

**REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET
POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA
RECHERCHE SCIENTIFIQUE
Ecole Nationale Polytechnique
Département de Génie Chimique**



**Projet de Fin d'Etudes
En vue de l'obtention du Diplôme d'Ingénieur d'Etat en Génie Chimique
Thème:**

**Simulation des procédés d'extraction
et de séparation par les réseaux de
neurones artificiels**

Proposé et Dirigé par :

Pr E.H.BENYOUSSEF

Présenté par :

**KEDDOUS Samer
HAOUI Imad Eddine**

Promotion : 2008

ملخص :

محاكاة ظواهر الاستخراج والفصل بواسطة الشبكات العصبونية الصناعية .
إن تجسيد عملية الاستخراج يتطلب حل العديد من المسائل أهمها البحث عن نموذج رياضي ممثل و حساب بعض المعاملات لهذا العديد من الباحثين حاولوا من خلال أعمالهم نمذجة ظاهرة انتقال المادة صلب- سائل بالاعتماد على العلاقات الكلاسيكية المحددة للظاهرة وبالإستعانة بفرضيات لاخترال التعقيد.
عملنا هذا يهدف إلى محاكاة و نمذجة ظاهرة انتقال المادة صلب-سائل وهذا بإتباع طريقة جديدة تدعى "الذكاء الصناعي" أو " الشبكات العصبونية الصناعية " وهذا للحصول على المعلومات الضرورية لحساب كلا من معامل الانتشار و انتقال المادة.
الكلمات المفتاحية : عملية الاستخراج صلب-سائل ، حبات الكسبرة ، الشبكات العصبونية الصناعية،معامل الانتشار،معامل انتقال المادة.

Abstract:

Simulation of the extraction and separation processes by artificial neural networks

The implementation of the extraction requires the resolution of the delicate problems such as the modelling and the determination of certain parameters, for that several authors tried to model the transfer of solid matter liquid while basing themselves on the traditional relations governing the phenomenon and while being based on simplifying assumptions to simplify the complexity of these relations.

In this context, our work aims at the simulation and the modelling of transfer of solid matter liquid while following a new approach called by "artificial intelligence" or " *artificial neural networks* " in order to collect the data necessary for the determination of coefficient of diffusion and that of transfer of matter.

Keys words: solide-liquide extraction, coriander seeds, diffusion coefficient, coefficient of transfer of matter

Résumé :

Simulation des procédés d'extraction et séparation par réseaux de neurones artificiels

La mise en oeuvre de l'extraction nécessite la résolution des problèmes délicats tels que la modélisation et la détermination de certains paramètres, pour cela plusieurs auteurs ont tenté de modéliser le transfert de matière solide liquide en se basant sur les relations classiques régissant le phénomène et en s'appuyant sur des hypothèses simplificatrices pour simplifier la complexité de ces relations.

Dans ce contexte, notre travail a pour objectif la simulation et la modélisation de transfert de matière solide liquide en suivant une nouvelle approche dite par « intelligence artificielle » ou

par « réseaux de neurones » afin de recueillir les données nécessaires pour la détermination de coefficient de diffusion et celui de transfert de matière.

Mots clés : extraction solide-liquide, grains de coriandre, réseaux de neurones artificiels, coefficient de diffusion, coefficient de transfert de matière.

Ce travail a été réalisé dans le cadre des activités de l'équipe « Valorisation des espèces végétales » du laboratoire valorisation des énergies fossiles sous la direction de professeur C.E CHITOUR.

Examiné par Le jury :

Présidente : F.SOUAHI

Directeur de thèse : E.H BENYOUSSEF

Examinatrice : R.DERRICH

Examineur : A.SELATNIA

Maître de conférence à l'ENP

Professeur à l'ENP

Professeur à l'ENP

Professeur à l'ENP

REMERCIEMENTS

Ce travail a été réalisé au département de génie chimique à l'école nationale polytechnique d'Alger sous la direction de Monsieur Le professeur E.H BENYOUSSEF.

Avant de l'aborder, il nous est particulièrement agréable d'adresser nos vifs remerciements à Monsieur le professeur E.H BENYOUSSEF et lui exprimer notre gratitude pour la sollicitude avec laquelle il nous a guidé et encouragé, pour les conseils pertinents et pour son dévouement permanent et inlassable dont il a fait preuve durant l'élaboration de ce travail.

Nous exprimons nos respectueux remerciements à Madame F .SOUAHI maître de conférence pour l'honneur qu'elle nous fait de présider le jury.

Que Madame R.DERRICH, Professeur à l'école nationale polytechnique,

Que Monsieur A.SELATNIA, professeur à l'école nationale polytechnique,

Trouvent ici l'expression de nos vifs remerciements pour avoir accepté de juger ce travail.

Aussi, nous tenons à remercier Monsieur S.FEKRACHE magister à l'école nationale polytechnique pour son aide et sa disponibilité.

Dédicace

À celle qui m'a donné un jour la vie et la lumière, et qui a beaucoup souffert pour que je sois ainsi, à ma chère mère comme témoignage de mon affection et de mon profond amour.

À mon père que dieu me le garde éternellement heureux,

À ma grand-mère qui ne cesse jamais de me prier et de m'encourager.

À mon grand frère Khalifa, qu'il m'a beaucoup aidé pour que je puisse continuer mes études.

À ma grande sœur «Nawel» et son mari «Chokri».

À mes sœurs et mon frère «Leila, Mona, Aicha, Farid»

À mon cher ami Salim

À tout mes amis

À tous

Je dédié ce modeste travail.

Imad

Dédicaces

*A celle qui a beaucoup sacrifié pour que je sois ainsi, à ma
chère mère.*

*A celui qui a été toujours derrière moi, me soutient et
m'encourage, au quel je dois ce que je suis aujourd'hui, à mon
père.*

*A Hanan, Linda, Hayette et Samiha mes exceptionnelles
sœurs, qu'elles trouvent ici le témoignage de ma gratitude
pour tous les sacrifices qu'ils m'ont apporté.*

*A mes frères Rachid, Hilal et Mahdi aux quels je passe à
travers ce travail mes vifs remerciements.*

*A celui qui ne cesse de me soutenir, de me donner la force et le
courage, à Salim.*

*A mon maître, Ahcen qu'il trouve ici l'expression de mes
remerciements profonds.*

A tous qui m'ont aidé à résister autant là bas.

Samer.

Liste des symboles utilisés

φ : Fonction d'activation	
x_i : Les entrées du neurone	
w_{ij} : Le paramètre de pondération	
y_k : La sortie du neurone	
θ_k : Le seuil du neurone	
th : Fonction tangente hyperbolique	
\exp : Fonction exponentielle	
E : Fonction d'énergie	
H : Matrice hessienne	
MSE : L'erreur quadratique moyen	
t : Le temps	
t_∞ : Le temps d'équilibre	
$C_s(0)$: Concentration du solide	
$C_s(t)$: Concentration du soluté dans le solide	(Kg/Kg)
D : Coefficient de diffusion	(m ² /s)
V_l : Volume du solvant	(ml)
m : Débit massique du soluté	(Kg/s)
A : la surface d'échange	(m ²)
k : Coefficient de transfert de matière	(s ⁻¹)
$C_l(t)$: La concentration du soluté dans le liquide	(Kg/Kg)
$C_l(\infty)$: La concentration du soluté dans le liquide à l'équilibre	(Kg/Kg)
α : Rapport des quantités de soluté dans le solide et dans le liquide à l'équilibre	
P : coefficient de partage	
q_n : les solutions caractéristiques de l'équation	
R : le rayon d'un grain de coriandre	(mm)
d : diamètre d'un grain de coriandre	(mm)
w : la vitesse d'agitation	(tr/min)
h : taux d'humidité	(Kg/m ³)
vm : L'hydromodule	
T : la température	(C°)

Liste des figures

Figure I.1 : un neurone formel

Figure I. 2 : Perceptron multicouches

Figure I.3 : réseau feedback

Figure II.1: neurone à fonction radiale.

Figure II.2 : carte de Kohonen

Figure II.3:Réseau de Hopfield

FigureII.4 : la séparabilité linéaire du Perceptron

FigureII.5 : le tableau de vérité et la représentation du XOR.

FigureII.6 : le Perceptron multicouches

Figure II.7 : la représentation de XOR par PMC

Figure. IV01 : l'évolution de l'erreur quadratique moyenne (MSE) lors de l'apprentissage en 150 itérations pour le réseau de 7 neurones cachés.

FigureIV02: prédiction de la concentration de l'expérience 6.

FigureIV03: prédiction de la concentration de l'expérience 3.

Figure IV05 : l'évolution de l'erreur quadratique moyenne (MSE) lors de l'apprentissage en 150 itérations pour le réseau de 7 neurones cachés.

FigureIV.6 : simulation de la courbe cinétique de l'expérience 11.

FigureIV.7 : simulation de la courbe cinétique de l'expérience 12.

Figure IV.5 : évolution de l'erreur quadratique moyenne (MSE) lors de l'apprentissage du réseau 1 [0, 90 min].

Figure IV6 : évolution de l'erreur quadratique moyenne (MSE) lors de l'apprentissage du réseau2 [95, 240 min].

FigureIV.10 : simulation de la courbe cinétique de l'expérience 4.

FigureIV.11 : simulation de la courbe cinétique de l'expérience 1.

Figure IV.12: la courbe obtenue en traçant l'équation (IV.5) pour l'expérience 7 avec $P=0.25$

Figure IV.13: la courbe obtenue en traçant l'équation (IV.5) pour l'expérience 7 avec $P=0.5$

Figure IV.14: la courbe obtenue en traçant l'équation (IV.5) pour l'expérience 7 avec $P=0.75$

Figure IV.15 : la courbe obtenue en traçant l'équation (IV.5) pour l'expérience 7 avec $P=1$.

Figure IV.16 : la courbe obtenue en traçant l'équation (IV.5) pour l'expérience 7 avec $P=1.25$

Figure IV.12 : graphe représentant $-\ln\left(1 - \frac{C_t(t)}{C_t(\infty)}\right) = f(t)$ de l'expérience 6

Figure IV.13 : graphe représentant $-\ln\left(1 - \frac{C_t(t)}{C_t(\infty)}\right) = f(t)$ de l'expérience 3

Liste des Tableaux

Tableau II.1 : Les principaux réseaux et leurs domaines d'application

Table III.1 : Les solutions de $\tan q_n = \frac{3q_n}{3 + \alpha q_n^2}$

Tableau IV.1 : les valeurs de concentration pour les vitesses d'agitation $w=500$ tr/min et 350tr/min

Tableau IV.2 : les valeurs de concentration pour différentes valeurs de l'hydromodule

Tableau IV.3 : les valeurs de concentration pour différentes valeurs du taux d'humidité

Tableau IV.4 : les valeurs de concentration pour différentes valeurs du diamètre.

Tableau IV.5 : les conditions opératoires des expériences utilisées

Tableau IV.6 : les écarts entre les valeurs expérimentales et celles prédites par le réseau des expériences 1 à 5

Tableau IV.7: les écarts entre les valeurs expérimentales et celles prédites par le réseau des expériences 6 à10

Tableau IV.8 : les résultats obtenus pour chaque expérience et pour chaque valeur de Coefficient de partage p..

Tableau IV.9 : les coefficients de diffusion obtenus pour chaque expérience.

Tableau IV.10 : les couples (C_{inf}, t_{inf}) pour chaque expérience

Tableau IV.11 : les résultats obtenus par le modèle rapporté par Crank, en se basant sur les valeurs simulées par le réseau.

Tableau IV.12 : les résultats obtenus par le modèle basé sur le coefficient de transfert de matière.

Sommaire

Introduction générale01

CHAPITRE I : Les réseaux de neurones une nouvelle approche de modélisation

I. 1 Introduction02

I. 2. Les RNA, une nouvelle approche de modélisation..... 02

I. 2. 1. Principe de base03

I. 2. 2. Propriété fondamentale.....03

I. 2. 2. 1. Approximation universelle04

I. 2. 2. 2. Parcimonie04

I. 3 Fondements théoriques05

I. 3. 1 La fonction d'activation du neurone.....06

I. 3. 1. 1 Exemples de fonction d'activation06

I. 3. 2 Réseaux de neurones07

I. 3.2.1. Les réseaux de neurones non bouclés.....07

I. 3. 2. 2 Les réseaux de neurones bouclés.....08

I. 3. 3 L'apprentissage des réseaux de neurones08

I. 3. 3.1 L'apprentissage supervisé.....09

I. 3. 3.2 L'apprentissage non supervisé.....09

I. 3.3.3 Les règles d'apprentissage.....09

I. 4 Conclusion.....10

CHAPITRE II : Les principaux modèles de réseaux de neurones

II. 1 Les réseaux FEED-FORWARD 11

II. 1. 1 Le Perceptron 11

➤ Le Perceptron monocouche (simple).....11

➤ Le Perceptron multicouches (PMC).....11

II. 1. 2 réseaux à fonction radiale.....11

II. 2. Les réseaux FEED-BACK.....12

II. 2. 1 Les cartes auto-organisatrices de Kohonen.....12

II. 2. 2 Les mémoires associatives de Hopfield12

II.3. Comparaison entre PMC et RBF13

II. 3.1. Les principales différences.....14

II. 3.1.1. Selon leurs fonctions d'activation	14
II. 3.1.2. Selon la règle d'apprentissage	14
II. 4. Le Perceptron.....	15
II. 4. 1 Le Perceptron simple.....	15
II. 4. 2 Le Perceptron multicouches	16
II. 4. 3 L'apprentissage du PMC.....	17
II. 4.3.1 Le rétro propagation de l'erreur	17
II. 4.3.1.1 la rétropropagation avec Momentum	18
II. 4.3.1.2 Propagation à taux d'apprentissage adaptatif.....	18
II. 4.3.1.3 limites du rétro propagation	18
II. 4. 3. 2 Méthode de Levenberge-Marquardt	19
II. 5. Conclusion	20

CHAPITRE III : Etude cinétique et modélisation de l'extraction solide liquide

III .1.Généralités sur l'extraction solide liquide	21
III.1.1.Définition	21
III .1.2 Performances industrielles de l'extraction solide- liquide.....	22
III.1.2.1. Débit.....	22
III.1.2.2. Rendement	22
III.1.2.3. Sélectivité.....	22
III.1.2.4.Concentration de l'extrait	22
III.1.3. Facteurs intervenant dans l'extraction solide liquide.....	23
III.1.3.1. Nature et état du solide et du soluté.....	23
III.1.3.2. Nature du solvant	23
III.1.3.3. Température	23
III.1.3.4. Degré d'agitation	24
III.1.3.5. Taux d'humidité du solide	24
III.1.3.6. Durée d'extraction	24
III.1.3.7. Hydromodule	24
III.2. Modélisation du procédé d'extraction	24
III.2.1 Modèle rapporté par crank.....	25
III.2.1.1.Diffusion dans une sphère	25

III.2.1.2. Diffusion dans une sphère plongée dans une solution parfaitement agitée.....	25
III.2.2 Modèle base sur le coefficient de transfert matière	27
III.2.2.1 Présentation du modèle	27
III.2.2.2 Hypothèses simplificatrices	28
III.2.2.3 Développement mathématique du modèle.....	28
III. 3. Conclusion.....	29

CHAPITRE IV : Simulation et interprétations

IV. 1 Construction de base d'apprentissage	30
IV.1.1 Influence de la vitesse d'agitation ω	31
IV.1.2 Influence de la température	31
IV.1.3 Le taux d'humidité.....	32
IV.1.4 Le taux de l'hydromodule.....	34
IV.1.5 Influence du diamètre des grains	35
IV. 2 La détermination du modèle neuronal optimal	37
IV. 2. 1 Choix du nombre de couches cachées du PMC.....	37
IV. 2. 2 Choix de nombre de neurones cachés.....	38
IV. 2. 2. 1 Approche 1.....	38
Réseau entraîné à 120 min	38
Réseau entraîné jusqu'à 180 min	40
IV. 2. 2. 2 Approche 2.....	44
IV.3 Détermination de coefficient de diffusion.....	47
IV .3. 1 Modèle de Crank.....	47
IV.3.1.1 Sur la base des résultats expérimentaux	47
IV.3.1.2 Sur la base des résultats simulés par le réseau.....	52
IV.4 calcul de coefficient de transfert de matière	54
Conclusion générale	56

Bibliographie

ANNEXES

ANNEXE 1 : généralités sur les réseaux de neurones

ANNEXE 2 : les règles d'apprentissage

ANNEXE 3 : La coriandre

ANNEXE 4 : les propriétés des réseaux établis

Introduction générale

Lorsque apparaît une nouvelle technique, l'ingénieur se demande naturellement en quoi cette nouveauté peut lui être utile. Si elle est dotée d'un nom plus métaphorique que technique - ce qui est évidemment le cas pour les réseaux de neurones.

Depuis quelques années, on observe l'augmentation de l'utilisation des réseaux de neurones artificiels dans tous les domaines, cela à cause de sa capacité et sa précision de résoudre des problèmes complexes et difficiles à décrire sous forme analytique, qui est le cas du transfert de matière solide liquide.

Pour cela et à fin d'illustrer l'applicabilité de cet outil neuronal, nous allons tenter de modéliser et simuler de processus de transfert de matière solide liquide dans le cas de l'extraction des essences des graines de coriandre par l'éthanol.

Le présente travail est constitué de quatre chapitres dont le premier a pour objectif de répondre à la question : pourquoi les réseaux de neurones ? Et ce, par la présentation de ses avantages, ses propriétés et leur fondement théorique.

Le second chapitre illustre les différents types d'associations de neurones pour arriver à un choix d'une architecture adéquate à notre cas.

Afin de bien concevoir le modèle neuronal, une étude des phénomènes de transfert solide - fluide est présentée dans le troisième chapitre. Dans cette partie, nous nous intéresserons aux modèles fondamentaux rapportés dans la bibliographie et qui permettent d'atteindre les grandeurs tels que les coefficients de diffusion et de transfert de matière.

Enfin, dans le dernier chapitre, nous présentons les démarches suivies et menant à la détermination du réseau optimal, les résultats de la simulation et la détermination de la concentration à l'équilibre, le coefficient de diffusion et le coefficient de transfert de matière.

Chapitre I:

Les réseaux de neurones
une nouvelle approche de
modélisation

I. 1 Introduction

La *modélisation* est la simplification d'un phénomène réel par établissement d'un ensemble d'équations reliant les différents paramètres du phénomène entre eux.

Gérard Dreyfus [1] donne la définition suivante au modèle « Un modèle est une représentation de la réalité visible ou observable » et selon Jean Perrin, d'après [1], modéliser, c'est « remplacer le visible compliqué par l'invisible simple ».

On peut classer les modèles en trois catégories : les modèles de connaissance, les modèles « boîte noire », et les modèles semi -physiques. [1]

- Les modèles de connaissance consistent à mettre en oeuvre des équations décrivant les phénomènes étudiés (biologiques, physiques, chimiques...) en se basant sur des connaissances théoriques relatives à ces derniers. Ils ont une valeur prédictive et explicative.
- Par contre, les modèles boîte noire, qui ne nécessitent que la possession des mesures, n'ont qu'une valeur prédictive sur un domaine de validation.
- Entre ces deux types de modèles se situent les modèles semi physiques. C'est l'élaboration des équations basées sur une théorie, et des équations purement empiriques qui résultent d'une modélisation de type boîte noire. [1]

I. 2. Les réseaux de neurones artificiels, une nouvelle approche de modélisation

Au début des années quatre-vingt, plusieurs recherches se sont orientées vers le développement de nouvelles techniques basées sur l'approximation de fonctions. Parmi ces techniques figurent les réseaux de neurones dont le domaine d'application ne cesse de s'étendre. [1]

En effet, les réseaux de neurones artificiels (RNA) habituellement utilisés depuis une dizaine d'années, à côté des méthodes classiques de modélisation [2]. Il existe plusieurs raisons qui ont conduit à s'intéresser à cet outil :

- Contrairement aux méthodes statistiques classiques, les réseaux de neurones artificiels ne nécessitent aucune hypothèse sur les variables.
- Ils sont tout à fait adaptés pour traiter des problèmes complexes non structurés, c'est-à-dire des problèmes sur lesquels il est impossible a priori de spécifier la forme des relations entre les variables utilisées. C'est grâce à des algorithmes, que ces systèmes apprennent par eux-mêmes les relations entre les variables à partir d'un ensemble de

données, à l'image de ce que ferait le cerveau humain. Ainsi, le réseau se paramètre de lui-même à partir des exemples qui lui sont fournis [1].

- Il ne répète pas bêtement le passé et s'il existe effectivement une relation de cause à effet entre les descriptions en entrée et les valeurs à prédire, le réseau l'identifiera. [1]
- Le réseau de neurones est robuste. Il n'est pas handicapé par quelques exemples bruités ou faux : ceux-ci sont écartés car ils sont incohérents avec le reste. Les valeurs manquantes sont également gérées habilement et ne nuisent pas à la construction du modèle [1].
- L'arrêt du processus itératif lorsque le réseau produit les meilleurs résultats sur l'échantillon de validation lui confère une certaine robustesse. On peut en effet considérer que de cette manière, seule l'information pertinente est intégrée dans le système. Les bruits qui pourraient concourir à la spécialisation du réseau ne sont pas intégrés [1].

I.2.1. Principe de base :

L'idée principale des réseaux de neurones consiste à modéliser une grandeur mesurable (physique, chimique, économique, ...), dont on soupçonne qu'elle dépend d'un ou plusieurs facteurs qui constituent les composantes du vecteur d'entrée.

L'objectif de la modélisation par les réseaux de neurones est de trouver, à partir des mesures disponibles, une relation, si elle existe, entre les variables d'entrée et la grandeur à modéliser. Le cadre conceptuel adapté à cette tâche est celui des statistiques [1].

En effet, le réseau n'a aucune idée a priori sur le modèle : on choisit une forme d'équation aussi générale que possible, et l'on ajuste les paramètres de cette équation de manière à lui conférer la meilleure capacité de généralisation (estimation des valeurs dans des situations qu'il n'a pas rencontrées) sans s'intéresser particulièrement aux valeurs ou au sens de ces paramètres. On parle alors de modélisation « boîte noire » [1].

I.2.2. Propriété fondamentale :

Une propriété fondamentale qui distingue un type particulier de réseaux de neurones est l'approximation universelle parcimonieuse. Cette expression traduit deux propriétés distinctes : d'une part les réseaux de neurones sont des approximations universels, et d'autre part, une approximation à l'aide de réseaux de neurones nécessite, en général, moins de paramètres ajustables que les approximations usuels. [1]

I.2.2.1. Approximation universelle :

Toute fonction bornée suffisamment régulière peut être approchée uniformément, avec une précision arbitraire, dans un domaine fini de l'espace de ses variables, par un réseau de neurones avec une certaine architecture [1].

Ainsi, on dit que les réseaux de neurones sont des approximateurs universels. Mais, cette propriété n'est pas spécifique aux réseaux de neurones, puisque les polynômes jouissent aussi de cette qualité. Par contre, dans la pratique, on fait toujours en sorte que le nombre de paramètres ajustables soit le plus faible possible. On cherche ainsi, l'approximation la plus parcimonieuse.

I.2.2.2. Parcimonie :

Pour obtenir un modèle non linéaire de précision donnée, un réseau de neurones a besoin de moins de paramètres ajustables. Or le nombre de données nécessaires pour ajuster le modèle est directement lié au nombre de ses paramètres.

En ce sens, le nombre de paramètres, pour une précision donnée, croît exponentiellement avec le nombre de variables dans le cas des approximateurs linéaires par rapport à leurs paramètres (polynômes), alors qu'il croît linéairement avec ce nombre pour les approximateurs non linéaires par rapport à leurs paramètres (réseaux de neurones) [1].

Mais en quoi cette propriété est-elle avantageuse ?

Au cours de l'apprentissage à partir d'exemples, les paramètres du réseau sont ajustés pour que celui-ci reproduise la relation non linéaire qui existe entre les entrées et les sorties. Or il est clair que le nombre d'exemples nécessaires pour ajuster les paramètres de manière significative doit être grand devant le nombre de ces paramètres. Les réseaux parcimonieux en paramètres sont donc également parcimonieux en nombre d'exemples ; or ceux-ci, très souvent, sont coûteux (notamment si ce sont des résultats d'expériences sur un processus industriel), et leur collecte peut être longue [1].

L'avantage pratique des réseaux de neurones par rapport aux techniques classiques de modélisation **non linéaire** réside donc dans leur capacité à réaliser des modèles de précision équivalente avec moins de données expérimentales (ou dans leur capacité à constituer des modèles plus précis à partir du même nombre d'exemples). De manière générale, les réseaux de neurones permettent de tirer le meilleur parti des données numériques disponibles, pour construire des modèles. [1]

En résumé :

Les réseaux de neurones sont donc un outil puissant d'analyse qui permet d'approximer tout ce qui a un sens et de construire, par la suite, un modèle de prévision à la fois robuste et flexible.

Avant de mettre à profit cette propriété puissante, nous allons, à travers ce qui va suivre, présenter le principe de fonctionnement de cette technique.

I. 3 Fondements théoriques :

Les réseaux de neurones artificiels sont à l'origine une tentative de modélisation mathématique du cerveau humain (cf. *Annexe 1*). Les premiers travaux datent de 1943 et sont l'œuvre de *Mac Culloch et Pitts*. Ils ont présenté un modèle assez simple pour les neurones et ont proposé la définition suivante : « Un neurone formel est un automate qui fait une somme pondérée des potentiels d'action qui lui parviennent (chacun de ces potentiels est une valeur numérique qui représente l'état du neurone qui l'a émis), puis s'active suivant la valeur de cette sommation pondérée. Si cette somme dépasse un certain seuil, le neurone est activé et transmet une réponse (sous forme de potentiel d'action) dont la valeur est celle de son activation. Si le neurone n'est pas activé, il ne transmet rien. » [3].

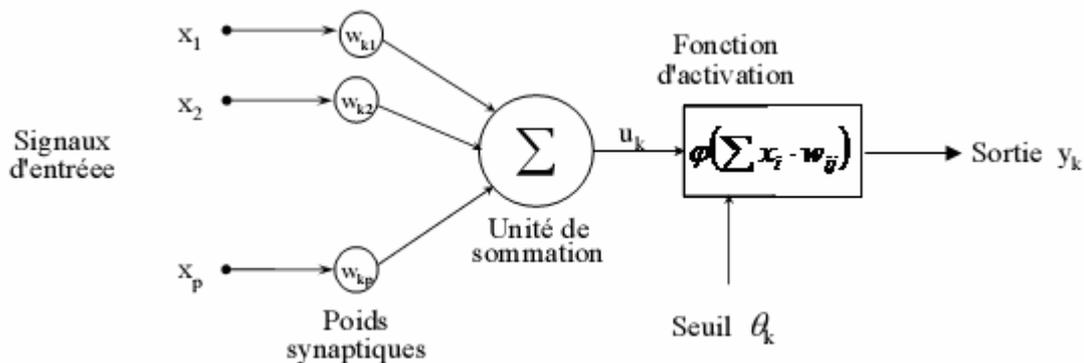


Figure I.1 : un neurone formel

Avec :

x_i : les entrées du neurone.

w_{ij} : le paramètre de pondération (le poids de la connexion qui relie le neurone i à son entrée j).

y_k : la sortie du neurone.

$\varphi(\sum x_i \cdot w_{ij})$: La fonction d'activation du neurone.

θ_k : Le seuil du neurone.

I.3. 1 La fonction d'activation du neurone :

La fonction d'activation joue un rôle prépondérant dans le comportement du neurone, et de là, du réseau entier. Toute fonction scalaire peut en principe être utilisée mais relativement peu d'entre elles sont employées en pratique.

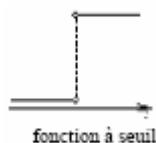
Les fonctions d'activation utilisées dans les réseaux de neurones sont variées, certains modèles de réseau emploient même plusieurs fonctions dans un même réseau, selon le rôle du neurone. Malgré cela un ensemble de caractéristiques est souvent trouvé dans les modèles de réseaux de neurones [4] :

- *Monotonie* : f est en général une fonction monotone croissante.
- *Seuillage* : f peut posséder une valeur en deçà de laquelle sa valeur est négligeable.
- *Saturation* : f peut posséder une valeur maximale au-dessus de laquelle sa valeur de réponse est essentiellement fixe. Ceci permet d'éviter de propager des grandes valeurs dans le réseau.
- *Dérivabilité* : f est souvent dérivable, ce qui est nécessaire dans certains modèles, surtout dans la définition des fonctions d'apprentissage.

I.3. 1. 1 Exemples de fonction d'activation [4] :

Fonction seuil : Net_i représente la somme pondérée des entrées des neurones.

$$a_i = \begin{cases} 1 & \text{si } Net_i \geq \theta_i \\ 0 & \text{autrement} \end{cases} \quad (\text{I.1})$$



Le seuillage introduit une non linéarité dans le comportement du neurone, cependant, il limite la gamme de réponse à deux valeurs.

Fonction linéaire bornée :

$$a_i = \begin{cases} Max_i, & Net_i > \theta_i \\ Net_i, & \theta_i \leq Net_i \leq \theta_i \\ Min_i, & Net_i < \theta_i \end{cases} \quad (\text{I.2})$$



Cette fonction représente un compromis entre la fonction linéaire et la fonction seuil, entre ses deux bornes de saturation elle confère au neurone une gamme de réponse.

Fonction sigmoïde tangentielle :

$$a_i = th(Net_i) \quad (I.3)$$

Certains auteurs ont observé que l'apprentissage des réseaux multicouches converge plus rapidement quand cette fonction est employée plutôt que la fonction sigmoïde exponentielle.

La dérivée de cette fonction est aussi simple à calculer

$$\frac{d a_i}{d Net_i} = (1 - a_i^2) \quad (I.4)$$

I.3.2 Réseaux de neurones :

Un neurone formel ne réalise donc rien d'autre qu'une somme pondérée suivie d'une non linéarité. C'est l'association de tels éléments simples sous la forme de réseaux qui permet de réaliser des fonctions utiles pour des applications industrielles. On distingue deux architectures différentes de réseaux : des réseaux non bouclés et bouclés.

I.3.2.1. Les réseaux de neurones non bouclés :

On dit d'un réseau qu'il est **non bouclé** (feed forward), ou encore à sens unique, lorsqu'il ne contient pas de boucle interne, c'est à dire lorsque l'information qui le traverse ne circule que de l'entrée vers la sortie. Les neurones sont organisés en couches successives, les neurones d'une couche étant reliés à ceux de la suivante

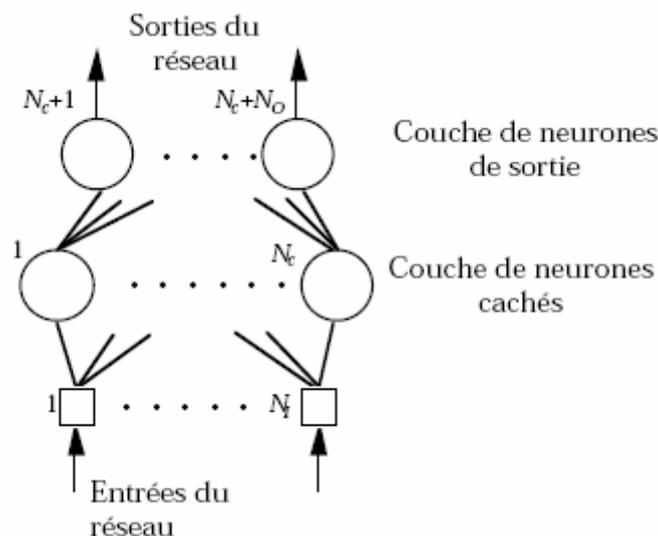


Figure I. 2 : Perceptron multicouches

Les réseaux de neurones non bouclés sont des objets statiques : si les entrées sont indépendantes du temps, les sorties le sont également. Ils sont utilisés principalement pour

effectuer des tâches d'approximation de fonction non linéaire, de classification ou de modélisation de processus statiques non linéaires [1].

I.3. 2. 3 Les réseaux de neurones bouclés (récurrents) :

Contrairement aux réseaux de neurones non bouclés, les réseaux de neurones bouclés peuvent avoir une topologie de connexions quelconque, comprenant notamment des boucles qui ramènent aux entrées la valeur d'une ou plusieurs sorties [1].

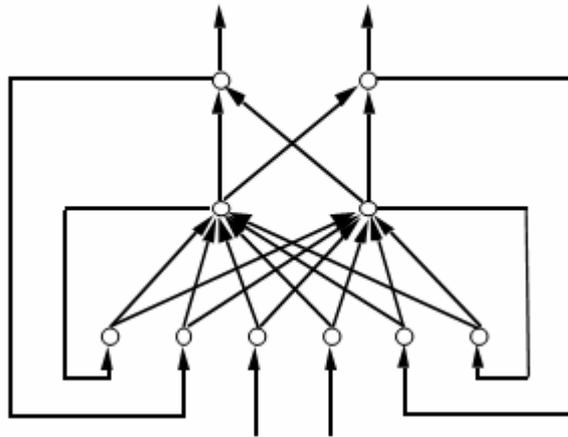


Figure I.3 : réseau feedback

Propriété :

Tout réseau de neurones bouclé, aussi complexe soit-il, peut être mis sur une forme canonique, comportant un réseau de neurone non bouclé dont certaines sorties sont ramenées aux entrées par des bouclages de retard unité [1].

I.3.3 L'apprentissage des réseaux de neurones :

La notion d'apprentissage est claire et intuitive pour les humains et les animaux : c'est une procédure cognitive qui doit faire en sorte que l'individu réalise, de manière autonome, une tâche donnée. Typiquement, cette procédure s'effectue à partir d'exemples : ainsi, pour apprendre à lire à un enfant, on lui présente des exemples de lettres et de chiffres, écrits avec des écritures et des fentes différentes. A la fin de l'apprentissage, on attend de l'enfant qu'il soit capable de lire non seulement tous les chiffres et lettres de son livre de lecture, mais également tous les chiffres et lettres qu'il est susceptible de voir : en d'autres termes, on attend de lui qu'il ait une capacité de *généralisation* à partir des exemples qui lui ont été

présentés, sans qu'il ne soit jamais nécessaire de lui fournir une description analytique et discursive de la forme et de la topologie des chiffres et des lettres [1].

Dans ce contexte on distingue deux types d'apprentissage à savoir « l'apprentissage supervisé » et « l'apprentissage non supervisé ».

I.3. 3. 1 L'apprentissage supervisé :

Il consiste à apprendre une association particulière entre deux ensembles de vecteurs. Durant une première phase, appelée phase d'apprentissage, on présente au réseau des vecteurs d'entrées en lui imposant, comme des contraintes, les vecteurs de sortie que l'on désire lui faire associer à ces entrées. Pendant une seconde phase, appelée phase de test, on présente au réseau de nouveaux vecteurs qui n'ont pas servi à l'apprentissage et l'on observe ses réponses. Les performances du réseau sont alors évaluées par le pourcentage d'erreurs.[5]

L'intérêt des réseaux neuronaux réside précisément dans leur capacité intrinsèque à la généralisation, c'est-à-dire leur aptitude à bien se comporter sur des vecteurs qu'ils n'ont pas appris [5].

I.3. 2. 2 L'apprentissage non supervisé :

Il consiste à découper l'ensemble des vecteurs d'entrées en classes d'équivalence, sans qu'il soit nécessaire de donner au réseau neuronal les noms de classes pour chaque exemple. Cette classification automatique ressemble à certaines méthodes utilisées en analyse de données.

La séparation en classes d'équivalence s'opère par mesure de ressemblance entre les vecteurs proposés. La contrainte que l'on doit imposer au réseau neuronal est le nombre de classes d'équivalence. Pendant la phase d'apprentissage, le réseau neuronal construit une topologie de l'espace des vecteurs d'entrées [5].

I.3. 2. 3 Les règles d'apprentissage :

Nous venons de voir les deux principaux types d'apprentissage. Pour chacun de ces types, il existe plusieurs règles qui se distinguent par la manière dont les changements de paramètre surviennent (*cf. annexe 2*). Les règles les plus utilisées sont :

- La règle de HEBB.
- La règle de correction des erreurs.

I.4 Conclusion :

Après avoir présenté l'approche neuronale de modélisation avec ses fondements de base, dans le chapitre suivant nous allons présenter les différents assemblages de ses neurones, comparer leurs performances, pour terminer par un choix d'architecture ainsi que l'algorithme d'apprentissage le plus fiable.

Chapitre II:

Les principaux modèles de réseaux de neurones

Comme nous l'avons précisé dans le chapitre précédent, les réseaux de neurones peuvent être classés, selon leurs architectures, en deux grandes familles : les réseaux de neurones non bouclés (FEED-FORWARD), et les réseaux bouclés (FEED-BACK).

II. 1 Les réseaux FEED-FORWARD :

II. 1. 1 Le Perceptron :

- *Le Perceptron monocouche (simple) :*

Le Perceptron proposé par Rosenblatt, représente le premier réseau de neurones artificiels. Ce modèle devait modéliser le fonctionnement de la perception humaine. Le Perceptron est un réseau simple puisqu'il ne se compose que d'une couche d'entrée et d'une couche de sortie. [5]

- *Le Perceptron multicouches (PMC) :*

C'est une extension du précédent, avec une ou plusieurs couches cachées entre l'entrée et la sortie. Chaque neurone dans une couche est connecté à tous les neurones de la couche précédente et de la couche suivante. Les fonctions d'activation utilisées dans ce type de réseaux sont principalement les fonctions sigmoïdes [1].

II. 1. 2 réseaux à fonction radiale :

Ce sont les réseaux que l'on nomme aussi RBF ("Radial Basic Functions"). L'architecture est la même que pour les PMC cependant, les fonctions de base utilisées ici sont généralement des fonctions Gaussiennes. Les RBF seront donc employés dans les mêmes types de problèmes que les PMC [2].

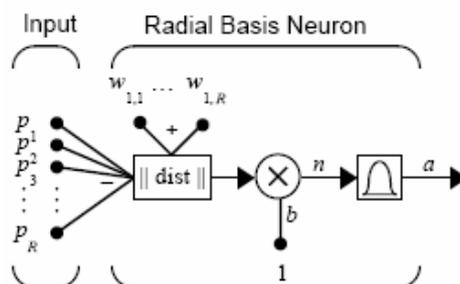


Figure II.1: neurone à fonction radiale.

II. 2. Les réseaux FEED-BACK :

II. 2. 1 Les cartes auto-organisatrices de Kohonen :

C'est probablement le type de réseau de neurones le plus utile, si le processus d'éducation du cerveau humain est à être simulé. Le coeur de ce type de réseaux est la mappe où des neurones s'organisent eux-mêmes par rapport à certaines données à l'entrée. Ce type de réseau est "feed-forward" ainsi que "feedback". Il est utilisé dans le codage d'informations.

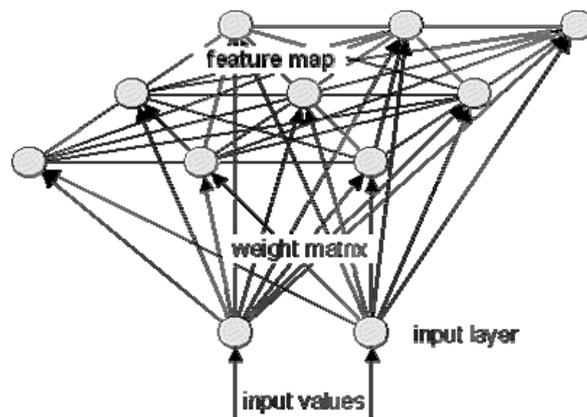


Figure II.2 : carte de Kohonen

II. 2. 2 Les mémoires associatives de Hopfield :

Ce type de réseau est un ensemble de neurones, où chaque neurone est connecté à chaque autre neurone. Il n'y a aucune différenciation entre les neurones d'entrées et de sorties. L'application principale d'un réseau Hopfield est l'entrepôt et la reconnaissance des patrons, exemple l'optimisation des images. La méthode d'apprentissage est non supervisée

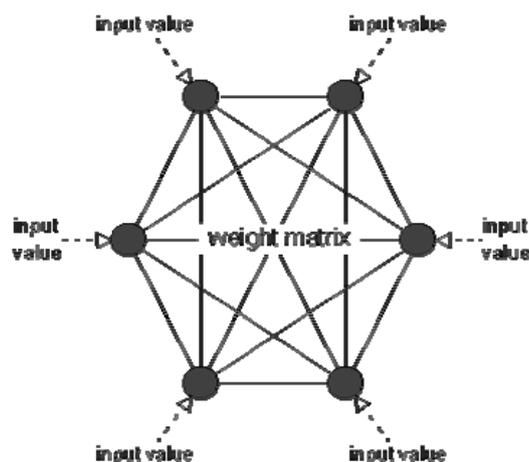


Figure II.3: Réseau de Hopfield

Dans le tableau récapitulatif suivant, sont présentés les différents réseaux et leurs domaines d'application :

Tableau II.1 : Les principaux réseaux et leurs domaines d'application [6].

	Types de réseaux	Domaines d'application
Feed-Forward	Perceptron monocouche	– Classification linéaire.
	PMC	– Classification non linéaire. – Approximation de fonction. – Contrôle. – Reconnaissance de formes et de paroles.
	RBF	– Classification non linéaire. – Approximation de fonction. – Contrôle. – Reconnaissance de formes et de paroles.
Feed-Back	Cartes auto-organisatrices	– Projection non linéaire des données. – Compression des données.
	Mémoires associatives	– Problèmes d'optimisation. – Reconnaissance de formes et de paroles.

Etant donné que la modélisation et la prédiction n'est rien d'autre qu'une approximation de fonction généralement non connue, les PMC et les RBF qui sont des réseaux à couches et à apprentissage supervisé sont, à partir du tableau II.1, les mieux placés pour le traitement de ce genre de problèmes.

La question qui se pose est : quelles sont les différences entre ces deux architectures et laquelle choisir ?

II.3. Comparaison entre le PMC et le RBF :

Bien que le PMC et le RBF permettent de résoudre les mêmes types de problèmes, il existe néanmoins quelques différences entre ces derniers. Dans ce qui suit, nous allons essayer de présenter les principaux points de divergence ainsi que quelques études comparatives qui vont guider notre choix du réseau le mieux adapté à la modélisation et la prédiction :

II. 3.1. Les principales différences :

II. 3.1.1. Selon leurs fonctions d'activation :

La principale différence entre ces deux types de réseaux est la manière dont interviennent leurs paramètres [1] :

- Le PMC est un réseau dont les composantes de base sont des neurones à entrées paramétrées. Pour cette catégorie de neurones, un paramètre (poids) est associé à chaque entrée du neurone. Enfin, une fonction d'activation sigmoïde sera appliquée au potentiel obtenu (somme pondérée des entrées). Pour une fonction d'activation sigmoïde telle que la tangente hyperbolique, nous aurons :

$$a = \text{th} \left(\sum_{i=1}^R w_i p_i \right) \quad (\text{II.1})$$

- Pour le RBF, les paramètres sont attachés à la non linéarité des neurones : ils interviennent directement dans la fonction d'activation. On utilise pour cela des fonctions radiales qui sont généralement des fonctions gaussiennes :

$$a = \exp \left[- \frac{\sum_{i=1}^R (x_i - w_i)^2}{2\sigma^2} \right] \quad (\text{II.2})$$

La conséquence directe de cette distinction est que les RBF ont des non linéarités locales qui tendent vers zéro dans toutes les directions de l'espace des entrées ; leur zone d'influence est donc limitée dans l'espace, ce qui n'est pas le cas des neurones à fonction d'activation sigmoïde qui possède de ce fait, un meilleur pouvoir de généralisation.

II. 3.1.2. Selon la règle d'apprentissage :

Pour les deux réseaux, l'apprentissage est de type supervisé. Cependant, la différence réside dans la règle d'apprentissage utilisée :

- Pour le PMC, l'apprentissage consiste à estimer les paramètres de la matrice \mathbf{W} en utilisant l'algorithme de la rétropropagation basé sur la règle de correction des erreurs, ainsi que l'algorithme de Levenberge-Marquardt
- Pour le RBF, étant donné que deux sortes de paramètres interviennent, l'apprentissage se fait en deux étapes :
 - La première consiste à déterminer les paramètres de la gaussienne (centre et rayons) en calculant la matrice variance/covariance des entrées.

- La deuxième, quant à elle, calcule les poids de la couche de sortie par correction de l'erreur.

Ainsi, un avantage considérable du RBF par rapport au PMC est que son entraînement est plus simple et plus rapide car il se décompose en deux actions locales combinées ; sauf que cet avantage est dépassé par la puissance des calculateurs actuels. De plus, un inconvénient majeur des réseaux RBF est que, comme les modèles polynomiaux, leur taille croît vite avec la dimension de l'entrée [7].

Les principales différences entre PMC et RBF, ont dévoilé la supériorité du Perceptron multicouches en terme de généralisation et de précision. En conséquence, nous adopterons ce type de réseaux pour notre application.

Mais avant d'appliquer l'architecture choisie, il est nécessaire de présenter, avec un peu plus de détail, la première forme du Perceptron, son évolution vers le multicouches et les règles utilisées pour l'apprentissage de ce dernier.

II. 4. Le Perceptron :

Le «Perceptron » appartient la famille générale des réseaux « non bouclé » dont l'information se propage dans un sens unique, des entrées vers les sorties sans aucune rétroaction. Son apprentissage est de type supervisé par la règle de correction des erreurs.

II. 4. 1 Le Perceptron simple :

Le perceptron simple est constitué d'une seule couche de neurones dont les fonctions d'activation sont linéaires de type seuils, suivant ce type de fonction les neurones constitutifs réalisent une partition des vecteurs d'entrée qui lui sont soumis en entrée en deux régions. La frontière entre ces deux régions est définie par la condition :

$\sum w_i x_i = \theta$, ou θ représente le seuil du neurone. En effet pour $\sum w_i x_i > \theta$ le neurone répond 1 et il répond 0 (ou -1) pour $\sum w_i x_i < \theta$. La frontière séparant les deux régions est un hyperplan [4].

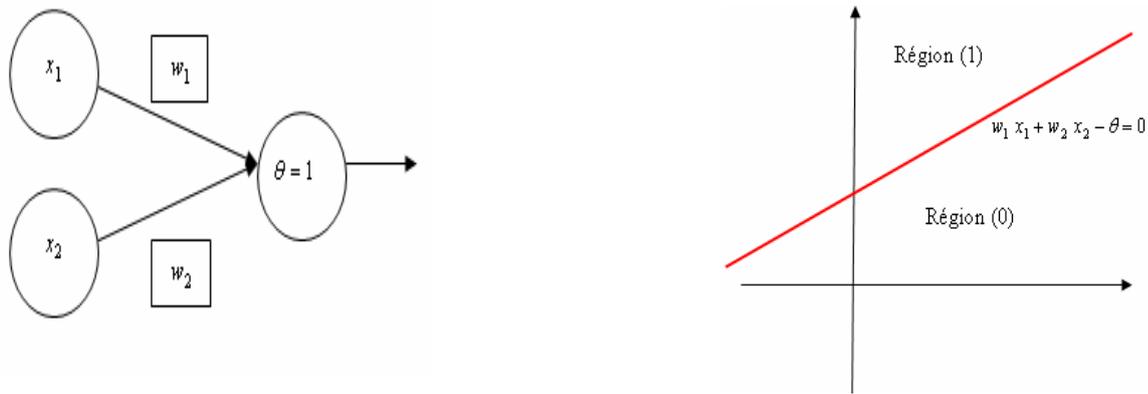


Figure II.4 : la séparabilité linéaire du Perceptron

Cette propriété des neurones à seuil constitue une très forte limite à leurs possibilités de classification. En effet un neurone à seuil ne peut pas représenter le XOR (le ou exclusif) quels que soient les poids des ses connexion.

D'après le tableau de vérité de la fonction XOR ainsi représentée, si un neurone linéaire à seuil réalisait cette fonction, il devrait donc répondre 0 pour les deux couples (0,0) et (1,1) et 1 pour les couples (1,0) et (0,1). Donc il doit séparer linéairement les deux classes (régions) $\{(0,0), (1,1)\}$ et $\{(1,0), (0,1)\}$; [4]

On peut représenter $XOR(x_1, x_2)$ comme suit :

entrée x1	entrée x2	sortie
0	0	0
0	1	1
1	0	1
1	1	0

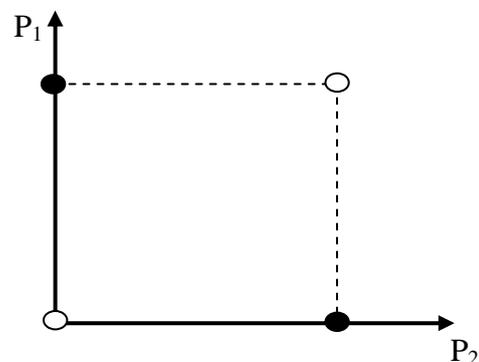


Figure II.5 : le tableau de vérité et la représentation du XOR.

II. 4. 2 Le Perceptron multicouches

Les problèmes que présentent les réseaux à deux couches ont été résolus en ajoutant une deuxième couche entre celle d'entrée et celle de sortie, cette couche appelée *couche cachée*, mais à condition d'utiliser une fonction d'activation non linéaire.

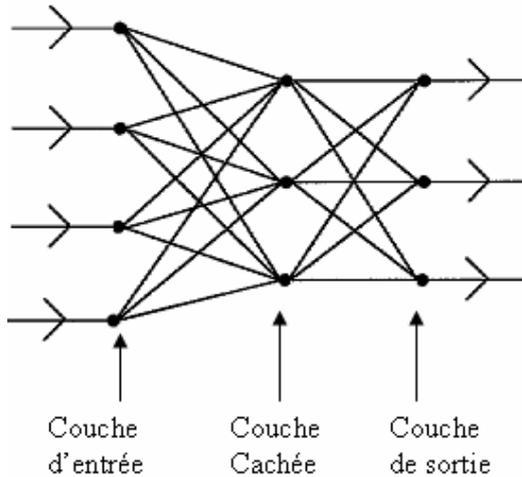


Figure II.6 : le Perceptron multicouches

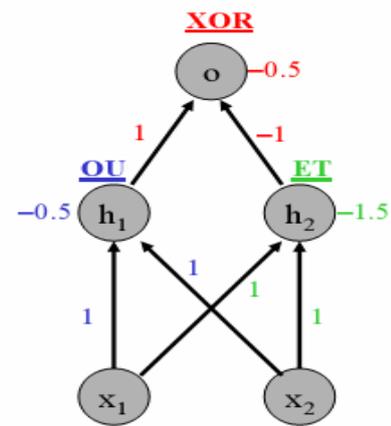


Figure II.7 : la représentation de XOR par PMC

II. 4. 3 L'apprentissage du PMC :

II. 4.3.1 Rétro propagation de l'erreur [8] :

Pour appliquer cet algorithme, la fonction de transfert du neurone doit être continue, non linéaire et différentiable. L'algorithme comprend six étapes qui sont :

1. initialisation des poids à des valeurs aléatoire de faible grandeur généralement entre -1 et 1.
2. lecture d'un premier exemple d'apprentissage dans la base d'apprentissage ;
3. calcul de la sortie du réseau ainsi que l'erreur entre la sortie du réseau et celle désirée

$$\xi = \frac{1}{2}(y_d - y)^2. \quad (\text{II. 3})$$

4. calcul du gradient de l'erreur et des taux de réadaptation des poids inter couches

$$\Delta W_{ij} = \alpha \delta_{w_i} f(x_i(t)) \quad (\text{II.4})$$

Où :

α est le taux d'apprentissage.

δ_{w_i} est le taux de réadaptation partielle des poids. Il est défini par : $\delta_{w_{is}} = (y_{d_i} - y_i)$ pour les neurones de la couche de sortie et par : $\delta_{w_i} = \sum \delta_{w_{st}} W_{ij}$ pour les neurones des autres couches.

5. réadaptation des poids ;
6. tant que l'erreur est trop importante : retourner à l'étape 2.

Le taux d'apprentissage α joue un rôle important dans la régulation de la vitesse de convergence de la méthode. En générale, sa valeur est prise entre 0.1 et 0.7 [9]. Selon la

détermination de ce paramètre, deux variantes améliorent sensiblement l'algorithme d'apprentissage qui sont :

II. 4.3.1.1 la rétro propagation avec Momentum [8] :

C'est une méthode efficace, qui assure une convergence rapide. Elle consiste à utiliser les changements précédents des poids pour la réadaptation des poids actuels. Les nouveaux poids sont donnés par la relation :

$$W_{ij}(t+1) = W_{ij}(t) + \alpha \Delta W_{ij}(t) + \alpha \Delta W_{ij}(t-1) \quad (\text{II.5})$$

Le terme : $\alpha \Delta W_{ij}(t-1)$ est appelé le *Momentum*

II. 4.3.1.2 Propagation à taux d'apprentissage adaptatif [8]

Si le taux d'apprentissage est fixe le long du programme deux problèmes peuvent surgir. Si α est très petit la convergence est très lente alors que si elle est très grande l'algorithme risque diverger (le système oscille). La méthode d'apprentissage adaptative consiste à varier le taux d'apprentissage selon l'erreur calculée pour chaque étape. Cette technique est très simple. Elle augmente la rapidité de la convergence, et œuvre un algorithme qui a pour rôle de contrôler l'erreur d'entraînement à chaque étape.

II.4.3.1.3 limites du rétro propagation [8] :

La mise au point de l'algorithme du rétro-propagation en 1985, a amplifié de façon phénoménale l'utilisation des réseaux neuronaux multicouches. Cet algorithme est appelé Perceptron multicouches. L'algorithme de la rétro-propagation et ses variantes présentent des caractéristiques très intéressantes qui servent de manière efficace dans le domaine de la classification. Néanmoins ces algorithmes ont des limites.

Parmi les limites que présente l'algorithme de la rétro propagation de l'erreur : la convergence n'est pas toujours assurée et le temps de calcul est important. D'autres structures ont été proposées, parmi lesquelles la méthode de Levenberg-Marquardt.

II.4.3.2 Méthode de Levenberge-Marquardt [8] :

La méthode de Levenberge-Marquardt est très souvent utilisée en raison de sa rapidité de convergence .Elle utilise la fonction énergie donner par la relation :

$$E(k+1) = E(k) + \dot{E}(k) \Delta W(k) + \frac{1}{2} (\Delta W(k))^T \ddot{E}(k) \Delta W(k) \quad (\text{II.6})$$

Elle consiste à minimiser cette fonction, ce qui revient à écrire : $\frac{\partial E(K+1)}{\partial \Delta W(K)} = 0$ ou : $\partial \Delta W(K)$

est le vecteur global des taux d'apprentissage.

La dérivée de l'erreur par rapport à ses poids est donnée par la relation :

$$\frac{\partial E(k+1)}{\partial \Delta W(k)} = F_i^l = \sum_{p=1}^K \sum_{i=1}^{nL} (d_i^p - y_i^l) \left[-\frac{\partial y_i^p}{\partial W_i^l} \right] \quad (\text{II.7})$$

Où : F_i^l sont les vecteurs des dérivées de l'erreur à la sortie du réseau par rapport à ses vecteurs

poids W_i^l . Ils sont rassemblés dans un vecteur Jacobien F ;

L est le nombre de couches cachées et $l=1 \dots L$;

K est le nombre d'exemples d'entraînement ;

nL est le nombre de neurones de la dernière couche.

A fin de trouver l'optimum, on utilise la relation : $H \Delta W = -F$ où H représenter la matrice hessienne de la fonction d'erreur en sortie du réseau.

Pour un système de n variables, cette matrice est symétrique .Elle est donnée par :

$$H = \frac{\partial F}{\partial W} = \begin{bmatrix} \frac{\partial F_{11}}{\partial W_1} & \dots & \frac{\partial F_{1n}}{\partial W_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_{n1}}{\partial W_1} & \dots & \frac{\partial F_{nn}}{\partial W_n} \end{bmatrix} \quad (\text{II.8})$$

Avec :

$$\frac{\partial F_{ij}}{\partial W_i} = \frac{\partial^2 E(W)}{\partial W_i \partial W_j} = \frac{\partial}{\partial W_i} \left(\frac{\partial E(W)}{\partial W_j} \right) \quad (\text{II.9})$$

Cette équation est un système linéaire dont la résolution peut être faite par la méthode de Gauss-Newton. Toutefois, son calcul est très lent. La méthode de Levenberge-Marquardt remplace le calcul du Hessien par une approximation numérique donnée par :

$$H = FF^T + \lambda \Omega \quad (\text{II.10})$$

Où :

FF est une matrice positive ;

λ est un coefficient positif ;

Ω est la matrice choisie.

Plusieurs possibilités sont envisageables pour les valeurs de Ω et de λ , le choix doit être fait judicieusement. La matrice Ω est généralement une matrice identité diagonale ou une matrice diagonale dont les éléments sont égaux à ceux de la matrice FF^T .

II. 5. Conclusion :

A travers ce chapitre nous venons de présenter une comparaison entre deux architectures et deux algorithmes afin de justifier notre choix. Et nous avons vu que le réseau et l'algorithme les mieux adaptés pour notre problème sont respectivement le PMC et Levenberge-Marquardt.

Dans le chapitre suivant nous allons présenter quelques aspects théoriques du procédé à simuler, ce qui nous aidera, par la suite, pour la conception du modèle neuronal. Aussi nous allons détailler les modèles théoriques sur la base desquels nous pourrions déterminer quelques paramètres relatifs à l'extraction solide liquide.

Chapitre III:

L'extraction solide liquide,
étude cinétique et
modélisation

Dans le cadre de notre application, nous sommes intéressés à la simulation de l'extraction des graines de coriandre (*cf.annexe3*) par le méthanol. Pour cela, et afin de bien concevoir notre réseau de simulation il est impératif d'analyser le procédé étudié et déterminer les paramètres qui lui sont pertinents, et de détailler les modèles sur la base desquels nous allons essayer de déterminer le coefficient de diffusion.

III .1.Généralités sur l'extraction solide liquide :

III.1.1.Définition :

L'extraction solide liquide, encore appelée extraction par solvant, est une opération de transfert de matière destinée à séparer les principes solubles d'un substrat solide par leur diffusion dans un solvant. En règle générale, la phase noble est constituée par la solution d'extrait tandis que la phase insoluble constitue le résidu. [10]

En principe, l'extraction solide liquide est réalisée, par contact intime du solide avec le solvant suivi d'une séparation mécanique de la solution et le solide. Toutefois, une certaine quantité de solution est retenue dans le solide, ce qui oblige parfois à procéder à des lavages ou à l'opération nommée désolvatisation, consistant à chasser, par évaporation, le solvant imbibant le résidu. [11]

L'extraction solide liquide est régie par des phénomènes très complexes, liés aussi bien au processus de transfert de matière lui-même qu'à la nature du solide à extraire. Le choix de la méthode d'extraction est déterminé par la proportion de matière soluble présente, sa distribution au sein du solide, la nature du solide ainsi que sa taille. [12]

En pratique, l'extraction solide- liquide se déroule en trois étapes :

1. Le passage du soluté dans le solvant qui comprend la dissolution et le transfert.
2. Le lavage du solide par le solvant.
3. La séparation mécanique du solide au contact de la solution.

De plus, l'extraction solide- liquide est influencée, à la fois par des paramètres thermodynamiques (température, pression, débit...etc.) et par des paramètres cinétiques (durée de traitement, vitesse de cheminement du solvant à l'intérieur de la matrice du solide,...etc.). Dans de nombreux cas, le soluté n'est pas distribué de façon homogène dans les pores du solide. Alors qu'une fraction de l'extrait se trouve très accessible, elle est donc extraite en priorité, l'extraction du reste de l'extrait est à la fois limité par la solubilité et la diffusion dans le solide. [12]

III.1.2 Performances industrielles de l'extraction solide- liquide [10]

Une extraction idéale satisferait aux quatre conditions suivantes :

- ✓ Débit élevé,
- ✓ Rendement maximum,
- ✓ Sélectivité élevée ou faible suivant l'objectif visé,
- ✓ Concentration de l'extrait élevée.

Ces quatre impératifs sont évidemment difficiles à concilier et l'on se tient dans la pratique à un compromis.

III.1.2.1. Débit

Le débit peut s'exprimer par la masse de l'extrait obtenue par unité de temps.

III.1.2.2. Rendement

Le rendement représente la masse de la fraction recherchée obtenue dans l'extrait, par rapport à la masse de cette fraction initialement présente dans la phase solide.

Un simple bilan de matière de l'extraction permet de déterminer son rendement, il suffit de connaître :

- La masse de solide traité,
- Sa teneur en composé que l'on veut extraire,
- Le volume d'extrait obtenu,
- La concentration de l'extrait en composé désiré.

III.1.2.3. Sélectivité

Dans certaines industries, le souci de sélectivité se réduit à la non extraction de certains composés indésirables (café, thé, etc.). Dans d'autres, on cherche à extraire un seul composé ou une fraction bien définie (extraction des huiles essentielles). La sélectivité est caractérisée par la teneur en extrait en substances recherchées rapportée à son extrait sec total.

III.1.2.4. Concentration de l'extrait

L'extrait doit être généralement non seulement concentré mais totalement séché. Il est clair que les frais de séchage, liés en particulier à la consommation d'énergie, seront d'autant plus lourds que l'extrait sera plus dilué.

III.1.3. Facteurs intervenant dans l'extraction solide liquide. [11,12]

Les principaux facteurs intervenant dans l'extraction solide liquide sont :

III.1.3.1. Nature et état du solide et du soluté

La nature et l'état physique du solide et du soluté ont une importance déterminante pour le transfert de matière. Le soluté contenu dans ces corps est réparti plus ou moins régulièrement à des teneurs variables dans le solide.

Au début de l'extraction, le solvant dissout le soluté situé à la surface, en laissant un solide poreux. Il pénètre ensuite cette couche extérieure avant d'atteindre le soluté situé en profondeur. Son cheminement devient de plus en plus difficile et se traduit par une diminution de la vitesse d'extraction.

Le soluté dans les matières végétales est généralement occlus dans des cellules d'où il est extrait par un mécanisme de dialyse ou de diffusion capillaire à travers les parois cellulaires.

III.1.3.2. Nature du solvant

Le choix du solvant d'extraction doit tenir compte des critères suivants :

- être sélectif, par rapport aux produits à extraire,
- posséder une grande capacité de dissolution
- avoir une température d'ébullition peu élevée, pour éviter toute dégradation des constituants thermolabiles,
- être non toxique, ininflammable et non explosif,
- avoir une volatilité élevée, pour éviter les opérations d'évaporation,
- avoir une faible viscosité et une masse volumique peu élevée, pour faciliter la diffusion, l'agitation et la séparation mécanique,
- être disponible et peu coûteux,

Les principaux solvants utilisés sont l'eau, les alcools (éthanol), et de moins au moins les hydrocarbures (hexane) et les solvants halogénés.

III.1.3.3. Température

L'élévation de la température permet généralement l'accroissement de la solubilité et de la diffusivité du soluté et la diminution de la viscosité de la solution. Elle est limitée par les risques d'extraire des composés nuisibles, et par les risques de dégradation thermique du soluté.

III.1.3.4. Degré d'agitation

L'agitation mécanique des particules dans le solvant qui permet leur maintien en suspension et l'homogénéisation du milieu à un effet toujours favorable sur l'opération d'extraction.

III.1.3.5. Taux d'humidité du solide

Lors de l'utilisation de solvants hydrophobes, la diffusivité varie inversement avec la teneur en eau du solide.

III.1.3.6. Durée d'extraction

Il est conseillé d'arrêter l'extraction à l'instant où le rendement en produit de meilleure qualité soit maximal.

III.1.3.7. Hydromodule

L'hydromodule, H , est le rapport entre le volume du solvant et la masse de solide à extraire. Il est exprimé en m^3/kg

L'augmentation de l'hydromodule crée un gradient de concentration du soluté dans les deux phases, favorisant ainsi sa diffusion vers la phase liquide. Toutefois, au-delà de certaines valeurs, son influence est peu significative vu que dans ces conditions, la diffusion du soluté ne pouvant pas être freinée par une saturation de la phase liquide où un gradient de concentration insuffisant entre les deux phases.

III.2. Modélisation du procédé d'extraction

L'extraction solide liquide est une opération unitaire présente pratiquement dans toutes les industries chimiques commençant par l'agro-alimentaire jusqu'à la pétrochimie, ce qui fait d'elle un procédé très important. [13]

Son importance nous oblige d'en connaître tous les détails et de déterminer les paramètres difficilement accessibles expérimentalement.

Plusieurs modèles ont été élaborés pour décrire la cinétique d'extraction des essences végétales. Nous avons choisi de nous intéresser et de présenter dans ce qui suit, ceux qui sont les solutions analytiques des équations bilans de transfert de matière, et qui ont été rapportés par Crank [23] et Mafarat et Béliard [10]

III.2.1 Modèle rapporté par Crank [14]

Dans ce modèle, nous n'allons présenter que la diffusion dans le cas d'une géométrie sphérique, étant donné que c'est la géométrie qui représente le mieux le substrat végétal (graines de coriandre), objet de notre application.

III.2.1.1 Diffusion dans une sphère :

Soit C la concentration de l'espèce diffusante En coordonnées sphériques (r , θ et φ) et en absence de réaction chimique, l'équation de continuité s'écrit :

$$\frac{\partial C}{\partial t} = \frac{1}{r^2} \left(\frac{\partial}{\partial r} \left(Dr^2 \frac{\partial C}{\partial r} \right) + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(D \sin \theta \frac{\partial C}{\partial \theta} \right) + \frac{D}{\sin^2 \theta} \frac{\partial^2 C}{\partial \varphi^2} \right) \quad (\text{III.1})$$

Si on admet maintenant, que le coefficient de diffusion est indépendant des coordonnées spatiales et que la diffusion est radiale, l'équation devient :

$$\frac{\partial C}{\partial t} = D \left(\frac{\partial^2 C}{\partial r^2} + \frac{2}{r} \frac{\partial C}{\partial r} \right) \quad (\text{III.2})$$

En posant

$$U = C.r \quad (\text{III.3})$$

L'équation (III.2) devient :

$$\frac{\partial U}{\partial t} = D \frac{\partial^2 U}{\partial r^2} \quad (\text{III.4})$$

III.2.1.2 Diffusion dans une sphère plongée dans une solution parfaitement agitée

Une sphère de volume $\frac{4}{3}.\pi.R^3$, dans laquelle la concentration en substance diffusante est uniforme et initialement égale à $C_s(0)$, est immergée dans une solution de volume V_1 et dont la concentration initiale en substance diffusante est nulle. La concentration de la solution en substance diffusante au cours du processus ne dépend que du temps, puisque la solution est supposée parfaitement agitée donc homogène.

Le problème consiste à résoudre le système d'équation suivant :

$$\begin{cases} \frac{\partial C_s}{\partial t} = D \left(\frac{\partial^2 C_s}{\partial r^2} + \frac{2}{r} \frac{\partial C_s}{\partial r} \right) & 0 < r < R \\ V_1 \frac{\partial C_s}{\partial t} = -4\pi R^2 PD \frac{\partial C_s}{\partial r} & r = R \end{cases} \quad (\text{III.5})$$

La deuxième équation de ce système exprime le fait que le flux de matière traversant la surface de la sphère est égal à celui entrant dans la solution.

La solution générale pouvant être appliquée quelque soit le sens du transfert du soluté : de la solution vers une sphère initialement exempte de soluté ou de la sphère vers un solvant initialement pur s'écrit :

$$\frac{C_s(t)}{C_s(\infty)} = 1 + \sum_{n=1}^{\infty} \frac{6\alpha(1+\alpha)e^{-Dq_n^2 t/R^2}}{9+9\alpha+q_n^2\alpha^2} \frac{R \sin(q_n r/R)}{r \sin q_n} \quad (\text{III.6})$$

La solution correspondante au transfert solide- liquide dans les temps longs écrite en fonction de la masse totale du soluté dans la phase liquide à l'instant t rapportée à la même masse après un temps infini et en remplaçant $r=R$ dans l'équation (III.6), on aura :

$$\frac{M_l(t)}{M_l(\infty)} = 1 - \sum_{n=1}^{\infty} \frac{6\alpha(1+\alpha)e^{-Dq_n^2 t/R^2}}{9+9\alpha+q_n^2\alpha^2} \quad (\text{III.7})$$

Les q_n représentent les solutions de l'équation caractéristique suivante :

$$\tan q_n = \frac{3q_n}{3 + \alpha q_n^2} \quad (\text{III.8})$$

Les valeurs de q_n sont données par la table III.1 pour différentes valeurs de α .

Table III.1 : Les solutions de $\tan q_n = \frac{3q_n}{3 + \alpha q_n^2}$ [14]

α	q_1	q_2	q_3	q_4	q_5	q_6
∞	3.1416	6.2832	9.4248	12.5664	15.7080	18.8496
9.0000	3.2410	6.3353	9.4598	12.5928	15.7207	18.8671
4.0000	3.3485	6.3979	9.5029	12.6254	15.7054	18.8891
2.3333	3.4650	6.4736	9.5567	12.6668	15.7888	18.9127
1.5000	3.6909	6.5665	9.6255	12.7265	15.8326	18.9341
1.0000	3.7264	6.6814	9.7156	12.7928	15.8924	19.6048
0.6667	3.8711	6.8246	9.8369	12.8948	15.9779	19.0784
0.4286	4.0236	7.0019	10.0039	13.6424	16.1082	19.1032
0.2500	4.1811	7.2169	10.2153	13.2953	16.3211	19.2098
0.1111	4.3395	7.4645	10.5437	13.6133	16.6821	19.2366
0	4.4934	7.7253	10.9041	13.6662	17.2208	20.2723

Et α est donné par la relation suivante :

$$\alpha = \frac{PV_1}{\frac{4}{3}\pi R^3} = \frac{PM_1}{M_s} \quad (\text{III.9})$$

Le bilan de matière établi à t_∞ , nous permet d'écrire :

$$\frac{M_l(\infty)}{M_s(0)} = \frac{1}{1+1/\alpha} \quad (\text{III.10})$$

Avec $M_s(0)$ représente la quantité de matière diffusante présente initialement dans la sphère.

Pour les temps courts Crank propose l'utilisation de l'équation suivante :

$$\frac{M_l(t)}{M_l(\infty)} = 6 \left(\frac{Dt}{R^2} \right)^{1/2} \left\{ \pi^{-1/2} + 2 \sum_{n=1}^{\infty} i \operatorname{erfc} \frac{nR}{\sqrt{(Dt)}} \right\} - 3 \frac{Dt}{R^2} \quad (\text{III.11})$$

En négligeant le deuxième terme du second membre devant sa racine et l'apport de la série, l'équation (III.11) s'écrit donc :

$$\frac{M_l(t)}{M_l(\infty)} = 6 \left(\frac{Dt}{R^2} \right)^{1/2} \frac{1}{\sqrt{\pi}} \quad (\text{III.12})$$

III.2.2 Modèle basé sur le coefficient de transfert matière [10]:

III.2.2.1 Présentation du modèle

Ce modèle est fondé sur le fait que le transfert de matière s'effectue par convection dans la phase liquide à cause du mouvement d'agitation. Le bilan de matière s'écrit:

$$\dot{m} = \frac{dm}{dt} = -kA(C'_s - C'_l) \quad (\text{III.13})$$

\dot{m} : Débit massique du soluté (kg/s).

A : aire d'échange, (m²).

C_s : la concentration de soluté dans la phase solide (kg/kg).

C_l : la concentration de soluté dans la phase liquide en (kg/kg).

k : coefficient de transfert de matière (m/s).

Si l'on considère que le volume de la phase solide ne varie pas, la variation de concentration de cette phase en soluté est proportionnelle à sa variation de masse et l'on peut écrire:

$$\frac{dC_s}{dt} = -k'A(C_s - C_l) \quad (\text{III.14})$$

Les constantes k' et A étant pratiquement impossibles à déterminer séparément, autant les globaliser en une seule constante $K=k'A$, dès lors

$$\frac{dC_s}{dt} = -K(C_s - C_l) \quad (\text{III.15})$$

Il est difficile de déterminer dans ces différentes étapes laquelle constitue le facteur limitant la vitesse d'extraction.

III.2.2.2 Hypothèses simplificatrices

Les calculs qui suivent impliquent les hypothèses simplificatrices de départ suivantes:

- ❖ la phase initiale achevée, on se trouve en phase de diffusion,
- ❖ le processus limitant est le processus de diffusion et non celui de solubilisation,
- ❖ le coefficient de partage du soluté entre les deux phases est égal à l'unité,
- ❖ la perte de masse de la phase solide consécutive à la diffusion de matière sèche est compensée par un gain de masse équivalent de solvant migrant dans la phase solide,
- ❖ la phase solide est suffisamment divisée pour que l'on puisse négliger les gradients internes de concentration en soluté.

III.2.2.3 Développement mathématique du modèle:

Soient C_s et C_l les concentrations (en masse par masse) d'un soluté respectivement dans la phase solide et dans la phase liquide.

Conformément à ce que nous avons écrit précédemment (III.14):

$$dC_s = -K(C_s - C_l)dt \quad (\text{III.16})$$

Soient « M » la masse totale de la suspension et « s » la masse de la phase solide rapportée à la masse totale de la suspension.

« C_0 » la concentration initiale de soluté dans la phase solide, un bilan matière permet d'écrire:

$$sMC_0 = sMC_s + (1-s)MC_l \quad (\text{III.17})$$

D'où

$$C_l = \left[\frac{s(C_0 - C_s)}{1-s} \right] \quad (\text{III.18})$$

Après une combinaison entre les équations (III.16), (III.17) et (III.18), nous aurons:

$$\begin{aligned} dC_s &= -K \left[C_s - \frac{s(C_0 - C_s)}{1-s} \right] dt \\ &= \frac{-K}{(1-s)} (C_s - sC_0) dt \end{aligned}$$

Et

$$\int_{C_0}^{C_s} \left(\frac{dC_s}{C_s - sC_0} \right) = \ln \frac{(C_s - sC_0)}{C_0(1-s)} = - \left(\frac{K}{1-s} \right) t \quad (\text{III.19})$$

Ce qui conduit à,

$$C_s(t) = C_0 \left[s + (1-s)e^{-\left(\frac{K}{1-s}\right)t} \right] \quad (\text{III.20})$$

L'équation (III.18) permet d'en déduire l'évolution au cours du temps de la concentration en soluté de l'extrait liquide:

$$C_l(t) = sC_0 \left[1 - e^{-\left(\frac{K}{1-s}\right)t} \right] \quad (\text{III.21})$$

Cette dernière équation est souvent vérifiée expérimentalement et peut s'écrire sous la forme plus condensée:

$$\boxed{C_l(t) = C_{l\infty} (1 - e^{-\lambda t})} \quad (\text{III.22})$$

A la fin, il suffit de déterminer la valeur de $C_{l\infty}$ par le réseau, pour en déduire, par calcul, la valeur de λ qui est égale $\frac{K}{1-s}$.

III. 3. Conclusion :

Après avoir présenté quelques aspects de l'extraction solide liquide, il s'avère qu'un tel procédé dépend de plusieurs paramètres, ce qui rend la modélisation de sa cinétique par l'outil neuronal conditionnée par une base de données la plus large possible.

Pour cela, dans le chapitre suivant nous allons présenter les données cinétiques sur la base des quelles nous allons construire l'ensemble d'apprentissage qui va servir à l'entraînement du modèle neuronal de simulation, et en fin terminer par l'exploitation du modèle afin d'atteindre les coefficients de transfert de matière et de diffusion.

Chapitre IV:

Simulation et Interprétation

Dans ce chapitre, nous allons décrire la méthodologie adoptée menant à la détermination du modèle neuronal optimal, pour en fin l'exploiter dans le cadre de la simulation et la modélisation. Cependant les étapes de notre démarche se résument à :

1. Construction de la base d'apprentissage : sur la base des données cinétiques, et après une analyse de celles ci à fin de déterminer les entrées du modèle, choisir les exemples d'apprentissage.
2. Détermination du réseau optimal : se qui revient à
 - Choisir le nombre de couches cachées du réseau.
 - Choisir le nombre de neurones cachés.
3. Simulation et prédiction
4. Extrapolation pour l'évaluation de la concentration à l'équilibre.
5. Calcul du coefficient de diffusion et de transfert de matière sur la base des modèles détaillés dans le chapitre précédent.

IV. 1 Construction de base d'apprentissage :

Plusieurs travaux traitant de l'extraction de l'essence de la coriandre par le méthanol ont été entrepris au laboratoire de valorisation des énergies fossiles du département de Génie chimique de l'Ecole Nationale Polytechnique. [12-15]

Les travaux d'Amokrane [12], ont porté sur :

- L'échantillonnage des lots de graines a été effectué selon la méthode d'usage dont le but d'utiliser des échantillons représentatifs de la population mère pour l'ensemble des essais d'extraction.
- Une étude granulométrique du solide a montré que l'on peut assimiler la graine de coriandre à une sphère de diamètre moyen de 3.65mm.
- Le taux d'humidité naturel des graines a été évalué à 7.0%.
- Le soluté est l'extrait méthanolique des graines de coriandre renfermant les constituants de l'huile essentielle ainsi que d'autres produits solubles dans ce solvant, tels ; les colorants, résinoïdes, cires, acides,...etc.
- L'évaluation de l'influence de quelques paramètres la vitesse d'agitation, l'hydromodule, le taux d'humidité et la température, sur le rendement en extrait.

IV.1.1 Influence de la vitesse d'agitation ω :

Tableau IV.1 : les valeurs de concentration (mg/ml) pour les vitesses d'agitation $w=500$ tr/min et 350tr/min

t (min)	d=3,65 T=20 $v_m=3$ (Kg/m ³) %H =7	
	w=500 tr/min	w=350tr/min
5	1.58	1.40
10	2.18	1.80
20	2.72	2.05
30	3.22	2.60
45	3.95	3.05
60	3.95	3.40
90	4.75	3.90
130	5.55	4.45
185	6.35	4.97
240	6.90	5.31

IV.1.2 Influence de la température:

Tableau IV.2 : les valeurs de concentration (mg/ml) pour différentes valeurs de l'hydromodule

t (min)	d=3.65 mm $v_m=3$ Kg/m ³ %H=7	
	T=40 °C	T=20 °C
5	1.7	1.4
10	2.23	1.8
15	2.775	2.05
20	3.1	2.3
25	3.4	2.45
30	3.7	2.6
35	3.85	2.78
40	4.1	2.9
45	4.35	3.05
50	4.6	3.2
55	4.8	3.3
60	5.1	3.4
65	5.2	3.5
70	5.25	3.6
75	5.45	3.68
80	5.6	3.78
85	5.8	3.85
90	5.7	3.9

95	6	4
100	6.1	4.1
105	6.2	4.18
110	6.23	4.22
115	6.4	4.3
120	6.45	4.36
125	6.5	4.4
130	6.6	4.45
135	6.7	4.5
140	6.8	4.55
145	6.85	4.6
150	6.9	4.65
155	6.97	4.7
160	7.05	4.75
165	7.1	4.8
170	7.15	4.84
175	7.2	4.88
180	7.25	4.9
185	7.3	4.97
190	7.35	5
195	7.4	5.05
200	7.45	5.08
205	7.5	5.1
210	7.55	5.15
215	7.58	5.2
220	7.6	5.23
225	7.65	5.26
230	7.7	5.28
235	7.77	5.3
240	7.8	5.31

IV.1.3 Influence de taux d'humidité

Tableau IV.3 : les valeurs de concentration (mg/ml) pour différentes valeurs du taux d'humidité.

t (min)	d=3,65mm T=20°C v _m =3Kg/m ³			
	%H = 14,27	%H = 27,9	%H = 29.9	%H = 17.21
5	3,51	4,72	5	4,31
10	4,1	5,22	5,51	4,65
15	4,43	5,47	5,88	5,12
20	4,8	5,7	6,15	5,3

25	5	5,9	6,45	5,59
30	5,27	6,22	6,63	5,84
35	5,5	6,3	6,9	5,9
40	5,7	6,4	7,1	6,1
45	5,84	6,5	7,3	6,19
50	6	6,75	7,48	6,4
55	6,15	6,9	7,66	6,55
60	6,29	7,1	7,75	6,75
65	6,4	7,15	7,9	6,8
70	6,55	7,25	8	6,9
75	6,65	7,35	8,1	7
80	6,8	7,45	8,2	7,1
85	6,85	7,5	8,3	7,2
90	6,9	7,6	8,41	7,25
95	7,1	7,68	8,5	7,34
100	7,2	7,75	8,57	7,4
105	7,25	7,8	8,67	7,49
110	7,33	7,85	8,7	7,54
115	7,39	7,9	8,75	7,6
120	7,45	7,95	8,77	7,7
125	7,58	8	8,82	7,79
130	7,6	8,07	8,85	7,81
135	7,65	8,1	8,87	7,83
140	7,7	8,13	8,9	7,9
145	7,785	8,18	8,92	7,95
150	7,8	8,2	8,98	8
155	7,83	8,26	9	8,05
160	7,86	8,3	9,02	8,1
165	7,9	8,33	9,05	8,13
170	7,95	8,36	9,065	8,15
175	7,98	8,39	9,08	8,18
180	8	8,41	9,09	8,21
185	8,07	8,43	9,1	8,26
190	8,11	8,45	9,11	8,29
195	8,13	8,47	9,12	8,31
200	8,15	8,485	9,14	8,33
205	8,17	8,5	9,16	8,35
210	8,175	8,504	9,165	8,38
215	8,179	8,506	9,17	8,4
220	8,18	8,509	9,175	8,42
225	8,18	8,51	9,18	8,43
230	8,18	8,51	9,185	8,44
235	8,18	8,51	9,19	8,45
240	8,18	8,51	9,2	8,45

IV.1.4 Influence d'hydromodule :

Tableau IV.4 : les valeurs de concentration (mg/ml) pour différentes valeurs d'hydromodule.

t(min)	D=3.65 mm T=20°C %H=7		
	vm=5.16(Kg/m ³)	vm=3(Kg/m ³)	vm=4.17(Kg/m ³)
5	0.6	1.4	0.8
10	0.7	1.8	1
15	0.84	2.05	1.2
20	0.94	2.3	1.25
25	0.99	2.45	1.37
30	1.06	2.6	1.46
35	1.14	2.78	1.55
40	1.22	2.9	1.64
45	1.24	3.05	1.75
50	1.33	3.2	1.8
55	1.38	3.3	1.88
60	1.46	3.4	1.94
65	1.5	3.5	1.98
70	1.53	3.6	2.04
75	1.56	3.68	2.1
80	1.59	3.78	2.15
85	1.63	3.85	2.2
90	1.65	3.9	2.25
95	1.7	4	2.3
100	1.74	4.1	2.35
105	1.78	4.18	2.38
110	1.82	4.22	2.42
115	1.84	4.3	2.45
120	1.85	4.36	2.46
125	1.9	4.4	2.51
130	1.94	4.45	2.56
135	1.96	4.5	2.6
140	1.98	4.55	2.65
145	2.02	4.6	2.68
150	2.04	4.65	2.7
155	2.06	4.7	2.74
160	2.085	4.75	2.78
165	2.1	4.8	2.8
170	2.135	4.84	2.83
175	2.16	4.88	2.87
180	2.17	4.9	2.9
185	2.195	4.97	2.93
190	2.215	5	2.97
195	2.23	5.05	3.01
200	2.25	5.08	3.05
205	2.28	5.1	3.07
210	2.29	5.15	3.09
215	2.3	5.2	3.11

220	2.305	5.23	3.14
225	2.31	5.26	3.16
230	2.31	5.28	3.18
235	2.31	5.3	3.19
240	2.31	5.31	3.2

Les travaux de BOUZID [15] ont concerné le même système solide- liquide que celui étudié par Amokrane [12], et se sont intéressés à l'étude de l'influence de la granulométrie sur le rendement en extrait. Ainsi un broyage des graines a été opéré suivi d'un tamisage pour la préparation d'échantillons de différentes classes granulométriques. Les résultats sont consignés dans le tableau IV.4 suivant :

IV.1.5 Influence du diamètre des grains :

Tableau IV.5 : les valeurs de concentration (mg/ml) pour différentes valeurs du diamètre.

t (min)	T=25° C $\nu_m=3(\text{Kg/m}^3)$ %H =7					
	d=1,25mm	d=1,85mm	d=2,25mm	d=2,285mm	d=3,25mm	d=3,675mm
5	12.28	6.96	7.7	8.44	4.48	5.98
10	14.02	8.18	8.5	8.94	6.26	6.86
15	15.6	9.2	9.8	10.3	7.3	8.5
20	16.34	10.16	10.2	10.8	8.5	9
25	17.2	10.8	11.1	11.45	9.7	10
30	17.6	11.2	11.7	11.96	10	10.2
35	18.4	12.2	12	12.6	11.3	11.3
40	19	12.8	12.6	13.1	12	11.8
45	19.6	13.3	12.85	13.85	12.7	12.25
50	20.2	13.8	13.4	14.1	13.3	12.75
55	20.7	14.4	13.8	14.4	13.7	13.2
60	21.6	14.49	14.2	14.6	14.1	13.4
65	21.7	15.3	14.4	15.2	14.5	13.8
70	22	15.8	14.7	15.5	14.8	14
75	22.4	16.3	15	15.9	15.1	14.3
80	22.8	16.5	15.2	16.1	15.5	14.6
85	23	16.9	15.4	16.4	15.8	14.85
90	23.25	17.5	15.4	16.5	16.1	14.95
95	23.4	17.6	15.8	16.9	16.3	15.4
100	24	17.9	15.9	17.2	16.5	15.65
105	24.35	18	16	17.45	16.8	15.8
110	24.55	18.3	16.2	17.6	17	16
115	24.7	18.4	16.4	17.85	17.3	16.22
120	24.95	18.65	16.6	17.95	17.5	16.4

125	25.2	18.8	16.75	18	17.7	16.6
130	25.4	18.95	16.8	18.3	17.88	16.7
135	25.6	19.1	16.9	18.5	18	16.8
140	25.8	19.2	17.1	18.7	18.18	16.95
145	25.95	19.45	17.2	18.85	18.25	17.1
150	26	19.5	17.25	19	18.5	17.2
155	26.3	19.6	17.28	19.2	18.65	17.3
160	26.4	19.64	17.35	19.3	18.74	17.4
165	26.6	19.67	17.4	19.4	18.85	17.49
170	26.8	19.79	17.45	19.5	19	17.55
175	26.9	19.89	17.55	19.6	19.03	17.65
180	27	19.9	17.62	19.7	19.1	17.78
185	27.1	19.96	17.65	19.8	19.28	17.8
190	27.2	19.98	17.7	19.9	19.35	17.82
195	27.35	20	17.8	20	19.4	17.87
200	27.5	20.05	17.82	20.1	19.5	17.9
205	27.6	20.07	17.85	20.2	19.6	18
210	27.65	20.09	17.9	20.3	19.68	18.1
215	27.75	20.1	17.93	20.33	19.7	18.45
220	27.8	20.11	17.96	20.35	19.73	18.2
225	27.92	20.115	18	20.4	19.75	18.23
230	28	20.118	18.05	20.45	19.77	18.25
235	28.05	20.12	18.08	20.5	19.79	18.27
240	28.1	20.12	18.09	20.55	19.8	18.3

L'analyse des données citées ci-dessus nous permet de faire les constatations suivantes :

- Il y a que deux expériences qui traitent l'influence de la vitesse d'agitation, à cause de cela nous n'allons pas tenir compte de ce paramètre. D'autant que ces vitesses sont suffisamment grandes pour négliger la résistance au transfert dans la phase liquide par rapport à celle qui existe dans le solide.
- Les valeurs des concentrations obtenues avec un hydromodule de 5.16 Kg/m^3 sont nettement plus petites que celles obtenues dans les autres essais, nous avons exclu cette série la base et ce, afin de normaliser l'ensemble d'apprentissage.
- Un grand écart entre les valeurs des températures, (20, 25 et 40 °C), est à noter. Une interpolation dans un large domaine pourrait être hasardeuse. En effet les calculs préliminaires ont montré que l'essai réalisé à 40 °C perturbe l'apprentissage et pour assurer un bon apprentissage du réseau nous l'avons exclu de la base.

En tenant compte des résultats dont nous disposons et de leur structure, le modèle que nous proposons comportera 5 entrées qui sont: le temps, le diamètre (d), la température (T), le taux d'humidité (H) et l'hydromodule (vm) et une sortie qui est bien entendue , la concentration en extrait .

IV. 2 La détermination du modèle neuronal optimal :

La détermination du réseau optimal est fondée sur deux approches différentes du processus étudié :

- **Première approche** : elle consiste à considérer que la cinétique de l'extraction suit une fonction unique dans tous le domaine du temps. Et sur cette approche deux bases d'apprentissage ont été utilisées, la première en entraînant le réseau de 0 à 120 min puis on prédit les valeurs de la concentration dans le domaine restant du temps, tandis que la deuxième est construite pour entraîner le réseau de 0 à 180 min puis les autres valeurs de la concentration sont prédites.
- **Deuxième approche** : en se basant sur le modèle rapporté par Crank, qui considère que la cinétique de l'extraction est régie par deux fonctions, une dans un premier intervalle de temps « temps courts » et l'autre dans un deuxième intervalle du temps, à l'approche de l'équilibre des phases appelé « temps longs ». pour cela deux réseaux ont été établis le premier est entraîné de 0 à 90 min et le deuxième de 95 à 240 min, puis les exemples du test sont simulés par les deux réseaux.

IV. 2. 1 Choix du nombre de couches cachées du PMC :

En 1989, K. Hornik et ses collègues ont démontré qu'un réseau à une seule couche cachée est capable d'approximer avec une bonne précision toute fonction continue à condition que la fonction d'activation ne soit pas polynomiale et que le nombre de neurones de la couche cachée soit suffisant[4] .

De plus, on montre qu'on peut améliorer les performances en augmentant le nombre de couches cachées, tout en diminuant le nombre de neurones par couche. En revanche, si l'on n'a pas obtenu des résultats satisfaisants avec une couche cachée, il est illusoire d'espérer en obtenir en augmentant le nombre de couches cachées [1].

Pour cela et comme notre application n'est qu'une tentative d'approximation de fonction, nous allons à juste titre utiliser un PMC avec une seule couche cachée pour tous

les modèles établis, avec une fonction d'activation de type tangente sigmoïde utilisée pour la couche cachée et une fonction linéaire pour la couche de sortie.

IV. 2. 2 Choix de nombre de neurones cachés:

Comme nous l'avons précisé, pour chaque modèle à établir nous avons 5 entrées qui sont le temps, le diamètre, la température, le taux d'humidité et l'hyromodule, et une seule sortie qui est la concentration. Ce qui se traduit par une entrée à 5 neurones et une sortie avec un seul neurone. Il reste à déterminer le nombre de neurones de la couche cachée pour chaque modèle. Pour cela nous sommes basés sur l'erreur quadratique moyenne générée lors de l'apprentissage des modèles tout en respectant le dilemme précision– temps de calcul.

IV. 2. 2. 1 Approche 1 :

a. Réseau entraîné à 120 min :

La détermination du nombre de neurones de la couche cachée se fait sur la base des essais, en augmentant à chaque fois le nombre de neurones cachés et on note l'erreur quadratique moyenne (MSE) générée lors de l'apprentissage.

Suivant cela, nous avons entraîné des réseaux constituaient successivement 3, 4, 5, 6, 7 neurones cachés et celui qui a donné le minimum d'erreur était celui de 7 neurones cachés (*cf. annexe 4*).

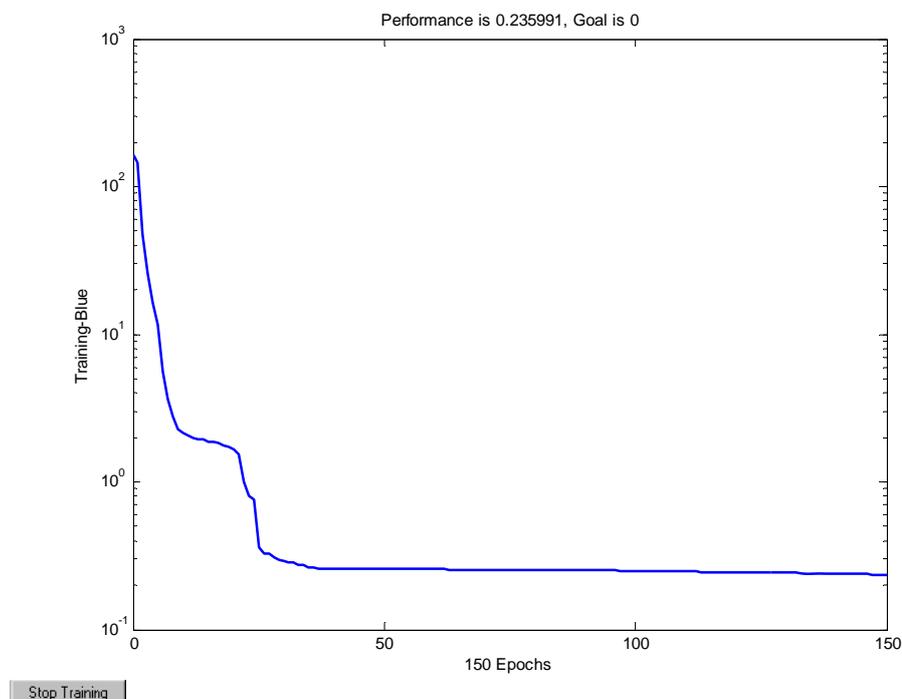
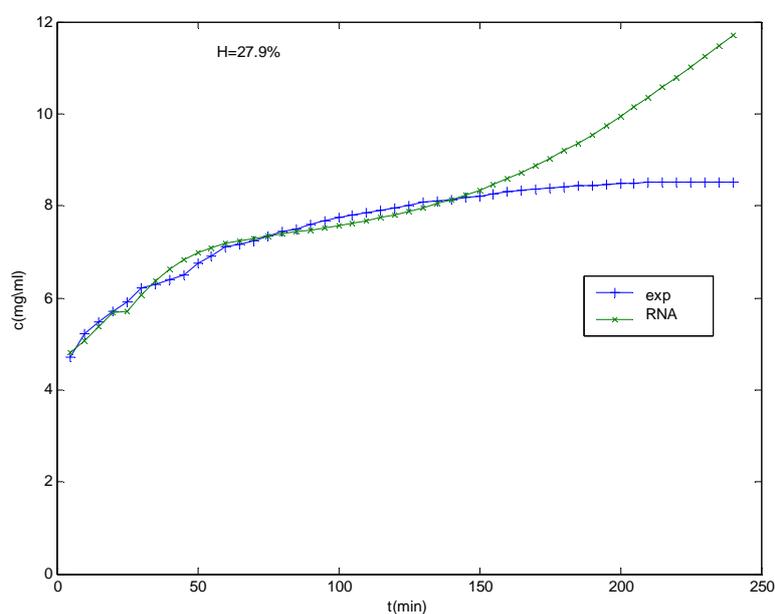


Figure. IV.1 : l'évolution de l'erreur quadratique moyenne (MSE) lors de l'apprentissage en 150 itérations pour le réseau de 7 neurones cachés.

Tests et Résultats de simulation :

Après avoir testé le réseau ainsi conçu, il s'avère que le réseau a bien maîtrisé l'évolution de la concentration dans le domaine où il a été entraîné, par contre en ce qui concerne la prédiction des valeurs de la concentration dans l'intervalle restant [120, 240 min], le réseau n'était pas capable de donner des bonnes valeurs et même pas de suivre la même forme de la courbe cinétique, ce qui influe, évidemment par la suite sur le calcul de la concentration en extrait à l'infini.



FigureIV.2: prédiction de la concentration de l'expérience 6.

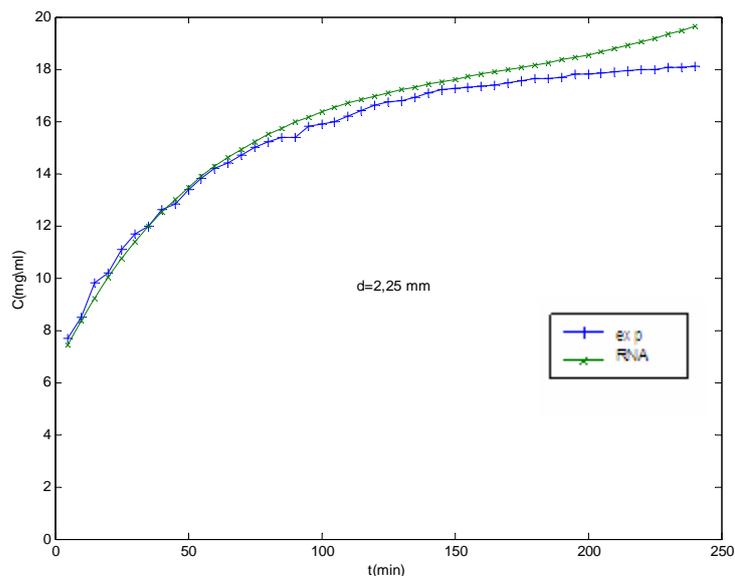


Figure IV.3: prédiction de la concentration de l'expérience 3.

On conclut, que la base d'apprentissage ainsi utilisée ne suffit pas pour que le réseau apprenne bien la forme de la courbe cinétique d'un tel processus.

b. Réseau entraîné jusqu'à 180 min :

Cette fois ci, en augmentant l'intervalle d'apprentissage, en se basant sur le même critère qui est l'erreur quadratique moyenne et en suivant la même démarche qu'avant, le réseau optimale se trouve constituer de 7 neurones cachés (*cf. annexe 4*).

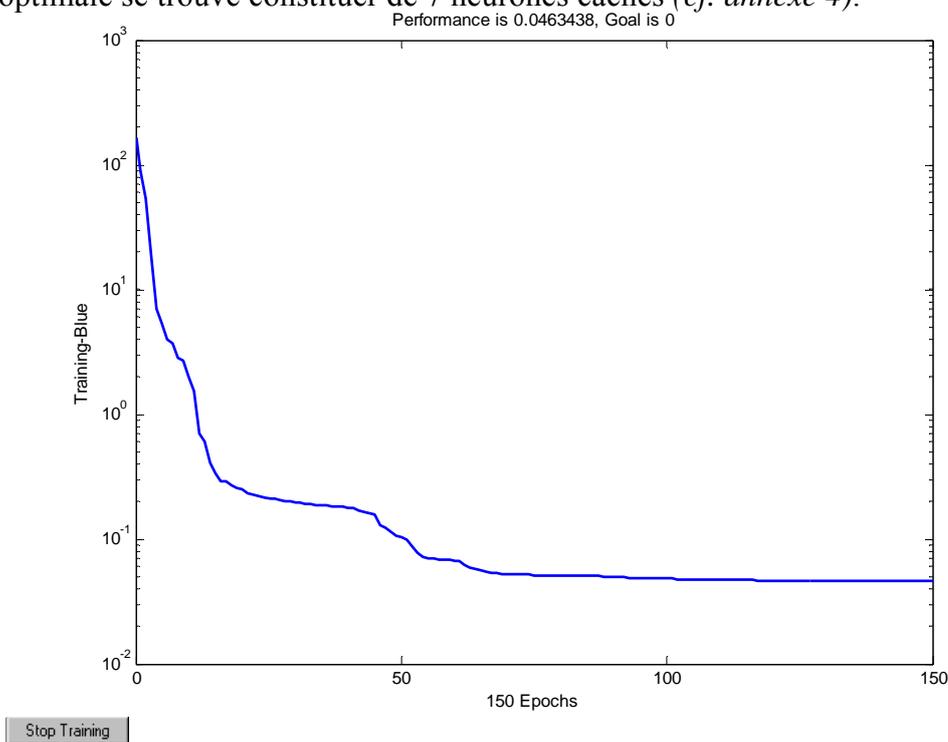


Figure IV.4 : l'évolution de l'erreur quadratique moyenne (MSE) lors de l'apprentissage en 150 itérations pour le réseau de 7 neurones cachés.

Test et Résultats de simulation :

Tableau IV.6 : les conditions opératoires des expériences utilisées

N° d'expérience	conditions opératoires
1	d=3,25mm T=25°C $v=3 \text{ kg/m}^3$ H =7%
2	d=3,675mm T=25°C $v=3 \text{ kg/m}^3$ H =7%
3	d=2,25mm T=25°C $v=3 \text{ kg/m}^3$ H =7%
4	d=3,65mm T=20°C $v=4,17 \text{ kg/m}^3$ H =7%
5	d=3,65mm T=20°C $v=3 \text{ kg/m}^3$ H =14,27%
6	d=3,65mm T=20°C $v=3 \text{ kg/m}^3$ H =27,9%
7	d=3,65mm T=20°C $v=3 \text{ kg/m}^3$ H =29,9%
8	d=3,65mm T=20°C $v=3 \text{ kg/m}^3$ H =7%
9	d=1,25mm T=25°C $v=3 \text{ kg/m}^3$ H =7%
10	d=1,85mm T=25°C $v=3 \text{ kg/m}^3$ H =7%
11	d=2,285mm T=25°C $v=3 \text{ kg/m}^3$ H =7%
12	d=3,65mm T=25°C $v=3 \text{ kg/m}^3$ H =17,21%

En utilisant le réseau ainsi établi nous avons essayé de prédire la concentration dans l'intervalle [185, 240 min] pour les expériences cités avant (voir tableau IV.6, IV.7). Les résultats des écarts entre les valeurs expérimentales et celles données par le réseau à chaque instant de [185 à 240 min] sont donnés dans le tableau suivant :

Tableau IV.7 : les écarts entre les valeurs expérimentales et celles prédites par le réseau des expériences 1 à 5

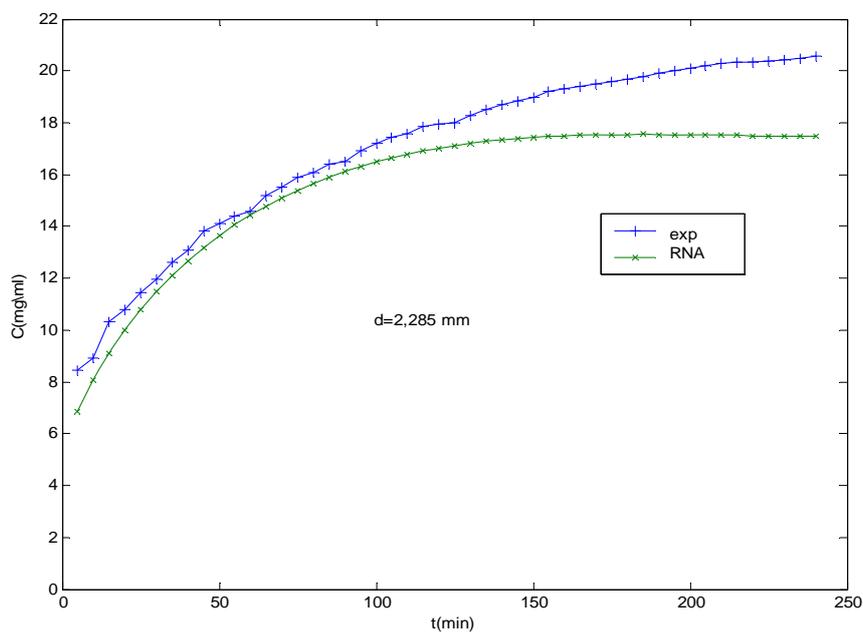
t (min)	expérience 6	expérience 7	expérience 8	expérience 9	expérience 10
185	0,08	0,21	0,19	-0,01	-0,56
190	0,10	0,24	0,21	-0,02	-0,66
195	0,11	0,27	0,24	0,03	-0,77
200	0,13	0,32	0,26	0,08	-0,83
205	0,14	0,36	0,27	0,08	-0,93
210	0,15	0,39	0,31	0,04	-1,02
215	0,15	0,42	0,36	0,05	-1,12
220	0,16	0,45	0,38	0,01	-1,22
225	0,16	0,48	0,40	0,04	-1,33
230	0,17	0,51	0,42	0,04	-1,43
235	0,17	0,54	0,43	0,01	-1,53
240	0,17	0,58	0,44	-0,02	-1,63

Tableau IV.8: les écarts entre les valeurs expérimentales et celles prédites par le réseau des expériences 6 à 10

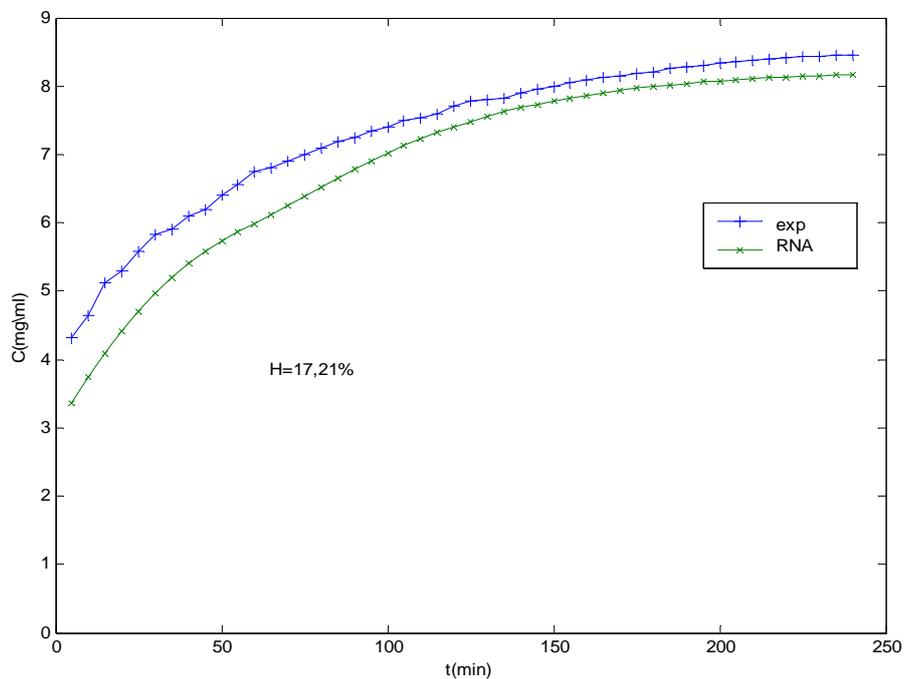
t (min)	expérience 1	expérience 2	expérience 3	expérience 4	expérience 5
185	0,24	-0,25	0,37	0,04	0,03
190	0,22	-0,32	0,41	0,05	0,05
195	0,18	-0,36	0,51	0,06	0,05
200	0,19	-0,41	0,53	0,08	0,06
205	0,21	-0,39	0,56	0,07	0,06
210	0,21	-0,38	0,61	0,06	0,06
215	0,15	-0,10	0,65	0,05	0,05
220	0,10	-0,43	0,68	0,05	0,04
225	0,04	-0,48	0,72	0,04	0,03
230	-0,01	-0,53	0,77	0,03	0,02
235	-0,06	-0,58	0,80	0,01	0,01
240	-0,13	-0,63	0,81	-0,01	0,01

A côté de cela deux expériences (celle qui traite $d=2,285$ mm et celle qui traite $H=17,21\%$) n'appartenaient pas à la base d'apprentissage et que nous avons testé notre réseau pour

reconstruire la courbe cinétique dès le début, et les résultats obtenus sont illustrés par les figures suivantes :



FigureIV.5 : simulation de la courbe cinétique de l'expérience 11.



FigureIV.6 : simulation de la courbe cinétique de l'expérience 12.

Nous voyons que le réseau, a donné une bonne simulation pour $H=17,21\%$ et surtout que l'écart entre les valeurs expérimentales et celles prédites diminue avec le temps, ce qui donnera une bonne détermination de la concentration à l'infinie.

Par contre pour le deuxième exemple ($d=2,285$ mm) l'écart entre les valeurs expérimentales et celles simulées était entre $[0.17, 3.08]$ qui est un intervalle large mais, malgré cela il a bien suivie la même allure de la courbe cinétique d'une façon générale.

IV. 2. 2. 2 Approche 2 :

Cette approche, comme nous l'avons déjà précisé consiste à découper la courbe cinétique en deux parties, la première correspondant à un l'intervalle du temps $[0, 90$ min] tandis que la deuxième à $[95, 240$ min]. Pour cela deux réseaux ont été établis dont un est entraîné de $[0$ à 90 min] et le deuxième de $[95$ à $240]$, pour chaque partie un réseau constitué de 6 neurones cachés a donné le meilleur performance (cf. annexe 4).

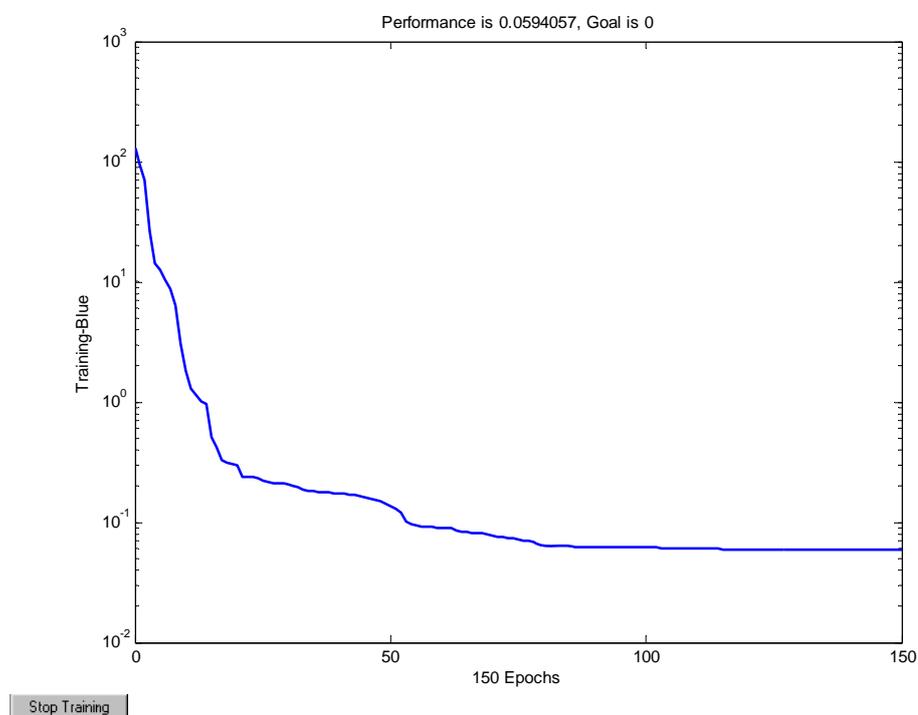


Figure IV.7 : évolution de l'erreur quadratique moyenne (MSE) lors de l'apprentissage du réseau 1 $[0, 90$ min].

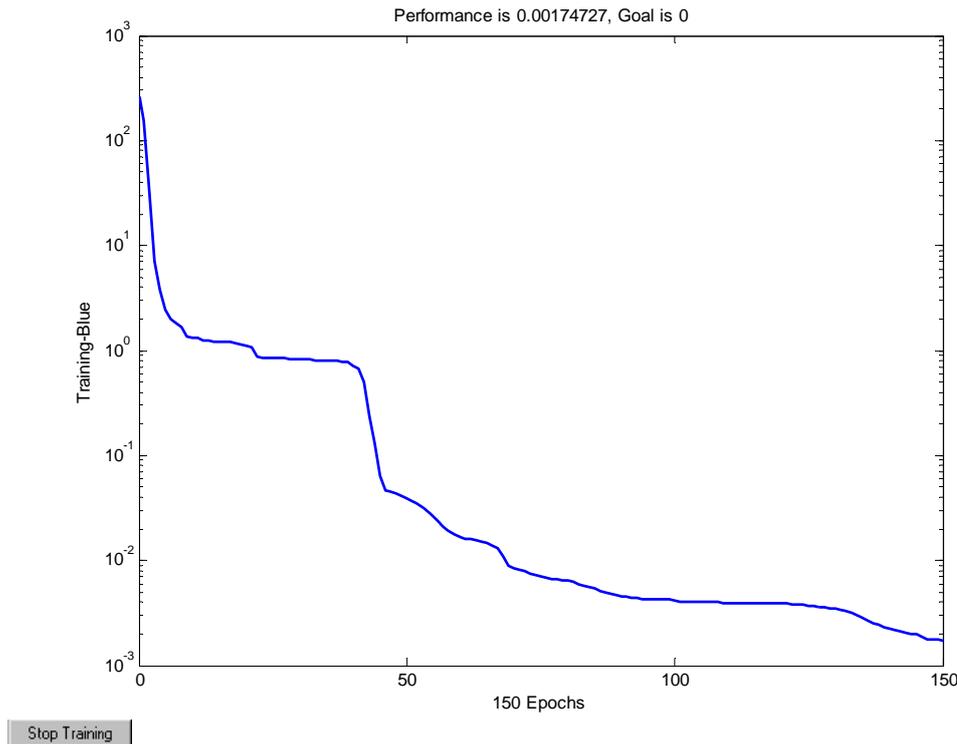
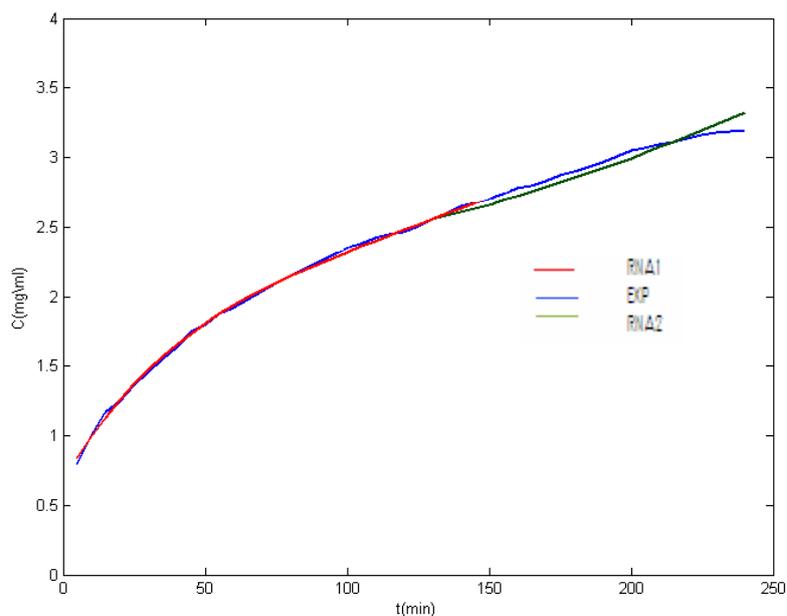


Figure IV.8 : évolution de l'erreur quadratique moyenne (MSE) lors de l'apprentissage du réseau2 [95, 240 min].

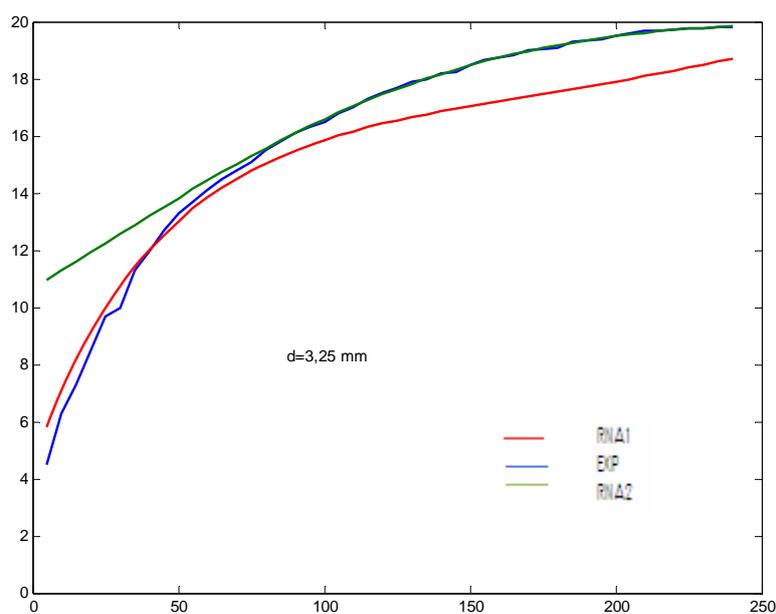
Test et résultat de simulation :

Après avoir testé les deux réseaux ainsi entraînés, il s'avère que cette approche présente quelques inconvénients à savoir :

- Pour certains cas les courbes cinétiques ainsi simulés sont bien espacés se qui rendre leur sommation non pratique.
- l'établissement de deux réseaux bien qu'il ait amélioré l'erreur pour certains cas, mais cette amélioration est marginal devant le temps de calcul inhérent à la manipulation des deux réseaux.



FigureIV.9 : simulation de la courbe cinétique de l'expérience 4.



FigureIV.10 : simulation de la courbe cinétique de l'expérience 1.

Donc, après toutes ces comparaisons, il est bien claire que le réseau optimal est celui entraîné jusqu'à 180 min, sur la base du quel nous allons dans ce qui suit essayer d'évaluer la concentration et le temps à l'infini, afin de déterminer le coefficient de diffusion et de transfert de matière.

IV.3 Détermination de coefficient de diffusion :**IV .3. 1 Modèle de Crank (1^{er} modèle) :****IV.3.1.1 Sur la base des résultats expérimentaux :**

Prenant l'équation III.7 en considérant que le premier terme de la série

$$\frac{C_l(t)}{C_l(\infty)} = 1 - \frac{6\alpha(1+\alpha)}{9+9\alpha+\alpha^2q_1^2} \exp\left(-\frac{Dq_1^2}{R^2}t\right) \quad (\text{IV.1})$$

$$\text{Ln}\left(1 - \frac{C_l(t)}{C_l(\infty)}\right) = -\frac{Dq_1^2}{R^2}t + \text{Ln}\frac{6\alpha(1+\alpha)}{9+9\alpha+\alpha^2q_1^2} \quad (\text{IV.2})$$

Prenons différentes valeurs du coefficient de partage $P = 0.25- 0.5- 0.75- 1- 1.25$, les valeurs de $C_l(\infty)$ et α sont calculés respectivement par :

$$C_l(\infty) = \frac{C_s(0)}{P + \frac{1-s}{s}} \quad (\text{IV.3})$$

$$\alpha = \frac{C_l(\infty)(1-s)}{C_s(0)s - C_l(\infty)(1-s)} \quad (\text{IV.4})$$

Pour le calcul de la valeur de q_1 , la fonction $q_1 = f(\alpha)$ est corrélé sur la base des valeurs données par le tableau de Crank, on obtient : $q_1 = 0.03272 \alpha^2 - 0.4057 \alpha + 4.273$ avec un coefficient de corrélation égale à 0.9037.

Après cela, selon chaque valeur de $C_l(\infty)$ on trace la courbe

$$\text{Ln}\left[\frac{\left(1 - \frac{C_l(t)}{C_l(\infty)}\right)[9+9\alpha+\alpha^2q_1^2]}{6\alpha(1+\alpha)}\right] = f(t) \quad (\text{IV.5})$$

Celle qui aura la meilleure tendance droite, sera utilisée pour calculer le coefficient de

diffusion sur la base de sa pente par la relation: $D = \frac{R^2 \text{ pente}}{q_1^2}$.

Exemple : les figures ci-dessous montrent la courbe obtenue en traçant l'équation (IV.5) pour chaque valeur de P pour l'expérience 7.

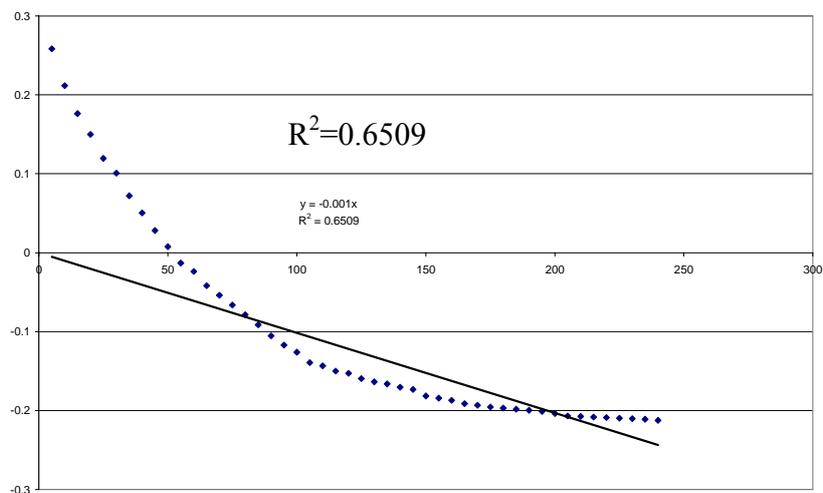


Figure IV.11: la courbe obtenue en traçant l'équation (IV.5) pour l'expérience 7 avec P=0.25

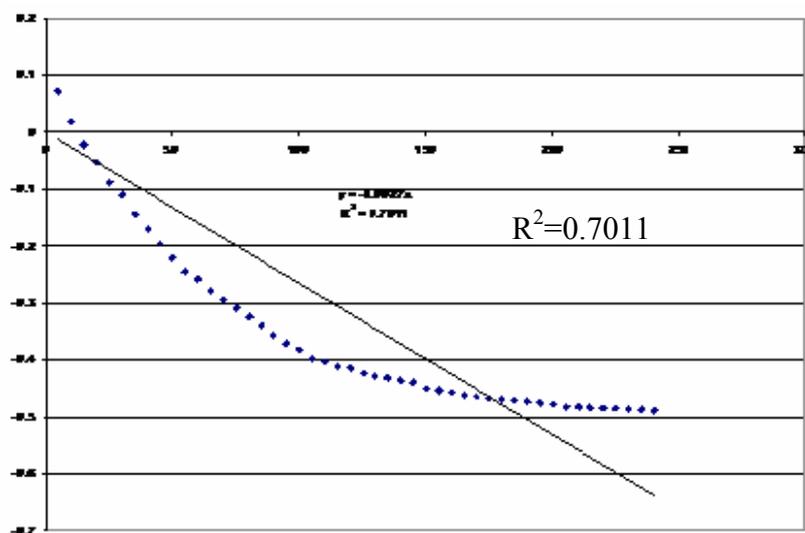


Figure IV.12: la courbe obtenue en traçant l'équation (IV.5) pour l'expérience 7 avec P=0.5

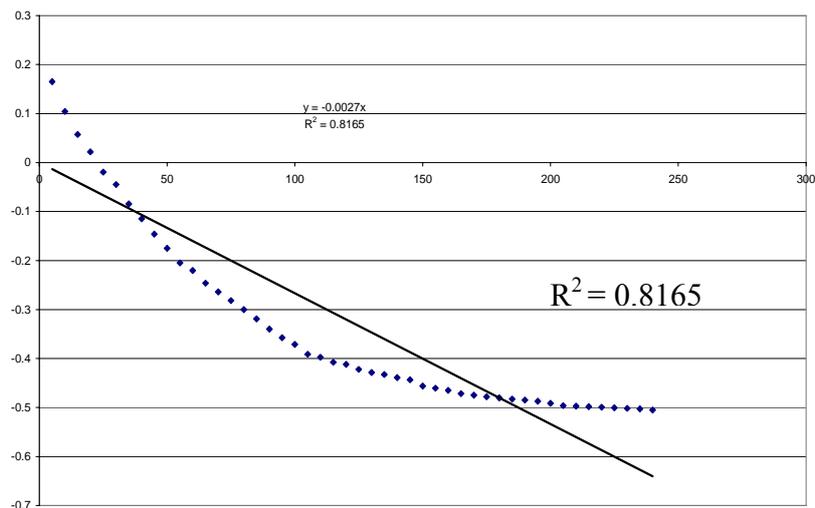


Figure IV.13: la courbe obtenue en traçant l'équation (IV.5) pour l'expérience 7 avec $P=0.75$

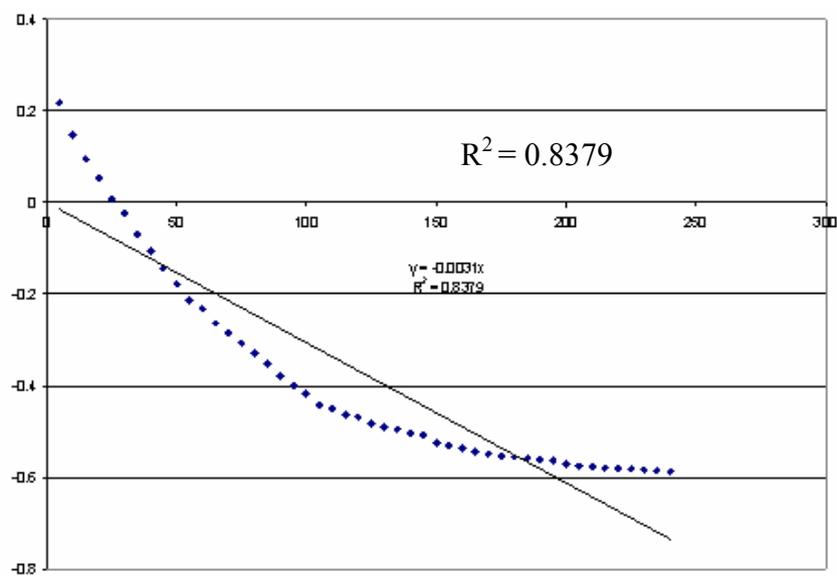


Figure IV.14 : la courbe obtenue en traçant l'équation (IV.5) pour l'expérience 7 avec $P=1$.

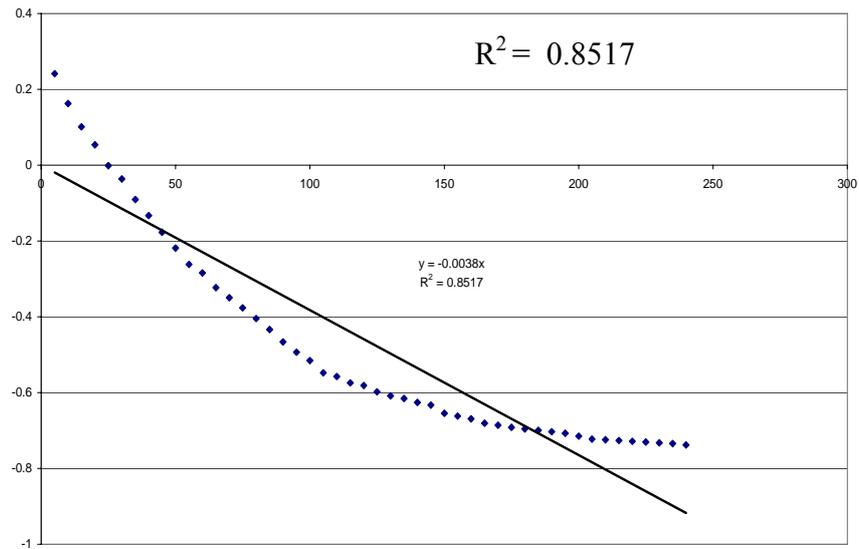


Figure IV.15 : la courbe obtenue en traçant l'équation (IV.5) pour l'expérience 7 avec $P=1.25$

En suivant ces démarches, on a obtenu les résultats suivants :

Tableau IV.9 : les résultats obtenus pour chaque expérience et pour chaque valeur de Coefficient de partage $p..$

exp	p	cinf	alpha	q	R ²
1 $C_s(0)=107$ Kg/Kg	0.25	40.71	9.51	3.37	-0.5316
	0.5	37.17	4.76	3.08	0.5961
	0.75	34.20	3.17	3.32	0.3486
	1	31.67	2.38	3.49	0.2704
	1.25	29.49	1.90	3.62	0.3121
2 $C_s(0)=107$ Kg/Kg	0.25	40.71	9.51	3.37	-1.2055
	0.5	37.17	4.76	3.08	0.4806
	0.75	34.20	3.17	3.32	0.0996
	1	31.67	2.38	3.49	-0.0339
	1.25	29.49	1.90	3.62	0.0137
3 $C_s(0)=107$ Kg/Kg	0.25	40.71	9.51	3.37	-1.3594
	0.5	37.17	4.76	3.08	0.5671
	0.75	34.20	3.17	3.32	0.1619
	1	31.67	2.38	3.49	0.0119
	1.25	29.49	1.90	3.62	0.0631
5 $C_s(0)=42.53$ Kg/Kg	0.25	16.19	9.51	3.37	-0.1319
	0.5	14.78	4.76	3.08	0.8303
	0.75	13.60	3.17	3.32	0.6673
	1	12.59	2.38	3.49	0.6083
	1.25	11.72	1.90	3.62	0.6412
6 $C_s(0)=42.54$ Kg/Kg	0.25	16.19	9.51	3.37	0.2919
	0.5	14.78	4.76	3.08	0.8418
	0.75	13.60	3.17	3.32	0.8419
	1	12.59	2.38	3.49	0.8204
	1.25	11.72	1.90	3.62	0.8375
7 $C_s(0)=42.54$ Kg/Kg	0.25	16.19	9.51	3.37	0.6509
	0.5	14.78	4.76	3.08	0.7011
	0.75	13.60	3.17	3.32	0.8165
	1	12.59	2.38	3.49	0.8379
	1.25	11.72	1.90	3.62	0.8517
9 $C_s(0)=107$ Kg/Kg	0.25	40.71	9.51	3.37	0.8803
	0.5	37.17	4.76	3.08	0.9225
	0.75	34.20	3.17	3.32	0.9636
	1	31.67	2.38	3.49	0.9773
	1.25	29.49	1.90	3.62	0.9832
10 $C_s(0)=107$ Kg/Kg	0.25	40.71	9.51	3.37	-0.2656
	0.5	37.17	4.76	3.08	0.7068
	0.75	34.20	3.17	3.32	0.5148
	1	31.67	2.38	3.49	0.453
	1.25	29.49	1.90	3.62	0.4898
11 $C_s(0)=107$ Kg/Kg	0.25	40.71	9.51	3.37	-0.3402
	0.5	37.17	4.76	3.08	0.7673
	0.75	34.20	3.17	3.32	0.5463
	1	31.67	2.38	3.49	0.4729
	1.25	29.49	1.90	3.62	0.5109
12 $C_s(0)=42.54$ Kg/Kg	0.25	16.19	9.51	3.37	0.6509
	0.5	14.78	4.76	3.08	0.7011
	0.75	13.60	3.17	3.32	0.8165
	1	12.59	2.38	3.49	0.8379
	1.25	11.72	1.90	3.62	0.8517

Tableau IV.10 : les coefficients de diffusion obtenus pour chaque expérience.

L'expérience	P	$C_1(\infty)$ (mg/ml)	α	q	R^2	D (m ² /s)
1	1.25	29.5	1.9	3.5	0.3121	1.58E-11
2	0.5	37.17	4.76	3.083	0.4800	1.89E-12
3	0.5	37.17	4.76	3.083	0.5671	7.10E-12
5	0.5	14.78	4.76	3.083	0.8303	2.90E-11
6	0.75	13.6	3.17	3.31	0.8588	3.65E-11
7	1.25	11.72	1.9	3.5	0.517	6.89E-11
9	1.25	29.5	1.9	3.5	0.9832	1.96E-11
10	0.5	37.13	4.76	3.083	0.70	8.40E-12
11	0.5	37.17	4.76	3.083	0.7673	1.28E-11
12	0.5	14.78	4.76	3.083	0.9007	4.21E-11

IV.3.1.2 Sur la base des résultats simulés par le réseau:

- évaluation de la concentration à l'équilibre

On définit le t_{inf} le temps correspondant à C_∞ comme étant l'instant après lequel l'augmentation du rendement en extrait n'est plus significative, et dans le cadre de notre application nous avons considéré une évolution de 1% comme étant in significative.

Pour cela, et à l'aide d'un programme Matlab nous avons déterminé la concentration à l'infini et l'instant correspondant pour chaque expérience, Les résultats obtenus sont illustrés dans le tableau suivant :

Tableau IV.11 : les couples (C_{inf} , t_{inf}) pour chaque expérience

N° d'expérience	C_{inf} (mg/ml)	t_{inf} (min)
1	22,4	505
2	23,92	710
3	17,28	185
4	29,33	3380
5	8,16	225
6	8,35	185
7	8,9	185
8	4,83	205
9	39,23	1950
10	39,27	2120

11	17,55	185
12	8,16	230

Sur la base de l'équation (IV.2) et en utilisant les valeurs de $C_i(\infty)$ calculées par le réseau (tableau IV.11), On trace pour chaque expérience $-\ln\left(1 - \frac{C_i(t)}{C_i(\infty)}\right) = f(t)$, la tendance de

la courbe se trouve de la forme $y = a + b.x$ où :

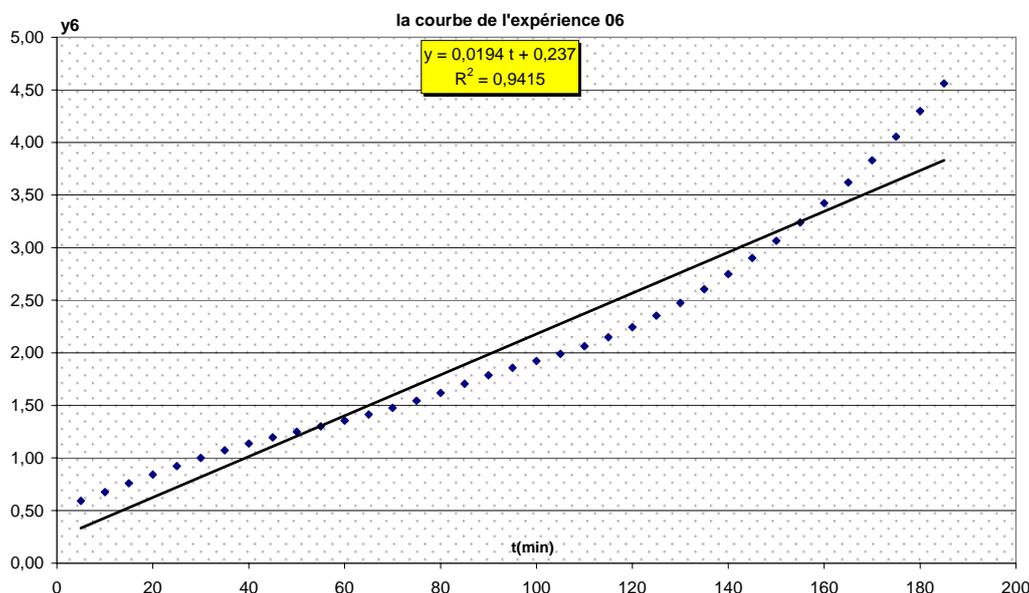
$$\begin{cases} b = \frac{Dq_1^2}{R^2} \\ a = -\ln \frac{6\alpha(1+\alpha)}{9+9\alpha+\alpha^2q_1^2} \end{cases}$$


Figure IV.16 : graphe représentant $-\ln\left(1 - \frac{C_i(t)}{C_i(\infty)}\right) = f(t)$ de l'expérience 6

Après cela, les paramètres α et q_1 , s'obtiendront en solvant le système d'équations suivant :

$$\begin{cases} -\ln \frac{6\alpha(1+\alpha)}{9+9\alpha+\alpha^2q_1^2} = a \\ \tan q_1 = \frac{3 \cdot q_1}{3 + \alpha \cdot q_1^2} \end{cases}$$

Une fois, déterminer les couples (α, q_1) , on calcule le coefficient de diffusion par la relation :

$$D = \frac{b \cdot R^2}{q_1^2} \text{ .le tableau (IV.12) indique les résultats de calcul :}$$

Tableau IV.12 : les résultats obtenus par le modèle rapporté par Crank, en se basant sur les valeurs simulées par le réseau.

N° de l'expérience	q_1	α	$D(m^2 / s)$
1	2,05	2.21	1,05772E-10
2	2,21	2.42	7,02827E-11
3	1,81	5.51	2,32436E-10
4	1,35	2.25	2,74126E-11
5	2,1	3.12	2,94545E-10
6	2,23	2.13	2,16554E-10
7	2,41	1.71	2,93413E-10
8	1,82	5.12	4,34751E-10
9	2,6	1.42	2,31139E-12
10	2,58	1.45	8,56945E-13
11	1,71	10.42	2,96109E-10
12	1,92	3.73	3,5989E-10

La différence entre les valeurs de coefficient de diffusion donnée par le modèle rapporté par Crank sur la base des données expérimentales et celles obtenues par le réseau est due essentiellement à la différence entre les valeurs de $C_l(\infty)$ évaluées, l'apprentissage suggère une projection à l'infini de la base alors qu'en pratique nous ne pouvons pas exclure le fait qu'il ait un changement de tendance au cours de processus, la seule façon de le vérifier est de mener l'extraction jusqu'à atteindre une vitesse de transfert nulle, voire négligeable.

IV.4 calcul de coefficient de transfert de matière

Sur la base des valeurs données dans le tableau IV.4, on trace pour chaque expérience.

$-\ln\left(1 - \frac{C_l(t)}{C_{l\infty}}\right) = f(t)$. La pente de la droite correspond à la valeur de λ , puis la valeur de K est

tirée par : $K = \lambda (1-s)$.

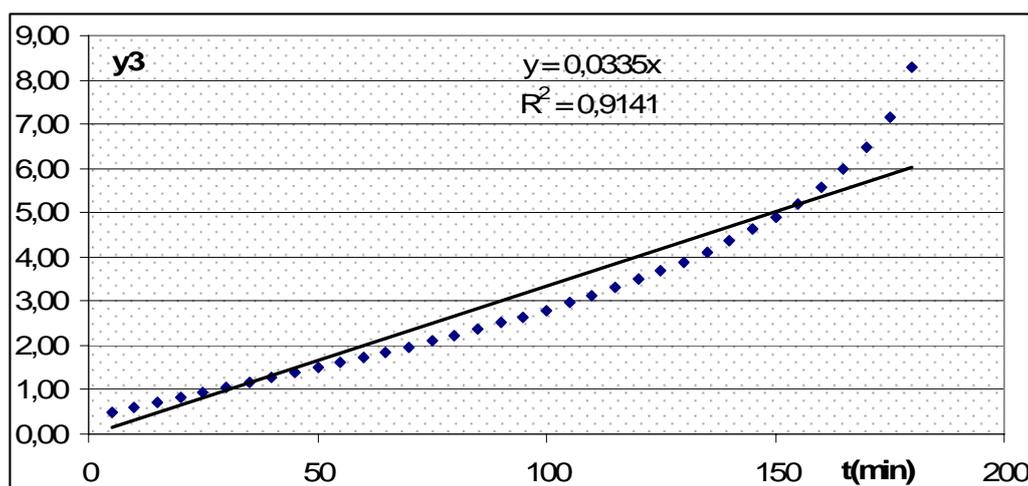


Figure IV.17 : graphe représentant $-\ln\left(1 - \frac{C_l(t)}{C_{l\infty}}\right) = f(t)$ de l'expérience 3

Tableau IV.13 : les résultats obtenus par le modèle basé sur le coefficient de transfert de matière.

N° d'expérience	$\lambda(\text{min}^{-1})$	$K(\text{s}^{-1})$	R^2 (coefficient de corrélation)
1	0,0106	5,23E-05	0,8857
2	0,0065	3,21E-05	0,7962
3	0,0335	1,65E-04	0,9141
4	0,001	3,88E-06	0,6545
5	0,0236	1,16E-04	0,9392
6	0,0213	1,05E-04	0,9297
7	0,0338	1,67E-04	0,8265
8	0,0237	1,17E-04	0,895
9	0,0029	1,43E-05	0,7394
10	0,0011	5,43E-06	0,1655
11	0,0355	1,75E-04	0,8882
12	0,0232	1,14E-04	0,9174

Conclusion générale

L'outil neuronal, par son entraînement, nous a permis de construire un modèle simple et pratique pour la simulation de la cinétique d'extraction des essences végétales (coriandre) et d'atteindre les valeurs des coefficients de transfert de matière et de diffusion par l'évaluation de la concentration en extrait à l'équilibre, or la valeur de celle-ci résulte souvent d'hypothèse relative au coefficient de partage.

En revanche, l'écart constaté entre les valeurs données par la simulation et celles données en se basant sur l'expérience était de 10 fois plus grand pour certains et moins pour d'autres. Cela est justifié par le grand écart entre les valeurs de la concentration à l'équilibre évaluées par les deux méthodes, vérifier cela revient à mener l'extraction jusqu'à atteindre une vitesse de transfert nulle, voire négligeable.

Le modèle neuronal bien qu'il présente des propriétés intéressantes à savoir l'approximation universelle et la parcimonie, quelques désavantages sont à noter :

- Son performance est directement liée à la qualité de son entraînement, ce qui est assuré par la disposition d'un ensemble d'apprentissage large et représentatif du processus à modéliser, c'est à dire des séries d'essais menées jusqu'au voisinage de l'équilibre.
- La bonne conception d'un modèle neuronal n'est pas fondée sur des lois et des relations, mais plutôt sur la base des constatations expérimentales, comme c'est le cas du nombre de neurones cachés, le choix des fonctions d'activations, la base d'apprentissage et les valeurs initiales des poids synaptiques, ce qui conduit à envisager plusieurs schémas avant d'aboutir à un modèle et des résultats satisfaisants.
- Les RNA restent encore, en partie, une boîte noire car ils ne permettent pas d'expliquer les relations pertinentes entre les variables.

Nous avons tenté lors de cette étude, de relier la forme des solutions analytiques à la base d'apprentissage, c'est à dire l'élaboration d'un modèle conjuguant les avantages des RNA tout en tenant compte des lois physiques régissant le phénomène, mais les résultats obtenus n'ont pas été satisfaisants, ce qui peut être dû au choix des bases d'apprentissage. Cette approche mérite d'être approfondie, ce qui peut faire l'objet de travaux futurs.

BIBLIOGRAPHIE

- [1] : G.DREYFUS.
Réseaux de neurones : méthodologie et application.
Ed. Eyrolles, 2002.
- [2] : R.LADJADJ.
Les réseaux de neurones.
Ed. Ingénieur 2000, 2003
- [3]: A.FREEMAN & M.SKAPURA
Neural networks algorithms and programing techniques
Addison Wesley 1992.
- [4] : J. F JODOUIN.
Les réseaux de neurones,
Ed Hermès (1994), Paris
- [5] : S. AMROUCHE
Prévision des variables climatologiques aux pas de temps mensuel et
décadaire par les réseaux de neurones et l'Analyse en Composantes
principales.
Mémoire de fin d'études, Département d'Hydraulique Ecole Nationale
Polytechnique, Alger 2006.
- [6] : A. MILOUDI & M. MEGHOUFEL.
Conception d'un modèle de prévision basé sur les réseaux de neurones
artificiels.
Mémoire de fin d'études, Département G. Industriel Ecole Nationale
Polytechnique, Alger 2005.
- [7]: C.DECAESTECKER & M.SAERENS.
Réseaux de neurones artificiels.
LINF, 2003.
- [8] : N.AOUCHAR & C.BENKHLED
Application des systèmes hybrides neuro-flous au diagnostic des huiles de
transformateurs.
Mémoire de fin d'études, département de G. Electrique Ecole Nationale
- [9] : L.MOKHNACHE.
Application des réseaux de neurones dans le diagnostic et la prédiction des
isolations de haute tension.
Thèse de doctorat. Département de génie électrique, Ecole Nationale
Polytechnique,Alger 2004.
- [10] : P. Mafarat et E. Béliard,
Génie Industriel Alimentaire Tome I,
Ed Tech et doc ; Lavoisier, Paris 1992.
- [11] : F.KOLAI
Modélisation du transfert de matière lors de l'extraction de l'essence de
coriandre par l'éthanol et le binaire éthanol eau .

Mémoire de fin d'études, Département de génie chimique Ecole Nationale Polytechnique, Alger 2002.

[12] : F.AMOKRANE

Modélisation du transfert de matière lors de l'extraction de l'essence des graines de coriandre par le méthanol.

Mémoire de Magistère, Département de génie chimique Ecole Nationale Polytechnique, Alger 2002.

[13] : N.CHIBANE

Modélisation du transfert de matière lors de l'extraction des essences des graines de coriandre par l'éthanol.

Mémoire de fin d'études, Département de génie chimique Ecole Nationale Polytechnique, Alger 2007

[14]: J.CRANK,

The Mathematics of Diffusion.

Clarendon Press, Oxford, London 1956

[15] : .BOUZID

Détermination du coefficient de diffusion lors de l'extraction des graines de coriandre par le méthanol.

Mémoire de fin d'études, Département de génie chimique Ecole Nationale Polytechnique, Alger 2002.

[16] : E.DAVALO & P.NAIM.

Des réseaux de neurones.

Ed. Eyrolles, 1992.

Annexes

ANNEXE 1 : « Généralités sur les réseaux de neurones »**I. Historique des réseaux de neurones [11] :**

Bien que le concept de la mémoire associative était introduite pour la première fois par le célèbre psychologue américain **W. James** en 1890, qui propose ce qui deviendra une loi de fonctionnement pour l'apprentissage sur les réseaux de neurones connues plus tard sous le nom de **la loi de Hebb**, l'étude des réseaux de neurones, proprement dite, date des années 1940-50. Héritière de la révolution scientifique qui mobilisait les esprits à cette période, cette discipline trouve ses sources à la fois dans les connaissances neurobiologiques de l'époque et dans un nouvel outil permettant d'explorer par simulation le comportement des modèles théoriques proposés, *l'ordinateur*.

Cette décennie vit aussi plusieurs travaux importants en particulier les réflexions sur la nature du raisonnement humain de **John Von Neumann**, d'**Alan Turing**, de **Alonzo Church**, d'**Alan Newell** et de **Herbert Simon** qui galvanisèrent les recherches sur le cerveau et l'esprit humain et posèrent les fondations de ce qui allait devenir *l'intelligence artificielle*.

Alan Turing et **John Von Neumann** étaient parmi les premiers à s'être intéressés au cerveau humain. Ils ont l'idée d'une machine basée sur le modèle de neurone artificiel proposé en 1943 par le neurophysiologiste **Warren McCulloch** et le logicien **Walter Pitts**. Bien qu'il soit essentiellement connu pour la machine qui porte son nom, **Alan Turing** est aussi un des pères des réseaux de neurones : dès 1948 il les étudie dans un article intitulé « *Intelligent Machinery* » publié en 1968, 14 ans après sa mort où il décrit les principes du connexionnisme. Il introduit une machine constituée de neurones artificiels associée à des dispositifs permettant de modifier leurs connexions. La publication tardive de ces travaux fait que l'on ne redécouvre que plus tard les réseaux de neurones.

En 1949, **Hebb** souligna l'importance du couplage synaptique dans le processus d'apprentissage en se fondant son argumentation sur le comportement psychophysiologique: première règle d'apprentissage qui il la publié dans un livre intitulé « *Organization of Behavior* ».

Ainsi, les premiers réseaux de neurones artificiels basés sur la vision des systèmes biologiques apparaissent sous le terme *perceptron* proposé en 1958 par **Frank Rosenblatt**. Inspiré des idées de **Hebb**, **Mc Culloch** et **Pitts**, *le perceptron* était en mesure d'apprendre à calculer certaines fonctions logiques en modifiant ses connexions synaptiques.

Le *perceptron*, qui est le réseau le plus simple possible, a été à l'origine développé pour la reconnaissance de formes. Par la suite, **Rosenblatt** en 1960, donne la preuve de convergence du *perceptron*. Dans la même époque, des réseaux similaires appelés **ADALINE** (*Adaptive Linear Neuron networks*) furent inventés.

En 1969, une critique violente du perceptron a été publiée dans un livre intitulé « *perceptrons : an introduction to computational geometry* » par **M.L Minsky et S.A Papert** où ils ont montré toutes les limites d'un tel modèle, son incapacité de résoudre les problèmes linéairement inséparables et le problème de XOR. Cela avait pour effet de réduire drastiquement la recherche sur les réseaux de neurones artificiels qui sont alors délaissés pendant une dizaine d'années.

Cependant, certains tels **Wilkie, Stonham et Aleksander** persistent et parviennent en 1980 à résoudre les problèmes d'apprentissage difficiles en utilisant, malgré tout, des réseaux de neurones multicouches. D'autres études telles celles de **Grossberg** (Grossberg1972, Grossberg1976a, et Grossberg1976b) et **Kohonen** en 1982 contribuent elles aussi au regain d'intérêt pour les réseaux de neurones formels et leurs dérivés. Dans la même période **werbos** invente l'idée de rétropropagation en 1974.

Par la suite se développent de nouveaux types de réseaux de neurones formels (**Hopfield**, 1982) et des méthodes d'apprentissage (algorithme de rétro propagation de l'erreur). Destiné aux réseaux avec des connexions non bouclées, l'algorithme de rétro propagation a été développé indépendamment par **Le Cun** (1985 et 1988), **Rumelhart** (1986b et 1986a), et **Parker** (1987).

Cela permettra alors la diffusion de l'utilisation des réseaux de neurones dans de nombreuses applications (reconnaissance de caractères, de la parole...).

Dans les années 90, d'autres modèles et algorithmes ont été développés avec l'attention d'étendre les possibilités des réseaux statiques non bouclés, introduction progressive d'une mémoire interne et de la récurrence au sein du réseau.

II. Fondements biologiques [10] [24] :

III.1 Structure et fonctionnement d'un neurone :

Les cellules nerveuses, appelées neurones, sont les éléments de base du système nerveux central. Ce dernier en posséderait environ cent milliards. Bien que les neurones possèdent des caractéristiques communes avec les autres cellules du corps, ils présentent cependant des caractéristiques qui leur sont propres et se retrouvent au niveau des fonctions spéciales qu'ils assurent :

- Recevoir des signaux en provenance des autres neurones.

- Engendre un influx nerveux.
- Le conduire, et le transmettre à un autre neurone.

La figure 1 dépeint les composants principaux d'une cellule nerveuse typique de dans le système nerveux central.

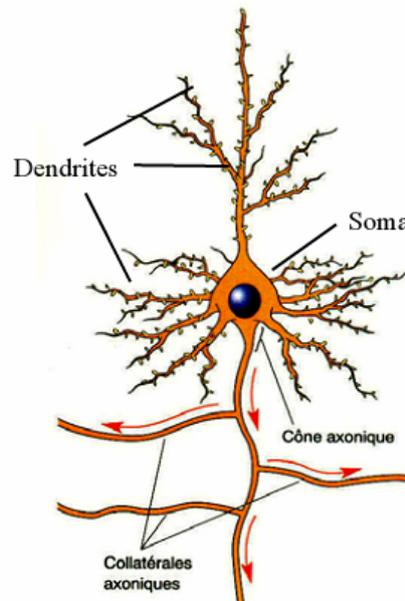


Figure1 : structure de base d'un neurone

Un neurone donc est constitué de trois parties :

1. **Le corps cellulaire (Soma)** : il contient le noyau et assurent la vie du neurone en synthétisant des enzymes et molécules nécessaire.
2. **Les dendrites** : elles sont les récepteurs principaux d'un neurone pour capter les signaux qui lui proviennent.
3. **L'axone** : sert de moyen de transport pour les signaux émis par le neurone. Il se distingue des dendrites par sa forme et par les propriétés de sa membrane externe.

La membrane du neurone agit pour maintenir une différence de potentielle dite d'équilibre (-70 mV) entre le milieu extérieur et l'intérieur de la cellule nerveuse. elle accomplit cette tâche par l'action de la pompe Sodium – Potassium .ce mécanisme transport les ions du Sodium à l'extérieur de la cellule et l'inverse pour celles de Potassium, il existent aussi des ions organiques négatives au plasma nerveuse et des ions de chlore à l'extérieur de celle-ci .

L'arrivé d'un influx nerveux généré par le corps provoque une diminution du potentiel de la membrane. Une telle dépolarisation change la perméabilité de la membrane aux ions de sodium en laissant pénétrer une quantité de celles-ci au plasma intérieur de la cellule.

Et comme les fibres nerveuses sont des mauvais conducteurs, la transmission de l'action potentielle se fait par des dépolarisations successives des nodes de Ranvier. Une

dépolarisation d'un node provoque la dépolarisation du suivant .une fois l'action potentiel passe un point donné, ce point rétablie son potentiel du repos en déclanchant les pompes de potassium.

4. La jonction synaptique :

Pour former le système nerveux, les neurones sont connectés les uns aux autres dans des endroits appelés **synapses** ou ils sont séparés par un petit espace synaptique de l'ordre de centième de microns.

Les synapses constituent les zones de jonction entre deux neurones ou entre un neurone et une cellule effectrice (cellule musculaire par exemple), et permettent la transmission d'une information de l'un à l'autre.

Dans une synapse chimique, l'extrémité de l'axone du neurone présynaptique contient des vésicules (petits organites entourés d'une membrane) remplies de neuromédiateurs.

L'arrivée d'un influx nerveux (potentiel d'action) provoque la migration des vésicules et leur fusion avec la membrane du neurone. Cette fusion permet la libération des molécules de neuromédiateurs dans l'espace synaptique. Ces molécules vont ensuite se fixer sur les récepteurs membranaires spécifiques de la cellule post-synaptique, ce qui déclenche chez cette dernière une réponse appropriée (nouvel influx nerveux s'il s'agit d'un neurone). Selon les synapses il existe plusieurs types de neurotransmetteurs :

- **Les neurotransmetteurs excitateurs** : ils excitent le neurone cible et produisent dans celui-ci une nouvelle impulsion.
- **Les neurotransmetteurs inhibiteurs** : ils bloquent l'impulsion nerveuse aux endroits synaptiques.

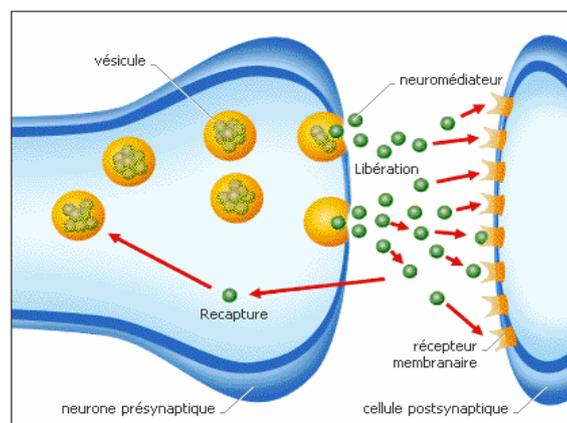


Figure2 : représentation du fonctionnement de la jonction synaptique

ANNEXE 2 : « Les Règles d'Apprentissage »

1. Apprentissage à correction d'erreur

Au cours de l'apprentissage à correction d'erreur, les poids des liens de communication du réseau sont ajustés en essayant de minimiser la fonction dite de coût qui dépend de la différence entre la sortie désirée et l'entrée exposée au réseau. La modification des poids est déterminée sur la base de l'erreur commise entre la sortie effective du réseau et celle désirée suivant plusieurs algorithmes, le plus connue peut être représenté comme suit:

$$\Delta w_{ki} = \eta x_{ik} (d_k - y_k)$$

Avec : Δw_{ki} la modification à affecter au neurone.

η le taux d'apprentissage.

d_k : la sortie désirée du neurone k.

y_k la sortie effective du neurone k .

x_{ik} : entrée du neurone i du neurone k.

Donc, le nouveau poids est proportionnel à la fois de l'erreur commise et l'ancien poids :

$$w_{ki}^{(nouveau)} = w_{ki}^{(ancien)} + \Delta w_{ki}$$

Un exemple d'une fonction d'erreur connue sous le nom « règle de Widrow-Hoff » est exprimée comme suit :

$$E_{Global} = \frac{1}{2p} \sum_{p=1}^p \sum_{k=1}^N (y_k^p - d_k^p)^2$$

Avec N : le nombre du neurone de la couche de sortie et P le nombre d'exemples dans l'échantillon d'apprentissage.

En générale les règles d'apprentissages sont des formes de descende de gradient

$$\Delta w_{ki} = -\eta \nabla_{w_{ki}} (E_{global})$$

La règle de Hebb [12] [24]:

«When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic changes take place in one or both cells such that A's efficiency as one of the cells firing B, is increased.»

Donald HEBB. [24]

Dans un contexte neurobiologique, Hebb cherchait à établir une forme d'apprentissage associatif au niveau cellulaire. Dans le contexte des réseaux artificiels, on peut reformuler l'énoncé de Hebb sous la forme d'une règle d'apprentissage en deux parties :

1. Si deux neurones de part et d'autre d'une synapse (connexion) sont activés simultanément (d'une manière synchrone), alors la force de cette synapse doit être augmentée.
2. Si les deux neurones sont activés d'une manière asynchrone, alors la synapse correspondante doit être affaiblie ou carrément éliminée.

Mathématiquement, on peut exprimer la règle de Hebb sous sa forme la plus simple par la formule suivante :

$$w_j(t) = w_j(t - 1) + \Delta w_j(t - 1)$$

Avec :
$$\Delta w_j(t - 1) = \eta p_j(t)a(t)$$

Où η est une constante positive qui détermine la vitesse de l'apprentissage, $p_j(t)$ correspond à l'activité pré-synaptique (l'entrée j du neurone) au temps t , et $a(t)$ à l'activité post-synaptique (sortie du neurone) à ce même temps t . Cette formule fait ressortir explicitement la corrélation entre le signal qui entre et celui qui sort. Sous une forme vectorielle, on écrit :

$$\Delta \mathbf{w}(t - 1) = \eta \mathbf{p}(t)\mathbf{a}(t)$$

Un problème immédiat avec la règle de l'équation 2.3 est que les changements de poids $\Delta w_j(t)$ peuvent croître de façon exponentielle si, par exemple, l'entrée et la sortie demeurent constantes dans le temps. Pour pallier à cette croissance exponentielle qui provoquerait invariablement une saturation du poids, on ajoute parfois un facteur d'oubli qui retranche de la variation de poids, une fraction α du poids actuel. On obtient ainsi :

$$\Delta w_j(t - 1) = \eta p_j(t)a(t) - \alpha w_j(t - 1)$$

Où $0 \leq \alpha \leq 1$ est une nouvelle constante. Sous forme vectorielle, on écrit :

$$\Delta \mathbf{w}(t - 1) = \eta \mathbf{p}(t)\mathbf{a}(t) - \alpha \mathbf{w}(t - 1)$$

Mais cette règle ne résout pas tous les problèmes. A cause du terme d'oubli, il est primordial que les stimuli soient répétés régulièrement, sinon les associations apprises grâce à la règle de l'équation 2.5 seront éventuellement perdues car complètement oubliées. Une autre variante de la règle de Hebb s'exprime donc de la manière suivante :

$$\Delta w_j(t - 1) = \eta p_j(t)a(t) - \alpha a(t)w_j(t - 1)$$

Et si l'on fixe $\alpha = \eta$ pour simplifier (on impose un rythme d'apprentissage égal à celui de l'oubli), on obtient la règle dite «instar» :

$$\Delta w_j(t - 1) = \eta a(t) [p_j(t) - w_j(t - 1)]$$

Que l'on peut réécrire sous forme vectorielle de la façon suivante :

$$\Delta \mathbf{w}(t - 1) = \eta \mathbf{a}(t) [\mathbf{p}(t) - \mathbf{w}(t - 1)]$$

ANNEXE 3 : « La coriandre »[15]

Les caractéristiques de cette espèce sont :

Nom:Coriandre

Nom latin:Coriandrum sativum.

Nom commun anglais : Coriandre.

Famille botanique : Umbellifère.

Cycle : annuel

Largeur : 20 cm

Hauteur : 75 cm

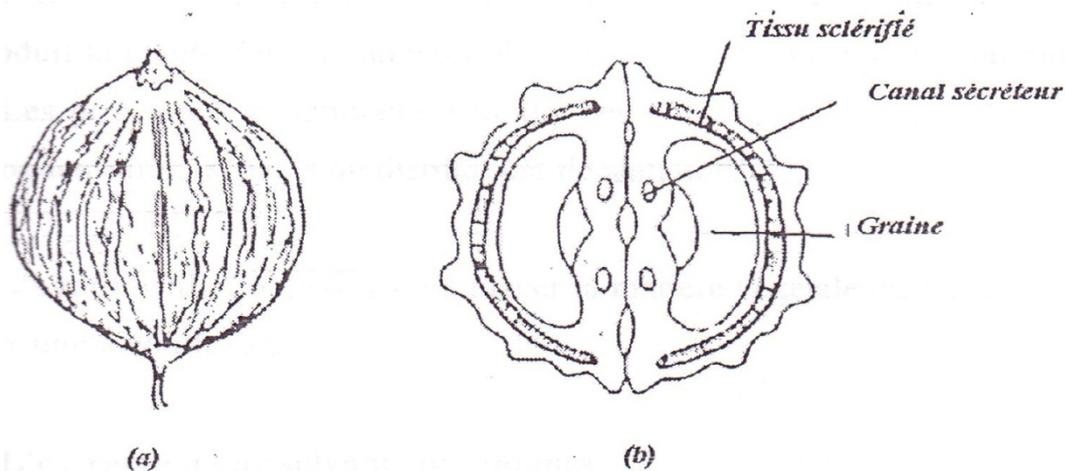


Figure1 : (a)Fruit, (b) Coupe transversale.

- **Culture et récolte**

Cette plante pousse dans toutes les régions tempérées. Elle préfère les sols sains, compacts, qu'ils soient sableux ou argileux, mais aussi des sols calcaires, perméables, frais, bien préparés et ensoleillés. On la trouve aussi dans des terrains acides. [3]

La culture est influencée principalement par le site de location, précisément la latitude. Il a été rapporté que dans un froid extrême et une courte saison de végétation, on obtient un bon rendement

- **Composition chimique :**

La graine de coriandre contient de l'eau, des protéines, des fibres, de l'amidon, des sucres, des minéraux, des lipides et de l'huile essentielle, leur teneur en pourcentage varie selon la terre et le climat où la plante été cultivé. A titre d'exemple, nous présentons dans le tableau I.1. la composition d'huiles obtenues par hydrodistillation des grains de coriandre

Tableau1 : composition d'huiles essentielles obtenues par hydrodistillation des grains de coriandre [13,14]

Composé	Teneur	Composé	Teneur
α-pinène	0.31	Octanal	3.07
Myrcène	0.16	Dodecane	0.53
p-cymene	0.13	Geranyl acetate	1.75
Cineol-1,8	0.21	Dodecanal	Tr
γ-terpinene	0.08	β-caryophyllene	0.24
Linalool	80.33	α-humulene	0.33
Camphre	1.36	Nerolidol	0.15
Terpinene 4-ol	0.60	Myritic acid	4.04
α-terpineol	0.42		

ANNEXE 4 : « Propriétés des Réseaux établis »

Réseau entraîné à 120 :

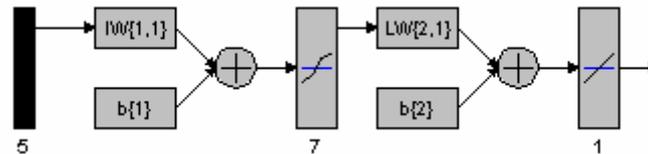


Schéma représentatif du réseau.

Les propriétés de la couche cachée :

- Les poids synaptique des connexions entrantes:

[0.79709 0.92989 2.6286 -0.80883 0.003495; -0.48722 0.22125 0.30634 0.044029 0.042028;
-3.8404 0.12479 -0.76949 0.77164 -0.012692; 0.78332 -1.7494 -0.21122 2.2616 -0.58286;
-0.40809 -1.9491 0.20881 2.9846 0.027461; 3.1879 -0.23128 3.6354 -0.30105 -0.010675;
-0.039853 0.045679 0.030272 1.3466 0.012175]

- Le bais chaque neurone: [-3.111; -5.7211; 10.36; -17.7709; 17.8135; -7.0342; -10.5374]

Les propriétés de la couche de sortie :

- Poids synaptique des connexions entrantes :

[-3.7004 1.7003 2.1604 0.19683 -5.3061 -4.2652 9.745]

- Le bais : [6.961].

Réseau entraîné jusqu'à 180 min :

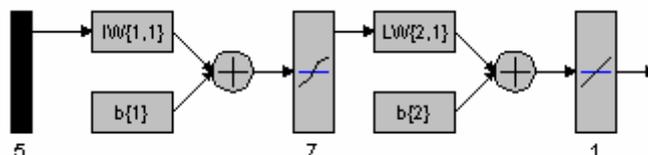


Schéma représentatif du réseau.

La couche cachée :

- Poids de connexions entrantes :

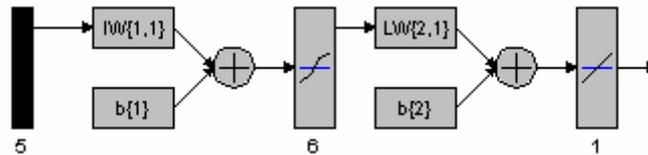
[-0.051715 0.27329 0.019078 -0.0082447 -0.02472; 0.94894 0.4626 1.7741 -0.040694 0.04205;
-0.71068 1.1027 2.466 -0.23614 0.0080442; -0.15334 0.74202 2.5333 -5.6339 0.00077746;
-0.12181 0.077839 -0.45274 0.025007 -0.010555; -3.2435 -0.042586 5.2457 1.4071 0.006563;
51.6857 -0.31566 -2.6913 -13.5145 -0.0014191]

- Le bais de chaque neurone :

[-7.3179; -19.439; -20.7205; 13.5371; -0.2241; -18.1645; 15.3167]

La couche de sortie :

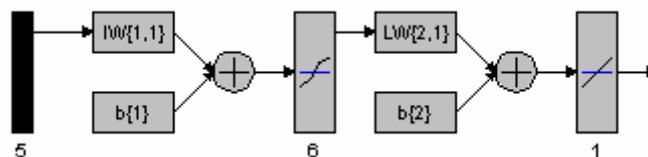
- Poids des connexions : [-13.5479 -0.5106 -2.1907 16.1578 -6.2124 -3.2789 -4.9481]
- Le biais : [5.6732]

Approche 2 :**Réseau entraîné de [95 à 240 min] :***Schéma représentatif du réseau.***Les propriétés de la couche cachée :**

- Les poids synaptique des connexions entrantes:
[13.7353 -1.2712 4.6238 -2.2376 0.001071; 1.1339 0.18281 -1.0066 0.043875 0.0042924;
-0.042996 1.2691 -5.6359 -1.8184 0.0080171; 1.3668 0.73895 -1.3065 -0.34557 0.0080177;
-30.3662 1.2765 2.7123 3.0658 0.00088331; -2.3044 -0.0060847 1.3331 0.069462 -0.0047281]
- Le biais : [16.1802; -5.6368; -2.1049; -16.5841; -6.5757; 4.534]

Les propriétés de la couche de sortie :

- Les poids synaptique des connexions entrantes:
[12.1567 5.6787 8.5998 -3.3376 10.7694 2.6844]
- Le biais chaque neurone : [8.7281].

Réseau entraîné de [5 à 240 min] :**Les propriétés de la couche cachée :**

- Les poids synaptique des connexions entrantes:
[0.09594 0.68878 1.9381 -1.899 -0.01215; 0.88032 10.6045 -35.2613 -6.0228 0.57693;
-0.068467 -0.066742 4.0041 -0.0046071 0.035668; 2.6256 1.7436 0.39908 0.073233 0.0050437;
0.12956 -0.25263 -2.9457 0.14336 0.0087325; 66.8808 -4.3287 -2.0996 -1.5949 -0.01522]
- Le biais : [-10.3616; -1.9861; -9.564; -48.8288; 11.2806; 4.1839]

Les propriétés de la couche de sortie :

- Les poids synaptique des connexions entrantes:

[-14.095 1.4513 5.1923 7.4724 2.8129 -6.7511]

- Le bais : [-1.7415].