprentissage. Nous montrerons en particulier que, les réseaux de neurones ne sont pas nécessairement des boîtes noires : en effet, il est parfaitement possible, et même vivement recommandé, d'introduire dans le réseau de neurones, dès sa conception, toutes les connaissances mathématiques disponibles concernant le processus à modéliser ou à commander. Nous présenterons aussi, d'une manière générale, les étapes de conception d'un réseau de neurones : le choix des entrées et sorties, l'élaboration de la base de données, de la structure du réseau etc.…. Nous terminerons ce chapitre par la présentation de diverses applications qui sont destinées à montrer que dans le domaine des procédés chimiques, les réseaux de neurones sont susceptibles d'apporter des solutions efficaces et élégantes notamment dans la modélisation et simulation des réacteurs et réactions chimiques.

II.2 Historique sur les réseaux de neurones

Les recherches menées dans le domaine du connexionnisme ont démarré avec la présentation en 1943 par MCCulloch. W et Pitts. W d'un modèle simplifié de neurone biologique communément appelé neurone formel. Ils montrèrent également théoriquement que des réseaux de neurones formels simples peuvent réaliser des fonctions logiques, arithmétiques et symboliques complexes. En 1949, Hebb. D initie, dans son ouvrage "The Organization of Behavior", la notion d'apprentissage. Deux neurones entrant en activité simultanément vont être associés (c'est-à- dire que leurs contacts synaptiques vont être renforcés). On parle de loi de Hebb et d'associationnisme. En 1958, F. Rosenblatt développe le modèle du Perceptron. C'est un réseau de neurones inspiré du système visuel. Il possède deux couches de neurones : une couche de perception (sert à recueillir les entrées) et une couche de décision. C'est le premier modèle pour lequel un processus d'apprentissage a pu être défini. S'inspirant du perceptron, Widrow et Hoff, développent, dans la même période, le modèle de l'Adaline (Adaptive Linear Elément). Ce dernier sera, par la suite, le modèle de base des réseaux de neurones multicouches. En 1969, Les recherches sur les réseaux de neurones ont été pratiquement abandonnées lorsque

Minsky .M et Papert .S ont publié leur livre « Perceptrons » (1969) et démontré les limites théoriques du perceptron, en particulier, l'impossibilité de traiter les problèmes non linéaires par ce modèle. En 1982, Hopfield développe un modèle qui utilise des réseaux totalement connectés basés sur la règle de Hebb pour définir les notions d'attracteurs et de mémoire associative. En 1984 c'est la découverte des cartes de Kohonen avec un algorithme non supervisé basé sur l'auto-organisation et suivi une année plus tard par la machine de Boltzman (1985). Une révolution survient alors dans le domaine des réseaux de neurones artificiels : une nouvelle génération de réseaux de neurones, capables de traiter avec succès des phénomènes non-linéaires : le perceptron multicouche ne possède pas les défauts mis en évidence par Minsky. Proposé pour la première fois par Werbos, le Perceptron Multi-Couche apparaît en 1986 introduit par Rumelhart, et, simultanément, sous une appellation voisine, chez Le Cun (1985). Ces systèmes reposent sur la rétropropagation du gradient de l'erreur dans des systèmes à plusieurs couches, chacune de type Adaline de Bernard Widrow, proche du Perceptron de Rumelhart. De nos jours, l'utilisation des réseaux de neurones dans divers domaines ne cesse de croître. Les applications en sont multiples et variées.

II.3 Les réseaux de neurones

II.3.1. Principe du neurone artificiel

Chaque neurone artificiel est un processeur élémentaire. Il reçoit un nombre variable d'entrées en provenance de neurones en amont ou des capteurs composant la machine dont il fait partie.

A chacune de ses entrées est associé un poids représentatif de la force de la connexion. Chaque processeur élémentaire est doté d'une sortie unique, qui se ramifie ensuite pour alimenter un nombre variable de neurones en aval. A chaque connexion est associé un poids. Il est commode de représenter graphiquement un neurone comme indiqué sur la figure 1.1. Cette représentation est à l'origine de la première vague d'intérêt pour les neurones formels, dans les années 1940 à 1970 (Mcculloch et al., 1943), (Minsky et al., 1969).

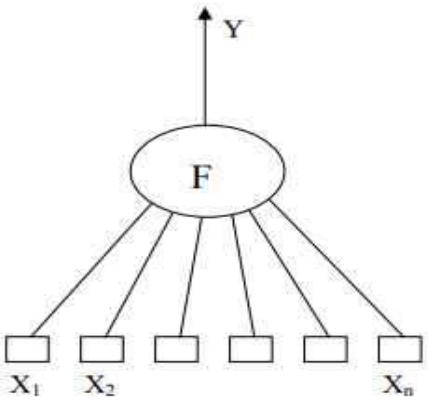
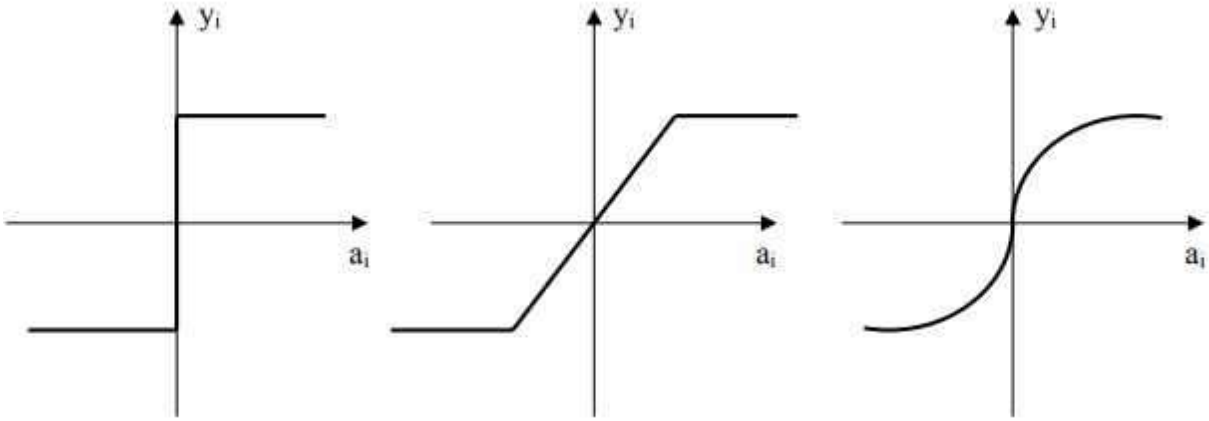


Figure II.1. Neurone artificiel.

Le neurone réalise alors trois opérations sur ses entrées :

- ✓ Pondération : multiplication de chaque entrée par un paramètre appelé poids de connexion
- ✓ Sommation : une sommation des entrées pondérées est effectuée
- ✓ Activation : passage de cette somme dans une fonction, appelée fonction d'activation. La valeur calculée est la sortie du neurone qui est transmise aux neurones suivants.



Fonction à seuil Fonction linéaire par morceaux Fonction de type sigmoïde

Figure II.2. Différents types de fonction de transfert pour le neurone artificiel.

La fonction f est appelée fonction d'activation (figure 2). Elle peut être une fonction à seuil, une fonction linéaire ou non linéaire. La fonction sigmoïde se présente comme une approximation continûment dérivable de la fonction d'activation linéaire par morceaux ou de

la fonction seuil. Elle présente l'avantage d'être régulière, monotone, continûment dérivable, et bornée entre 0 et 1 :

$$f(x) = \frac{1}{1+e^{-x}} \quad \text{(II.1)}$$

La fonction f peut être paramétrée de manière quelconque. Deux types de paramétrages sont fréquemment utilisés :

Les paramètres sont attachés aux entrées du neurone : la sortie du neurone est une fonction non linéaire d'une combinaison des entrées {xi} pondérées par les paramètres {wi}, qui sont alors souvent désignés sous le nom de poids

$$Y = th[w_0 + \sum_{i=1}^{n=1} w_i\, x_i] \quad \text{(II.2)}$$

Les paramètres sont attachés à la non-linéarité du neurone : ils interviennent directement dans la fonction y.

II.3.2. Définition

Un réseau de neurones peut être considéré comme un modèle mathématique de traitement réparti, composé de plusieurs éléments de calcul non linéaire (neurones), opérant en parallèle et connectés entre eux par des poids.

Les réseaux de neurones artificiels sont des réseaux fortement connectés de processeurs élémentaires fonctionnant en parallèle. Chaque processeur élémentaire calcule une sortie unique sur la base des informations qu'il reçoit.

Les neurones artificiels sont souvent utilisés sous forme de réseaux qui diffèrent selon le type de connections entre les neurones, une cinquantaine de types peut être dénombrée. En guise d'exemples nous citons : le perceptron de Rosemblat, les réseaux de Hopfield etc.….

Ces derniers sont les plus utilisés dans le domaine de la modélisation et de la commande des procédés. Ils sont constitués d'un nombre fini de neurones qui sont arrangés sous forme de couches. Les neurones de deux couches adjacentes sont interconnectés par des poids.

L'information dans le réseau se propage d'une couche à l'autre, on dit qu'ils sont de type « feed-forward ». Nous distinguons trois types de couches :

Couche d'entrée : les neurones de cette couche reçoivent les valeurs d'entrée du réseau et les transmettent aux neurones cachés. Chaque neurone reçoit une valeur, il ne fait pas donc de sommation.

Couches cachées : chaque neurone de cette couche reçoit l'information de plusieurs couches précédentes, effectue la sommation pondérée par les poids, puis la transforme selon sa fonction d'activation qui est en général une fonction sigmoïde. Par la suite, il envoie cette réponse aux neurones de la couche suivante.

Couche de sortie : elle joue le même rôle que les couches cachées, la seule différence entre ces deux types de couches est que la sortie des neurones de la couche de sortie n'est liée à aucun autre neurone.

II.3.3. Architecture des réseaux de neurones

On distingue deux structures de réseau, en fonction du graphe de leurs connexions, c'est-à-dire du graphe dont les nœuds sont les neurones et les arêtes les « connexions » entre ceux-ci :

- Les réseaux de neurones statiques (ou acycliques, ou non bouclés).
- Les réseaux de neurones dynamiques (ou récurrents, ou bouclés).

II.3.3.1. Les réseaux de neurones non bouclés

Un réseau de neurones non bouclé réalise une (ou plusieurs) fonction algébrique de ses entrées par composition des fonctions réalisées par chacun de ses neurones. Dans un tel réseau (figure1.3), le flux d'information circule des entrées vers les sorties sans retour en arrière. Si l'on représente le réseau comme un graphe dont les nœuds sont les neurones et les arêtes les «

Tout neurone dont la sortie est une sortie du réseau est appelé « neurone de sortie ». Les autres, qui effectuent des calculs intermédiaires, sont des « neurones cachés ».

Il existe deux types de réseaux de neurones : les réseaux complètement connectés et les réseaux à couche. Le réseau de neurones à une couche cachée et une sortie linéaire est un cas particulier de ce dernier type.

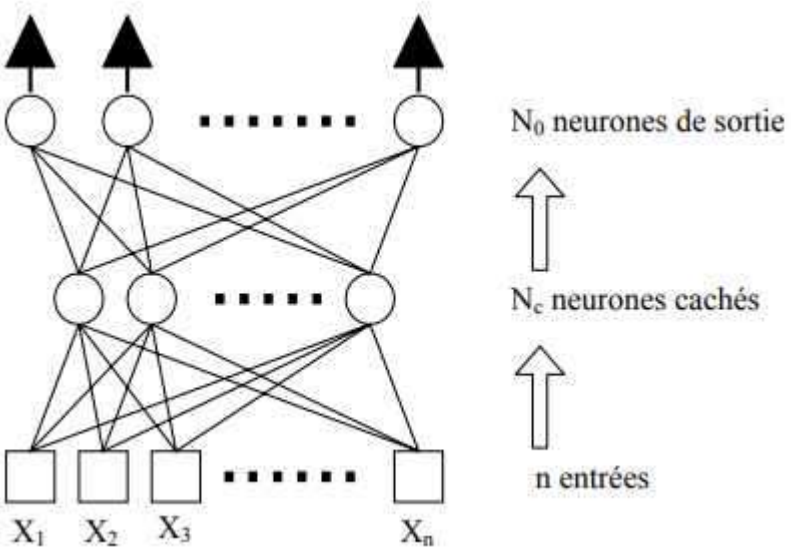


Figure II.3. Réseau de neurones à n entrées, une couche de Nc neurones Cachés et N_0 neurones de sortie.

- **Les réseaux de neurones complètement connectés**

Dans un réseau complètement connecté, les entrées puis les neurones (cachés et de sortie) sont numérotés, et, pour chaque neurone :

a) Ses entrées sont toutes les entrées du réseau ainsi que les sorties des neurones de numéro inférieur.
b) Sa sortie est connectée aux entrées de tous les neurones de numéro supérieur.

- **Les réseaux de neurones à couches**

Dans une architecture de réseaux à couches, les neurones cachés sont organisés en couches, les neurones d'une même couche n'étant pas connectés entre eux. De plus les connexions entre deux couches de neurones non consécutives sont éliminées.

Une telle architecture est historiquement très utilisée, surtout en raison de sa pertinence en classification.

II.3.3.2. Les réseaux de neurones bouclés

L'architecture la plus générale pour un réseau de neurones est le « réseau bouclé », dont le graphe des connexions est cyclique : lorsqu'on se déplace dans le réseau en suivant le sens des connexions, il est possible de trouver au moins un chemin qui revient à son point de départ

(un tel chemin est désigné sous le terme de « cycle »). La sortie d'un neurone du réseau peut donc être fonction d'elle-même ; cela n'est évidemment concevable que si la notion de temps

Ainsi, à chaque connexion d'un réseau de neurones bouclé (ou à chaque arête de son graphe) est attaché, outre un poids comme pour les réseaux non bouclés, un retard, multiple entier (éventuellement nul) de l'unité de temps choisie. Une grandeur, à un instant donné, ne pouvant pas être fonction de sa propre valeur au même instant, tout cycle du graphe du réseau doit avoir un retard non nul.

Les connexions récurrentes ramènent l'information en arrière par rapport au sens de propagation défini dans un réseau multicouche. Ces connexions sont le plus souvent locales. Pour éliminer le problème de la détermination de l'état du réseau par bouclage, on introduit sur chaque connexion « en retour » un retard qui permet de conserver le mode de fonctionnement séquentiel du réseau (figure1.4).

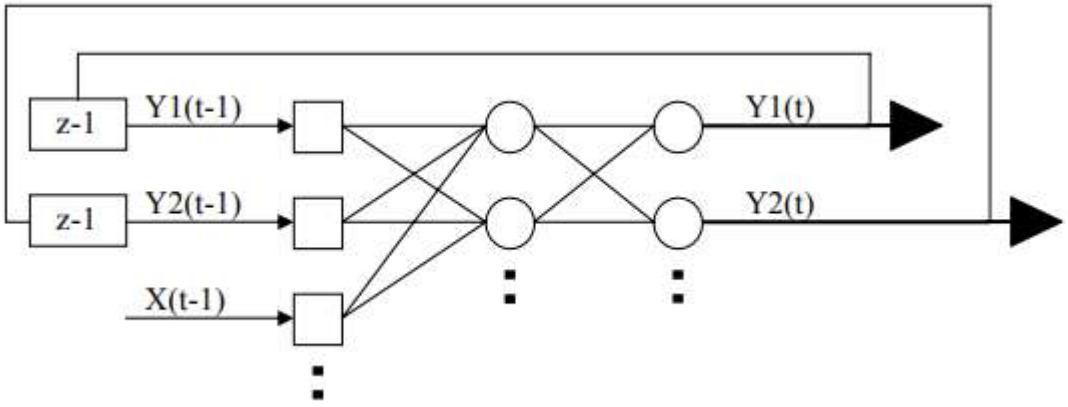


Figure II.4. Réseau de neurone bouclé

Le graphe des connexions de réseaux récurrents est cyclique. Ces réseaux sont décrits par un système d'équations aux différences.

- **Forme canonique des réseaux récurrents :**

Il a été montré [Nerrand et al., 1993] que tout réseau bouclé peut être mis sous une forme particulière, appelée forme canonique, qui est la représentation d'état minimale de la fonction réalisée par ce réseau. Cette forme canonique est constituée d'un graphe acyclique, et de connexions à retard unité reliant certaines sorties de ce graphe à ses entrées. La fonction réalisée par un réseau de neurones ayant cette structure particulière est décrite par les équations aux différences suivantes :

$$x(k+1) = \Phi\big(x(k), u(k+1)\big) \quad \text{(II.3)}$$

$$y = \Psi\big(x(k+1), u(k+1)\big) \quad \text{(II.4)}$$

Où x(k) est le vecteur d'état à l'instant k, u(k) est le vecteur des variables de commande exogènes, y(k) le vecteur des sorties, Ψ et Φ sont deux fonctions qui dépendent de la structure de la partie acyclique du réseau.

II.3.4. Apprentissage des réseaux de neurones

Le point crucial du développement d'un réseau de neurones est son apprentissage. Il s'agit d'une procédure adaptative par laquelle les connexions des neurones sont ajustées face à une source d'information (Hebb, 1949 ; Grossberg, 1982 ; Rumelhart et al., 1986).

Dans le cas des réseaux de neurones artificiels, on ajoute souvent à la description du modèle l'algorithme d'apprentissage. Le modèle sans apprentissage présente en effet peu d'intérêt.

Dans la majorité des algorithmes actuels, les variables modifiées pendant l'apprentissage sont les poids des connexions. L'apprentissage est la modification des poids du réseau dans l'optique d'accorder la réponse du réseau aux exemples et à l'expérience. Les poids sont initialisés avec des valeurs aléatoires. Puis des exemples expérimentaux représentatifs du fonctionnement du procédé dans un domaine donné, sont présentés au réseau de neurones. Ces exemples sont constitués de couples expérimentaux de vecteurs d'entrée et de sortie. Une méthode d'optimisation modifie les poids au fur et à mesure des itérations pendant lesquelles on présente la totalité des exemples, afin de minimiser l'écart entre les sorties calculées et les sorties expérimentales. Afin d'éviter les problèmes de surapprentissage, la base d'exemples est divisée en deux parties : la base d'apprentissage et la base de test. L'optimisation des poids se fait sur la base d'apprentissage, mais les poids retenus sont ceux pour lesquels l'erreur obtenue sur la base de test est la plus faible. En effet, si les poids sont optimisés sur tous les exemples de l'apprentissage, on obtient une précision très satisfaisante sur ces exemples mais on risque de ne pas pouvoir généraliser le modèle à des données nouvelles. A partir d'un certain nombre d'itérations, le réseau ne cherche plus l'allure générale de la relation entre les entrées et les sorties du système, mais s'approche trop près des points et « apprend » le bruit (Pollard et al, 1992).

Sur la figure 1.5 ci-dessous, on peut observer qu'au début de l'apprentissage, pour les premières itérations, l'erreur sur la base d'apprentissage est grande et peut légèrement

augmenter étant donné que les poids initiaux sont choisis aléatoirement. Ensuite, cette erreur diminue avec le nombre d'itérations. L'erreur sur la base de test diminue puis augmente à partir d'un certain nombre d'itérations. Les poids retenus sont ceux qui minimisent l'erreur sur la base de test.

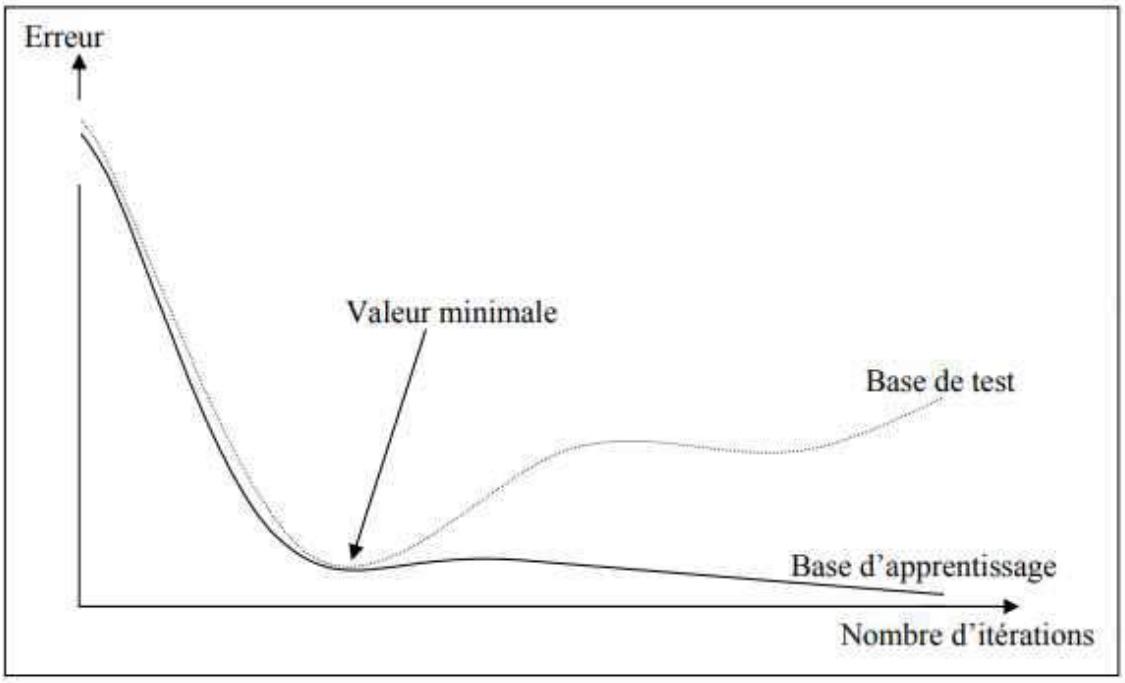


Figure II.5. Erreur moyenne sur la base d'apprentissage en fonction du nombre d'itérations.

- **Surapprentissage**

Il arrive qu'à faire apprendre un réseau de neurones toujours sur le même échantillon, celui-ci devient inapte à reconnaître autre chose que les éléments présents dans l'échantillon. Le réseau ne cherche plus l'allure générale de la relation entre les entrées et les sorties du système, mais cherche à reproduire les allures de l'échantillon. On parle alors de surapprentissage : le réseau est devenu trop spécialisé et ne généralise plus correctement.

Ce phénomène apparaît aussi lorsqu'on utilise trop d'unités cachées (de connexions), la phase d'apprentissage devient alors trop longue (trop de paramètres réglables dans le système) et les performances du réseau en généralisation deviennent médiocres.

II.3.4.1. Type d'apprentissage

Il existe de nombreux types de règles d'apprentissage qui peuvent être regroupées en trois catégories (Hassoum, 1995) : les règles d'apprentissage supervisé, non supervisé, et renforcé. Mais l'objectif fondamental de l'apprentissage reste le même : soit la classification, l'approximation de fonction ou encore la prévision (Weiss et Kulikowski, 1991). Dans l'optique de la prévision, l'apprentissage consiste à extraire des régularités (à partir des exemples) qui peuvent être transférées à de nouveaux exemples.

- **Apprentissage supervisé**

Un apprentissage est dit supervisé lorsque l'on force le réseau à converger vers un état final précis, en même temps qu'on lui présente un motif. Ce genre d'apprentissage est réalisé à l'aide d'une base d'apprentissage, constituée de plusieurs exemples de type entrées-sorties (les entrées du réseau et les sorties désirées ou encore les solutions souhaitées pour l'ensemble des sorties du réseau).

La procédure usuelle dans le cadre de la prévision est l'apprentissage supervisé (ou à partir d'exemples) qui consiste à associer une réponse spécifique désirée à chaque signal d'entrée. La modification des poids s'effectue progressivement jusqu'à ce que l'erreur (ou l'écart) entre les sorties du réseau (ou résultats calculés) et les résultats désirés soient minimisés.

Cet apprentissage n'est possible que si un large jeu de données est disponible et si les solutions sont connues pour les exemples de la base d'apprentissage.

- **Apprentissage renforcé**

L'apprentissage renforcé est une technique similaire à l'apprentissage supervisé à la différence qu'au lieu de fournir des résultats désirés au réseau, on lui accorde plutôt un grade (ou score) qui est une mesure du degré de performance du réseau après quelques itérations. Les algorithmes utilisant la procédure d'apprentissage renforcé sont surtout utilisés dans le domaine des systèmes de contrôle (White et Sofge, 1992 ; Sutton, 1992).

- **Apprentissage non supervisé**

L'apprentissage non supervisé consiste à ajuster les poids à partir d'un seul ensemble d'apprentissage formé uniquement de données. Aucun résultat désiré n'est fourni au réseau. Qu'est-ce que le réseau apprend exactement dans ce cas ? L'apprentissage consiste à détecter

les similarités et les différences dans l'ensemble d'apprentissage. Les poids et les sorties du réseau convergent, en théorie, vers les représentations qui capturent les régularités statistiques des données (Fukushima, 1988 ; Hinton, 1992). Ce type d'apprentissage est également dit compétitif et (ou) coopératif (Grossberg, 1988). L'avantage de ce type d'apprentissage réside dans sa grande capacité d'adaptation reconnue comme une autoorganisation, «self-organizing» (Kohonen, 1987). L'apprentissage non supervisé est surtout utilisé pour le traitement du signal et l'analyse factorielle.

II.3.4.2. Algorithme d'apprentissage

L'algorithme d'apprentissage est la méthode mathématique qui va modifier les poids de connexions afin de converger vers une solution qui permettra au réseau d'accomplir la tâche désirée. L'apprentissage est une méthode d'identification paramétrique qui permet d'optimiser les valeurs des poids du réseau. Plusieurs algorithmes itératifs peuvent être mis en œuvre, parmi lesquels on note : l'algorithme de rétropropagation, la Méthode Quasi-Newton, Algorithme de BFGS etc…

L'algorithme de rétropropagation (ARP) ou de propagation arrière « backpropagation » est l'exemple d'apprentissage supervisé le plus utilisé à cause de l'écho médiatique de certaines applications spectaculaires telles que la démonstration de Sejnowski et Rosenberg (1987) dans laquelle l'ARP est utilisé dans un système qui apprend à lire un texte. Un autre succès fut la prédiction des cours du marché boursier (Refenes et al, 1994 ; Lee et al. 1996) et plus récemment la détection de la fraude dans les opérations par cartes de crédit (Dorronsoro et al, 1997).

La technique de rétropropagation du gradient (Backpropagation en anglais) est une méthode qui permet de calculer le gradient de l'erreur pour chaque neurone du réseau, de la dernière couche vers la première. L'historique des publications montre que l'ARP a été découvert indépendamment par différents auteurs mais sous différentes appellations (Grossberg, 1998). Le principe de la rétropropagation peut être décrit en trois étapes fondamentales : acheminement de l'information à travers le réseau ; rétropropagation des sensibilités et calcul du gradient ; ajustement des paramètres par la règle du gradient approximé. Il est important de noter que l'ARP souffre des limitations inhérentes à la technique du gradient à cause du risque d'être piégé dans un minimum local. Il suffit que les gradients ou leurs dérivées soient nuls pour que le réseau se retrouve bloqué dans un minimum local. Ajoutons à cela la lenteur de convergence surtout lorsqu'on traite des réseaux de grande taille (c'est à dire

pour lesquels le nombre de poids de connexion à déterminer est important). Pour rendre l'optimisation plus performante, on peut utiliser des méthodes de second ordre telles que les méthodes dites de Quasi-Newton ou de Newton modifiée.

- ✓ **Méthodes Quasi-Newton**

Dans notre travail nous avons utilisé la technique de Quasi-Newton (Dennis et al., 1983 ; Thibault et Grandjean, 1991). Cet algorithme nécessite la connaissance des valeurs de la fonction objectif à minimiser et du vecteur des dérivées premières au point solution courant. Le vecteur des dérivées est évalué analytiquement.

Avec les méthodes Quasi-Newton, on veut généraliser la formule itérative de Newton.

$$x^{k+1} = x^k - \alpha^k [\nabla^2 f(x^k)]^{-1}. \quad (f x^k) \qquad \text{(II.5)}$$

Le principe des méthodes de résolution de type Quasi-Newton est de générer une séquence de matrices symétriques définies positives qui soient des approximations, toujours améliorées, de la matrice Hessienne réelle ou de son inverse. Désignons par Hk la suite d'approximation de l'inverse de la matrice Hessienne $[\nabla^2 f(x^k)]^{-1}$ et par Bk la suite des estimations de la matrice Hessienne $\nabla^2 f(x^k)$.

On recherche une méthode telle que, dans le cas d'un problème quadratique, la matrice Bk converge vers la valeur exacte des dérivées secondes (constantes dans ce cas), de sorte qu'en fin de convergence, on retrouve une convergence de type Newton. Si l'on applique la méthode à une fonction quelconque, Bk peut être considéré, à chaque instant, comme une approximation (définie positive) du Hessien.

- ✓ **Algorithme de BFGS**

L'algorithme de BFGS (du nom de ses inventeurs : Broyden, Fletcher, Goldfarb et Shanno) (Press et al., 1988) prend implicitement en compte les dérivées secondes et s'avère donc nettement plus performante que la méthode de rétropropagation. Le nombre d'itérations est nettement plus faible et les temps de calcul réduits d'autant.

L'algorithme de BFGS est une règle d'ajustement des paramètres qui a l'expression suivante :

$$\theta^k = \theta^{k-1} \pm \mu_k M_k \quad (\theta^{K-1}) \qquad \text{(II.6)}$$

Où Mk est une approximation, calculée itérativement, de l'inverse de la matrice Hessienne.

II.4. Conception d'un réseau de neurones

- Les réseaux de neurones réalisent des fonctions non linéaires paramétrées. Leurs mises en œuvre nécessitent :
- La détermination des entrées et des sorties pertinentes, c'est à dire les grandeurs qui ont une influence significative sur le phénomène que l'on cherche à modéliser.
- La collecte des données nécessaires à l'apprentissage et à l'évaluation des performances du réseau de neurones.
- La détermination du nombre de neurones cachés nécessaires pour obtenir une approximation satisfaisante.
- La réalisation de l'apprentissage.
- L'évaluation des performances du réseau de neurones à l'issue de l'apprentissage.

II.4.1. Détermination des entrées/sorties du réseau de neurones

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

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

II.4.2. Choix et préparation des échantillons

Le processus d'élaboration d'un réseau de neurones commence toujours par le choix et la préparation des échantillons de données. La façon dont se présente l'échantillon conditionne le type de réseau, le nombre de cellules d'entrée, le nombre de cellules de sortie et la façon dont il faudra mener l'apprentissage, les tests et la validation (Bishop, 1995). Il faut donc déterminer les grandeurs qui ont une influence significative sur le phénomène que l'on cherche à modéliser.

Lorsque la grandeur que l'on veut modéliser dépend de nombreux facteurs, c'est-à-dire lorsque le modèle possède de nombreuses entrées, il n'est pas possible de réaliser un « pavage » régulier dans tout le domaine de variation des entrées : il faut donc trouver une méthode permettant de réaliser uniquement des expériences qui apportent une information significative

pour l'apprentissage du modèle. Cet objectif peut être obtenu en mettant en œuvre un plan d'expériences. Pour les modèles linéaires, l'élaboration de plans d'expériences est bien maîtrisée, par ailleurs, ce n'est pas le cas pour les modèles non linéaires.

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

Notons qu'il n'y a pas de règle pour déterminer ce partage d'une manière quantitative, néanmoins chaque base doit satisfaire aux contraintes de représentativité de chaque classe de données et doit généralement refléter la distribution réelle, c'est à dire la probabilité d'occurrence des diverses classes (Nascimento, 2000).

II.4.3. Elaboration de la structure du réseau

La structure du réseau dépend étroitement du type des échantillons. Il faut d'abord choisir le type de réseau : un perceptron standard, un réseau de Hopfield, un réseau à décalage temporel (TDNN), un réseau de Kohonen, un ARTMAP etc...

Par exemple, dans le cas du perceptron multicouches, il faudra aussi bien choisir le nombre de couches cachées que le nombre de neurones dans cette couche.

- ✓ **Nombre de couches cachées :**

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

- ✓ **Nombre de neurones cachés :**

Chaque neurone peut prendre en compte des profils spécifiques de neurones d'entrée. Un nombre plus important permet donc de mieux ″coller″ aux données présentées mais diminue la capacité de généralisation du réseau. Il faut alors trouver le nombre adéquat de neurones cachés nécessaire pour obtenir une approximation satisfaisante. Il n'existe pas, à ce jour, de résultat théorique permettant de prévoir le nombre de neurones cachés nécessaires pour obtenir une performance spécifique du modèle, compte tenu des modèles disponibles. Il faut donc nécessairement mettre en œuvre une procédure numérique de conception de modèle.

II.4.4. Apprentissage

L'apprentissage est un problème numérique d'optimisation. Il consiste à calculer les pondérations optimales des différentes liaisons, en utilisant un échantillon. La méthode la plus utilisée est la rétropropagation, qui est généralement plus économe que les autres en termes de nombres d'opérations arithmétiques à effectuer pour évaluer le gradient.

Pour rendre l'optimisation plus performante, on peut utiliser des méthodes de second ordre. Le calcul est très efficace, mais lourd. Elles ont de nombreuses limitations, quant aux conditions de convergence, sur les dérivées secondes. Des corrections sont proposées pour éviter ce problème, et sont prises en compte par les méthodes dites de Quasi-Newton ou de Newton modifiée.

Il a été observé que les poids calculés par la méthode de rétropropagation sont plus faibles que ceux obtenus par la technique de Quasi-Newton, ce qui semblerait montrer que la recherche d'un minimum par rétropropagation est restreinte à un voisinage immédiat des poids initiaux, d'où une dépendance plus forte de cette méthode par rapport à l'initialisation (Thibault, 1991). C'est pour ces raisons que l'optimisation par une méthode Quasi-Newton a été largement utilisée dans certains travaux récents de modélisation de procédés de génie chimique (Piron et al. 1997 ; Fakhr-Eddine et al, 1999), et a été choisie également dans la suite de nos travaux.

II.4.5. Validation et Tests

Alors que les tests concernent la vérification des performances d'un réseau de neurones hors échantillon et sa capacité de généralisation, la validation est parfois utilisée lors de l'apprentissage. Une fois le réseau de neurones développé, des tests s'impose afin de vérifier la qualité des prévisions du modèle neuronal.

Cette dernière étape doit permettre d'estimer la qualité du réseau obtenu en lui présentant des exemples qui ne font pas partie de l'ensemble d'apprentissage. Une validation rigoureuse du modèle développé se traduit par une proportion importante de prédictions exactes sur l'ensemble de la validation.

Si les performances du réseau ne sont pas satisfaisantes, il faudra, soit modifier l'architecture du réseau, soit modifier la base d'apprentissage.

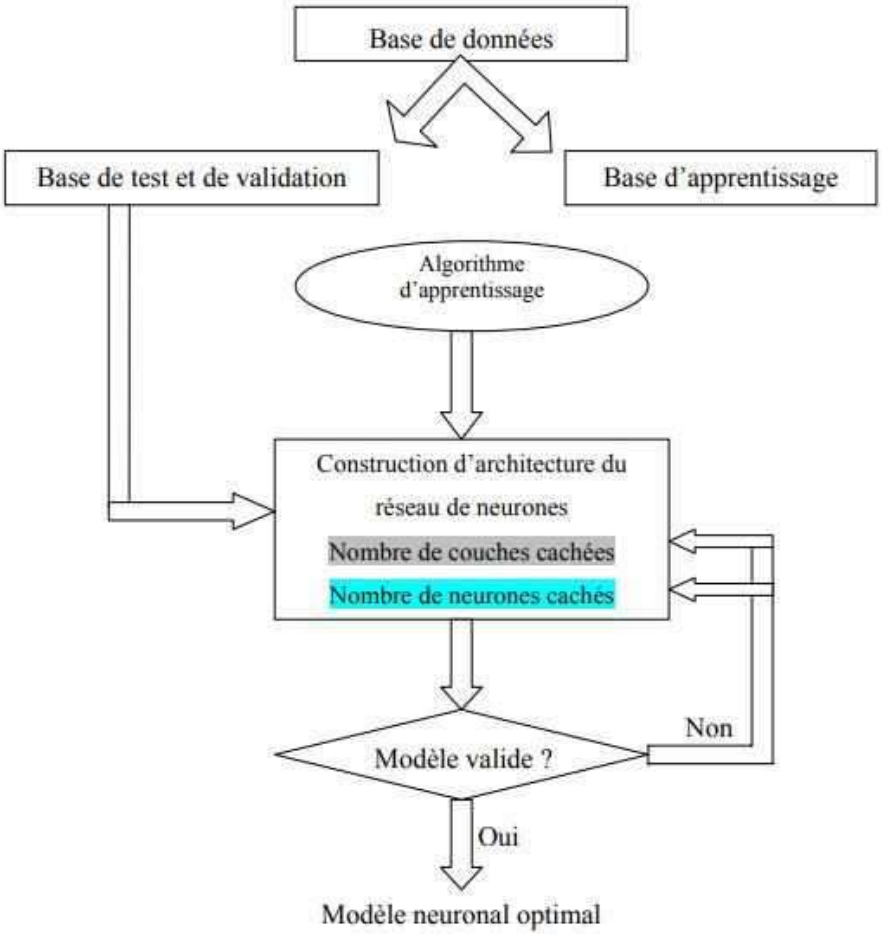


Figure II.6. Organigramme de conception d'un réseau de neurones.

II.5. Conclusion

Dans ce chapitre, nous avons introduit les définitions essentielles relatives aux réseaux de neurones. Nous avons notamment distingué entre les réseaux de neurones non bouclés, statiques, qui réalisent des fonctions non linéaires et ceux bouclés, dynamiques, qui réalisent des équations aux différences non linéaires.

Nous avons aussi mis l'accent sur l'utilisation des réseaux de neurones comme outils de modélisation par apprentissage. Ces derniers permettent d'ajuster des fonctions non linéaires très générales à des ensembles de points. Comme toute méthode qui s'appuie sur des techniques statistiques, l'utilisation de réseaux de neurones nécessite que l'on dispose de données suffisamment nombreuses et représentatives.

Enfin, nous avons présenté les concepts fondamentaux de la modélisation à l'aide de réseaux de neurones et une méthodologie complète de conception et de mise en œuvre de modèles neuronaux.

Partie Modélisation

III.1 Introduction

De nombreux composés pharmaceutiques sont des molécules peu solubles dans l'eau ce qui limite très fortement leurs biodisponibilités (assimilation par le corps humain) et donc leurs commercialisations. L'industrie pharmaceutique cherche donc à développer de nouveaux procédés afin d'augmenter cette biodisponibilité.

Les propriétés thermodynamiques du CO_2 dans les conditions supercritiques et la facilité de récupération du CO_2 sont responsables de l'utilisation répandue du solvant CO_2 pour séparer différents types de substances (Asiabi et al., 2015; Henley et al., 2011)

Étant donné que les corrélations basées sur la densité disponibles et les équations d'états ne sont pas suffisamment précis pour prédire les données de solubilité de différents solutés dans le CO_2 supercritique.

La modélisation thermodynamique concerne plusieurs méthodes parmi ces méthodes on a la méthode de Van Der Waals c'est une équation d'état mais elle ne peut prétendre représenter correctement à la fois la zone du point critique et le comportement des phases liquides et elle suggère pour un corps pure la coexistence des deux phases liquide et gaz sur une certaine plage de pression pour une même température. Il y'a aussi Les méthodes semi-empiriques sont des techniques de résolution de l'équation de Schrödinger de systèmes à plusieurs électrons, elles utilisent des données ajustées sur des résultats expérimentaux afin de simplifier les calculs mais elles restent plus difficile à appliquer et insuffisante lorsqu'il s'agit de prévoir la géométrie de molécules plus complexes , on a la loi empirique c'est une formule issue de faits expérimentaux mais dont on ne connaît pas de base théorique ou qu'on ne peut pas relier à une base théorique simple, Ce type de loi, bien qu'empirique peut cependant souvent alimenter l'élaboration d'une théorie. Pour éviter tous ces problèmes on a travaillé sur la modélisation par le RNA et SVM :

Le RNA en tant que systèmes capables d'apprendre, mettent en œuvre le principe de l'induction, c'est-à-dire l'apprentissage par l'expérience. Par confrontation avec des situations ponctuelles, ils infèrent un système de décision intégré dont le caractère générique est fonction du nombre de cas d'apprentissages rencontrés et de leur complexité par rapport à la complexité du problème à résoudre.

La modélisation par la méthode de SVM Les machines à vecteurs de support ou séparateurs à vaste marge sont un ensemble de techniques d'supervisé destinées à résoudre des problèmes de discrimination et de régression. Les SVM sont une généralisation des classifieurs linéaires.

Notre travail est basé sur la solubilité des solutés solides dans le CO_2, on a choisis le CO_2 car c'est un fluide supercritique et un excellent solvant apolaire de plus il est plus sain que les solvants organiques.

Dans la contribution 1, nous avons construire des modèles basés sur le réseau de neurone artificiel (RNA) sur de larges fissures de température et de pression, on compare l'effet de l'utilisation de descripteurs moléculaires (la méthode QSAR) comme entrés pour prédire les valeurs de solubilités de 25 solutés solides dans le dioxyde de carbone supercritique.

Dans la contribution 2, on a étudié notre base de données expérimentale de la solubilité de 25 médicaments dans le CO_2 avec la méthode des machines à vecteurs de support.

L'objectif de la modélisation de la solubilité de ces 25 composés pharmaceutiques c'est de découvrir un modèle par les méthodes de RNA et SVM et distinguer le meilleur parmi les deux après la comparaison des résultats obtenus à la fin de l'étude.

CONTRIBUTION 1 : Modélisation par réseaux de neurones artificiels.

III.2 La modélisation neuronale

Les réseaux neuronaux artificiels sont bien connus pour leurs structures dans lesquelles une grande quantité d'informations traitées via un réseau interconnecté comprenant plusieurs nœuds dans des couches successives, (Graupe, 2007).

Le flux d'informations dépend de l'architecture du réseau qui imite celle présente dans les neurones des corps humains. Sur la base de la structure biologiquement inspirée, la classification des motifs dans plusieurs domaines, y compris la reconnaissance, l'optimisation, le contrôle et la prédiction, est possible une fois que le modèle de réseau a été formé et testé bien pour différents ensembles de données. La sélection des facteurs d'influence les plus importants, la structure du réseau, le processus d'apprentissage et le problème des minimums locaux sont responsables du processus de développement du modèle. Après avoir sélectionné les variables d'entrée par voie théorique ou semi-empirique, le processus d'apprentissage accompagné de phases de formation et de test peut être effectué après la normalisation et l'élimination des valeurs aberrantes des variables d'entrée, (Passino, 2005).

III.2.1 Méthodologie de travail

La modélisation par RNA regroupe plusieurs procédures sont les suivant :

1. Sélection et collecte des données thermodynamiques.
2. Calcul des descripteurs et prétraitement de la base de données.

3. Séparation de la base de données.
4. Normalisation des données.
5. Sélection des paramètres de réseau de neurones.
    a. Algorithme d'apprentissage.
    b. Fonction de transfert de la couche cachée de sortie
    c. Optimisation du nombre de neurones dans la couche cachée.
6. Validation et généralisation.
7. Exploitation des résultats
    a. Nombre de couches cachées
    b. Equation du RNA
    c. Interprétation des résultats

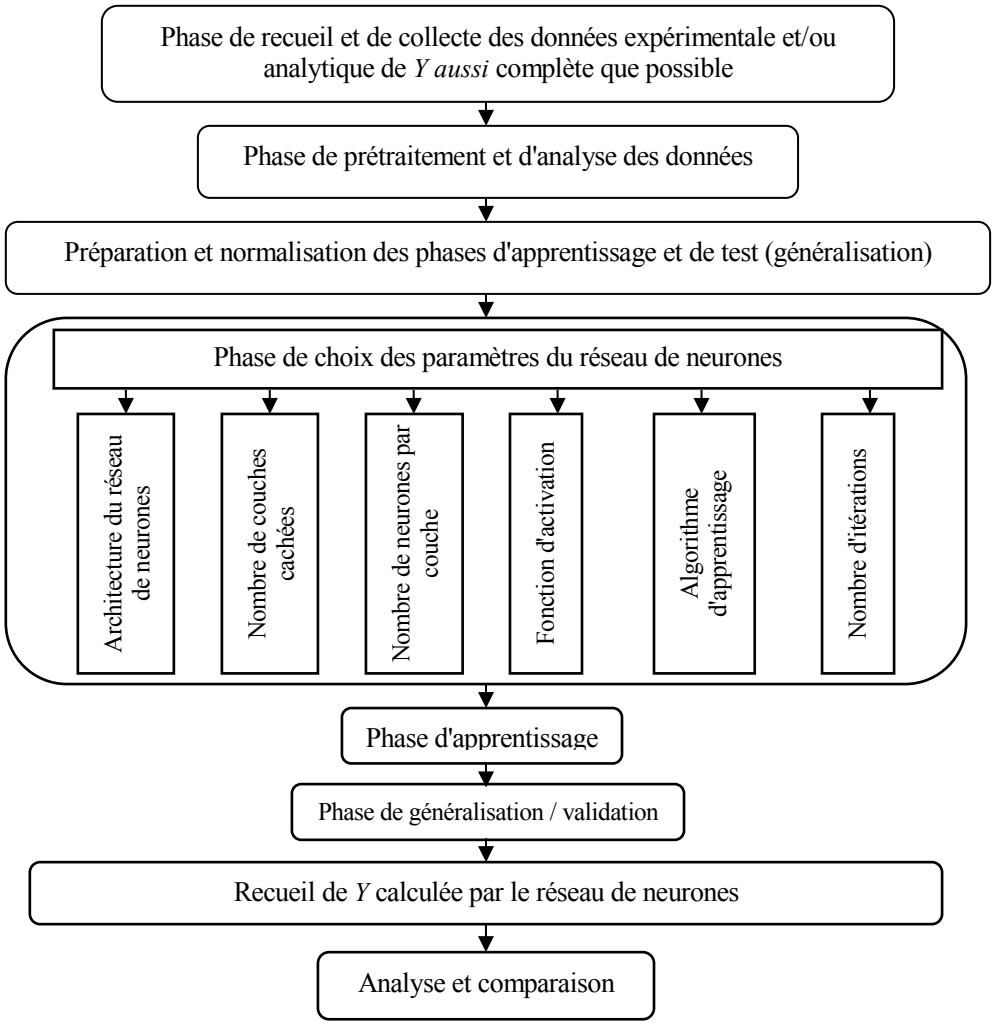


Figure III.1 : Méthodologie de calcul de la solubilité par les réseaux de neurones.

Ces étapes sont programmées sous MATLAB. Nous découvrons ensuite les résultats obtenus de calcul et leurs interprétations.

III.2.1.1 Sélection et collecte des données thermodynamiques

Dans cette partie concernant la modélisation par RNA, nous avons sélectionné 25 composés pharmaceutiques qui étudie la solubilité dans le CO_2 supercritique par RNA les données expérimentales concernent la température (k), la pression (bar), la masse volumique (kg/m^3), la masse molaire (g/mol) et la solubilité du soluté solide (en fraction molaire). Le tableau III.1 représente les données expérimentales suivantes :

Tableau III.1 : Source et domaine des données utilisées pour l'apprentissage et la validation du modèle de RNA pour le système de 25 solutés solides qui sont solubles dans le CO_2 supercritiques.

N°	Noms des composés	T (K)	P (bar)	ρ (Kg/m3)	Taill	Référence
1	Simvastatin	308-348	121.6-354.6	327-955	45	Hojjati.M et al.2006
2	Lovastatine	308-348	121,6-354,6	327-955	45	Hojjati.M et al.2006
3	Anastrazole	308-348	122-355	327-955	45	Hojjati.M et al.2009
4	Letrozole	308-348	122-355	327-955	45	Hojjati.M et al.2009
5	Exemestane	308-348	122-355	327-955	45	Hojjati.M et al.2009
6	Atropine	308-348	122-355	327-955	45	Yamini.Y et al 2001
7	Diazepam	308-348	122-355	327-955	45	Yamini.Y et al 2001
8	Codeine	308-348	122-355	327-955	45	Yamini.Y et al 2001
9	Carbamazepine	308-348	122-355	477-955	39	Yamini.Y et al 2001
10	Benzocaine	308-348	122-355	327-955	40	Garmroodi.A et al 2004
11	Naproxen	308-348	122-355	327-955	40	Garmroodi.A et al 2004
12	Metronidazole	308-348	122-355	327-955	40	Garmroodi.A et al 2004
13	Methylparaben	308-348	122-355	327-955	40	Asghari-Khiavi.M et al 2003
14	Bisacodyl	308-348	122-355	396-955	39	Asghari-Khiavi.M et al 2003
15	Methimazole	308-348	122-355	396-955	39	Asghari-Khiavi.M et al 2003
16	Lamotrigine	318-348	121,6-354,6	327-919	36	Haji Hosseini.M et al 2010
17	Clozapine	318-348	121,6-354,6	327-919	27	Haji Hosseini.M et al 2010
18	Zopiclone	313-333	100-250	291,41-880,24	21	Medina.I et al 2001
19	Nimodipine	313-333	100-250	291,416-880,243	21	Medina.I et al 2001
20	Acetanalide	308,2-323,2	90-400	383,928-972,039	29	P. Coelho.J et al 1959-007
21	Propanamide	308,2-323,2	90-400	284,752-972,039	30	P. Coelho.J et al 1959-007
22	Butanamide	308,2-323,2	90-400	284,752-972,039	30	P. Coelho.J et al 1959-007
23	DA	308-348	122-355	327-955	45	Asiabia H et al 2013
24	Ketoprofen	312,5-331,9	100-220	404-859	10	J. Macnaughton.S et al 1996
25	Piroxicam	312,5-331,5	100-220	494-859	9	J. Macnaughton.S et al 1996

III.2.1.2 Calcul des descripteurs et prétraitement de la base de données

Les produits chimiques sont des éléments essentiels de la civilisation humaine. Les applications des produits chimiques à l'ère moderne couvrent une large gamme, des environnements industriels aux ménages. L'objectif estimé d'un chimiste réside dans le développement de produits chimiques avec le profil désiré de leur meilleure manifestation comportementale, c'est-à-dire activité / propriété. En chimie, le chimiste veut améliorer l'efficacité de la molécule tout en réduisant les effets toxiques exercés. Maintenant, suivant l'axiome général, le développement d'un nouveau produit chimique peut être réalisé soit en concevant une nouvelle entité complète, soit en modifiant une ancienne. Dans les deux cas, les chimistes doivent posséder suffisamment de connaissances sur le produit chimique et son potentiel d'interaction avec le système biologique. Il est évident que l'activité biologique (y compris la toxicité) et les propriétés de tout produit chimique (p. Ex., Médicaments, produits pharmaceutiques et cancérogènes) dépendent de son interaction avec le système biologique concerné. Par conséquent, l'objectif principal d'un concepteur chimique réside dans l'établissement d'une explication rationnelle du mécanisme d'action du produit chimique, ce qui peut conduire à une dérivation théorique et donc à l'adaptation de sa structure. Parfois, le manque d'action appropriée pour la drogue limite la dérivation hypothétique. Afin d'établir une base théorique appropriée pour l'action des produits chimiques, il faudrait avoir un aperçu profond de la chimie fondamentale du système qui contrôle les comportements physico-chimiques et structurels (Hansch et al., 1995).

La chimie a été une science centrale essentielle qui corrèle et explore différents problèmes de discipline scientifique (Balaban and Klein, 2006).

Compte tenu du nombre important de produits chimiques qui influent sur notre vie du point de vue de l'application industrielle, de la recherche en laboratoire, et la consommation des ménages, il sera intéressant de développer une base qui donne une raison appropriée pour expliquer le comportement des produits chimiques. Une fois que cette base est développée, elle nous permettra également de modifier le comportement de ces produits chimiques en incorporant des changements structurels raisonnables. La notion relation activité-structure (SAR) tente d'établir une relation entre les différentes caractéristiques comportementales des composés chimiques (à savoir, activité / propriété / toxicité et informations chimiques dérivables des structures chimiques). Un tel processus implique une connaissance significative de la chimie et de la physique (définition des composés), des mathématiques (permettant une analyse de modélisation) et de la biologie (donner des raisons pour le profil d'activité biologique

des produits chimiques) pour développer une base rationnelle. En d'autres termes, l'étude SAR offre une option pour construire une équation mathématique pour un ensemble de produits chimiques pour leur comportement spécifique d'activité / propriété / toxicité en utilisant des informations sur leur structure chimique. On peut noter que, étant donné que nous parlons d'équations mathématiques, cela implique des nombres ou des quantités. Si l'information comportementale des produits chimiques est exprimée en termes qualitatifs, l'étude est appelée SAR qualitative, alors que la présence de nombres quantitatifs représentant l'activité des produits chimiques définit le processus comme une étude QSAR (Todeschini et al., 2009).

Cependant, les études qualitatives et quantitatives sont appelées collectivement QSAR. QSAR est une étude théorique pour construire des conclusions rationnelles et réduit le fardeau d'expérimentation pratique ainsi que le temps. Il utilise comparativement moins de composés pour établir une relation mathématique qui peut être employée pour caractériser un grand nombre de produits chimiques appartenant à une classe identique ou différente selon la nature du modèle (local ou global). Cette caractéristique est particulièrement importante pour évaluer l'activité biologique et / ou la toxicité des produits chimiques qui impliquent aux animaux de laboratoire. La technique QSAR peut conduire à l'utilisation d'une expérimentation moins animale (puisqu'elle utilise comparativement moins de grandeur de données expérimentales et peut donner la priorité aux composés pour une autre expérimentation), ce qui constitue une alternative précieuse aux études sur les animaux. Dans la section suivante, nous expliquerons brièvement les principaux objectifs du développement de modèles prédictifs quantitatifs (Todeschini et al., 2009; Tute MS, 1990).

Les descripteurs moléculaires dont leur utilisation dans la modélisation neuronale représentent l'originalité de ce travail, jouent un rôle fondamental dans la chimie, les sciences pharmaceutiques, la politique de protection de l'environnement, et les recherches sur la santé.

Les programmes de calcul de descripteurs acceptent les données d'entrées des molécules sous forme de code. Différentes formes de codifications de la structure moléculaires existent en chimio-informatique. La codification (notation) SMILES est la plus utilisée pour sa simplicité.

La représentation numérique de la structure moléculaire a été évaluée en termes de descripteurs moléculaire. La notion SMILES nécessaire au calcul des descripteurs a été tirée en ligne par le logiciel « pubchem » à partir du site : https://pubchem.ncbi.nlm.nih.gov.

Le Tableau III.2 suivant représente la formule SMILE des composés.



Figure III.2 : Capteur d'écran du site intitulé pubchem (https://pubchem.ncbi.nlm.nih.gov).

Tableau III.2 : Les composés et ses formules SMILE (https://pubchem.ncbi.nlm.nih.gov).

N°	Noms des composés	Formule SMILE des composés
1	Simvastatin	CCC(C)(C) C(=O) OC1CC(C=C2C1C(C(C=C2) C) CCC3CC(CC(=O) O3) O) C
2	Lovastatine	CCC(C)C(=O) OC1CC(C=C2C1C(C(C=C2) C) CCC3CC(CC(=O) O3) O) C
3	Anastrazole	CC(C)(C#N) C1=CC(=CC(=C1) CN2C=NC=N2) C(C)(C) C#N
4	Letrozole	C1=CC(=CC=C1C#N) C(C2=CC=C(C=C2) C#N) N3C=NC=N3
5	Exemestane	CC12CCC3C(C1CCC2=O) CC(=C) C4=CC(=O) C=CC34C
6	Atropine	CN1C2CCC1CC(C2) OC(=O) C(CO)C3=CC=CC=C3
7	Diazepam	CN1C(=O) CN=C(C2=C1C=CC(=C2) Cl) C3=CC=CC=C3
8	Codeine	CN1CCC23C4C1CC5=C2C(=C(C=C5) OC) OC3C(C=C4) O
9	Carbamazepine	C1=CC=C2C(=C1) C=CC3=CC=CC=C3N2C(=O) N
10	Benzocaine	CCOC(=O) C1=CC=C(C=C1) N
11	Naproxen	CC(C1=CC2=C(C=C1) C=C(C=C2) OC) C(=O) O
12	Metronidazole benzoate	CC1=NC=C(N1CCOC(=O) C2=CC=CC=C2) [N+] (=O) [O-]
13	Methylparaben	COC(=O) C1=CC=C(C=C1) O
14	Bisacodyl	CC(=O) OC1=CC=C(C=C1) C(C2=CC=C(C=C2) OC(=O) C) C3=CC=CC=N3
15	Methimazole	CN1C=CNC1=S
16	Lamotrigine	C1=CC(=C(C(=C1) Cl) Cl) C2=C(N=C(N=N2) N) N
17	Clozapine	CN1CCN(CC1) C2=C3C=CC=CC3=NC4=C(N2) C=C(C=C4) Cl
18	Zopiclone	C1CN(CCN1) C(=O) OC2C3=NC=CN=C3C(=O) N2C4=NC=C(C=C4) Cl. Cl
19	Nimodipine	CC1=C(C(C(=C(N1) C) C(=O) OCCOC) C2=CC(=CC=C2) [N+] (=O) [O-
20	Acetanalide	CC (=O) NC1=CC=CC=C1
21	Propanamide	CCC(=O) N
22	Butanamide	CCCC(=O) N
23	DA	CC(=O) OCC(=O) C1CCC2C1(CCC3C2CCC4=CC(=O) CCC34C) C
24	Ketoprofen	CC(C1=CC=CC(=C1) C(=O) C2=CC=CC=C2) C(=O) O
25	Piroxicam	CN1C(=C(C2=CC=CC=C2S1(=O) =O) O) C(=O) NC3=CC=CC=N3

Tous les descripteurs ont été obtenus par le programme « PaDel Descriptor Calculator » le programme calcule 1875 descripteurs (1444 :1D&2D descripteurs et 431 :3D descripteurs) à partir du site : http:/www.scbdd.com/padel_desc/index (voir tableau III.3)

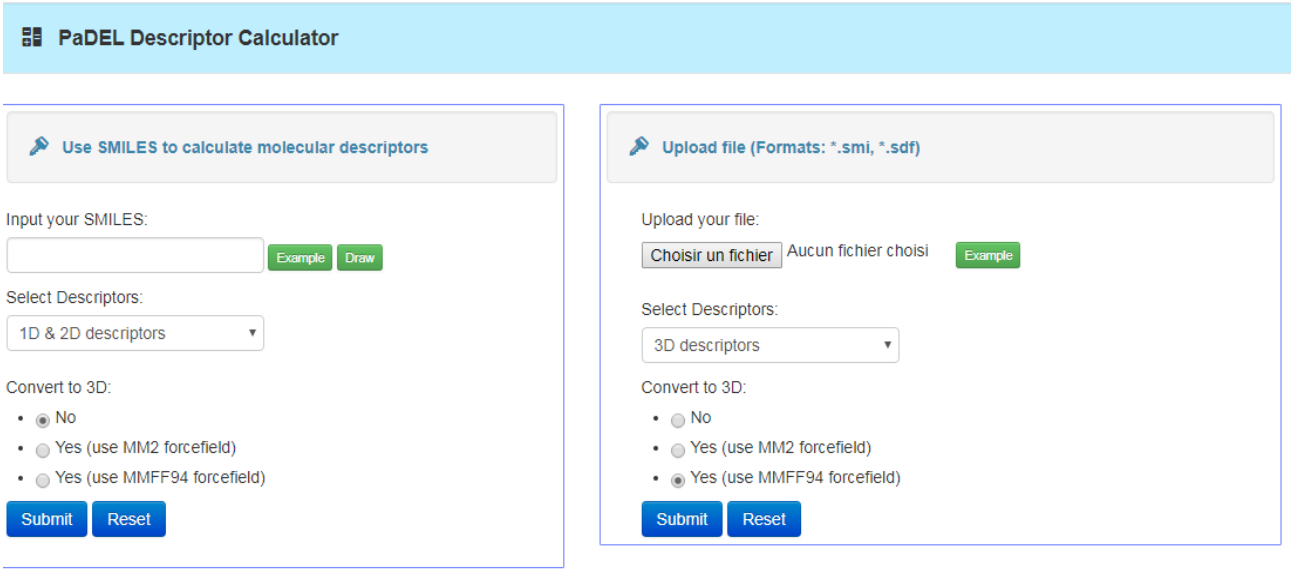


Figure III.3 : Capteur d'écran du site intitulé PaDel (www.scbdd.com).

Tableau III.3 : Nombre de descripteurs calculés (www.scbdd.com).

Catégorie de descripteurs	Nombre par catégorie
Autocorrelation descriptors	346
Basak descriptors	42
BCUT descripteurs	6
Burden descriptors	96
Connectivity descriptors	56
Constitutional descriptors	120
E-state descriptors	489
Kappa descriptors	3
Molecular property descriptors	15
Quantum chemical descriptors	6
Topological descriptors	265
CPSA descriptors	29
RDF descriptors	210
Geometrical descriptors	21
WHIM descriptors	91
3D Autocorrelation descriptors	80
Total	1875

Le prétraitement de la base de données a pour objectif d'éliminer les descripteurs qui ne sont pas pertinents afin d'éviter le phénomène de sur-apprentissage. Il s'agit alors de la réduction de la dimension de l'espace des variables (descripteurs) du modèle avec l'élimination entrées qui n'ont pas ou possèdent peu d'influence sur la sortie.

Une procédure basée sur un pré-traitement de la base de données a été utilisée pour la sélection des descripteurs les plus pertinents suivants les étapes suivantes (Hamdache et al. 2014) :

1. On calcule le minimum et le maximum pour chaque descripteur, puis on enlève les descripteurs qui ont le maximum et le minimum égaux.
2. On élimine le descripteur qui possède la même valeur de plus de 75% des échantillons.
3. On calcule la déviation standard relative (RSD) pour chaque descripteur et on élimine ceux ayant des valeurs de RSD inférieure à 0,05.
4. On obtient alors une matrice qui représente la corrélation entre la f_b ($f_b = a_0 + \sum a_i d_i$) et les descripteurs retenus.

   On classe les descripteurs selon la valeur décroissante du coefficient de corrélation. On prend le descripteur ayant la plus grande corrélation et on le compare aux autres descripteurs de matrice. Ceux dont la valeur du coefficient de corrélation est supérieure à 0,75 sont éliminés à leur tour. On répète la même chose avec le descripteur classé juste après le premier et ainsi de suite.
5. On utilise un programme pour la sélection des descripteurs les plus pertinents parmi ceux obtenus à l'étape 4.

Le tableau III.4 regroupe les résultats obtenus après le prétraitement des descripteurs car il y'a plusieurs types et des class différents ça dépend l'autocorrélation et de la topologie des descripteurs etc.

Tableau III.4: Les descripteurs et leurs types et leurs descriptions (www.scbdd.com).

	Descripteur	Type	Description	Dimension	Class Entende
1D&2D	GATS5c	Geary AutocorrelationDescriptor	Geary autocorrelation - lag 5 / weighted by charges	2	Autocorrelation descriptors
	ATSC2e	Centered Broto-Moreau Autocorrelation Descriptor	Centered Broto-Moreau autocorrelation - lag 2 / weighted by Sanderson electronegativities	2	Autocorrelation descriptors
	Zagreb	Zagreb Index Descriptor	Sum of the squares of atom degree over all heavy atoms I	2	Topological descriptors
3D	De	WHIMDescriptor	D total accessibility index / weighted by relative Sanderson electronegativities	3	WHIM descriptors
	E3e	WHIMDescriptor	3rd component accessibility directional WHIM index / weighted by relative Sanderson electronegativities	3	WHIM descriptors
	L2i	WHIMDescriptor	2nd component size directional WHIM index / weighted by relative first ionization potential	3	WHIM descriptors

III.2.1.3 Séparation de la base de données

La base des données expérimentales concernant les 5 paramètres thermodynamiques des équilibres de phases (Masse volumique ρ_a et moléculaire M_a adimensionnelle, Température T_r et Pression $P_{r\ réduite}$ et la solubilité Y en fraction molaire) du système mais pour améliorer le model thermodynamique on a utilisées les descripteurs qui influent plus sur la solubilité.

Tableau III.5 : Les différentes entrées de la matrice.

N	Les entrées	Descriptions
1 2 3 4	T_r P_r ρ_a M_a	Propriétés thermodynamique
5 6 7	GATS5c ATSC2e Zagreb	1D&2D
8 9 10	De E3e L2i	3D
11	Y	Solubilité

La base de données est réalisée par une matrice rectangulaire répartie de la façon suivante :

- Entrée : l'entrée variable (1 ; 10).
- Sortie : la sortie fixée c'est la solubilité Y (11).

La base de données constituée de l'ensemble des données qui comportent les propriétés thermodynamiques réduites et les paramètres adimensionnels et les descripteurs. Elle est divisée en trois sous bases : une base d'apprentissage (60%) ; validation (20%) qui permet l'optimisation de l'architecture du RNA et une base test (20%). Cette division c'est une division aléatoire utilisée pour déterminer ses performances prédictives.

Dans cette étude, sept architectures de réseaux neuronaux multicouches sont utilisées afin d'étudier la performance de chaque modèle pour une estimation précise les solubilités sont cités dans le tableau III.6, Par souci de comparaison, on utilise le MAARD (III.1) et R^2 (III.2) pour déterminer les performances des réseaux.

$$MAARD\ \ \% = \frac{1}{n}\sum_{i=1}^{n}\left(\left|\frac{y_{i,\ \ a}-y_{i,exp}}{y_{i,exp}}\right|\right)\times 100\% \qquad \text{(III.1)}$$

$$R^2 = 1-\sum_{i=1}^{n}\left(\frac{\left(y_{pred\,;i}-y_{exp,i}\right)^2}{\left(y_{pred\,,i}-y_m\right)^2}\right) \qquad \text{(III.2)}$$

Tableau III.6 : Les sept différents systèmes de réseau de neurones.

Noms	Entrées	Sortie
RNA 1	T_r , P_r , ρ_a , M_a	Y
RNA 2	(1D&2D)	Y
RNA 3	3D	Y
RNA 4	(1D&2D) +3D	Y
RNA 5	(T_r , P_r , ρ_a , M_a) + (1D&2D) +3D	Y
RNA 6	(T_r , P_r , ρ_a , M_a) + 3D	Y
RNA 7	(T_r , P_r , ρ_a , M_a) + (1D&2D)	Y

III.2.1.4 Normalisation des données

Tous les paramètres d'entrée et de sortie ont été normalisés entre -1 et +1 par l'équation suivante dans le but d'accélérer la convergence du modèle neuronal :

$$x_{in} = \frac{2\{x_i - \min(x_i)\}}{\{\max(x_i) - \min(x_i)\}} - 1 \qquad \text{(III.3)}$$

x_{in} : valeur normalisée

x_i : valeur réelle

La fonction **Premnmx** de MATLAB réalise la normalisation selon l'équation (III.1.1), la fonction inverse de la dénormalisation des données pour la sortie (x_{in}) est réalisée par la fonction **Postmnmx** de MATLAB par l'expression suivante :

$$x_i = \frac{x_{in}\{\max(x_i) - \min(x_i)\} + \{\max(x_i) + \min(x_i)\}}{2} \quad \text{(III.4)}$$

III.2.1.5 Sélection des paramètres de réseau de neurones

a. Algorithme d'apprentissage

Pour la modélisation statique, le réseau de neurones le plus employé est de type perception multicouche, ce dernier est caractérisé par son abtitude a modéliser les phénomènes physique non linéaire toute un utilisant une couche cachée avec un nombre de neurones minimal et une fonction de transfert de type "tansig" pour la couche cachée et le "purelin" pour la sortie, ce réseau est programmé dans Matlab sous le nom "trainlm"

b. Fonction de transfert de la couche cachée de sortie

Comme il n'existe pas de méthode permettant de prévoir a priori quelle sera la fonction de transfert (d'activation) qui donnera les meilleurs résultats. Plusieurs travaux montrent qu'un MLP est capable a modélisé n'importe quel phénomène physique non linéaire avec un nombre minimal de neurones cachée ont utilisant la fonction sigmoïde pour les neurones cachées et généralement une fonction linéaire pour le neurone de la couche sortie.
La relation sigmoïde et linéaire définie par les relations suivantes :

$$f(x) = \frac{1}{1+e^{-x}} \quad \text{(III.5)}$$

$$f(x) = y \quad \text{(III.6)}$$

c. Optimisation du nombre de neurones dans la couche cachée

La performance d'un modèle neuronal est généralement mesurée par la règle test-erreur, le réseau le plus performent correspond à une erreur vers 0 et R vers 1 raison pour laquelle, le

nombre de neurone cachée a été varié aléatoires de 4 à 40 et chaque architecture a été répétée 30 fois dans le but d'éviter le problème de sous et sur apprentissage, plusieurs modèle ont été testé avec et sans descripteur et le résultat montre que le RNA 5 est le plus performent avec un erreur et R optimale (voir le tableau III.6).

III.2.1.6 Validation et généralisation

La validation consiste à tester le réseau après ou au cours de l'apprentissage avec des données proche des données utilisées lors de la phase d'apprentissage. Les résultats de ce test permettront au modèle d'éviter les situations de sous ou de sur apprentissage. En revanche la généralisation s'effectue avec des données tout à fait nouvelle et différentes de la base de données de départ. Elle permet de savoir si le réseau a développé une capacité de généralisation assez performante.

III.2.1.7 Exploitation des résultats

Nous avons essayé avec une couche cachée, d'après les résultats un RNA avec une couche cachée semble meilleur, raison pour laquelle on donne l'organigramme pour une couche cachée, le nombre de neurones dans la couche cachée a été sélectionné aléatoirement.

Il n'y a pas de formule magique pour sélectionner le nombre optimal de neurones dans la couche cachée. Nous avons donc suivi un moyen de trouver le meilleur nombre de neurones dans la couche cachée en utilisant un algorithme basé sur le langage MATLAB. Nous avons atteint le nombre de neurones dans la couche cachée comme indiqué dans le tableau III.7.

Tableau III.7 : Structure du réseau de neurones optimisé (RNSO).

Type du réseau	Couche d'entrée	1ière couche cachée		Couche de sortie		Algorithme d'apprentissage (Nb. D'itérations) {Erreur de performance
	Nb. De neurones	Nb. De neurones	Fonction d'activation	Nb. De neurones	Fonction d'activation	
RNS MLP	10	16	Tangente hyperbolique TANSIG	1	Identité Purline	LM+RB (TRINBR) (1000)

Dans notre travail on exploite la technique de RNA, en utilisant le logiciel MATLAB R2009b afin de prédire la solubilité.

L'optimisation de l'architecture du RNA consiste à la détermination :

- Du nombre des neurones dans la couche cachée,
- De la fonction d'activation de la couche de sortie,
- De l'algorithme d'apprentissage

Pour comparer la précision de chaque RNA on à calculer deux constantes : MAARD et R^2 le tableau illustre les résultats trouvés.

Tableau III.8 : Architecture des RNA utilisés

Noms	Architecture	MAARD (%)	R^2 (-)
RNA1	De 7 à 10 → 11	7,8363	0,91644
RNA2	De 1 à 3 → 11	7,8177	0.91605
RNA3	De 4 à 6 → 11	7 ,8731	0,91532
RNA4	De 1 à 6 → 11	6,2932	0,93627
RNA5	De 1 à 10 → 11	**1,4223**	**0,99692**
RNA6	De 4 à 10 → 11	**1,5429**	**0,99628**
RNA7	De 1 à 3+7 à 10) → 11	**1,6687**	**0,99575**

Le tableau montre plusieurs architectures qui contiennent des différents paramètres comme suit :

- De 7 à 10 → 11 qui contient seulement les propriétés.
- De 1 à 3 → 11 qui contient les descripteurs 2D
- De 4 à 6 → 11 qui contient les descripteurs 3D
- De 1 à 6 → 11 qui contient les descripteurs 2D et 3D
- De 1 à 10 → 11 qui contient les descripteurs 2D et 3D et les propriétés.
- De 4 à 10 → 11 qui contient les descripteurs 3D et les propriétés.
- De 1 à 3+7 à 10) → 11 qui contient les descripteurs 2D et les propriétés.

Et montre aussi la gamme des valeurs de MAARD est entre 1.4223% et 7.873% mais quand nous prendrons les meilleures des MAARD mentionné en gras, nous trouverons une gamme des valeurs entre 1.4223% et 1.6687% avec une moyenne de 1.5455%. On a remarqué que les descripteurs 3D jouent un rôle très important afin d'assurer une erreur minimale dans la base des données.

MAARD minimum (1.4223%) est obtenu par le RNA5, et MAARD maximum (7.8731%) est obtenu par le RNA3.

La comparaison globale entre la performance des RNA indique que le RNA5 a donné la meilleure corrélation ou de MARRD est 1.4223%. Le second RNA qui a la meilleure corrélation est le RNA6 ou de MAARD est 1.5429%. Le troisième avec une meilleure corrélation est le RNA7 ou de MAARD est 1.6687%. Ces trois réseaux ensemble représente une meilleure corrélation des 25 composés considéré.

a. Nombre de couches cachées

Le nombre de neurones dans la couche d'entrée et la couche de sortie est fixé par le nombre de variables indépendantes et dépendantes respectivement. Il n'y a pas de règles générales qui fixent le nombre de couches cachées d'un MLP pour un problème donné. Par conséquent la stratégie d'essai et erreur est la plus raisonnable. Le nombre de neurones dans chaque couche cachée est fixé par le modélisateur à partir d'un programme Matlab par rapport à l'erreur qu'on a fixé.

b. Equation du RNA

D'après la théorie du réseau de neurone artificiel on peut tirer une relation entre la sortie et les entrés. L'équation représente la relation sortie-entrés du RNA5.

$$y = \sum_{e=1}^{n} \frac{1}{1+e^{-(w_{e,i}\, x_i)}} + b_i \qquad \text{(III.6)}$$

Où : i, e et n sont l'entré, la sortie, le nombre de neurones dans la couche d'entrée et la couche cachée 1.

Tableau III.9 : les poids et les biais

	Wj1	Wj2	Wj3	Wj4	Wj5	Wj6	Wj7	Wj8	Wj9	Wj10	B
1	2.0549	-0.4763	-04989	-0.5290	5.0982	-0.0541	-1.2980	0.2643	-0.1601	0.6545	0.8809
2	-0.8931	0.0458	0.0529	1.2023	-0.4598	0.1625	-0.8505	0.3611	-0.2299	0.5485	0.9549
3	2.5994	0.0852	-0.8001	2.9202	-0.7223	-1.3968	0.2575	0.3891	0.0061	1.4363	-0.1781
4	-1.5625	0.8788	-3.3051	-0.1785	-1.0369	0.2192	2.7620	0.8177	-0.5354	1.5232	-0.5755
5	-5.5106	-0.4214	0.6680	4.1510	1.3817	-0.5613	-2.8476	-0.5998	0.2717	1.8899	-0.1098
6	-1.0746	-0.9321	-0.2954	-0.2938	-0.5228	-0.8031	0.8528	-0.3630	1.5031	-0.1612	-0.1044
7	0.5847	1.7158	-1.2660	0.0483	-2.5245	2.2804	2.5044	-0.3756	-0.0927	0.1911	0.5533
8	-0.6884	-0.3070	-1.7003	1.8755	1.3551	1.6931	2.4326	-0.2403	0.1637	-0.5112	-0.7525
9	-0.2916	-1.4341	2.5214	1.6979	0.7721	-0.6820	-2.1381	0.2222	-0.1197	-0.0624	-1.0473
10	5.0190	-0.3031	-0.7063	-0.9968	2.6377	0.8735	-1.7054	0.9345	-0.1308	0.6944	-0.2475
11	0.1633	-1.1071	-1.9244	-1.3208	0.9661	-0.1551	3.1074	-0.1865	0.2527	-0.5895	0.8389
12	-2.0056	-0.4529	-0.2821	-1.5374	-0.1950	0.0369	0.7423	-1.8299	0.3581	-1.6042	0.1160
13	-0.2054	0.1824	-0.1505	0.2131	-0.0485	0.2051	1.2797	0.5345	-1.1932	-0.0969	0.7510
14	1.4849	-0.2781	-0.0715	2.3052	1.2637	2.0228	-2.2732	-0.8632	0.2665	-1.1416	0.2466
15	-0.6327	0.4307	2.7286	2.4709	-1.5179	-3.5777	1.1819	0.5604	-0.1437	0.7443	0.4088
16	-1.6767	-2.6424	4.3895	-0.3064	-4.6227	2.3742	-4.2256	-0.3071	0.4396	0.1940	-0.2604

c. Interprétation des résultats

Les résultats de performances du modèle du réseau de neurones établies pour les phases d'apprentissage, test et validation pour l'architecture de 1 à 10 → 11 est présenté sur la figure I montre bien que tous les points sont situés sur la première bissectrice avec un coefficient de corrélation approchant l'idéal (0,99692). La figure II montre la courbe de régression linéaire de l'architecture de 4 à 10 →11 calculée par le RNA avec l'architecture expérimentale que tous les points sont situés sur la première bissectrice avec un coefficient de corrélation approchant l'idéal (0,99628). La figure III montre le même calcul pour la troisième architecture de 1 à 6+7 à 10 → 11 avec un vecteur de corrélation linéaire égal (0,99575). Ce qui confirme la robustesse du RNA.

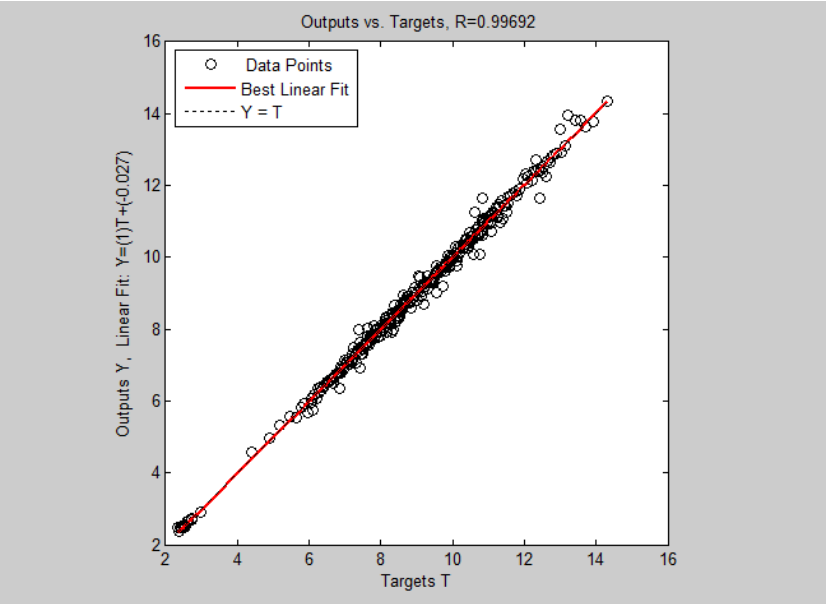


Figure III.4 : Régression linéaire de l'architecture de 1 à 10 → 11 calculée par le RNA optimisé pour les phases d'apprentissage, test et validation.

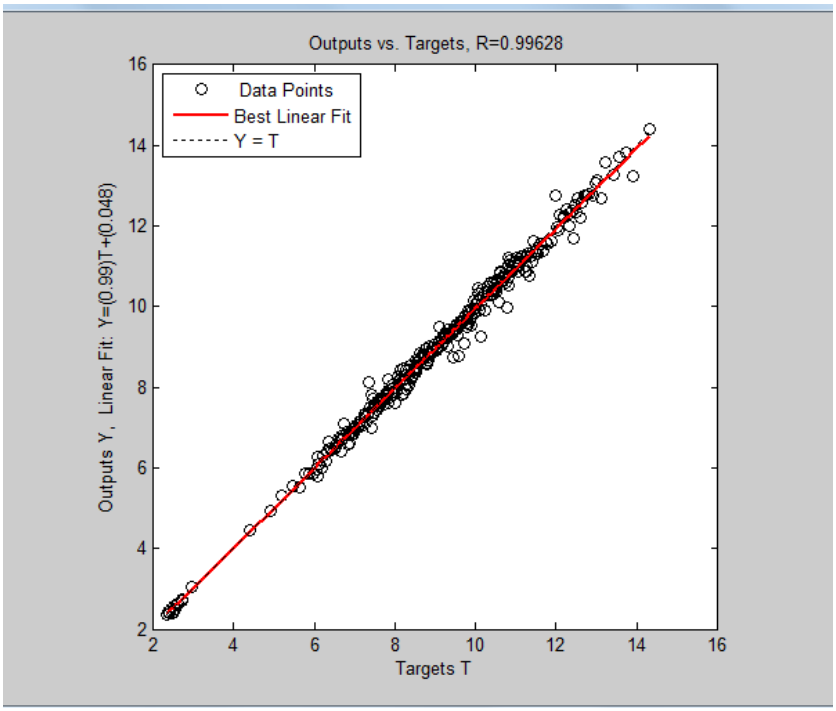


Figure III.5 : Régression linéaire de l'architecture de 4 à 10 → 11 calculée par le RNA optimisé pour les phases d'apprentissage, test et validation.

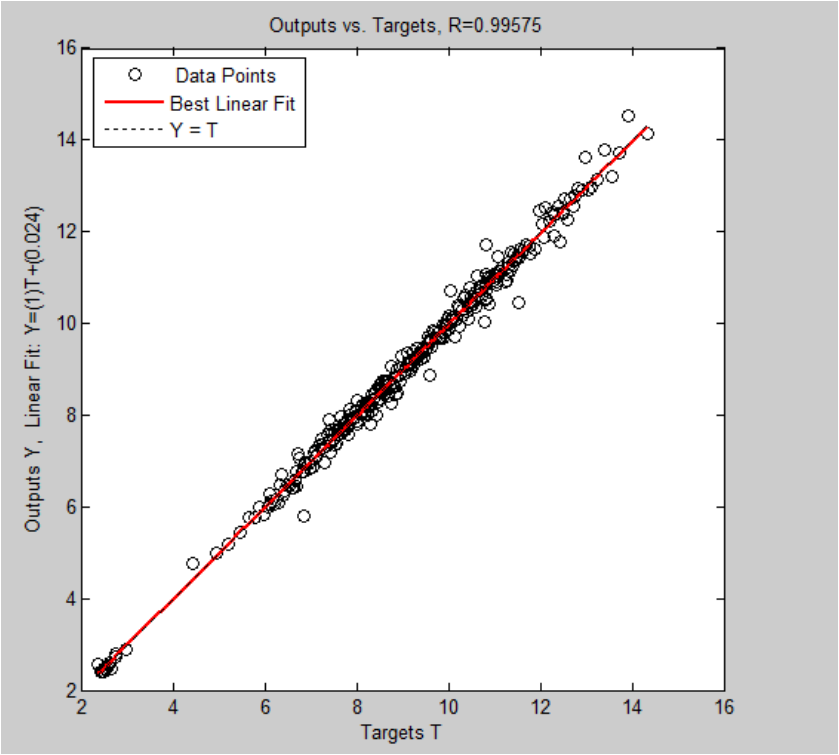


Figure III.6 : Régression linéaire de l'architecture de 1 à 3+7 à 10 → 11 calculée par le RNA optimisé pour les phases d'apprentissage, test et validation.

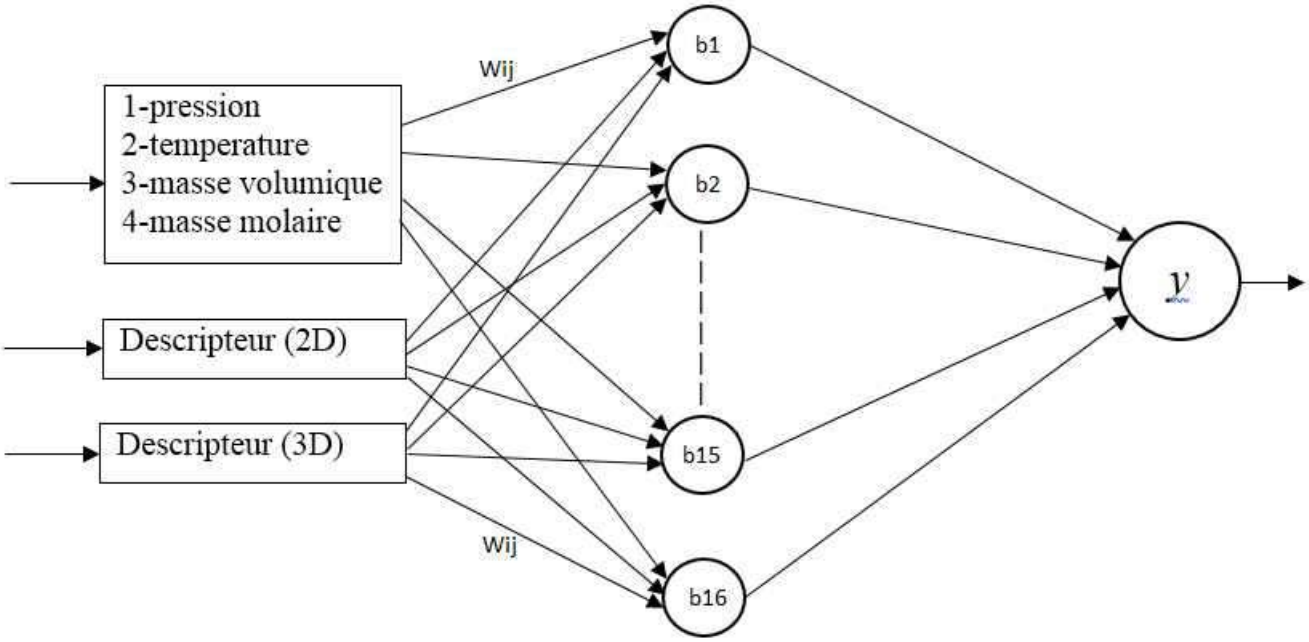


Figure III.7 : Réseau de neurones artificiels

III.2.1.8 Conclusion

La modélisation neuronale de la solubilité du 25 composés solides dans le CO_2 supercritique choisis les descripteurs moléculaires et les variables intensive comme des entrés a fait sortir les remarques suivantes :

- Les résultats de la modélisation neuronale montrent que les RNA, comme il était montré dans plusieurs études passés peut-être utilisé avec fiabilité pour représenter les données expérimentales de la solubilité des solutés solides dans le CO_2 supercritique.
- La modélisation neuronale par les descripteurs choisi à partir des structures 2D et 3D donne un meilleur résultat par rapport aux autres.

Les résultats trouvés montrent que les descripteurs moléculaires peuvent influencer sur les résultats de prédiction de la solubilité. Mais reste que la modélisation neuronale basé sur les propriétés thermodynamique et descripteur donne un meilleur résultat pour la prédiction de la solubilité.

CONTRIBUTION 2 : Modélisation par machines à vecteurs supports.

III.3 La modélisation par machines à vecteurs supports

Les Support Vecteur Machines souvent traduit par l'appellation de Séparateur à Vaste Marge (SVM) sont une classe d'algorithmes d'apprentissage initialement définis pour la discrimination c'est-à-dire la prévision d'une variable qualitative binaire. Ils ont été ensuite généralisés à la prévision d'une variable quantitative. Dans le cas de la discrimination d'une variable dichotomique, ils sont basés sur la recherche de l'hyperplan de marge optimale qui, lorsque c'est possible, classe ou sépare correctement les données tout en étant le plus éloigné possible de toutes les observations. Le principe est donc de trouver un classifieur, ou une fonction de discrimination, dont la capacité de généralisation (qualité de prévision) est le plus grand possible.

Cette approche découle directement des travaux de Vapnik en théorie de l'apprentissage à partir de 1995. Elle s'est focalisée sur les propriétés de généralisation (ou prévision) d'un modèle en contrôlant sa complexité.

Concernant la dimension de Vapnik Chernovenkis qui est un indicateur du pouvoir séparateur d'une famille de fonctions associé à un modèle et qui en contrôle la qualité de prévision. Le principe fondateur des SVM est justement d'intégrer à l'estimation le contrôle de la complexité c'est-à-dire le nombre de paramètres qui est associé dans ce cas au nombre de vecteurs supports. L'autre idée directrice de Vapnik dans ce développement, est d'éviter de substituer à l'objectif initial : la discrimination, un ou des problèmes qui s'avèrent finalement plus complexes à résoudre comme par exemple l'estimation non-paramétrique de la densité d'une loi multidimensionnelle en analyse discriminante.

Le principe de base des SVM consiste de ramener le problème de la discrimination à celui, linéaire, de la recherche d'un hyperplan optimal. Deux idées ou astuces permettent d'atteindre cet objectif :

_ La première consiste à définir l'hyperplan comme solution d'un problème d'optimisation sous contraintes dont la fonction objective ne s'exprime qu'à l'aide de produits scalaires entre vecteurs et dans lequel le nombre de contraintes "actives" ou vecteurs supports contrôle la complexité du modèle.

_ Le passage à la recherche de surfaces séparatrices non linéaires est obtenu par l'introduction d'une fonction noyau (kernel) dans le produit scalaire induisant implicitement une transformation non linéaire des données vers un espace intermédiaire (feature space) de

plus grande dimension. D'où l'appellation couramment rencontrée de machine à noyau ou kernel machine.

Sur le plan théorique, la fonction noyau définit un espace hilbertien, dit auto-reproduisant et isométrique par la transformation non linéaire de l'espace initial et dans lequel est résolu le problème linéaire. 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 spam, 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 (WikiStat).

III.3.1 Le principe des SVM

L'idée principale des SVM consiste à projeter les données dans un espace de plus grande dimension appelé, espace de caractéristiques, afin que les données non linéairement séparables dans l'espace d'entrée deviennent linéairement séparables dans l'espace de caractéristiques. En appliquant dans cet espace la technique de construction d'un hyperplan optimal séparant les deux classes, on obtient une fonction de classification qui dépend d'un produit scalaire des images des données de l'espace d'entrée dans l'espace des caractéristiques.

Ce produit scalaire peut être exprimé, sous certaines conditions, par des fonctions définies dans l'espace d'entrée, qu'on appelle les noyaux. Ce multiple choix de noyaux rend les SVM plus intéressantes et surtout plus riches puisqu'on peut toujours chercher de nouveaux noyaux qui peuvent être mieux adaptés à la tâche qu'on veut accomplir. Les trois noyaux les plus utilisés sont : le noyau linéaire, le noyau polynomial et le noyau gaussien noté aussi RBF (Radial Basis Fonction).

Les Support Vecteur Machines (SVM) sont des nouvelles techniques discriminantes dans la théorie de l'apprentissage statistique. Elles ont été proposées en 1995 par V. Vapnik dans son livre _The nature of statistical learning théorie.

Elles permettent d'aborder plusieurs problèmes divers et variés comme la régression, la classification, la fusion etc.

Les SVM fournissent une approche très intéressante de l'approximation statistique. Souvent, le nombre des exemples pour l'apprentissage est insuffisant pour que les estimateurs fournissent un modèle avec une bonne précision. D'un autre côté, l'acquisition d'un grand nombre d'exemples s'avère être souvent très coûteuse et peut même mener à des problèmes de sur-apprentissage dans le cas où la capacité du modèle est très complexe. Pour ces deux raisons, il faut arriver à un compromis entre la taille des échantillons et la précision recherchée. Dans ces cas spécifiques comme la reconnaissance de formes, il serait intéressant de trouver une mesure de la fiabilité de l'apprentissage, et d'avoir une mesure du taux d'erreur qui sera commis durant la phase de test. Ces nouvelles techniques unifient deux théories : Minimisation du risque empirique et Capacité d'apprentissage d'une famille de fonctions. C'est la Minimisation du Risque Structurel. (Kharroubi.2002)

La méthode SVR mappe les données dans une grande dimension l'espace caractéristique (F) via un mappage non linéaire (φ), puis la régression linéaire est faite dans cet espace (Vapnik 1995).

La fonction d'estimation SVR générique est : $f(x) = (w.\ \varphi(x)) + b$ où $w \subset R^n$, b est le seuil, et φ dénote la transformation non-linéaire de R^n en dimensionnelle espace. Il est ciblé pour trouver la valeur de w et b de sorte que les valeurs de x peuvent être déterminées en minimisant le risque de régression :

$$R_{reg}(f) = \sum_{i=0}^{l} C(f(x_i) - y_i) + \lambda\|w\|^2 \qquad \text{(III.7)}$$

Où l'indice la taille de l'échantillon, C(.) est la fonction de coût, λ est la constante de régularisation, et $(y_1; \ldots; y_l)$ sont les mesures. Equation (III.8) peut être minimisée pour un grand ensemble de fonctions de coût considérant le problème comme solvable problème de programmation quadratique. Le vecteur w peut être écrit comme :

$$w = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)\ \varphi(x_i) \qquad \text{(III.8)}$$

Où α_i et α_i^* sont les solutions de ce qui précède problème de programmation quadratique, et α_i et α_i^* peut être expliqué que les forces poussantes et tirant l'estimation $f(x_i)$ vers les mesures y_i (Vapnik 1995). Le problème complet peut être réécrit comme un produit interne dans le bas de l'espace d'entrée :

$$f(x) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)\big(\varphi(x_i).\varphi(x)\big) + b \qquad \text{(III.9)}$$
$$= \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)\ k(x_i, x) + b$$

Dans l'équation ci-dessus, une fonction du noyau k (x_i ; x) est introduit et toute fonction noyau symétrique satisfaisant la condition de Mercer correspond à un produit intérieur dans certaines caractéristiques espace. Un noyau commun pour la régression est la base radiale Fonction (RBF) comme suit :

$$k(x_i, x) = exp\{-\gamma|x - x_i|^2\} \quad \text{(III.10)}$$

La fonction de perte ε-insensible de Vapnik est un coût largement utilisé fonction (Vapnik 1995) et sa forme peut être la suivante :

$$C(f(x) - y) = \begin{cases} |f(x)-y|-\varepsilon, \text{ for } |f(x)-y| \geq \varepsilon \\ 0 \quad \text{otherwise} \end{cases} \quad \text{(III.11)}$$

Le problème quadratique correspondant peut être exprimé comme suit :

$$\longrightarrow \frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i^* - \alpha_j)(\alpha_i^* - \alpha_j)k(x_i, x_j)$$

Minimiser : $-\sum_{i=1}^{l}\alpha_i^*(y_i - \varepsilon) - \alpha_i(y_i + \varepsilon)$ (III.12)

Sujet à : $\sum_{i=1}^{l}\alpha_i - \alpha_i^* = 0, \ \alpha_i, \alpha_i^* \in \left[0, \frac{1}{\lambda}\right]$ (III.13)

Pour tous les points seulement si l'exigence $|f(x) - y| \geq \varepsilon$ est remplie, alors les multiplicateurs de Lagrange peuvent être des valeurs non nulles et utilisé comme vecteurs de support. Dans cet article, nous avons utilisé l'environnement STATISTICA et le Noyau RBF défini par : $k(x_i, x_j) = exp\left(-\gamma\left\|-x_i - x_j\right\|^2\right)$, où γ est un paramètre positif défini par l'utilisateur. Plusieurs essais ont été faits pour sélectionner les paramètres appropriés définis par l'utilisateur. Nous avons commencé avec ε égal à 0,01, et utilisé la recherche de grille pour sélectionnez les valeurs appropriées pour C et γ. Nous avons changé les valeurs de γ et C de sorte que chacun prendra l'une des valeurs dans la séquence [10-100]. Nous avons sélectionné paire (γ, C) qui a donné la meilleure performance sur la validation des données. Ce processus a continué pendant quelques cycles jusqu'à ce que les valeurs finales des paramètres (γ et C) aient été sélectionnées. Ensuite, nous avons essayé plusieurs valeurs de ε dans la plage de [0.01- 0.5], et sélectionné la valeur qui a fait que SVM effectue le meilleur sur les données de validation (Bakhashwain.2016).

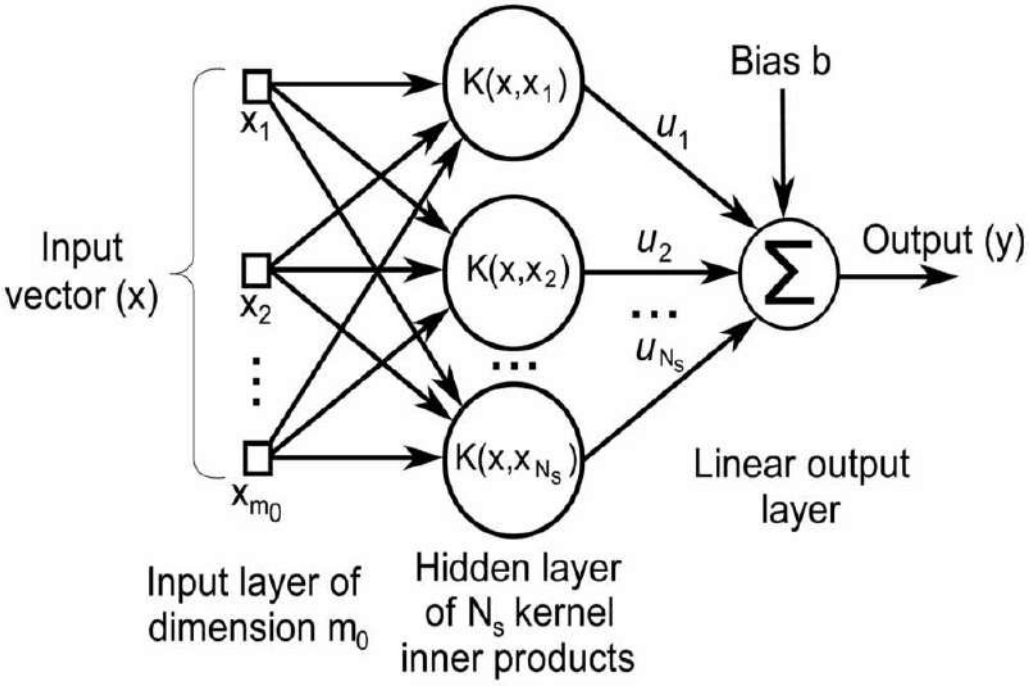


Figure III.8: Architecture d'une machine de régression construite par le SV Algorithme. (Bakhashwain.2016).

III.3.2 Exploitation des résultats (SVM)

Le SVM construit avec la fonction noyau RBF a été appliqué à l'ensemble de données sélectionné pour prédire la solubilité de quelques soluté solides dans le CO_2 supercritique par SVM utiliser l'ensemble des variables comme estimateurs (des paramètres d'entrée).

Méthode de recherche avec une validation croisée décuplée a été utilisée pour dériver les paramètres du modèle SVM.

La totalité ensemble de données composées de 895 analyses sont divisées au hasard en jeu de données d'apprentissage (671 points de données) et l'ensemble de données de test (224 points de données) et l'ensemble de données de validation (895 points de données)

Le modèle de vecteur de support est évalué sur la base de paramètres statistiques et ensuite comparés avec le modèle MLR.

On peut noter que les valeurs prédites par le modèle de la variable dépendante une forte corrélation avec les valeurs expérimentales (R^2 était 0.922 pour l'apprentissage et 0.911, 0.919 pour test et validation).

Les résultats de performances du modèle support à vecteur machine établies pour les phases d'apprentissage, test et validation est présenté sur la figure II.5. Montre bien que tous

les points sont situés sur la première bissectrice avec un coefficient de corrélation approchant l'idéal (**0.910806**) et avec une valeur optimale de l'erreur est **0.024.**

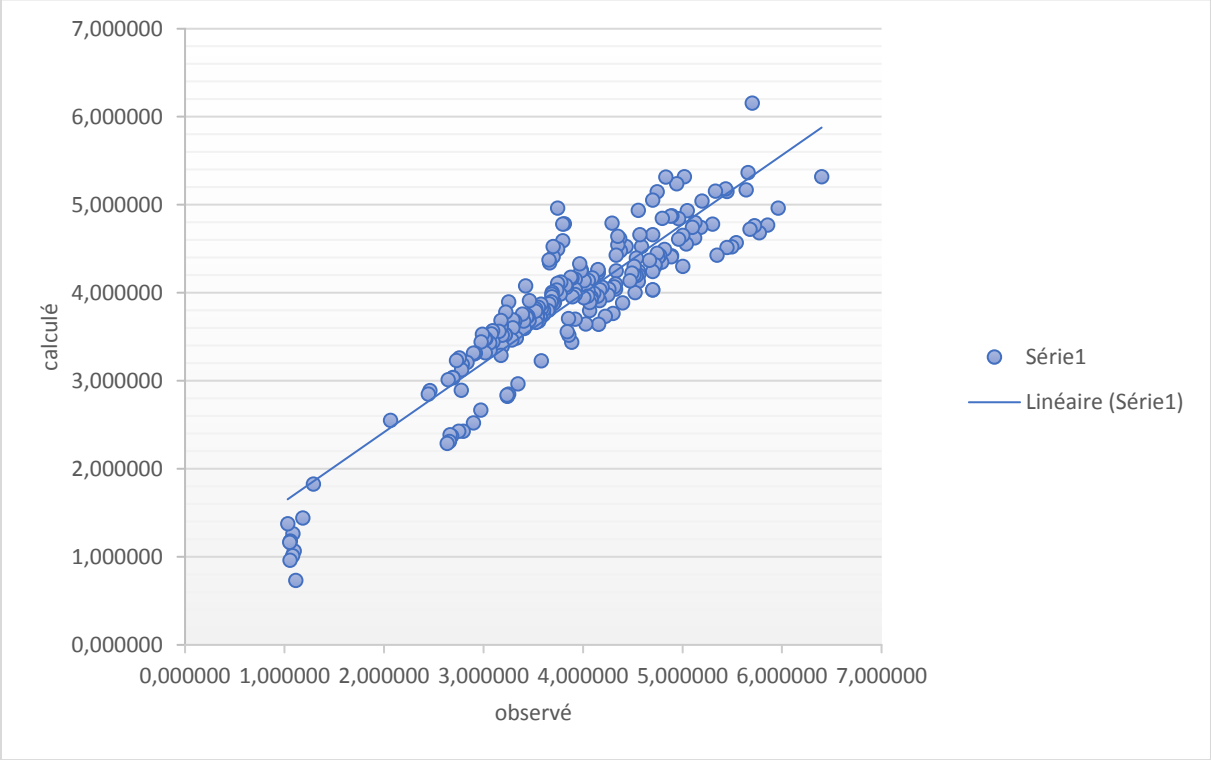


Figure III.9 : La régression linéaire de support vecteur machine

III.3.3 Analyse du modèle de régression de support vecteur machine

Dans cet article, le noyau de la fonction de base radiale (RBF) a été utilisé tel qu'il a été trouvé pour donner une bonne performance générale et qu'un seul paramètre noyau (γ) doit être ajusté (Zaidi, 2015). La méthodologie de recherche de grille avec validation croisée 10 a été appliquée pour obtenir les paramètres du modèle optimal, C, ε, γ. Ces paramètres ont été variés dans une large gamme comme : C [10, 100], γ [0.01, 2.5] et ε [0.01, 0.5].

Tableau 10 : Paramètres du modèle SVM pour la solubilité de quelques solutés solides dans le CO_2 supercritique.

Model	C	γ	Ɛ	Type Kernel	Quantité de vecteurs supports
La solubilité	10,000	0,100	0,100	Radial basis Fonction	279

III.3.4 Comparaison entre RNA et SVM

Tableau 11 : Évaluation de SVM model par rapport au modèle MLR pour les données de test.

	MSE(Train)	**MSE(Test)**	**MSE(val)**	**R^2(Train)**	**R^2(Test)**	**R^2 (val)**
RNA	0.0476	0.0472	0.0325	0.99571	0.99587	0.99709
SVM	0.162	0.174	0.165	0.922	0.911	0.919

Une comparaison a été faite de la performance du modèle MLR avec le modèle de régression de support vecteur machine Tableau 11 présente les paramètres d'évaluation statistique pertinents pour les deux montre clairement la performance de modèle de régression de réseaux de neurones artificiels. On voit ainsi que le MLR dépasse le modèle SVM.

III.3.5 Conclusion

Dans notre cas, Sur la base des paramètres d'évaluation statistique et de la comparaison avec Modèle MLR, les résultats prouvent que le modèle de calcul de régression de support à vecteur machine est un modèle inferieure au modèle MLR en termes de performance de prédiction, de précision et capacité de généralisation. Ainsi, SVM peut être appliqué dans le génie chimique et ses domaines connexes.

Conclusion générale

La connaissance de la solubilité des solutés solides dans les phases supercritiques est importante pour le design des procédés de séparation. De plus la disponibilité des résultats expérimentaux de solubilité des nouveaux composés intéressants de point de vue pharmaceutique nous offre l'opportunité de modéliser et voir les capacités des modèles existants à prédire cette propriété dans des conditions opératoires bien déterminées.

Les réseaux de neurones et support vecteur machine ont été comparés pour la prédiction de la solubilité de 25 composés différents dans le dioxyde de carbone supercritique. Les SVM fournissent une approche très intéressante de l'approximation statistique. Souvent, le nombre des exemples pour l'apprentissage est insuffisant pour que les estimateurs fournissent un modèle avec une bonne précision. Le réseau neuronal est un puissant modèle de données informatiques capable de capturer et de représenter des relations d'entrées/sorties complexes. La motivation pour le développement de la technologie des réseaux de neurones découle du désir de développer un système artificiel capable de réaliser des tâches « intelligentes » semblables à celles réalisées par le cerveau humain.

La base de données utilisée dans cette étude a été recueillis à partir de la littérature publiée au cours des 20 derniers années. La base contient les données expérimentales d'un ensemble de 25 solutés solides représenté par un ensemble de 895 points expérimentaux. Les résultats de la modélisation neuronale montrent que les RNA, comme il était montré dans plusieurs études passés peut-être utiliser avec fiabilité pour représenter les données expérimentales de la solubilité des solutés solides dans le CO_2 supercritique.

Les résultats de la première contribution qui a été fait pour déterminer l'effet des propriétés thermodynamiques avec les descripteurs moléculaire sur la prédiction de la solubilité montre que les descripteurs ont une influence sur les modèles des réseaux de neurones artificiels. Le RNA qui comporte les propriétés thermodynamiques avec tous les descripteurs qu'on 'a cités avant, a donné une meilleure corrélation (R^2=**0,99692**) parmi les sept réseaux proposés et une meilleure valeur de l'erreur absolue relative moyenne (MAARD=**1,4223**%).

Les résultats de la deuxième contribution qui a été fait pour la performance du modèle de régression de support vecteur machine. Les résultats prouvent que le modèle de calcul de régression de MLR est un modèle supérieur au modèle support vecteur machine en termes de performance de prédiction, de précision et capacité de généralisation. Ainsi, MLR peut être appliqué dans le génie chimique et ses domaines connexes.

REFERENCES BIBLIOGRAPHIQUES

REFERENCES BIBLIOGRAPHIQUES

Abdallah El Hadj A. (2009). "Prédiction de la solubilité dans les phases supercritiques à partir des équations basées sur la théorie associative", Univ Dr. Yahia Fares Médéa.

Anderson J. A. et Rosenfeld E. (1988). "Neuro Computing Foundations of Research", MIT Press, Cambridge.

Asghari-Khiavi M. et Yamini Y. (2003). "Solubility of the Drugs Bisacodyl, Methimazole, Methylparaben, and Iodoquinol in Supercritical Carbon Dioxide" J. Chem. Eng.48, 61-65.

Asiabi H., Yamini Y., Latifeh F et Vatanara A. (2015). "Solubilities of four macrolide antibiotics in supercritical carbon dioxide and their correlations using semi-empirical models". J. Supercrit. Fluids 104, 62–69.

Asiabia H., Yaminia Y., Tayyebia M., Moradib M., Vatanarac A. et Keshmirid K. (2013). "Measurement and correlation of the solubility of two steroid drugs in supercritical carbon dioxide using semi empirical models " J. of Supercritical Fluids 78,28– 33.

Bakhashwain M.J. (2016). **"**Prediction of global solar radiation using support vector machines", International J. of Green Energy, 13:14, 1467-1472.

Bishop C. (1995). "Neural networks for pattern recognition", Oxford University Press.

Bishop C.M. (1995). "Training with noise is equivalent to Tikhonov regularization". Neural Computation 7 (1), 108-116.

Coelho P.J., Naydenov P.G., Yankov S.D. et Stateva P.R. (1959)."Experimental Measurements and Correlation of the Solubility of Three Primary Amides in Supercritical CO2: Acetanilide, Propenamide, and Butanamide" J. of chemical & engineering.

Dennis J.E., Jr et Schnabel Robert B. (1983). "Numerical Methods for Unconstrained Optimization and Nonlinear Equations. Prentice-Hall", New York.

Dorronsoro., Ginel J., Sanchez F.C. et Cruz C. (1997). "Neural Fraud Detection in Credit Card Operations". IEEE Transactions on Neural Networks, 8; 827-834.

Fakhr-Eddine K., Cabassud M., Le Lann M.V. et Couderc J.P. (1999). "Application des réseaux de neurones à la modélisation des réacteurs de LPCVD" : dépôt de silicium dopé in situ au bore, J. of Chemical Engineering, 72, pp. 171-182,

Fukushima K. (1986). "A neural network model for selective attention in visual pattern recognition". Biological Cybernetics, 55(1), 5-15.

Furusjo E. et Danielsson L.-G. (1998).Anal. Chim. Acta 373, pp. 83– 94.

Garmroodi A., Hassan J. et Yamini Y., (2004). "Solubilities of the Drugs Benzocaine, Metronidazole Benzoate, and Naproxen in Supercritical Carbon Dioxide " J. Chem. Eng.49, 709-712.

Giddings J.C., Myers M.N., McLaren L et Keller R.A. (1968)."High-pressure gas chromatography of nonvolatile species. Compressed gas is used to cause migration of intractable solutes". Science 162 (3849), 67-73.

Grossberg S. (1988). "Nonlinear neural networks: principles, mechanisms, and architectures". Neural Networks, 1, 17–61.

Grossberg S. et McLoughlin N. (1997). "Cortical dynamics of 3-D surface perception: binocular and half-occluded scenic images". Neural Networks, 10, 1583–1605.

Haji Hosseini M., Alizadeh N. et Reza Khanchi A. (2010) " Solubility analysis of clozapine and lamotrigine in supercritical carbon dioxide using static system " J. of Supercritical Fluids 52,30–35.

Hamdache M., Khaouane L., Benkortbi O., Si Moussa C., Hanini S. et Amrane A. (2014). "Prediction of Acute Herbicide Toxicity in Rats from Quantitative Structure _Activity Relationship Modeling". Environnemental Engineering Science31(5) :243-252.

Hannay J.B., Hogarth J. (1879). "On the solubility of solids in gases. Proceedings of the Royal Society of London", 29, 324.

Hassoum M. H. (1995). "Fundamentals of artificial neural networks". Cambridge: MIT Press.

Hebb D. O. (1949). "The organisation of behavior, Wiley", New-York,

Henley E. J., Seader J. D. et Roper D.K. (2011). "Separation process Principles".

Hinton D. E., The S. J., Okihiro M. S., Cooke J.B. and Parker L.M. (1992). "Phenotypically altered hepatocyte populations in dietylnitrosamine-induced medaka liver carcinogenesis: Resistance, gouwth, and fate". Marine Environment. Res. 34, 1-5.

Hojjati M., Vatanara A., Yamini Y., Mort Moradi M. et Rouholamini Najafabadi A. (2009). "Supercritical CO2 and highly selective aromatase inhibitors: Experimental solubility and empirical data correlation" J. of Supercritical Fluids 50,203–209.

Hojjati M., Yamini Y., Khajeh M. et Vatanara A. (2006). "Solubility of some statin drugs in supercritical carbon dioxide and representing the solute solubility data with several density-based correlations", J. of Supercritical Fluids 41,187–194.

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

Hornik K. (1991). "Approximation capabilities of multilayer feedforward networks", Neural Networks, vol.4, pp. 251-257.

https://pubchem.ncbi.nlm.nih.gov (consulté le 28/01/2018).

KARELSON M. (2000). "Molecular Descriptors in Qsar/Qspr". John Wiley & Sons Inc, New York.

Kharroubi J. (2002). "Etude de techniques de classement ”Machines à vecteurs supports” pour la vérification automatique du locuteur". HAL Id: pastel-00001124.

Knutson B., Debenedetti P. et Tom J. (1996)."in Drugs and pharmaceuticals sciences series: Preparation of microparticules using supercritical fluids", Marcel Drekker, New-York, pp 89-125.

Kohonen T. (1982)."Self-organized formation of topologically correct feature maps", Biol Cybernetics, Vol. 43, pp. 59-69.

Lee, I.M. et Lee J.H. (1996)."Prediction of pile bearing capacity using artificial neural networks". Computers and Geotechnics, 18(3), 189-200.

Luque de Castro M.D., Valcarcel M. et Tena M.T. (1994)."In Analytical supercritical fluid extraction. Springer laboratory", Berlin, p 321.

Macnaughton S.J., Kikic I., Foster R.N., Alessi P., Cortesi A. et Colombo I. (1996) "Solubility of Anti-Inflammatory Drugs in Supercritical Carbon Dioxide " J. Chem. Eng.41, 1083-1086.

Mc Culloch W.S. et Pitts W. (1943). "A logical calculs of the ideas immanent in nervous activity", Bulletin of Math. Biophysics, vol.5, pp. 115-133.

McHugh M.A. et Krukonis V.J. (1986)."In Supercritical fluid extraction-Principles and practices". Butterworths, Boston, p 10.

Medina I. et Bueno L.J. (2001). "Solubilities of Zopiclone and Nimodipine in Supercritical Carbon Dioxide " J. Chem. Eng.46, 1211-1214.

Minsky M. et Papert S. (1988). "Perceptron: an introduction to computational geometry", MIT Press, expanded edition.

Nascimento G. G. F., Lacatelli J., Freitas P.C. et Silva G.L. (2000). "Antibacterial activity of plant extracts and phytochemicals on antibiotic-resistant bacteria". Braz. J. of Microbiol. 31(4): 886-891.

Nerrand P., Roussel d., urbain L., Personnaz G. et DREYFUS. (1994). "Training recurrent network: why and how? an illustration in process modeling", IEEE trans, on neural network 5 (2).

Piron E., Accominotti M. et Domard A. (1997). "Interaction between chitosan and uranyl ions role of physical and physicochemical parameters on the kinetics of sorption". Langmuir, 13: (6), 1653–1658

Pollard H.B., Guy H.R et Arispe N. (1992). "De la Fuente M, Lee G, Rojas EM, Pollard JR, Srivastava M, Zhang-Keck ZY, Merezhinskaya N, et al. Calcium channel and membrane fusion activity of synexin and other members of the Annexin gene family". Biophys J. Apr; 62(1):15–18.

Refenes A.N., Zapranis. A.S. and Francis G. (1994)."Stock Performance Modeling Using Neural Networks Comparative study with Regressive Model", Neural Networks, 7(2), 375-712.

Rosenblatt F. (1958)."The perceptron: a probalistic model for information storage and organisation in the brain", Psycological Review, vol.65, 386-408.

Rosenblatt F., The perceptron: a probalistic model for information storage and organisation in the brain, Psycological Review, vol.65, 386-408, (1958).

Rumelhart D. E., Mc Clelland. J. L. et Parallel (1986). "Distributed Processing: Exploration in the MicroStructure of Cognition", MIT Press, Cambridge.

Rumelhart D., Hinton G. et Williams R. (1986). "Learning internal representations by error propagation," Parallel Distributed Processing, Vol. 1, MIT Press, pp. 318-362.

Sejnowski T. J. and Rosenberg C.R. (1986). "Connectionist Models of Learning", in Perspectives in Memory Research and training, edited by M. S. Gazzaniga, (MIT Press).

Stahl E. et Quirin K.W. (1983)."Dense gas extraction on a laboratory scale: A survey of some recent results". Fluid Phase Equilibria 10, 269-278.

Thibault J. et Grandjean B.P.A. (1991). "Neural networks in process control "- a survey, FAC Inter. Symp. ADCHEM-9 1, Toulouse. France, 14- 16 October 199 1, pp. 295-304.

Thibault P.J. (1991). "Social semiotics as praxis: Text, social meaning making, and Nabokov's Ada. Minneapolis: University of Minnesota Press.

Weiss, S. M. et Kulikowski C.A. (1991). "Computer systems that learn". San Mateo, CA: Morgan Kaufmann.

Werbos P.J. (1986). "Generalized Information requirements of intelligence decision-making systems". SUGI 11 Proceedings, SAS Institute, Cary, NC. A version updated later in 1986 is available from the author.

White D.A. et Sofge, D.A. (1992). (eds):"Handbok og Intelligent Control". Van Nostrand Reinhold, New York.

White H. (1992). "artificial neural networks", Blackwell, New York.

Widrow B. and Hoff D.E. (1960)."Adaptative switching circuits, IRE Western Electric Show and convention Recor", vol.4, pp. 96-104.

Widrow B. et Hoff D.E. (1960)."adaptative switching circuits, IRE Western Electric Show and convention Record", vol.4, pp. 96-104.

WikiStat " Machines à vecteurs supports ".

www.scbdd.com (consulté le 05/02/2018).

Yamini Y., Hassan J. et Haghgo S. (2001). "Solubilities of Some Nitrogen-Containing Drugs in Supercritical Carbon Dioxide " J. Chem. Eng.46, 451-455.

Ziemons E. (2006)."Etude en temps réel du processus d'extraction de la tagitinine c en fonction des caractéristiques physico-chimiques du co_2 supercritique à l'aide de fibres optiques couplant un spectrophotomètre irtf a un extracteur a fluide supercritique", Faculté de Médecine, Laboratoire de Chimie Analytique, France.s