

**INSTITUT NATIONAL AGRONOMIQUE**

Pour obtenir le titre de Magister En SCIENCES AGRONOMIQUES

Spécialité : Génie Rural

Option : HYDRAULIQUE

# *Contribution à la modélisation de la qualité des eaux par les réseaux de neurones*

Mémoire présenté par :

**HEDDAM SALIM**

Directeur de thèse : Mr. N.DECHEMI Professeur ENP) Co-Directeur de thèse : Mr.

A.BERMAD Maître de conférences (ENP)

Promotion: 2006

Président : Mr. A.AIDAOUI Professeur (INA) Examineurs : Mr. T.HARTANI Maître de conférences  
(ina) Mr. F.SELAM Chargé de cours (ina) Invités : Mr. A.BOUDJNOUN Algérienne des eaux (Alger)  
Mr. N.ADIM SEAAL (Alger)



# Table des matières

Remerciements . .	1
..	3
Résumé .	5
قرئندملا ص خلم .	7
Abstract . .	9
Introduction générale . .	11
Chapitre I. Principes généraux de traitement des eaux .	13
.I.1. Introduction .	13
.I.2. Prétraitement .	14
I.3. Pré oxydation .	15
I.4. Clarification .	15
I.4.1 Introduction .	15
I.4.2 Coagulation-Floculation .	16
I.4.3. Essais de Coagulation – Floculation .	20
I.4.4. La séparation solide-liquide .	20
I.5. Désinfection . .	22
I.6. Affinage .	22
I.7. Réservoirs de stockage .	22
Chapitre II. Analyse en composantes principales . .	25
II.1.Historique de l'analyse des données .	25
.II.2. Définitions & contexte de base .	26
II.2.1. Notion d'individu et de caractère . .	27
II.2.2. Définition algébrique . .	27
II.2.3. Définition géométrique .	29
II.3. Formulation mathématique du problème .	29
II.4. Choix de la métrique .	32

II.5. Procédé d'application de l'ACP .	32
II.5.1. Calcul de la matrice de variance covariance .	32
II.5.2. Calcul des valeurs et vecteurs propres . .	32
II.5.3. Recherche des axes principaux . .	36
II.5.4. Principaux résultats : .	37
II.5.5. Représentation graphique : .	37
II.5.6. Programmation de l'ACP . .	37
II.5.7. Fichier de données (input) .	37
II.5.8. Organigramme de calcul . .	38
II.6. Organigramme du programme .	38
II.7. Conclusion .	39
Chapitre III.Théorie des réseaux de neurones . .	41
III.1. Présentation et historique .	41
III.2. Qu'est-ce qu'un neurone ? . .	42
III.2.1. Une morphologie particulière .	42
III.3. Définitions .	42
III.4. Historique .	43
III.5. Le modèle neurone formel . .	44
III.5.1. Structure .	45
III.5.2. Variables descriptives .	46
III.5.3.Architecture .	47
III.6. Processus d'apprentissage des réseaux de neurones .	47
III.6.1. Définition . .	48
III.6.2. Types d'Apprentissage .	48
III.7. Le perceptron et le perceptron multicouches .	49
III.7.1. Le perceptron simple .	49
III.7.2. Adaline et Widrow-Hoff .	50
III.8. Conclusion . .	51
Chapitre IV.Elaboration du modèle . .	53

<b>IV.1. Introduction . .</b>	<b>53</b>
<b>IV.2. Problématique . .</b>	<b>54</b>
<b>IV.3. Etat de l'art de la modélisation .</b>	<b>55</b>
<b>IV.4. Présentation de la station de traitement .</b>	<b>56</b>
<b>IV.5. Phases de l'élaboration du modèle . .</b>	<b>56</b>
<b>IV.5.1. Collecte des données .</b>	<b>57</b>
<b>IV.5.2. Analyse des données . .</b>	<b>57</b>
<b>IV.5.3. Séparation de la base de données .</b>	<b>57</b>
<b>IV.5.4. Choix de l'architecture du réseau . .</b>	<b>57</b>
<b>IV.5.5. Mise en forme des données .</b>	<b>58</b>
<b>IV.5.6. Apprentissage du réseau de neurones .</b>	<b>58</b>
<b>IV.5.7. Validation .</b>	<b>58</b>
<b>IV.6. Elaboration du modèle . .</b>	<b>59</b>
<b>IV.6.1. Logiciel, méthodologie d'implémentation .</b>	<b>59</b>
<b>IV.6.2. Conception des algorithmes . .</b>	<b>60</b>
<b>IV.6.3. Algorithme général du programme .</b>	<b>61</b>
<b>IV.6.4. Organigramme du programme . .</b>	<b>61</b>
<b>IV.6.5. L'algorithme du rétropropagation du gradient .</b>	<b>62</b>
<b>IV.6.6. Fonctions coût .</b>	<b>62</b>
<b>IV.7. Implémentation de l'algorithme du gradient .</b>	<b>63</b>
<b>IV.7.1. Formulation mathématique du problème . .</b>	<b>63</b>
<b>IV.7.2. Critères de validation. .</b>	<b>68</b>
<b>IV.7.3. Choix d'initialisation des paramètres . .</b>	<b>69</b>
<b>IV.8. Applications .</b>	<b>70</b>
<b>IV.8.1. Base de données. . .</b>	<b>70</b>
<b>IV.8.2. Statistique des paramètres étudiés .</b>	<b>70</b>
<b>IV.8.3. Graphes des variations annuelles .</b>	<b>71</b>
<b>IV.8.4. Application de l'ACP .</b>	<b>72</b>
<b>IV.8.5. Application des réseaux de neurones . .</b>	<b>78</b>

<b>IV.8.6 Etude de l'effet saisonnier .</b>	<b>80</b>
<b>IV.9.Tentatives d'amélioration .</b>	<b>83</b>
<b>IV.10. Prévission de la dose du coagulant par le modèle retenu .</b>	<b>89</b>
<b>Conclusion générale .</b>	<b>91</b>
<b>Références bibliographiques .</b>	<b>93</b>
<b>Annexe A Résultats des ACP (Fichiers saisonniers) .</b>	<b>99</b>
Annexe A.1: Résultats d'application de l'analyse en composantes principales Saison hiver (décembre-janvier-février). .	99
Annexe A.2 : Résultats d'application de l'analyse en composantes principales Saison: printemps (mars-avril-mai). . .	102
Annexe A.3 :Résultats d'application de l'analyse en composante principale Saison été (juin- juillet -août). .	104
Annexe A.4 :Résultats d'application de l'analyse en composantes principales Saison automne (septembre-octobre-novembre). .	107

## Remerciements

Comment débiter ces remerciements sans exprimer les gratitudes que j'ai envers mon directeur et co-directeur du mémoire qui m'ont accueilli chaleureusement ?

Je tiens à remercier en premier lieu mon directeur de mémoire de magistère, le professeur **N. DECHEMI**, d'avoir accepté l'encadrement de ce travail, sa disponibilité et sa patience a été pour moi d'un grand soutien pour l'achèvement de ce travail.

Mon co-directeur de mémoire de magistère, docteur **A. BERMAD**, mérite tous les remerciements; pour sa participation qui était très enrichissante; Je le remercie vivement pour l'intérêt qu'il a porté à ce travail, pour ses précieux conseils, sa disponibilité, et sa sympathie.

A l'issue de ce travail, je tiens à exprimer ma gratitude aux nombreuses personnes qui ont contribué à son élaboration d'une manière directe ou indirecte. Que ceux que j'aurais pu oublier m'en excusent.

Tout d'abord je présente mes chaleureux remerciements au professeur **AIDAOUI ABDELLAH**, non seulement parce que il a accepté de présider mon jury de soutenance mais pour tout ce qu'il fait pour la réussite de la **promotion de post-graduation hydraulique agricole (2004-2005)**, il mérite tous les remerciements.

Les membres du jury de l'intérêt manifesté à cette étude et d'avoir accepté de bien vouloir l'examiner ;

**M. HARTANI TAREK**, qui a accepté d'examiner ce mémoire de magistère, je le remercie vivement.

**M. SELLAM FOUAD** d'avoir accepté d'examiner ce travail de recherche, je lui en suis très reconnaissant.

J'exprime mes vifs remerciements à **M. ADIM NABIL** chef service du laboratoire de la station de traitement des eaux de **BOUDOUAOU (SEAAL)** qui est toujours disponible pour me faire profiter de ses connaissances en traitement des eaux et pour les nombreuses orientations qu'il m'a donné durant la réalisation de ce travail; à lui et tous le personnel de la station de **BOUDOUAOU**, je présente mes chaleureux remerciements.

**M. BOUDJNOUN ABEDLOUAHEB, (ADE ALGER )** pour m'avoir accepté en tant qu'étudiant stagiaire au sein de l'ADE, je le remercie vivement pour ses encouragements et pour sa sympathie.

Je tiens à remercier monsieur **BOUKHADRA ABDERAHMANE**, mon enseignant en graduation pour tous les encouragement qu'il m'a donné tous le long de mon cursus.

Je tiens à remercier tous mes **collègues de post graduation**; leurs connaissance m'a fait un grand plaisir.

Enfin, j'exprime un grand merci à toutes les personnes faisant vivre la science.





---

*Dédicaces*

A tous ceux qui m'aiment A toute ma famille A toute ma famille A toute ma famille A  
l'âme de mon père; son souci était de voir ce travail A eux tous je dédiece travail



## Résumé

La modélisation du procédé de coagulation dans le domaine de traitement des eaux potables en utilisant des approches traditionnelles s'avère complexe, ceci est dû à la quantité et à la complexité des informations et des tests à réaliser dans le procédé. Dans ce mémoire, les réseaux de neurones artificiels (RNA) ont été utilisés; dans le but d'optimiser le contrôle de la dose du coagulant utilisé, car les RN ont la capacité de trouver les rapports multi variés et non linéaires entre les variables d'un système. Deux objectifs principaux ont été poursuivis, à savoir, la détermination de la dose du coagulant à injecter en fonction des paramètres descripteurs de la qualité de l'eau brute à l'entrée de la station, ainsi que la prévision de cette dose. On a constitué une banque de données provenant de la station de traitement des eaux de Boudouaou. Nous avons retenu 06 paramètres comme entrée du réseau de neurones (température, PH, turbidité, conductivité, oxygène dissous, les UV), et la dose du coagulant comme sortie. L'analyse en composantes principales a été appliquée afin de déceler d'éventuels effets prépondérants des différents paramètres pris en considération. Le modèle élaboré dans ce mémoire a donné de très bons résultats que ce soit en modélisation de la dose du coagulant, ou dans la prévision.

Mots clé : traitements des eaux ; coagulation; réseaux de neurones.



## قرنكذملا ص خلم

توفد هذه الدراسة إلى جريدة "الوقت" بدمشق في 10/10/2008م. وقد تم نشرها في العدد 10/10/2008م. وقد تم نشرها في العدد 10/10/2008م. وقد تم نشرها في العدد 10/10/2008م.

عمية الفعالة إلى قضية تمت بواسطة الطريقة العلمية المبرورة بلم استجبت الحسوية الاصطناعية

### RESEAU X DE NEURONES APPLICABLES

قد تم أولاً بتحديد وإعداد تلك المعلومات. انظر إلى أبحاث مجلة معالجة العبد الفعالة بدمشق 10/10/2008م. وقد تم نشرها في العدد 10/10/2008م. وقد تم نشرها في العدد 10/10/2008م. وقد تم نشرها في العدد 10/10/2008م.



## Abstract

The modeling of the process of coagulation in the drinkable field of water treatment by using traditional approaches would prove to be complex and very expensive due to the quantity and the complexity of information and the tests which must be carried out in the process. In this thesis, the artificial neural networks (ARN) were used for this purpose; with an aim of optimizing the control of the amount of the coagulant used, because the ANN have the capacity to find the multi reports/ratios varied and nonlinear between the variables of a system. Two principal objects were continued, namely, to determine the parameters of the transformation (variable of the quality of raw water to the entry of the station to be used as entered of the network of neuron) which influence more on the variation of the amount of the coagulant, like predicting this amount with ARN. One initially worked out a bank of data coming from a station of water treatment (station of BOUDOUAOU), sufficient to explain the succession of the seasons. In the bank, one retained 06 parameters as entered of the network of neurons (temperature, pH, turbidity, conductivity, dissolved oxygen, UV); among several recorded parameters. We applied the principal components analyze.

Keywords: water treatment; coagulant; artificial neural network





## Introduction générale

Toutes les eaux de la nature ne sont pas bonnes à boire. Même une eau d'apparence limpide transporte en son sein toutes sortes de substances inertes et vivantes, dont certaines peuvent être nocives pour l'organisme humain. Ces substances proviennent soit du milieu physique dans lequel l'eau a évolué, soit des rejets de certaines activités humaines dont l'eau est devenue le réceptacle. L'eau est ainsi le vecteur de transmission privilégié de nombreuses maladies. Pour pouvoir être consommée sans danger, l'eau doit donc être traitée. Mais la pollution croissante des réserves rend cette opération de plus en plus délicate, obligeant les traiteurs d'eau à constamment innover. Les techniques ont d'ailleurs beaucoup évolué, faisant aujourd'hui du traitement de l'eau une industrie de pointe.

Presque tous les services publics d'eau cherchent à améliorer le fonctionnement de leurs stations pour contrôler leurs coûts et rencontrer les exigences réglementaires sévères de la qualité de l'eau. Une amélioration du contrôle des procédés et de l'automatisation des stations peut permettre d'atteindre ce but. Leur fonction est, schématiquement de permettre au système de production de suivre l'évolution de la qualité de l'eau depuis son entrée à la station, jusqu'à la dernière étape de traitement, afin d'assurer en permanence l'adéquation entre « la qualité de l'eau à atteindre à la sortie de la station », et les moyens à débloquer pour atteindre le dit objectif, en particulier les produits chimiques à utiliser qui coûtent souvent très chers.

Pour accroître la flexibilité et la réactivité du système de production, assurer sa pérennité et limiter le temps d'intervention au sein des usines de production d'eau potable,

nous proposons l'introduction des méthodes connexionistes, une nouvelle architecture pour la prévision de la dose du coagulant à injecter en fonction des paramètres de qualité de l'eau brute.

Bien que l'approche par réseau neuronal artificiel (RNA) ait été utilisée dans diverses disciplines d'ingénierie depuis les années 1980, son utilisation dans l'industrie d'eau potable est relativement nouvelle et présente un bon potentiel.

Le traitement de l'eau fait partie de cette famille des problèmes complexes exigeant des qualités pour être entièrement appréhendée, c'est pour ces raisons que notre attention s'est portée sur cette technique de modélisation.

Nous proposons dans le cadre de ce mémoire de modéliser, la dose du coagulant à injecter pendant la phase de clarification, par le biais des réseaux de neurones artificiels.

Et comme toute modélisation nécessite des données, afin d'en extraire les régularités de façon à comprendre le mécanisme qui leur a donnée naissance, nous avons recueilli une banque de données mise gracieusement, à notre disposition, par la société eau et assainissement d'Alger (SEAAL); de Boudouaou et qui est constituée de sept paramètres (température, turbidité, conductivité, oxygène dissous, les ultraviolets , le PH); et bien sur la variable à expliquer qui est la dose du coagulant sur une période d'observations de 13 années.

L'objectif de notre travail de recherche est de modéliser, et de prévoir la dose du coagulant à injecter lors de la phase de coagulation-floculation dans la station de traitement des eaux de Boudouaou en fonction des paramètres descripteurs de la qualité de l'eau brute mesurés. Pour mener à bien ce travail nous avons adopté la démarche suivante :

Le **premier chapitre** est consacré à une étude bibliographique générale sur Le principe de traitement des eaux. Nous avons ensuite recensé les connaissances acquises en matière de traitement des eaux, et présenté la méthodologie adoptée pour la détermination de la dose du coagulant utilisée lors de la phase coagulation floculation.

La théorie de l'analyse en composantes principales (ACP) a fait l'objet du **chapitre deux**.

Le **troisième** chapitre est consacré à des notions générales sur les réseaux de neurones et à une présentation des différents types de réseaux de neurones existants.

Nous avons traité dans le **quatrième** chapitre la modélisation et la prévision de la dose du coagulant à injecter par les réseaux de neurones.

Dans la conclusion générale nous avons mis l'accent sur l'importance de ce travail de recherche et les principaux résultats obtenus par le modèle élaboré.

# Chapitre I. Principes généraux de traitement des eaux

## .I.1. Introduction

Les hommes, ainsi que d'autres êtres vivants, sont très sensibles à la qualité de l'eau consommée. Malheureusement, les ressources en eau potable (sources, nappes phréatiques...) sont assez peu importantes, surexploitées, et souillées par une activité humaine croissante. L'homme a développé des techniques de traitement de l'eau pour éliminer les souches de maladies et les contaminants se retrouvant dans l'eau disponible. L'eau douce est majoritairement utilisée, même si elle ne représente qu'un dixième de la quantité globale de l'eau, pour des raisons économiques. Les techniques de production d'eau potable doivent, donc, utiliser tout un éventail de processus de traitements physiques, chimiques et biologiques adaptés à la nature des eaux superficielles utilisées et aux objectifs de qualité visés. Nous verrons d'abord les étapes par lesquelles passe une eau brute pour qu'elle soit potable et susceptible d'être consommée par la population; Toutefois, nous nous attarderons plus longuement sur le procédé de coagulation, dont la modélisation constitue l'objet principal de ce travail.

Une station de traitement peut comporter une ou plusieurs opérations (figure I-1),

suivant son importance et la qualité de l'eau brute

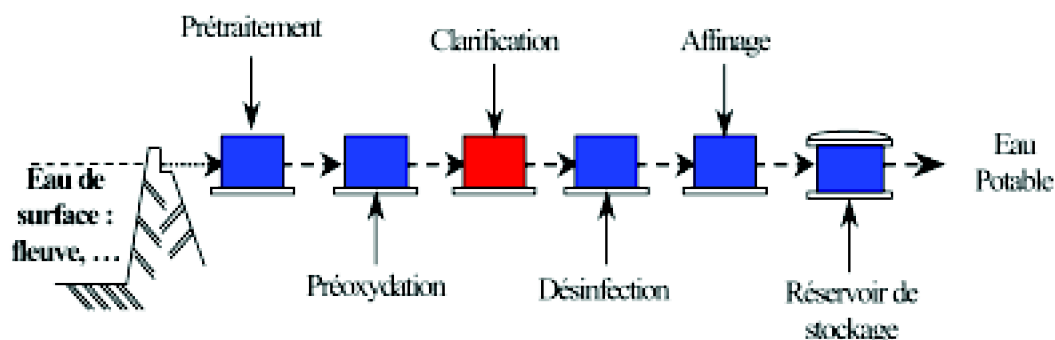


Figure I-1 : Principales opérations de traitement des eaux.

## .I.2. Prétraitement

Une eau, avant d'être traitée, doit être débarrassée de la plus grande quantité possible

d'éléments dont la nature et la dimension constitueraient une gêne pour les traitements ultérieurs. Pour cela, on effectue des pré-traitements de l'eau de surface [Anonym1, 2000]. Dans le cas d'une eau potable, les pré-traitements sont principalement :

- dégrillage,
- dessablage,
- débouage
- dégraissage, fréquemment associé au dessablage,
- déshuilage,
- tamisage,
- évacuation et traitement des sous-produits.

Le dégrillage, premier poste de traitement, indispensable aussi bien en eau de surface qu'en eau résiduaire, permet de protéger les ouvrages aval contre l'arrivée de gros objets susceptibles de provoquer des bouchages dans les différentes unités de l'installation, de séparer et d'évacuer facilement les matières volumineuses charriées par l'eau brute, qui pourraient nuire à l'efficacité des traitements suivants, ou en compliquer l'exécution. [Valentin ,2000].

Le dessablage a pour but d'extraire des eaux brutes les graviers, sables et particules minérales plus ou moins fines, de façon à éviter les dépôts dans les canaux et conduits, à protéger les pompes et autres appareils contre l'abrasion, et à éviter de perturber les stades de traitement suivants. Le domaine usuel du dessablage porte sur les particules de granulométrie égale ou supérieure à 200 mm; une granulométrie inférieure est en

général du ressort du débouage ou de la décantation.

Le débouage est une étape de séparation solide-liquide qui précède la clarification des eaux de surface particulièrement chargées lorsque la décantation classique en un seul étage n'est plus possible. Le but de cette pré-décantation est d'éliminer la majorité des matières en suspension de l'eau brute, d'en assurer l'évacuation sous forme de boues concentrées et de fournir à l'étape de décantation principale une eau de qualité acceptable.

Les opérations de dégraissage et de déshuilage consistent en une séparation de produits de densité légèrement inférieure à l'eau, par effet de flottation, naturelle ou assistée, dans une enceinte liquide de volume suffisant.

Le dégraissage est une opération de séparation liquide-solide réalisant un compromis entre une rétention maximale des graisses et un dépôt minimal de boues de fond fermentescibles. Son rendement est difficile à définir par suite des difficultés d'échantillonnage et d'analyse.

Le déshuilage est l'élimination d'huiles présentes en quantité notable dans les eaux résiduaires industrielles, en particulier dans les industries du pétrole (mais normalement absentes des eaux résiduaires urbaines, leur rejet en égout étant interdit). Le déshuilage est une opération de séparation liquide-liquide.

### I.3. Pré oxydation

A l'issue du pré-traitement, on a une eau relativement propre mais qui contient encore des particules colloïdales en suspension, et des matières organiques en solution. On les élimine dans la mesure du possible par une pré-oxydation. Celle-ci peut être faite de trois façons différentes :

- Ajout de chlore ou pré chloration ;
- Ajout de dioxyde de chlore ;
- Ajout d'ozone ou pré-ozonation.

### I.4. Clarification

#### I.4.1 Introduction

---

Les procédés de coagulation et de floculation facilitent l'élimination des matières en suspensions (MES) et des particules colloïdales.

Dans une chaîne de traitement conventionnelle, la coagulation consiste à destabiliser

---

les particules, tandis que la floculation consiste à faire agglomérer les particules déstabilisées [Bouchard et Serodes, 2002].

La coagulation est l'une des étapes les plus importantes dans le traitement des eaux de surface. 90 % des usines de production d'eau potable sont concernées. La difficulté principale est de déterminer la quantité optimale de réactif à injecter en fonction des caractéristiques de l'eau brute. [Valentin, 2000]

La floculation est l'agglomération de ces particules "déchargées" en micro floc, puis en flocons volumineux et décantables, le floc. Cette floculation peut être améliorée par l'ajout d'un autre réactif : le floculant ou adjuvant de floculation.

Un mauvais contrôle de ces procédés peut entraîner une augmentation importante des coûts de fonctionnement et le non-respect des objectifs de qualité en sortie. Cette opération a également une grande influence sur les opérations de décantation et de filtration ultérieures. En revanche, un contrôle efficace peut réduire les coûts de main d'œuvre et de réactifs et améliorer la conformité de la qualité de l'eau traitée.

En résumé, le contrôle de cette opération est donc essentiel pour trois raisons : la maîtrise de la qualité de l'eau traitée en sortie (abattement de la turbidité), le contrôle du coagulant résiduel en sortie (réglementation de plus en plus stricte relative à la présence de coagulant résiduel dans l'eau traitée) et la diminution des contraintes des coûts de fonctionnement (coûts des réactifs et des interventions humaines).

### I.4.2 Coagulation-Floculation

---

A l'heure actuelle, la coagulation – floculation est un procédé efficace pour l'élimination de la matière organique dissoute (MOD) de taille relativement importante, mais reste encore limitée pour les MOD de tailles réduites [Matilainen et al. 2002].

Il existe deux principaux types de coagulants [Lind, 1995] : les sels de fer (chlorure ferrique) et les sels d'aluminium (sulfate d'aluminium). Le choix du coagulant peut varier avec la température et donc avec la saison. Le sulfate d'aluminium, par exemple, est un coagulant utilisé pour une température d'eau supérieure à 10-12 °C. Rappelons également que plus un coagulant a de charges positives, plus son efficacité est grande.

Par la suite, nous allons énumérer l'ensemble des paramètres influençant le bon fonctionnement du procédé de coagulation [Anonym2 1989, Lind 1994].

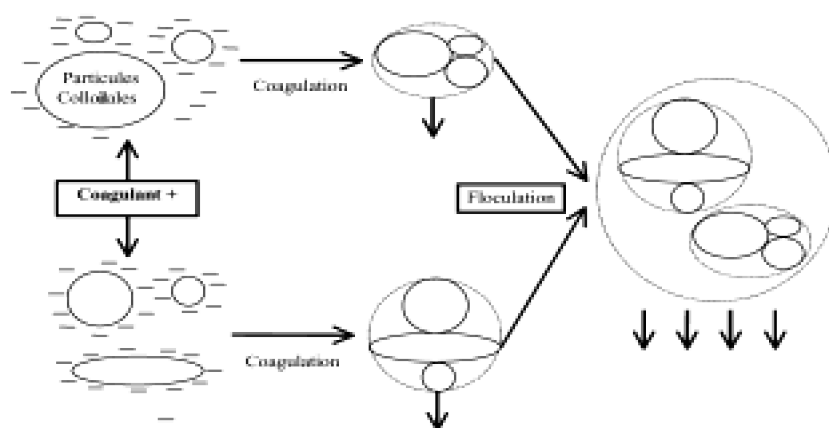


Figure I-2 : Principe de la coagulation-floculation

### I.4.2.1. Coagulation

#### I.4.2.1.1. Considérations physico-chimiques concernant la coagulation

Les particules présentes dans les eaux naturelles sont chargées négativement et sont entourées d'un nuage ionique. Cette charge électrostatique, ainsi que la couche ionique diffuse, constitue une barrière à l'agglomération de particules à cause de la répulsion électrostatique et parce que les forces attractives ne peuvent agir. [Adrian, 2004]

Les principaux facteurs influençant l'efficacité de la coagulation – floculation sont la dose de coagulant, le pH, la turbidité initiale et la température de l'eau.

#### I.4.2.1.2. Optimisation de la Coagulation

Lors de la coagulation, il faut chercher :

- La maximisation de la déstabilisation des particules et des colloïdes organiques pour faciliter leur agglomération et leur enlèvement subséquent, par un procédé de séparation solide-liquide;
- La minimisation de la concentration en coagulant résiduel;
- La minimisation de la production des boues;
- La minimisation des coûts d'opération.
- La coagulation optimale correspond au dosage de coagulant et au pH qui assurent l'atteinte de tous ces objectifs [Edzwald et Tobiason, 1999].

#### I.4.2.1.3. Coagulants usuels

Les coagulants usuels sont :

- Sulfate d'aluminium hydraté (appelé communément alun) :

$\text{Al}_2(\text{SO}_4)_3 \cdot 14\text{H}_2\text{O}$ ; dosages moyens 15 à 40 mg/l, chaque ppm d'alun produit 0,26 ppm de  $\text{Al}(\text{OH})_3$  et consomme 0,5 ppm d'alcalinité;

- Chlorure ferrique :  $\text{FeCl}_3$  ; moins sensible que l'alun au contrôle du pH, mais meilleur dans le traitement de la couleur; chaque ppm de  $\text{FeCl}_3$  produit 0,66 ppm de  $\text{Fe}(\text{OH})_3$  et consomme 0,9 ppm d'alcalinité;
- Poly-alumino-silico-sulfate (PASS):  $\text{Al}_a(\text{OH})_b(\text{SO}_4)_c(\text{SiOx})_d$ ;
- Aluminate de sodium :  $\text{NaAlO}_2$ .

Les deux derniers coagulants sont ceux qui sont les moins sensibles aux basses températures, mais sont également les plus coûteux [Adrian, 2004]. En Algérie on utilise souvent le premier type de coagulant.

### 1.4.2.1.4. Coagulation de la turbidité et de la couleur

Pour la coagulation des particules à l'origine de la turbidité, les conditions optimales qui correspondent à la coagulation par piégeage sont un pH de 7,5 et une dose d'alun, aux environs de 30 mg/l. Au contraire, pour la coagulation des colloïdes organiques naturels à l'origine de la couleur jaune-brun des eaux douces de surface, le pH optimal est plus acide, soit aux environs de 5-6. [Adrian, 2004]

### 1.4.2.1.5. Coagulation de la matière organique (MO)

Les substances humiques correspondent à une partie importante de la MO. Elles sont très présentes dans les eaux de surface. Ayant des origines naturelles et étant directement reliées à la couleur, celles-ci peuvent causer plusieurs types de problèmes pour les usines de production d'eau potable [Amirtharajah et O'Melia, 1990] :

- Coloration de l'eau;
- Occupation des sites d'adsorption du charbon actif en grain (CAG);
- Augmentation de la demande en chlore ou de la demande pour un autre oxydant;
- Agents de transport de polluants par complexation et adsorption.

Selon [Adrian, 2004], les études conduites par Black et Packham sur l'enlèvement de la couleur des eaux contenant des grandes concentrations (25 mg/l) en carbone organique dissous (COD), ont démontré une relation entre la concentration de la matière organique dans l'eau à traiter et la concentration de coagulant ( $\text{Fe}^{3+}$ ,  $\text{Al}^{3+}$ ) nécessaire pour assurer un enlèvement efficace de la couleur. Ces études ont aussi démontré un enlèvement plus poussé de la couleur pour des valeurs de pH plus acides de 4,5 à 5,5 pour  $\text{Fe}^{3+}$  et de 5 à 6 pour  $\text{Al}^{3+}$ . L'explication du phénomène serait une précipitation des substances humiques reliées à la couleur, par la formation des précipités métalo - humiques [Amirtharajah et O'Melia, 1990].

Pour le cas où des sels de Fe ( $\text{FeCl}_3$ ) sont utilisés comme coagulants pour



l'enlèvement du COD, les mécanismes d'enlèvement restent les mêmes, mais les gammes de pH qui leur correspondent sont légèrement plus basses. Les autres facteurs qui influencent la coagulation de la matière organique dissoute par les sels de Fe, sont la masse moléculaire, la charge électrique et la solubilité des composants organiques [Amirtharajah et O'Melia, 1990].

Étant donnée l'importance d'une bonne coagulation dans une chaîne de traitement conventionnelle, il est important d'avoir un contrôle très précis sur la dose de coagulant, le pH et la température, tout en tenant compte de la turbidité initiale et de la teneur en carbone organique total (COT) de l'eau à coaguler. En plus, les conditions optimales de la coagulation varient en fonction des paramètres qualitatifs de l'eau : turbidité, couleur ou les deux. [Adrian, 2004]

### **I.4.2.1.6. Agitation**

La coagulation requiert un mélange rapide pendant un temps qui varie habituellement de 0,5 à 4 min [Bouchard et Sérodes, 2002].

### **I.4.2.2. Flocculation**

#### **I.4.2.2.1. Objectifs de la flocculation**

Après la coagulation, il est nécessaire d'assurer les contacts entre les micros flocons formés et constitués des particules emprisonnées par les hydroxydes métalliques fraîchement précipités. Une fois agglomérés, les flocons seront plus gros et sédimenteront plus rapidement.

#### **I.4.2.2.2. Aides flocculants usuels**

Plusieurs aides flocculants sont utilisés pour faciliter la flocculation :

- Poly électrolytes : polymères de très haut poids moléculaire, qui donnent du corps et de la résistance aux flocons. Les dosages usuels sont de 0,05 à 0,5 mg/l;
- Silice activée : Silice activée par aluminat de sodium, qui alourdit les flocons et agit aussi comme adsorbant pour certains micro-polluants. Les dosages usuels sont de 0,5 à 4 mg/l  $\text{SiO}_2$ .

#### **I.4.2.2.3. Agitation**

La flocculation requiert une agitation moins forte que la coagulation. L'agitation doit être suffisamment forte pour favoriser les contacts entre les flocs mais pas trop forte pour briser les flocs déjà formés [Bouchard et Serodes, 2002]. La flocculation est un phénomène plus lent que la coagulation. De manière générale, la flocculation peut prendre de 6 à 30 min, tout dépend de la configuration du bassin de mélange, du niveau de mélange, du type d'aide flocculant utilisé et de son dosage. Quand les conditions ne sont pas propices

au grossissement des floccs, la taille des floccs demeure faible (quelques dizaines à quelques centaines de microns). On parle alors de micro-floccs comme dans le cas de la filtration directe.

### I.4.3. Essais de Coagulation – Flocculation

---

Afin d'évaluer les conditions optimales de coagulation et de flocculation, des essais dits de Jar-test (JT) sont conduits à l'échelle de laboratoire. Ces tests sont très courants et sont réalisés aussi bien dans les usines de production d'eau potable que dans les laboratoires de recherche. Le dispositif expérimental consiste en une série de six (06) bécchers de 1 litre chacun, qui contiennent les échantillons d'eau à tester. La température des échantillons est d'abord ajustée à la valeur choisie. Chaque béccher est muni d'un agitateur à pales. La vitesse d'agitation est variable mais elle est la même pour tous les bécchers. Des doses variables de coagulant, ou des doses variables d'acide ou de base pour obtenir différents pH, peuvent être ajoutées dans les bécchers, tandis qu'un béccher dans lequel il n'y a aucun ajout sert généralement de témoin. Un essai de JT comprend trois (3) phases :

- Coagulation : ajout de coagulant et, au besoin d'acide ou de base; courte durée et forte agitation;
- Flocculation : ajout d'un aide flocculant au besoin; durée moyenne avec agitation lente;
- Sédimentation : durée relativement longue et aucune agitation.

À la fin de la phase de sédimentation, un échantillon de surnageant est prélevé pour fins d'analyse. Des JT menés dans une large gamme de conditions d'opération permettent de déterminer le type de coagulant, son dosage, le pH et les conditions d'agitation qui maximisent la réduction de la turbidité et la réduction de la couleur pour des filières de traitement physico-chimiques conventionnelles [Adrian, 2004].

### I.4.4. La séparation solide-liquide

---

Une fois le flocc formé, il faut le séparer de l'eau. C'est ce qu'on appelle la séparation solide liquide. Elle peut s'effectuer par différents moyens :

- la coagulation sur filtre,
- la décantation,
- la flottation.

#### I.4.4.1. La décantation

Il existe deux types de décanteurs : les décanteurs statiques et les décanteurs à contact de boues. L'usage a consacré le terme « statique » pour désigner les décanteurs qui ne sont ni à circulation de boues, ni à lit de boues, bien que, dans ces appareils, la

décantation s'effectue en fait selon un processus dynamique.

Afin d'augmenter la probabilité qu'ont les particules de se décanter, on met l'eau additionnée de réactifs (eau brute + coagulant) en contact avec des boues concentrées provenant de la décantation. Deux procédés peuvent être employés pour assurer le contact avec les boues [Anonym2, 1989] :

- les appareils à re-circulation de boues: les boues sont séparées de l'eau clarifiée dans une zone de décantation. Elles sont ensuite re-circulées dans une zone de mélange munie d'une agitation mécanique ou hydraulique. L'eau brute, additionnée de réactifs, est également introduite dans cette zone de mélange.

- les appareils à lit de boues proprement dit : on ne cherche pas à faire circuler la boue. Il s'agit également de la maintenir sous forme d'une masse en expansion que l'eau peut traverser de bas en haut, d'une façon aussi régulière et uniforme que possible. Le brassage, très doux, est réalisé au niveau de l'introduction de l'eau à traiter. [Valentin, 2000]

### I.4.4.2. La filtration

La filtration est définie comme le passage d'un fluide à travers une masse poreuse pour en retirer les matières solides en suspension [Anonym2, 1989]. La filtration représente donc le moyen physique pour extraire de l'eau les particules non éliminées préalablement lors de la décantation. De façon générale, un filtre aura une longévité entre deux lavages d'autant plus importante que les traitements préalables auront été efficaces (coagulation - floculation - décantation).

- Le matériau : on peut utiliser du sable ou du charbon actif en grain. Ce dernier type de matériau est le plus souvent utilisé en traitement d'affinage (adsorption) en deuxième étage de filtration afin d'éliminer les pesticides et les sous-produits d'oxydation (odeurs et goûts) ;

- La taille de grain du milieu filtrant ;

- Le coefficient d'uniformité : si l'homogénéité des tailles n'est pas respectée, le lavage classe les grains selon leur taille, les grains les plus gros se trouvant au fond du filtre et les plus fins en surface. Dans ce cas, l'encrassement de surface entraîne un cycle de filtration plus court.

- La hauteur de la couche filtrante : il faut vérifier périodiquement si le filtre perd du matériau et compléter au besoin ;

- Le taux de filtration (en  $\text{m}^3/\text{m}^2 \cdot \text{h}$ ) ou vitesse de filtration.

Le choix de ces différents critères est guidé par les caractéristiques de l'eau à traiter et la qualité de l'eau que l'on souhaite obtenir. Un filtre doit produire une eau de qualité satisfaisante et constante tout au long d'un cycle de filtration.

## I.5. Désinfection

Le but de la désinfection est d'éliminer tous micro-organismes pathogènes présents dans l'eau afin d'empêcher le développement de maladies hydriques [Anonym1, 2000]. Au niveau de l'exploitant, les dangers liés à la présence de micro-organismes pathogènes dans les eaux distribuées sont multiples :

- Risque à très court terme amplifié par le délai d'analyse qui est au minimum de 24h ;
- Risque omniprésent, les problèmes pouvant intervenir sur tous types de ressource et dans n'importe quel réseau de distribution ;
- Gravité des maladies, qui peuvent être mortelles ;
- Ampleur de la contamination qui peut atteindre plusieurs milliers de personnes.

Le principe de la désinfection est de mettre en contact un désinfectant à une certaine concentration pendant un certain temps avec une eau supposée contaminée. Cette définition fait apparaître trois notions importantes : les désinfectants, le temps de contact et la concentration en désinfectant.

Les trois principaux désinfectants utilisés en eau potable sont les suivants :

- Le chlore ;
- L'ozone ;
- Le rayonnement UV.

## I.6. Affinage

Cette étape a pour effet l'oxydation et la biodégradation des matières organiques et l'élimination ou l'absorption de certains micropolluants. En outre, l'affinage améliore les qualités organoleptiques de l'eau (saveur, odeur, limpidité). Les traitements d'affinage de l'eau font intervenir des procédés d'ozonation, de filtration sur charbon actif ou de filtration sur membranes (ultrafiltration ou nano filtration).

## I.7. Réservoirs de stockage

A la fin de la chaîne traitement, l'eau est stockée avant d'être fournie pour le consommateur; pour cela des réservoirs de stockage sont conçus; leur but principal est le

stockage des eaux traitées prêtes à être consommées.

Les réservoirs constituent une sécurité en cas de problèmes de production, de demandes de pointe, ou de pollution des ressources.

A travers ce chapitre, nous avons présenté une description de la technique de traitement des eaux en relation avec l'objectif de notre mémoire.

Nous avons décrit la chaîne de traitement la plus complète possible depuis l'entrée de l'eau à la station jusqu'à sa sortie vers le consommateur.

Nous nous sommes attardés à détailler le problème de la coagulation- floculation, car c'est ce procédé qui est l'objet de notre travail.



# Chapitre II. Analyse en composantes principales

## II.1. Historique de l'analyse des données

L'analyse de données est apparue dans les années trente grâce à H. Hotelling, qui initialisait le principe de l'analyse en composantes principales et de l'analyse canonique en s'inspirant des travaux de C. Spearman et de K. Pearson datant du début du siècle, [Hotelling 1933, Pearson 1901, Spearman 1904].

Ces méthodes furent ensuite complétées et perfectionnées sans grand succès, car elles nécessitaient de très importants calculs. Le développement de l'informatique, dans les années soixante, a remis cette approche de la statistique au goût du jour et la multiplication de micro-ordinateurs scientifiques a permis sa vulgarisation.

La statistique classique ne s'intéresse qu'à un nombre limité de caractères mesurés sur un ensemble restreint d'individus. Elle s'appuie sur des estimations ou des tests très restrictifs. Les hypothèses émises sont souvent très approximatives et les lois de distribution très théoriques. La différence essentielle de l'analyse de données, par rapport à la statistique classique, est qu'elle réalise une étude globale du couple individu - caractère en s'appuyant sur des représentations graphiques beaucoup plus lisibles que

de volumineux tableaux de mesures. Les ressemblances ou les dissemblances entre individus ou entre caractères, leur corrélations positives ou négatives sont à la base de l'analyse, qui peut s'effectuer suivant plusieurs points de vue. Cette approche multidimensionnelle, depuis qu'elle est maîtrisée grâce à l'apport de l'informatique, a connu un très grand nombre d'applications pratiques pour l'observation et la compréhension des phénomènes complexes notamment en médecine, physique, sciences naturelles, sciences humaines, etc....

## .II.2. Définitions & contexte de base

Dans les situations réelles, les individus d'un échantillon sont le plus souvent décrits par un grand nombre de variables, ce qui en rend la visualisation par le cerveau humain impossible. Or il existe la plupart du temps une forte redondance entre ces variables. On peut donc imaginer qu'il soit possible de décrire ces individus par un petit nombre de variables « synthétiques », obtenues par combinaison des variables initiales, avec une faible perte d'informations. A l'extrême, si on ne retient que deux de ces variables synthétiques, il devient possible de donner de l'échantillon une représentation graphique plane, facilement interprétable.

Ces considérations ont donné naissance à l'analyse en composantes principales (ACP), technique très classique d'analyse des données, et qui permet la création de nouvelles variables ayant la propriété d'être décorréélées, ce qui permet de n'en retenir qu'un petit nombre pour décrire l'échantillon avec des pertes d'informations minimales [Jackson,1991].

L'ACP consiste à calculer un nombre réduit de nouvelles dimensions, qui sont des combinaisons linéaires des dimensions originales des données (c'est-à-dire des traits descriptifs) [Bouroche et Saporta, 1980]. Ces nouvelles dimensions sont non corrélées et expriment le maximum de variance des données. [Jolliffe, 1986], [Lebart et al. 1977]. Les nouveaux axes sont les vecteurs propres, ordonnés par valeurs propres décroissantes, de la matrice de covariance des données.

Autrement dit ce sont les principaux axes de dispersion du nuage de données, en ordre d'importance décroissante ; les valeurs propres correspondantes indiquent la part de variance exprimée par chaque axe. Les premiers axes rendent donc généralement compte de la plus grande partie de la variance. Les composantes principales sont les nouvelles valeurs des données sur chaque axe ainsi obtenu.

Cette méthode peut jouer un double rôle de compression des données et d'outil d'exploration dans des domaines fortement multidimensionnels. En effet les axes principaux ainsi calculés permettent à la fois une réduction des données et une interprétation plus facile du domaine traité, car les nouvelles dimensions sont souvent très significatives. [Lebart et Salem, 1994].

Le problème est que l'ACP demande le calcul préalable de la matrice carrée de covariance des données, qui est de taille  $(n \times n)$  pour des vecteurs de dimension  $n$ . Cette



matrice est déjà coûteuse à calculer, et sa taille et son traitement deviennent prohibitifs en grande dimension. Ainsi des données de dimension 1000 donneraient lieu à une matrice de un million d'éléments !

### II.2.1. Notion d'individu et de caractère

---

En analyse de données on distingue deux types d'informations :

- Les individus
- Les caractères

Les individus peuvent représenter des entités comme des personnes, des entreprises, des pays, des journaux, etc. Sur ces individus sont réalisés des observations ou des mesures. Les individus peuvent représenter l'ensemble d'une population ou simplement une partie (caractères) qui n'est pas nécessairement sensée être, comme en statistique inférentielle, un échantillon représentatif. Sur un individu, on relève un certain nombre de caractères (dits aussi variables).

L'individu "i" est décrit par le vecteur appartenant à  $\mathbb{R}^P$  :

$$X_i = (X_{ij} / j = 1 \text{ à } P)$$

Le caractère (ou variable) "j" est décrit par le vecteur de  $\mathbb{R}^N$  :

$$X_j = (X_{ij} / i = 1 \text{ à } N)$$

### II.2.2. Définition algébrique

---

L'analyse en composantes principales (ACP) décrit un tableau X de valeurs numériques continues du type " *individus - caractères* ".

On est en présence d'un tableau *individus caractères*, X (n,p) de n individus et p caractères.

Le tableau X se présente comme suit :

$$[X] = \begin{bmatrix} X_{11} & X_{12} & X_{1J} & X_{1P} \\ X_{21} & X_{22} & X_{2J} & X_{2P} \\ X_{31} & X_{32} & X_{3J} & X_{3P} \\ X_{M1} & X_{M2} & X_{MJ} & X_{MP} \end{bmatrix}$$

Pour chaque variable on calcule la moyenne et l'écart type

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad (II-1)$$

La variance est donnée par l'équation suivante :

$$\text{var}(x_i) = \sum_j p_j (x_{ij} - \bar{x}_i)^2 \quad (II-2)$$

L'écart type est la racine carrée de la variance :

$$S_i = \left[ \frac{1}{N} \sum_{i=1}^N (x_{ij} - \bar{x}_i)^2 \right]^{1/2} \quad (II-3)$$

La covariance de deux variables X, Y est donnée par :

$$s_{xy} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) \quad (II-4)$$

Le coefficient de corrélation entre de variable est :

$$r_{xy} = \frac{s_{xy}}{S_x S_y} \quad (II-5)$$

Le tableau [X] est remplacé par un tableau [Z] (individus x nouvelles variables) en réduisant le nombre de variables nécessaires pour décrire les individus, avec une perte minimale d'informations, ces nouvelles variables sont appelées composantes principales (ou CP).

Calculer les composantes principales notées  $C_j$  revient à déterminer N relations linéaires entre les variables  $X_j$  :



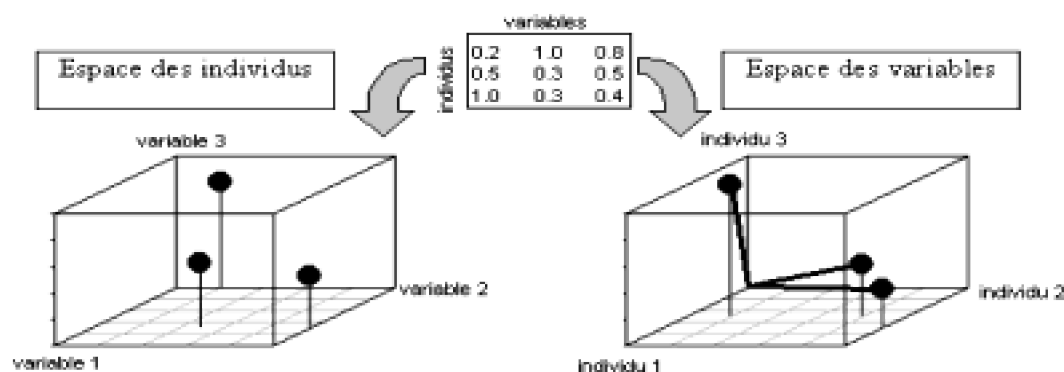


Figure N° II-1 : Représentation géométrique des individus et variables.

L'ACP se divise en étapes :

- \_ Normalisation des données pour être indépendant des unités des P paramètres;
- \_ Calcul d'une matrice de corrélation;
- \_ Recherche des éléments propres de C, qui donnent les axes principaux;
- \_ Représentation des individus dans le nouvel espace (en ne considérant que les valeurs propres expliquant une variance cumulée suffisante).

L'analyse en composantes principales effectue une projection de l'ensemble des données (qu'elle que soit la classe d'appartenance) sur un sous espace qui, tout en étant de dimension réduite, autorise une bonne représentation des données.

Le sous espace choisi de telle sorte que la variance du nuage de points projeté soit la plus grande possible. Ce qui en d'autres termes est équivalent à assurer une déformation minimale entre les points initiaux et leur projection.

[OUKHELLOU ,1997].

Soit B la matrice d'observations constituée des N réalisations (ou individus) représentées chacune par un vecteur paramètre (ou variable) de dimension d :

$$B = [X_1 X_2 \dots X_d] = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix} = \begin{bmatrix} X_{11} & & X_{1d} \\ X_{21} & & \\ \vdots & & \\ X_{N1} & & X_{Nd} \end{bmatrix}$$

$\leftarrow Y_i$  individu ou observation

$\leftarrow X_i$  variable ou paramètre

Le centre de gravité de l'ensemble des données est défini par :

$$m = [m_1 \quad m_2 \quad \dots \quad m_d]$$

Avec  $m_j = E[X_j] = \frac{1}{N} \sum_{j=1}^N x_j$  la moyenne sur le jème paramètre

La matrice de variance covariance de B s'écrit :

$$V_B = \left\{ E[(X_i - m_i)(X_j - m_j)] \right\} \text{ Avec } 1 \leq i, j \leq d$$

Afin d'éviter qu'un paramètre ne prenne trop d'importance par rapport aux autres simplement parce qu'il varie dans de plus grandes proportions, on procède généralement à une normalisation des données.

$$\left\{ \frac{x_j - m_j}{\sigma_j} \right\} \text{ Avec } 1 \leq i \leq d \text{ et } 1 \leq j \leq d$$

Où  $\sigma_j$  représente l'écart type du jème paramètre et  $m_j$  sa moyenne

On montre que pour une dimension  $d_r$  donnée, le sous-espace maximisant la variance du nuage projeté est engendré par les  $d_r$  vecteurs propres de la matrice de covariance  $V_B$  associé aux  $d_r$  premières valeurs propres rangées par ordre décroissant [Dubuisson,1990].

La nouvelle matrice d'observations C de dimension réduite  $N \times d_r$ , obtenue après projection des observations sur le sous espace engendré par les  $d_r$  premiers vecteurs propres associés à  $V_B$  s'écrit alors sous la forme :

$$C = BU = [C_1 \quad C_2 \quad \dots \quad C_{d_r}] \quad d_r < d$$

$$\text{Avec } U = [u_1 \quad u_2 \quad \dots \quad u_{d_r}]$$

$$V_B u_i = \lambda_i u_i \text{ Avec } (\lambda_1 > \lambda_2 > \dots > \lambda_{d_r})$$

$u_i$  étant le vecteur propre de la matrice  $V_B$  associé à la valeur propre  $\lambda_i$ .

Chaque vecteur colonne  $C_i$  de la matrice C est communément appelé composante principale et représente la projection des individus selon le ième axe  $u_i$  ces composantes présentent des propriétés intéressantes :

- La variance de la ième composante principale est égale à  $\lambda_i$ ;

- Et la covariance des composantes principales entre elle est nulle [Dubuisson, 1990].

En résumé, l'ACP cherche par combinaisons linéaires des d variables initiales ( $X_1, X_2, \dots, X_d$ ) , $d_r$  nouvelles variables ou composantes principales(  $C_1, C_2, \dots, C_{d_r}$ ) centrées, de variâhce

maximale et qui sont non corrélées entre elles.

## II.4. Choix de la métrique

Une métrique est une matrice permettant de définir un produit scalaire et donc des distances entre individus ou entre variables.

La métrique généralement utilisée en ACP normée est la métrique diagonale qui revient à pondérer les caractères en  $1/s^2$  représentant l'écart type du caractère, ainsi

$$M = D_{1/s^2} = \begin{pmatrix} 1/s_1^2 & & 0 \\ & 1/s_2^2 & \\ 0 & & \ddots \\ & & & 1/s_p^2 \end{pmatrix}$$

Ce qui revient à diviser chaque variable par son *écart-type*.

Ainsi cette métrique donne à chaque caractère la même importance (un caractère dispersé " n'absorbe " pas ainsi toute la variance) et d'autre part les distances entre individus sont adimensionnelles ce qui permet de traiter des caractères d'unités différentes.

## II.5. Procédé d'application de l'ACP

### II.5.1. Calcul de la matrice de variance covariance

---

La matrice des covariances, notée [R] est la base de l'ACP, elle est obtenue en appliquant la relation suivante :

$$[R] = (1/n)[X]^2 \cdot [M] \cdot [X]$$

[R] : Matrice de covariance de dimension (p\*p) ;

$[X]^2$  : Matrice de données ;

$[X]^2$  : Matrice transposée de [X] ;

[M]: Métrique.

### II.5.2. Calcul des valeurs et vecteurs propres

---

### II.5.2.1. Définition

Un scalaire  $\lambda$  est une valeur propre de la matrice  $A$  si et seulement si l'une des conditions équivalentes suivantes est vérifiée :

Le rang de la matrice  $A - \lambda I$  est strictement inférieur à  $n$ .

Le déterminant de la matrice  $A - \lambda I$  est nul  $\text{Det}(A - \lambda I) = 0$ .

On dit que  $V$  est un vecteur propre de  $A$  associé à la valeur propre  $\lambda$  si  $V$  est un vecteur non nul et  $AV = \lambda V \leftrightarrow (A - \lambda I)V = 0$

### II.5.2.2. Méthode de la puissance

Celle-ci permet de déterminer la plus grande valeur propre (en valeur absolue) de la matrice  $A$  et le vecteur propre associé. Partant d'un vecteur initial arbitraire  $x^{(0)}$  dont une composante est unitaire, on construit un vecteur intermédiaire  $y^{(1)}$  tel que  $A \cdot x^{(0)} = y^{(1)}$ , puis on construit  $x^{(1)}$  en "normalisant"  $y^{(1)}$  de manière à ce que la composante choisie soit à nouveau égale à l'unité. L'opération doit être répétée jusqu'à la convergence, en changeant éventuellement de composante de référence si celle-ci converge vers 0. Le facteur de normalisation converge vers la plus grande valeur propre et  $x$  converge vers le vecteur propre associé, sous les conditions que la plus grande valeur propre soit unique et que tous les vecteurs propres soient indépendants. La vitesse de convergence est proportionnelle au rapport entre les deux plus grandes valeurs propres (en valeur absolue).

### II.5.2.3. Principe de base

Le polynôme  $|A - \lambda I| = 0$  ; est appelé polynôme caractéristique, donc on aura par exemple la matrice  $A$  :

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

on introduisant la notion de vecteur et valeur propres on aura

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

Donc on introduisant la matrice unité on aura :

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{23} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

$$\left[ \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{23} & a_{33} \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right] \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Finalement

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{23} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

#### II.5.2.4. Algorithme

Soit A une matrice (n x n) symétrique réelle. On note  $(\lambda_1, \lambda_2, \dots, \lambda_n)$  ses valeurs propres (supposées réelles) et  $(v_1, v_2, \dots, v_n)$  les vecteurs propres correspondants (normés).

On suppose que les  $\lambda_i$  sont des modules distincts, et on les ordonne par modules décroissants :  $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$

Soit  $x^{(0)}$  un vecteur de  $\mathbb{R}^n$  et la suite  $x^{(k)}$  définie par la relation de récurrence

$$x^{(i+1)} = \frac{A x^{(i)}}{\|A x^{(i)}\|} \quad (II-6)$$

Où  $\| \cdot \|$  est la norme euclidienne sur  $\mathbb{R}^n$ .

On montre que  $x^{(k)} \rightarrow v_1$  quand  $k \rightarrow \infty$ .

On forme alors

$A^{(1)} = A - \lambda_1 v_1 v_1^t$  ( $v_1^t$  vecteur ligne transposé du vecteur colonne  $v_1$ ).

On montre que les valeurs propres de  $A^{(1)}$  sont  $(\lambda_2, \lambda_3, \dots, \lambda_n)$ .

$$x^{(i+2)} = \frac{A^{(1)} x^{(i)}}{\|A^{(1)} x^{(i)}\|} \quad (II-7)$$

On en déduit l'algorithme :

Pour i de 1 à n



Calcul de la limite  $v$  de la suite  $x^{(k)}$

Calcul du  $\lambda$  correspondant

Remplacement de  $A$  par  $A - \lambda_1 v v^t$

On décompose  $x^{(0)}$  sur la base  $v_1, v_2, \dots, v_n$  :

$$x^{(0)} = \sum_{i=1}^n c_i v_i \quad (\text{II-8})$$

On considère aussi la suite  $y^{(k)}$  définie par

$$y^{(k)} = x^{(k)} - A y^{(k-1)} \quad (\text{II-9})$$

On a

$$x^{(k+1)} = \frac{A y^{(k)}}{\|A y^{(k)}\|} \quad (\text{II-10})$$

On a alors

$$A y^{(0)} = \sum_{i=1}^n c_i \lambda_i v_i = \sum_{i=2}^n c_i \lambda_i v_i \quad (\text{II-11})$$

Donc

$$y^{(0)} = \sum_{i=1}^n c_i \lambda_i v_i \quad (\text{II-12})$$

Et par récurrence, à condition que  $c_1 \neq 0$  :

$$y^{(1)} = \sum_{i=1}^n c_i \lambda_i^2 v_i = c_1 \lambda_1^2 \left( v_1 + \sum_{i=2}^n \frac{c_i}{c_1} \frac{\lambda_i^2}{\lambda_1^2} v_i \right) \quad (\text{II-13})$$

$$|y^{(1)}| = c_1 \lambda_1^2 \left( 1 + \sum_{i=2}^n \frac{c_i}{c_1} \frac{\lambda_i^2}{\lambda_1^2} \right) \quad (\text{II-14})$$

Donc

$$x^{(1)} = \sum_{i=1}^n \alpha_i v_i \quad (\text{II-15})$$

Comme  $|\lambda_i| < |\lambda_1|$  on a  $\frac{\lambda_i^k}{\lambda_1^k} \rightarrow 0$  et ainsi

$$|\alpha_i^k| \rightarrow 1 \quad \alpha_i^k \rightarrow 0$$

Pour  $\mathbf{A}^{(1)}$  on a

$$\mathbf{A}^{(1)} \mathbf{v}_1 = (A - \lambda_1 \mathbf{v}_1 \mathbf{v}_1^T) \mathbf{v}_1 = A \mathbf{v}_1 - \lambda_1 \mathbf{v}_1 = 0$$

$$\mathbf{A}^{(1)} \mathbf{v}_i = (A - \lambda_1 \mathbf{v}_1 \mathbf{v}_1^T) \mathbf{v}_i = A \mathbf{v}_i = \lambda_i \mathbf{v}_i$$

Donc  $\mathbf{A}^{(1)}$  a les mêmes vecteurs propres que  $A$  avec les valeurs propres  $\lambda_2, \dots, \lambda_n$ , 0.

Si on forme alors une suite définie par la relation de récurrence :

$$x^{(k+1)} = \frac{A^{(1)} x^{(k)}}{\|A^{(1)} x^{(k)}\|} \quad (11-6)$$

On détermine  $v$ , vecteur propre associé à la plus grande valeur propre de  $\mathbf{A}^{(1)}$  en module, c'est-à-dire  $\lambda_2$ .

Pour calculer  $\lambda$  on peut simplement utiliser la définition  $\mathbf{A} \mathbf{v} = \lambda \mathbf{v}$  et prendre

$$\lambda = \frac{(A \mathbf{v})_k}{v_k}; \text{ il faut s'assurer que } v_k \neq 0 \text{ et que } \lambda \text{ est bien le même suivant } k.$$

Une vérification peut alors être faite en calculant  $\|A \mathbf{v} - \lambda \mathbf{v}\|$ .

Il est cependant préférable de calculer  $\lambda = \frac{(A \mathbf{v} | \mathbf{v})}{(\mathbf{v} | \mathbf{v})}$ . La colinéarité de  $\mathbf{v}$  et  $A \mathbf{v}$  étant

vérifiée par le fait que  $\frac{(A \mathbf{v} | \mathbf{v})}{\|A \mathbf{v}\| \|\mathbf{v}\|}$  vaut 1.

### II.5.3. Recherche des axes principaux

Les axes et les facteurs principaux  $v_1, v_2, \dots, v_n$  sont les vecteurs propres de la matrice associée aux valeurs propres. Le nombre de valeurs propres non nulles (car sinon il y a une relation linéaire entre les caractères initiaux) donne la dimension de l'espace dans

lequel sont réellement faites les observations.

Les composantes principales sont définies par les variables  $c$ . Les variances d'une composante principale sont les valeurs propres. L'ACP remplace les  $p$  caractères initiaux par des caractères non corrélés de variance maximale et d'importance décroissante. Les premières composantes principales  $C$  sont les combinaisons linéaires des variables centrées et réduites ayant une variance maximale.

### II.5.4. Principaux résultats :

---

En général une analyse en composantes principales fournit trois sources de renseignements, toutes nécessaires à l'interprétation :

- Un tableau de vecteurs et valeurs propres.
- Un tableau des corrélations des individus avec les axes factoriels.
- Un tableau des corrélations des variables aux axes principaux.

### II.5.5. Représentation graphique :

---

Une fois les résultats numériques obtenus nous passons à la représentation graphique :

Nous pouvons représenter chaque variable par un point dont les coordonnées sont les covariances avec les CP réduites.

### II.5.6. Programmation de l'ACP

---

Dans le cadre de ce travail de recherche nous avons développé un programme d'analyse en composantes principales sous **DELPHI**, qui nous permis de réaliser tous les calculs nécessaires pour obtenir les résultats de l'analyse en composantes principales; ainsi que la visualisation des résultats graphiques.

Le programme élaboré, est structuré en procédures et fonctions, la seule spécificité et qu'il utilise des fichiers de type **DELPHI**, ces derniers sont compilés séparément, leur lecture se fait sous format et accès libres.

### II.5.7. Fichier de données (input)

---

Les données sont organisées comme suit :

$$\begin{array}{ccccccc} X_{11} & X_{12} & X_{13} & X_{14} & X_{15} & X_{1j} & X_{1P} \\ X_{21} & X_{22} & X_{23} & X_{24} & X_{25} & X_{2j} & X_{2P} \\ X_{31} & X_{32} & X_{33} & X_{34} & X_{35} & X_{3j} & X_{3P} \\ X_{41} & X_{42} & X_{43} & X_{44} & X_{45} & X_{4j} & X_{4P} \\ X_{51} & X_{52} & X_{53} & X_{54} & X_{55} & X_{5j} & X_{5P} \\ X_{61} & X_{62} & X_{63} & X_{64} & X_{65} & X_{6j} & X_{6P} \\ X_{N1} & X_{N2} & X_{N3} & X_{N4} & X_{N5} & X_{Nj} & X_{NP} \end{array}$$

N : Entier désignant le nombre d'individus (lignes) ;

P : Entier désignant le nombre de variables (colonnes) ;

$X_{ij}$  : Données à introduire.

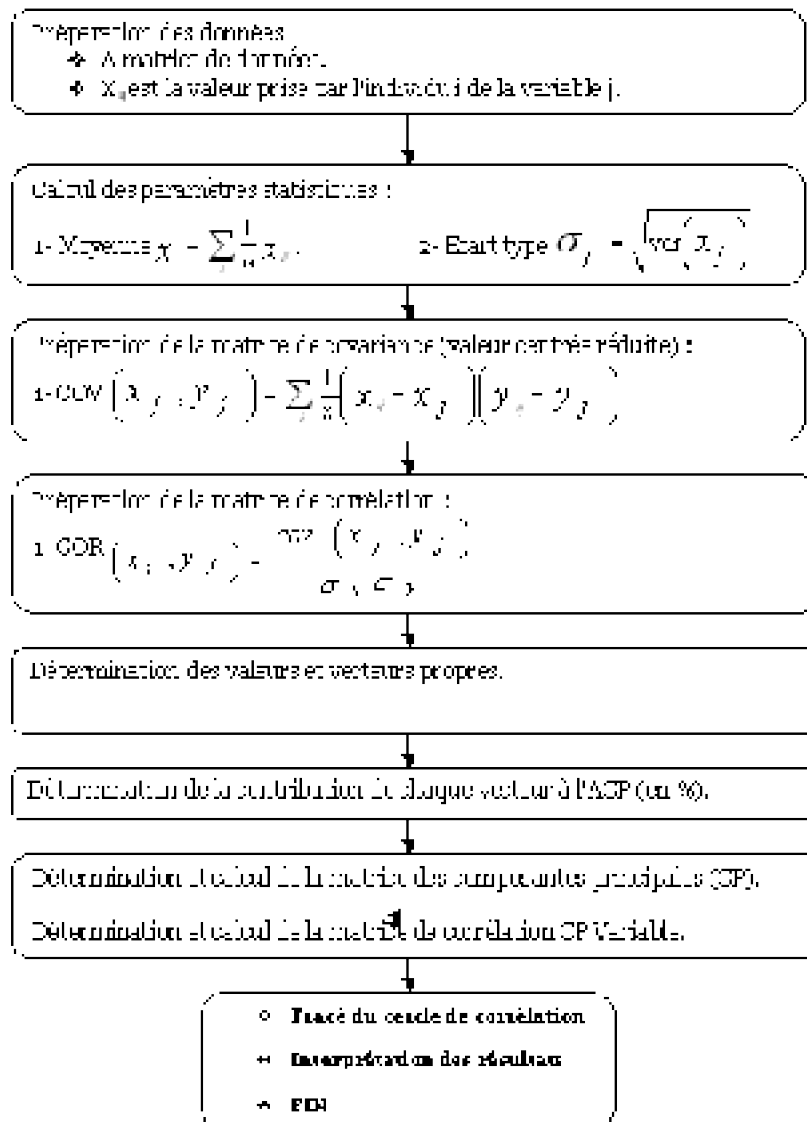
### II.5.8. Organigramme de calcul

---

Le programme que nous avons développé se compose de :

- Calcul des paramètres statistiques des données.
- Calcul de la matrice de covariance-variance.
- Calcul de la matrice de corrélation.
- Calcul des vecteurs et valeurs propres.
- Calcul des composantes principales.
- Visualisation graphique.

## II.6. Organigramme du programme



## II.7. Conclusion

L'analyse en composantes principales est une méthode factorielle d'analyse des données, elle permet d'étudier les corrélations entre les variables d'un tableau de données, tout en cherchant la réduction de sa taille, en remplaçant les variables originales par des nouvelles variables appelées composantes principales.

En utilisant l'analyse en composantes principales, on doit répondre à deux questions:

1. Est-ce que les différentes variables sont équivalentes et véhiculent toutes la même information ? Est-ce que l'information contenue dans un nombre de variables réduites peut expliquer le phénomène étudié?

2. Dans les cas où l'information véhiculée serait trop différente d'un indicateur à

l'autre, est il nécessaire d'utiliser toutes les variables originales ?

# Chapitre III. Théorie des réseaux de neurones

## III.1. Présentation et historique

Depuis une cinquantaine d'années, les chercheurs s'inspirent de la biologie dans le but de réaliser des constructions capables de résoudre divers problèmes complexes.

Développée pour la première fois dans les années 1940, [McCulloch et Pitts, 1943], (perceptron) [Rosenblatt, 1958], (Adaline) [Widrow et Hoff, 1960], l'élaboration des neurones artificiels a pour but de comprendre et de modéliser le système nerveux.

Cependant dans le milieu des années 1980, les recherches en méthodologie neuronales s'éloignent des justifications biologiques, et prennent une orientation mathématique. Une volonté de théorisation succède aux méthodes empiriques des premiers développements : les mathématiciens s'investissent pour comprendre ce que font les réseaux de neurones, et déterminent leur liens avec les méthodes statistiques déjà existantes. Les nouveaux réseaux de neurones s'avèrent capables de traiter des problèmes complexes de simulation de processus non linéaires et/ou dynamiques, ils trouvent donc de nombreuses applications en ingénierie, et pour les mêmes raisons, la recherche utilise ces nouveaux outils en physique appliquée, discipline dans laquelle le

nombre et la dimension des données, ainsi que la non linéarité des problèmes traités posent de nombreuses difficultés à l'emploi des outils statistiques classiques.

On évite souvent d'élaborer des réseaux de neurones artificiels de grande tailles, car il est en général plus efficace de découper les chaînes de traitement neuronales en une suite de petits réseaux de complexité relativement faible.

## III.2. Qu'est-ce qu'un neurone ?

### III.2.1. Une morphologie particulière

---

Le neurone biologique est une cellule spécialisée dans le traitement et la transmission de l'information. Il se compose d'un corps cellulaire (soma), et d'expansions : les dendrites d'une part, et l'axone d'autre part [Figure III.1]. Les dendrites forment des ramifications qui entrent en contact avec d'autres neurones, et dont le rôle est de recevoir des informations, électriques ou chimiques. L'axone est un prolongement de la cellule, typiquement long de quelques millimètres, qui conduit un signal électrique jusqu'à son arborisation terminale, où il peut alors entrer en contact avec les dendrites d'autres neurones. La jonction axone dendrite est appelée synapse [Romain ,2003].

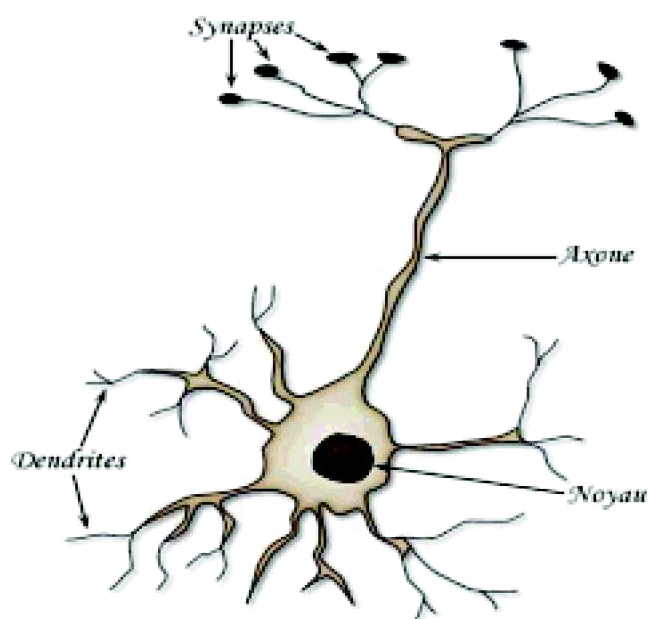


Figure III-1 : Schéma du neurone biologique

## III.3. Définitions

---



Les réseaux de neurones sont une des composantes importantes du domaine de l'intelligence artificielle. Par analogie aux neurones biologiques, les neurones artificiels ont pour but de reproduire des raisonnements « intelligents » d'une manière artificielle. La structure d'un neurone artificiel est en fait copiée sur la structure des neurones biologiques. Les principales structures biologiques des neurones ont toutes leur équivalent artificiel, et ce dans le but de reproduire leur fonctionnement de la meilleure façon possible (d'une manière logique, simple et facilement représentable).

Un neurone est une fonction non linéaire, paramétrée, à valeurs bornées. Cette représentation est le reflet de l'inspiration biologique qui a été à l'origine de la première vague d'intérêt pour les neurones formels, dans les années 1940 à 1970 [Dreyfus, 2004]. [McCulloch et Pitts, 1943], [Minsky et Papert, 1969].

Tableau III.1 : Analogie entre les neurones biologiques et artificiels

Neurone biologique	Neurone artificiel	Représentation symbolique
Soma	Neurone	
Dendrite	Intrant ( <i>Input</i> )	$x_j$ , $x$ (intranet net)
Axone	Extrant ( <i>Output</i> )	$Y$
Synapse	Poids	$w_j$

### III.4. Historique

- J. McCulloch et W. Pitts (1943), ont laissé leurs noms à une modélisation du neurone biologique (un neurone au comportement binaire). Ce sont les premiers à montrer que des réseaux de neurones formels simples peuvent réaliser des fonctions logiques, arithmétiques et symboliques complexes.

- D. Hebb (1949), physiologue américain explique le conditionnement chez l'animal par les propriétés des neurones eux-mêmes. Ainsi, un conditionnement de type pavlovien tel que, nourrir tous les jours à la même heure un chien, entraîne chez cet animal la sécrétion de salive à cette heure précise même en l'absence de nourriture. La loi de modification des propriétés des connexions entre neurones qu'il propose, explique en partie ce type de résultats expérimentaux.

- Von Neumann (1958), envisageait déjà la possibilité d'un réseau fiable formé de composants non fiables par l'utilisation de la redondance.

- Rosenblatt (1958), a développé une méthode analytique rigoureuse d'adaptation des poids au sein d'un modèle multicouche appelée le perceptron.

Il démontra la convergence d'un algorithme itératif d'adaptation des poids pour la classe élémentaire des perceptrons monocouches.

- **Minsky et Papert (1969)**, ont montré les limites du perceptron et son incapacité à résoudre les problèmes non linéairement séparables telle que le problème du ou exclusif

(XOR), un éclairage théorique sur la structure des problèmes qui était impossible à résoudre avec le perceptron. Cela a eu pour conséquences le ralentissement de la recherche dans le domaine.

– Réalisation d'une mémoire associative, – un formalisme plus riche a été développé par Hopfield **H opfield** , [1982], **H inton et Sejnowski**, [1986 ], **P retto** , [1984]. – **W erbos (1974)**, a inventé l'idée de la rétro propagation de l'erreur destiné aux réseaux avec des connexions non bouclées, l'algorithme de rétro propagation a été développé indépendamment en 1985 par **R umelhart** ,[1986],**P arker** , [1985] et **L e cun** ,[1985 ].

### III.5. Le modèle neurone formel

C'est dans l'article " A logical Calculus of the ideas immanent in nervous activity" paru dans le bulletin of mathematical en 1943 qu'est apparu pour la première fois la notion de neurone formel [**Mc Culloch et Pitts**, 1943].

- Il est binaire : il est soit actif ou inactif.
- Chaque neurone possède un seuil de déclenchement fixé.

Le neurone reçoit des entrées de synapses excitatrices, ayant des poids identiques ; le neurone peut aussi recevoir des entrées de synapses inhibitrices, dont l'action est absolue ; c'est à dire que, si la synapse est active, le neurone ne peut pas être actif.

Un quantum de temps est nécessaire à l'intégration des entrées synaptiques, basé sur le délai observé physiologiquement.

Le fonctionnement est simple si aucune synapse inhibitrice n'est active, le neurone ajoute ces entrées synaptiques et vérifie que la somme atteint ou excède son seuil. On dit alors que le neurone devient actif, sinon le neurone est inactif.

Voyons tout cela en détail sur un exemple. Le neurone reçoit les entrées  $x_1, x_2, x_3 \dots x_n$ .

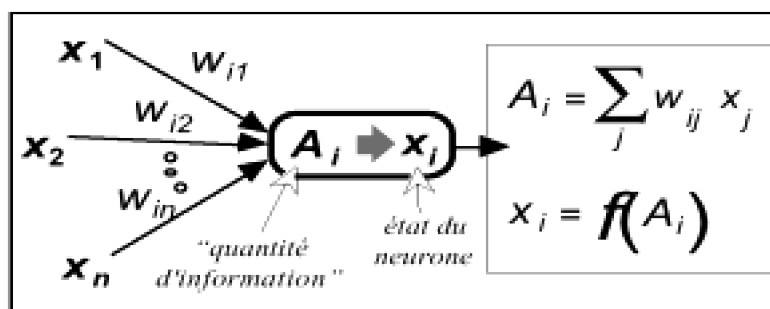


figure iii.2 : Le neurone formel

$x_1, x_2, x_3 \dots x_n$  sont les entrées, ou encore les excitations.  $w_1, w_2, w_3$  sont appelés les coefficients synaptiques ou poids synaptiques.

Le potentiel p est donné par  $A_i = x_1 * w_1 + x_2 * w_2 + x_3 * w_3 \dots + x_n * w_n$

La fonction de sortie est  $x_i = f(A_i)$  ou  $f()$  est la fonction de transfert. Le neurone créé par Mc Culloch et Pitts était un neurone utilisant une fonction binaire (**Figure. III.3**).

si  $p > a$  alors  $S=1$  sinon  $S = -1$

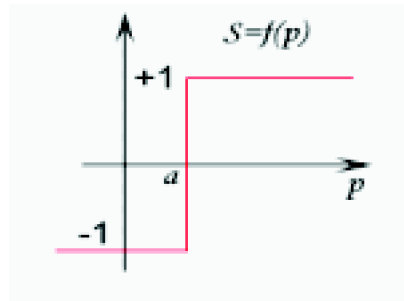


figure III.3 :Fonction de transfert binaire

### III.5.1. Structure

La **figure III.4** montre la structure d'un neurone artificiel. Celui ci est un processeur élémentaire. Il reçoit un nombre variable d'entrées en provenance de neurones précédents. A chacune de ces entrées est associée un poids  $w$  représentatif de la force de la connexion. Chaque processeur élémentaire est doté d'une sortie unique, qui se ramifie ensuite pour alimenter un nombre variable de neurones avals. à chaque connexion est associée un poids.

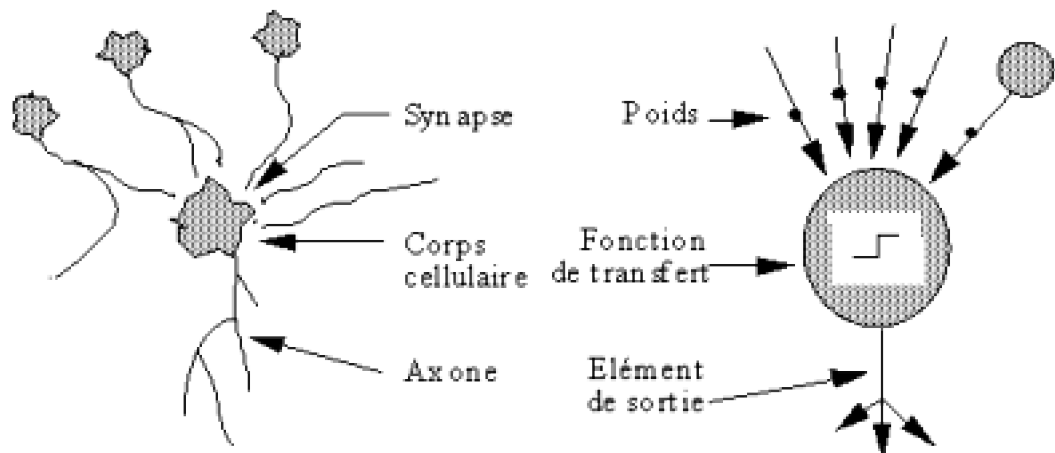


Figure III.4 : Mise en correspondance neurone biologique / neurone artificiel

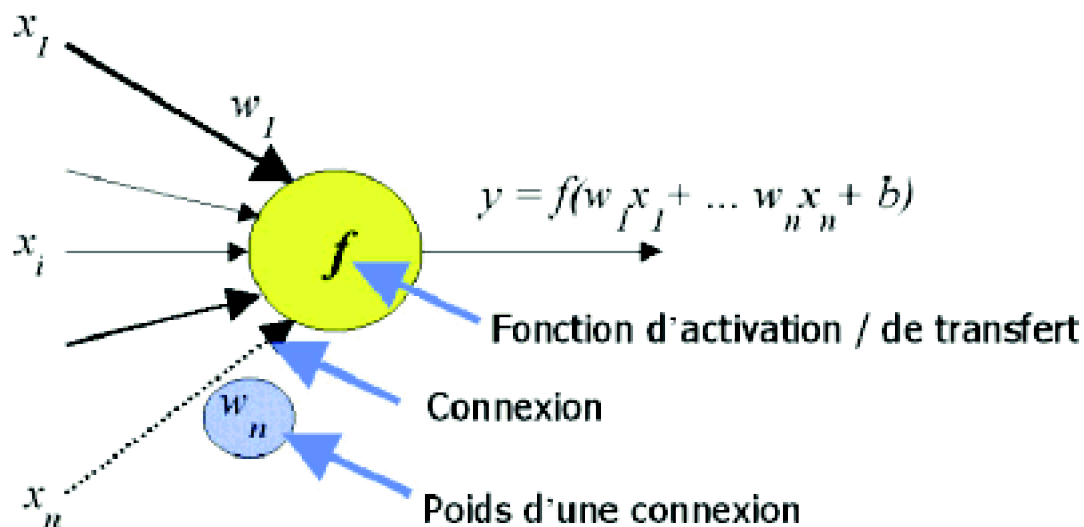


Figure III.5: Modélisation simplifiée d'un neurone artificiel

### III.5.2. Variables descriptives

Ces variables décrivent l'état du système. Dans le cas des réseaux de neurones qui sont des systèmes non autonomes, un sous-ensemble des variables descriptives est constitué par les variables d'entrées, variables dont la valeur est déterminée extérieurement au modèle.

Un réseau de neurones se compose de neurones qui sont interconnectés de façon à ce que la sortie d'un neurone puisse être l'entrée d'un ou plusieurs autres neurones. Ensuite il y a des entrées de l'extérieur et des sorties vers l'extérieur [Rumelhart et al. 1988 in Albert, 1992].

Rumelhart et al. donnent huit composants principaux d'un réseau de neurones

[Rumelhart et al., 1988 in Albert, 1992 ]:

- Un ensemble de neurones ;
- Un état d'activation pour chaque neurone (actif, inactif, ...) ;
- Une fonction de sortie pour chaque neurone ( $f(S)$ ) ;
- Un modèle de connectivité entre les neurones (chaque neurone est connecté à tous les autres, par exemple) ;
- Une règle de propagation pour propager les valeurs d'entrée à travers le réseau vers les sorties ;
- Une règle d'activation pour combiner les entrées d'un neurone (très souvent une somme pondérée) ;
- Une règle d'apprentissage ;
- Un environnement d'opération (le système d'exploitation, par exemple).

### III.5.3. Architecture

---

Un réseau de neurones est, comme son nom l'indique, un ensemble de neurones connectés en réseau, de sorte que les signaux sortant (outputs) des neurones deviennent des signaux entrant (inputs) dans d'autres neurones. L'architecture générale des réseaux de neurones consiste en la représentation des neurones en couches (layers) successives, la première représentant la couche d'intrants (input layer), appelée souvent la **rétine**, la dernière étant la couche de sortie (output layer), les couches intermédiaires étant les couches cachées (hidden layers) du réseau. Ces couches sont dites cachées, car de l'extérieur du réseau, on ne peut analyser clairement leur fonctionnement. On ne connaît vraiment que les signaux entrants et sortants du réseau. Les neurones de la couche d'intrants ne sont pas vraiment des neurones traitants (computing neurons), mais ont pour seule utilité de normaliser l'entrée des signaux ainsi que la distribution des signaux d'entrée. Dans cette architecture normalisée, les couches de neurones sont totalement interconnectées, c'est-à-dire que les neurones d'une couche sont tous reliés à tous les neurones des couches adjacentes. Cette architecture normalisée peut sembler rigide, mais elle permet une représentation correcte de la plupart du réseau de neurones, tout en permettant l'utilisation d'algorithmes d'entraînement plus généraux.

D'une manière générale, un réseau de neurones se comporte, d'un point de vue extérieur, comme une fonction  $S$  qui traite des données (inputs) et produit une réponse correspondante (output). Les données en entrée peuvent être de n'importe quel type de manière binaire ou numérique. Elles peuvent également être vues comme des vecteurs, et le réseau de neurones une application vectorielle.

## III.6. Processus d'apprentissage des réseaux de neurones

La notion d'apprentissage est claire et intuitive pour les humains ou les animaux : c'est une procédure cognitive qui doit faire en sorte que l'individu réalise, de manière autonome, une tâche donnée. [ **D reyfus, 2004**].

L'apprentissage numérique poursuit exactement le même objectif : il s'agit de faire en sorte, à l'aide d'une procédure numérique programmée et exécutée sur un ordinateur, d'inférer un modèle d'un processus que l'on observe et sur lequel on peut effectuer des mesures, c'est-à-dire un ensemble d'équations qui décrivent le processus observé et qui permet de faire des prédictions concernant le comportement de celui-ci. A cette fin, on fait l'hypothèse que le processus peut être décrit avec la précision désirée par une ou plusieurs fonctions qui contiennent des paramètres, et l'on ajuste ces derniers pour que cette ou ces fonctions s'ajustent aux données. L'apprentissage est un algorithme qui permet d'ajuster ces paramètres.

### III.6.1. Définition

---

C'est un processus dynamique et itératif permettant de modifier les paramètres d'un réseau en réaction avec les stimuli qu'il reçoit de son environnement. Le type d'apprentissage est déterminé par la manière dont les changements de paramètres surviennent.

L'objectif de l'apprentissage [Hinton, 1989], est de fournir une méthode au réseau afin qu'il puisse ajuster ces paramètres lorsqu'on lui présente des exemplaires à traiter.

### III.6.2. Types d'Apprentissage

---

Dans le cas des réseaux de neurones artificiels, on ajoute souvent à la description du modèle l'algorithme d'apprentissage. Dans la majorité des algorithmes actuels, les variables modifiées pendant l'apprentissage sont les poids des connexions.

L'apprentissage est la modification des poids du réseau dans l'optique d'accorder la réponse du réseau aux exemples et à l'expérience. Il est souvent impossible de décider à priori des valeurs des poids des connexions d'un réseau pour une application donnée. A l'issue de l'apprentissage, les poids sont fixés : c'est alors la phase d'utilisation.

Généralement il existe trois types d'apprentissages, l'apprentissage supervisé, non supervisé et l'apprentissage par renforcement .

Dans le cas de l'apprentissage non supervisé le réseau décide lui-même quelles sont les bonnes sorties. Cette décision est guidée par un but interne au réseau qui exprime une configuration idéale à atteindre par rapport aux exemples introduits.

Les cartes auto organisatrices de Kohonen sont un exemple de ce type de réseau [Kohonen, 1984].

L'apprentissage supervisé est le type d'apprentissage le plus utilisé. [Rumelhart et al. 1986].

#### III.6.2.1. L'apprentissage Supervisé

L'apprentissage supervisé est caractérisé par la présence d'un «professeur» qui possède une connaissance approfondie de l'environnement dans lequel évolue le réseau de neurones. En pratique, les connaissances de ce professeur prennent la forme d'un ensemble de  $Q$  couples de vecteurs d'entrée et de sortie que nous noterons  $\{(p_1, d_1), (p_2, d_2), \dots, (p_Q, d_Q)\}$ , où  $p_i$  désigne un stimulus (entrée) et  $d_i$  la cible pour ce stimulus, c'est-à-dire les sorties désirées du réseau. Chaque couple  $(p_i, d_i)$  correspond donc à un cas d'espace de ce que le réseau devrait produire (la cible) pour un stimulus donné. Pour cette raison, l'apprentissage supervisé est aussi qualifié d'apprentissage par des exemples [Marc, 2004].

L'apprentissage supervisé est illustré d'une manière conceptuelle à la figure III-6, l'environnement est inconnu du réseau. Celui-ci produit un stimulus  $p$  qui est acheminé à

la fois au professeur et au réseau. Grâce à ses connaissances intrinsèques, le professeur produit une sortie désirée  $d(t)$  pour ce stimulus. On suppose que cette réponse est optimale. Elle est ensuite comparée (par soustraction) avec la sortie du réseau pour produire un signal d'erreur  $e(t)$  qui est réinjecté dans le réseau pour modifier son comportement via une procédure itérative qui, éventuellement, lui permet de simuler la réponse du professeur. Autrement dit, la connaissance de l'environnement par le professeur est graduellement transféré vers le réseau jusqu'à l'atteinte d'un certain critère d'arrêt. Par la suite, on peut éliminer le professeur et laisser le réseau fonctionner de façon autonome [Marc, 2004].

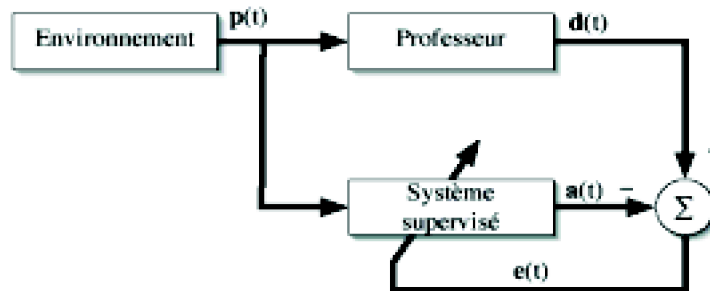


figure III.6 : Schéma bloc de l'apprentissage supervisé

L'apprentissage supervisé, tel que celui mis en oeuvre par l'algorithme de la rétro propagation du gradient d'erreur, nécessite la définition d'une base d'exemples d'apprentissage représentative. Chaque exemple est un couple (entrée, sortie désirée) et le principe de l'algorithme est de réduire l'erreur quadratique réalisée par le réseau. Cette erreur est la différence entre la valeur de sortie désirée et la valeur de sortie obtenue. La rétro propagation du gradient opérant sur un réseau multicouche a contribué de façon majoritaire à la popularité et à la diffusion des applications connexionnistes. Cependant, la difficulté inhérente à la construction d'une base d'apprentissage "représentative" limite son utilisation. Des recherches ont donc été conduites avec l'objectif de construire de manière automatique la base d'apprentissage.

## III.7. Le perceptron et le perceptron multicouches

### III.7.1. Le perceptron simple

Mis au point par Rosenblatt [1962], le perceptron se veut un modèle de l'activité perceptive. Il s'agit d'associer des configurations présentées en entrée à des réponses [figure III.7]. Il s'agit d'une seule couche de  $S$  neurones connectés totalement sur un vecteur  $P$  de  $R$  entrées. L'objectif est d'entraîner le réseau jusqu'à ce que la valeur obtenue en sortie soit égale à la valeur désirée. Pour cela on utilise la règle de Hebb qui consiste à renforcer les connexions entre deux cellules si elles répondent similairement,

sinon à diminuer cette valeur.

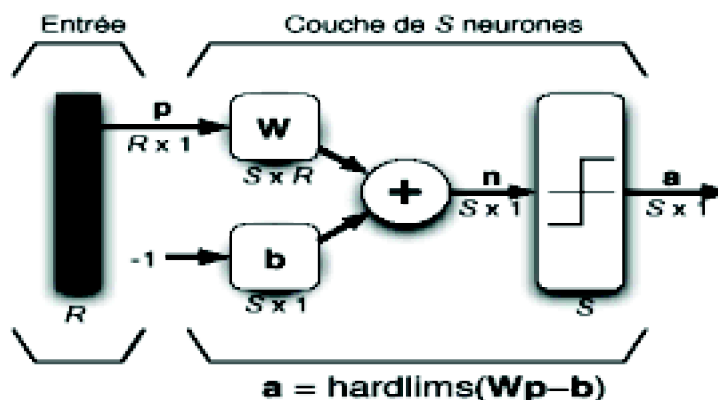


Figure III.7 : Perceptron à une seule couche avec fonction seuil

### III.7.2. Adaline et Widrow-Hoff

A la section précédente nous avons traité le cas du perceptron simple où les neurones utilisent une fonction de transfert de type "seuil".

Nous allons maintenant considérer la même architecture de réseau à une seule couche mais avec cette fois ci une fonction de transfert linéaire comme présentée à la [figure III.8]. Ce réseau s'appelle « ADALINE » [**AD**APtative **L**inear **NE**uron] à cause de sa fonction de transfert linéaire, il souffre des mêmes limitations que le perceptron simple ( il ne peut résoudre que des problèmes linéairement séparables). Cependant son algorithme d'apprentissage, la règle moindre carrée [**L**east **M**eans **S**quare] (LMS) est beaucoup plus puissante que la règle du perceptron original, bien que cette dernière peut ne pas converger dans certains cas vers une solution. En revanche la règle LMS minimise l'erreur quadratique moyenne, de sorte que la frontière de décision a tendance à se retrouver aussi loin que possible des prototypes. [Marc, 2004].

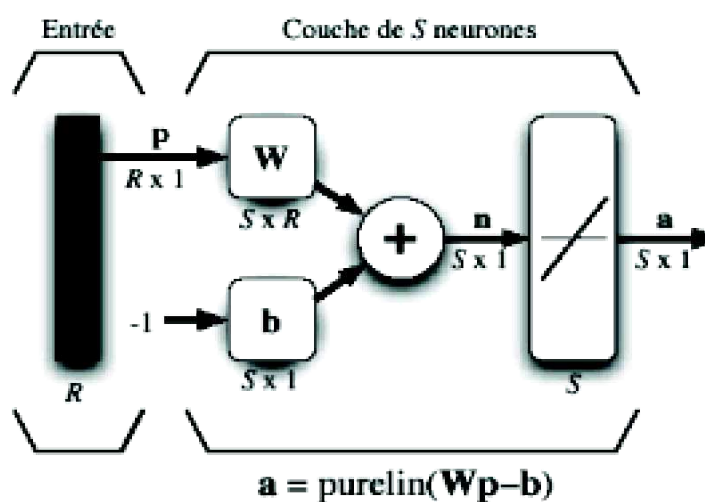


Figure III.8 : Réseau ADALINE



## III.8. Conclusion

A travers le présent chapitre nous avons exposé le principe de passage du neurone biologique à un neurone artificiel, Il s'agit tout simplement d'un ensemble d'entrées interconnectées d'une façon ou d'une autre pour former ce que l'on appelle un réseau de neurones par l'intermédiaire d'une fonction d'activation .

Les étapes de fonctionnement d'un réseau de neurones se subdivisent en phases de validation et d'apprentissage:

- Phase d'apprentissage : au cours de laquelle on présente au réseau la représentation du problème à résoudre et, dans le cas d'un apprentissage supervisé, la solution attendue. On aura donc au préalable codé le problème afin de le présenter au réseau. Le processus se répète sur tout les exemples à apprendre jusqu'à ce que les réponses données par le réseau soient jugées satisfaisantes.

- Phase de validation : au cours de laquelle on présente au réseau un problème qu'il a appris afin d'obtenir la solution. On peut aussi présenter un problème nouveau mais voisin des exemples appris afin d'obtenir une solution.



# Chapitre IV.Elaboration du modèle

## IV.1. Introduction

L'objectif de ce mémoire est l'élaboration d'un modèle de prévision de la dose du coagulant à injecter en usine de traitement des eaux lors de la phase de clarification (coagulation-floculation), et cela en fonction des paramètres descripteurs de l'eau brute (température, PH, turbidité, conductivité, oxygène dissous, et ultraviolets).

Le réseau de neurones utilisé est multicouche à sens unique et à sortie réelle. De manière plus précise, il comporte trois couches dont la dernière n'est constituée que d'un neurone (il représente la dose de coagulant à prévoir). Remarquons que chaque neurone est en relation avec tous les neurones des couches précédentes et de la couche suivante.

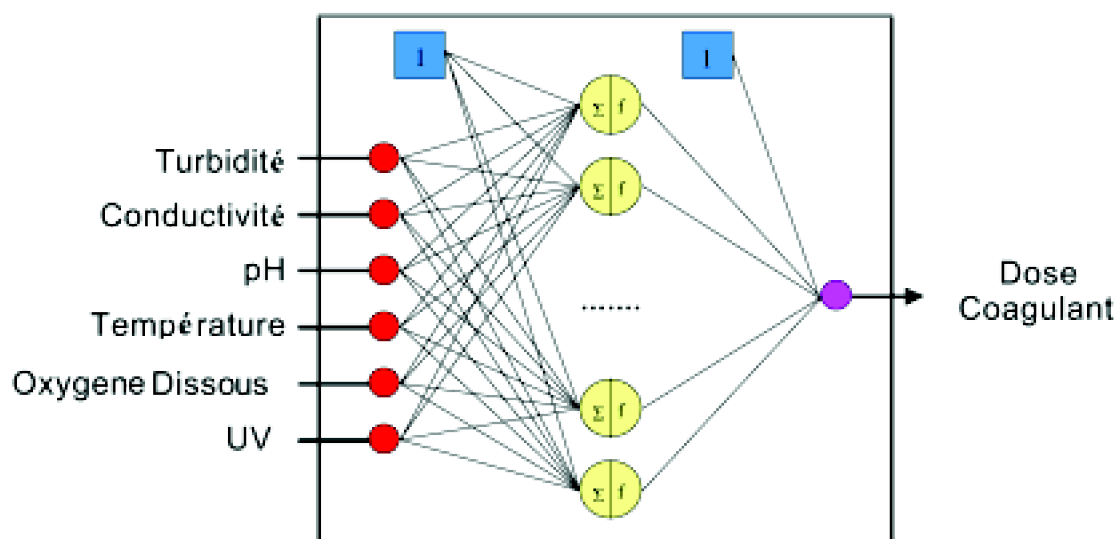


Figure IV.1 : Modèle de type boîte noire pour la prévision de la dose du coagulant.

Les entrées de ce réseau de neurones représentent la situation à étudier (les paramètres de l'eau brute); la fonction d'activation prise en considération est la fonction sigmoïde.

Il est important de signaler qu'il n'existe pas de relation directe entre les coefficients d'un réseau de neurones et son adaptation à une utilisation donnée.

## IV.2. Problématique

Les procédés de coagulation sont fortement non linéaires, non stationnaires et leur fonctionnement reste difficilement modélisable à cause de la mauvaise connaissance des phénomènes mis en jeu. Pour assurer une bonne conduite et contribuer au bon fonctionnement de ce procédé, il faudrait pouvoir exploiter toutes les informations disponibles sur le procédé lui-même, telles que les mesures des paramètres descripteurs de la qualité d'eau brute, et leurs évolutions qui peuvent provenir des anomalies imprévues. Pour ces raisons, nous avons choisi la modélisation de la dose du coagulant par la méthode des réseaux de neurones en utilisant une technique de prévision par apprentissage : le perceptron multicouches (MLP), qui permet d'exploiter et d'agrèger certaines informations afin de reconnaître des situations spécifiques et agir sur le procédé. L'idée générale est la suivante : Comment peut-on estimer la dose optimale du coagulant, en fonction des mesures faites sur l'eau brute à l'entrée de la station de traitement, afin de répondre aux exigences d'une bonne qualité de l'eau ?

Le choix des paramètres d'entrée a un effet direct sur la robustesse et la performance du réseau de neurones, ce dernier nécessite six paramètres comme données d'entrée :

- Température

- PH
- Conductivité
- Turbidité
- Oxygène dissous
- Ultras violets

Et la dose du coagulant comme paramètre de sortie.

L'apprentissage du réseau est réalisé à partir d'une base de données constituée de résultats de campagne des essais « Jar Test » réalisés au niveau du laboratoire de la station de traitement des eaux de Boudouaou.

### IV.3. Etat de l'art de la modélisation

Le contrôle permanent et automatique des processus de traitement implique la mesure d'un certain nombre de paramètres. Ceux-ci peuvent se classer en deux grandes familles : les paramètres usuels (débits, température, ...) et les paramètres spécifiques de l'eau (PH, conductivité, ...) [Valiron, 1988].

La mesure automatique d'un certain nombre de paramètres spécifiques permet d'optimiser les traitements en réduisant les temps de réponse.

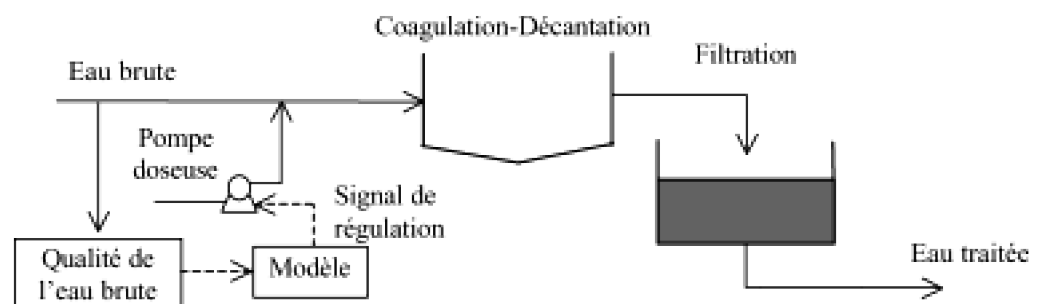


Figure IV.2: Procédé de clarification de l'eau brute.

La principale fonction d'une automatisation est l'amélioration des conditions d'exploitation. En effet, il est souhaitable de supprimer le maximum de tâches répétitives et pénibles pour l'opérateur, [Anonym2 1989, Valiron 1988].

L'intérêt du dispositif de contrôle du procédé est d'aboutir à une meilleure efficacité du traitement et une réduction des coûts d'exploitation.

Les approches développées actuellement sont basées sur les modèles intégrant les paramètres descripteurs de la qualité de l'eau brute.

Les premières études [Brodart et al. 1989, Critchley et al. 1990] sont basées sur un modèle linéaire reliant les paramètres descripteurs de l'eau, et le taux de traitement.

Brodart et al. (1989) proposaient d'utiliser un modèle linéaire pour relier la turbidité, la température et le débit de la rivière à la dose de coagulant. Ces études ont montré l'intérêt de cette approche mais également les limites de la modélisation linéaire.

Néanmoins, la modélisation de la dose du coagulant est redevenue très prometteuse depuis le développement de nouvelles techniques comme les réseaux de neurones artificiels [valentin ,2000].

Plusieurs études [Adgar et al .1995], [Cox et al. 1996], [Gagnon et al. 1997], [Mirsepassi et al. 1995], [Rodriguez et Sérodes. 1996] ont déjà montré le potentiel de ces techniques pour la modélisation du procédé de coagulation.

L'ensemble de ces études propose de relier la dose de coagulant à différents paramètres descripteurs de la qualité de l'eau brute.

Adgar et al. [1995] proposent d'utiliser la couleur, la conductivité et la turbidité de l'eau brute pour prédire la dose de coagulant. Les résultats obtenus sur un site pilote [Cox et al. 1996] semblent encourageants. Cependant, le manque de paramètres descripteurs en entrée ne permet pas de prendre en compte l'ensemble des variations de la qualité de l'eau brute. [Mirsepassi et al. 1995] propose d'utiliser beaucoup plus de paramètres en entrée en prenant en compte les valeurs des paramètres aux instants  $t-1$ ,  $t-2$ , ...,  $t-6$  ( $t$  représente le jour courant). Gagnon et al. (1997) montrent l'intérêt de construire des modèles saisonniers. Ils utilisent quatre paramètres descripteurs de la qualité de l'eau brute (pH, turbidité, conductivité, température), cette étude a permis la comparaison du modèle annuel, à quatre modèles saisonniers.

### IV.4. Présentation de la station de traitement

La station de traitement d'eau potable de Boudouaou (wilaya de Boumerdes), est prise en considération, celle-ci alimente en eau potable une grande partie de la wilaya d'Alger. L'eau à l'entrée de la station provient d'une prise d'eau sur le barrage de Keddara. Sa capacité maximale de traitement est de 6000 l/s. Compte tenue des caractéristiques de l'eau brute, la station utilise un procédé physico-chimique comportant les phases suivantes : pré-chloration, coagulation-floculation (clarification), décantation, filtration et désinfection.

L'eau traitée est finalement stockée dans deux réservoirs de 50 000 m<sup>3</sup> chacun.

### IV.5. Phases de l'élaboration du modèle

Elles se subdivisent en sept étapes :

La collecte des données ;

1.

L'analyse des données ;	2.
La séparation des bases de données (apprentissage-validation) ;	3.
Le choix de l'architecture du réseau de neurones ;	4.
La mise en forme des données ;	5.
L'apprentissage ;	6.
La validation ;	7.

### **IV.5.1. Collecte des données**

---

L'objectif de cette étape est de recueillir des données, à la fois pour développer et valider le modèle. Le but est de rassembler un nombre de données suffisant pour constituer une base représentative nécessaire à la modélisation.

### **IV.5.2. Analyse des données**

---

Celle-ci consiste à détecter d'éventuelles valeurs douteuses de la base de données recueillies qui peuvent avoir une répercussion sur les résultats de la modélisation.

### **IV.5.3. Séparation de la base de données**

---

La base de données est scindée en deux parties, la première servant comme base d'apprentissage (détermination des paramètres du modèle), et la seconde utilisée comme base de validation (performances du modèle). Il n'y a pas de règle pour déterminer ce partage de manière quantitative. Il résulte souvent d'un compromis tenant compte du nombre de données dont on dispose et du temps imparti pour effectuer l'apprentissage. Chaque base doit cependant satisfaire aux contraintes de représentativité de chaque classe de données, et doit généralement refléter la distribution réelle, c'est à dire la probabilité d'occurrence des diverses classes.

### **IV.5.4. Choix de l'architecture du réseau**

---

Il existe un grand nombre de types de réseaux de neurones, avec pour chacun des avantages et des inconvénients. Le choix de l'architecture du réseau peut dépendre :

- De la tâche à effectuer (classification, association, contrôle de processus, ...) ;
- De la nature des données (données présentant des variations au cours du temps) .

Ce choix est aussi fonction de la maîtrise, ou de la connaissance que l'on a de certains réseaux, ou encore du temps dont on dispose pour tester une architecture prétendue plus performante.

### IV.5.5. Mise en forme des données

---

De manière générale, la base de données doit subir un prétraitement afin d'être adaptée aux entrées et sorties du réseau de neurones.

Un prétraitement courant consiste à effectuer une normalisation des données :

$$Z_i = \frac{X_i - \bar{X}}{\sigma_x}$$

Avec :

$Z_i$  : Variable centrée réduite ;

$X_i$  : Variable d'entrée ;

$\bar{X}$  : Moyenne ;

$\sigma_x$  : Ecart type.

### IV.5.6. Apprentissage du réseau de neurones

---

La plupart des modèles de réseaux de neurones requièrent un apprentissage. Plusieurs types d'apprentissages peuvent être adaptés à un même type de réseau de neurones. Les critères de choix sont souvent la rapidité de convergence, ou les performances de généralisation (Chap. III-6).

### IV.5.7. Validation

---

Une fois le réseau de neurones entraîné, il est nécessaire de le tester sur une base de données différente de celle utilisée pour l'apprentissage. Ce test permet d'apprécier les performances du modèle élaboré. Si ces performances ne sont pas satisfaisantes, il faudra soit modifier l'architecture du réseau, soit modifier la base d'apprentissage. La démarche expérimentale exige cependant une étape de validation avant toute utilisation. Elle est en effet nécessaire, pour vérifier la fiabilité de la représentation interne du réseau. Elle consiste en fait à simuler l'utilisation du réseau.

Cette validation est effectuée à l'aide de couples entrées-sorties, autres que ceux utilisés lors de l'apprentissage. Évidemment, le réseau étant par définition un outil d'interpolation, les valeurs des entrées et des sorties de ces couples, tout comme celles des entrées fournies lors de l'utilisation, appartiennent aux mêmes intervalles que celles des expériences qui ont servies à l'apprentissage.

La validation consiste alors à fournir au réseau de neurones les entrées de ces



couples, à partir desquelles il calcule (grâce aux poids définis à l'apprentissage) les sorties correspondantes. Ce qui permet de constater l'erreur commise par le réseau de neurones. Cette étape de validation est particulièrement utilisée lors de l'ajustement d'un réseau de neurones à la situation à modéliser, puisqu'il quantifie l'erreur commise. Elle est donc systématiquement réalisée à chaque test effectué lors de la détermination de l'architecture optimale et des autres paramètres (fonction d'activation par exemple); Il s'agit d'une évaluation de l'erreur de généralisation. La différence entre l'erreur d'apprentissage et l'erreur de généralisation représente une mesure de la qualité de l'apprentissage effectué; [Vapnik 1982, Vapnik 1995, Boucheron 1992].

L'erreur de généralisation dépend avant tout de trois paramètres : le nombre d'exemples utilisé pour l'apprentissage, la complexité du problème sous-jacent et l'architecture du réseau.

Les approches statistiques de la généralisation [Vapnik et Chervonenkis 1971, Vapnik 1982] sont l'un des domaines d'investigation majeur pour optimiser les performances de l'apprentissage des réseaux de neurones. On constate qu'une amélioration de la généralisation peut être observée sous deux aspects imbriqués :

- Si la taille du réseau est fixée, quel est l'ensemble d'exemples d'apprentissage qui donnera la meilleure généralisation?
- Si le nombre des exemples est fixé, comment choisir le réseau pour avoir la meilleure généralisation?

Il est difficile de répondre à chacune de ces questions de manière individuelle tant la réponse à l'une nécessite de se pencher sur l'autre. Il est nécessaire ici d'introduire la notion de capacité : si on considère le réseau de neurones comme un système permettant de choisir une fonction parmi un ensemble déterminé par la structure du réseau, alors la capacité du système représente le nombre d'exemples que le réseau peut apprendre correctement.

## IV.6. Elaboration du modèle

Tous les neurones utilisés sont interconnectés entre eux, les paramètres (poids) reliant ces neurones sont initialisés aléatoirement, et on les laisse évoluer sans contraintes. Le neurone de sortie est linéaire, il reçoit les sommes pondérées de toutes les sorties activées des neurones de la couche cachée.

### IV.6.1. Logiciel, méthodologie d'implémentation

---

L'absence de théorèmes précis concernant les possibilités pratiques des réseaux de neurones entraîne souvent un recours massif à l'expérimentation. Pour faciliter ces expérimentations, de nombreux simulateurs neuronaux ont été proposés

[Azema-Barac et al. 1990, Pim et al. 1991, Benoît et al. 1989, Fuentes et al. 1993, Alexander et al. 1992, Edmond et al. 1992]. Ces simulateurs pèchent souvent par excès de généralité, ils implémentent de très nombreux modèles assez différents (perceptrons multicouches (MLP), modèle de Hopfield, Kohonen, etc. ...).

Les possibilités d'extension de ces simulateurs sont en général assez limitées par cette absence d'unification.

Nous avons développé une application implémentée sous le langage DELPHI. Le code de cette méthodologie est partagé suivant des procédures et des fonctions. Ainsi, dans l'implémentation nous avons pris en compte la structure du réseau de neurone. Un tel réseau se compose de :

- Une première couche composée de six paramètres, et souvent appelée rétine.
- Une couche d'entrée composée d'un nombre variable de neurones (entre 1 et 20).
- Une couche de sortie composée d'un seul neurone.

Les fichiers de données sont formés, par les données nécessaires à l'entraînement du réseau de neurones, qui ne sont autres que les six paramètres de la qualité de l'eau brute et la dose du coagulant à modéliser, ils sont structurés sous forme d'enregistrements en ligne et en colonne, chaque ligne correspond à un enregistrement complet, et chaque colonne correspond à une variable étudiée.

Pendant la phase d'exploitation du réseau de neurones, le programme fait appel au fichier et lit ainsi les enregistrements l'un après l'autre.

Il faut noter cependant qu'on peut accéder aux données sans respecter l'ordre de classement, ce qui permet de scinder la base de données selon un partage permettant une meilleure généralisation.

Notre méthodologie d'implémentation consiste à décrire le réseau de neurones, en tant qu'organigramme structuré (structure simple et logique), où l'on pourra spécifier et modifier chaque fonction selon les besoins, et enfin, à décrire les fonctions de propagation et de rétro propagation.

L'approche utilisée est parfaitement adaptée pour l'approximation de fonctions non linéaires par les réseaux de neurones et justifie l'intérêt de leur utilisation :

- A. Programmation effectuée par un neurone.
- B. Mise en place de contraintes sur l'architecture du réseau.

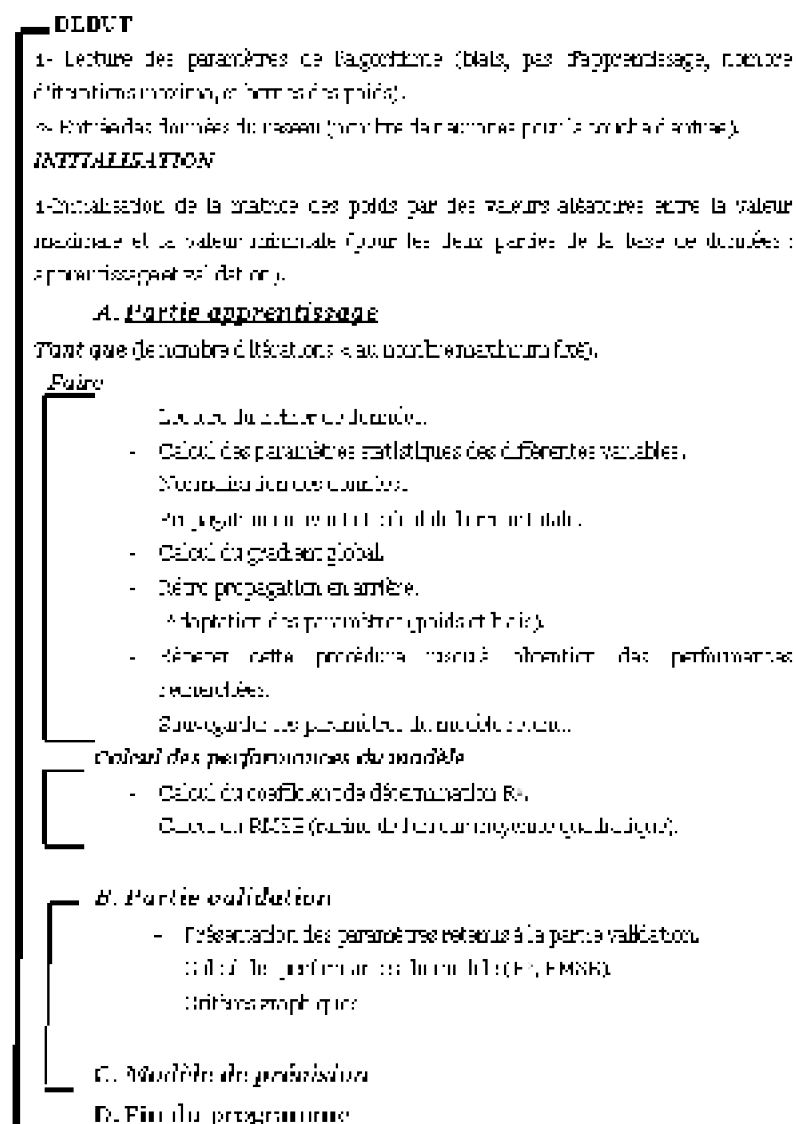
### IV.6.2. Conception des algorithmes

---

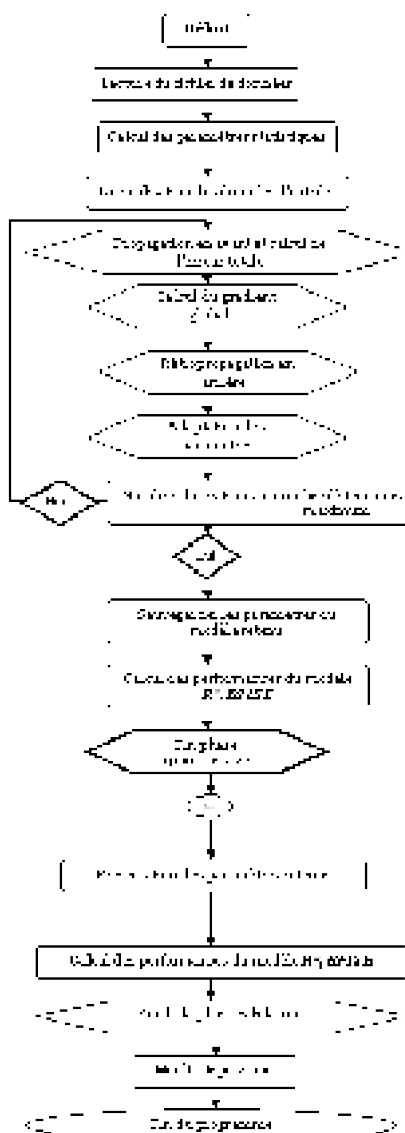
- |   |    |
|---|----|
| Ouverture et lecture des fichiers de données.   | 1. |
| Prétraitement des données et calcul des paramètres statistiques.  | 2. |
| Définition des paramètres du réseau (nombre de neurones, taux d'apprentissage, bornes des poids maximum et minimum, nombre d'itérations). | 3. |

- Initialisation aléatoire des poids. 4.
- Propagation en avant et calcul de l'erreur totale. 5.
- Rétro propagation en arrière. 6.
- Arrêt de l'apprentissage (critère de convergence atteint). 7.
- Vérification des performances du modèle (critères graphiques, critères numériques (coefficient de détermination, RMSE)). 8.
- Modèle de prévision. 9.

### IV.6.3. Algorithme général du programme



### IV.6.4.Organigramme du programme



### IV.6.5. L'algorithme du rétropropagation du gradient

L'algorithme de rétro propagation du gradient (RP) est à la base des premiers succès des réseaux de neurones. Sa mise en application a permis au domaine du connexionnisme de sortir de la période de silence qui a régné après la sortie du livre «Perceptrons» de Minsky et Papert (1969). Il figure aujourd'hui parmi les algorithmes d'apprentissage les plus utilisés. Il a été appliqué avec succès à une grande variété de problèmes [Sejnowski et Rosenberg, 1987], [Gorman et Sejnowski, 1988], [Canu et al. 1990, Canu et al. 1997, Ding et al. 1996], [Boisseau et al. 1995].

### IV.6.6. Fonctions coût

Le choix de la fonction coût (ou d'erreur) utilisée pour l'apprentissage des réseaux de

neurones multicouches a une certaine influence sur la rapidité d'apprentissage, et sur la qualité de généralisation du réseau [Gayner et Downs 1994, Moller 1993].

Le critère d'erreur le plus utilisé est la fonction d'erreur quadratique moyenne. Celle-ci a tendance à amplifier les erreurs les plus importantes, par conséquent, au cours de l'apprentissage, la mise à jour des poids est largement déterminée par la correction des grandes erreurs, ce qui est recherché en général [Hassibi et Stork 1993, Fahlman 1988].

## IV.7. Implémentation de l'algorithme du gradient

La technique la plus appropriée pour faire apprendre à un réseau certaines propriétés de calcul consiste en un algorithme *Essai-Erreur*, et correction sur une base d'apprentissage, constituée d'éléments pour lesquels on connaît les résultats auxquels devrait aboutir le réseau.

### IV.7.1. Formulation mathématique du problème

Soit le couple  $\{x(n), d(n)\}$ , désignant la  $n^{\text{ième}}$  donnée d'entraînement du réseau où

$$X(n) = \{x_1(n), \dots, x_p(n)\} \text{ et } d(n) = \{d_1(n), \dots, d_q(n)\} \quad (\text{IV-1})$$

Correspondent respectivement aux  $p$  entrées et aux  $q$  sorties désirées du système.

L'algorithme de rétro propagation consiste alors à mesurer l'erreur entre les sorties désirées  $d(n)$  et les sorties observées  $y(n)$  :

$$y(n) = \{y_1(n), \dots, y_q(n)\} \quad (\text{IV-2})$$

résultant de la propagation vers l'avant des entrées  $x(n)$ , et à rétropropager cette erreur à travers les couches du réseau en allant des sorties vers les entrées.

#### 1. Cas de la couche de sortie :

L'algorithme de rétro propagation procède à l'adaptation des poids (neurone par neurone) en commençant par la couche de sortie.

Soit l'erreur observée  $e_j(n)$  pour le neurone de sortie  $j$ , et la donnée d'entraînement  $n$

$$e_j(n) = d_j(n) - y_j(n) \quad (\text{IV-3})$$

Où  $d_j(n)$  correspond à la sortie désirée du neurone  $j$ , et  $y_j(n)$  à sa sortie observée.

La variable  $n$  représentera toujours la donnée d'entraînement c'est-à-dire le couple contenant un vecteur d'entrées et un vecteur de sorties désiré.

- L'objectif de l'algorithme est d'adapter les poids des connexions du réseau de manière à minimiser la somme des erreurs sur tous les neurones de sortie.
- L'indice j représentera toujours le neurone pour lequel on veut adapter les poids.

Soit  $E(n)$  la somme des erreurs quadratiques observées sur l'ensemble C des neurones-de-sortie.

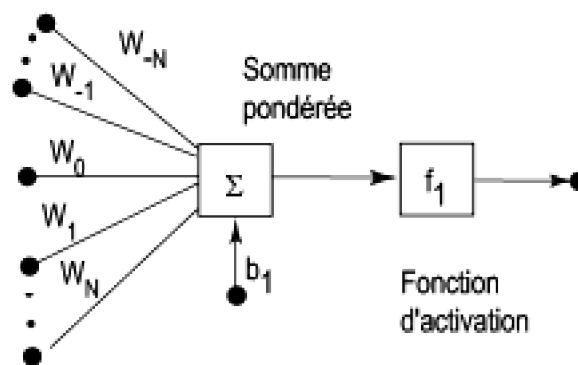


Figure IV.3: Architecture neuronale.

$$E(n) = \frac{1}{2} \sum_{j \in C} e_j^2(n) \quad (IV-4)$$

La sortie  $y_j(n)$  du neurone j est définie par :

$$y_j(n) = \varphi \left[ v_j(n) \right] = \varphi \left[ \sum_{i=0}^N w_{ji}(n) y_i(n) \right] \quad (IV-5)$$

Où  $\varphi$  : Fonction d'activation du neurone.

$v_j(n)$  : Somme pondérée des entrées du neurone j,

$w_{ji}(n)$  : Poids des connexions entre le neurone i de la couche précédente et le neurone j de la couche courante,

$y_i(n)$  : Sortie du neurone i.

On suppose ici que la couche précédente contient r neurones numérotés de 1 à r, que le poids  $w_{j0}(n)$  correspond au biais du neurone j, et que l'entrée  $y_0(n) = 1$  (entrée du biais).

Pour corriger l'erreur observée, il s'agit de modifier les poids  $w_{ji}(n)$  dans le sens opposé au

$$\text{gradient } \frac{\partial E(n)}{\partial w_{ji}(n)} \text{ de l'erreur.}$$

Cette dérivée partielle représente un facteur de sensibilité.

On a aussi le dérivé de la fonction de transfert  $\frac{\partial \sigma(x)}{\partial x} = \sigma'(x) = \sigma(x)(1-\sigma(x))$  obtenu :

$$\frac{\partial \sigma(x)}{\partial x} = \frac{\partial \sigma(x)}{\partial x} \frac{\partial x}{\partial x} = \frac{\partial \sigma(x)}{\partial x} \frac{\partial \sum_j W_{ji} x_j}{\partial x} = \frac{\partial \sigma(x)}{\partial x} \sum_j W_{ji} \quad (173-0)$$

Travaillons maintenant le cas de sortie :

$$\frac{\partial E(n)}{\partial W_{kj}} = \frac{\partial E(n)}{\partial \sigma(x_k)} \frac{\partial \sigma(x_k)}{\partial W_{kj}} \quad (173-1)$$

avec  $0 < \sigma(x_k) < 1$

Travaillons maintenant l'erreur de gradient :

$$\frac{\partial E(n)}{\partial \sigma(x_k)} = \frac{\partial \sum_j \sigma_j^2(n)}{\partial \sigma(x_k)} = \frac{2 \sigma(x_k) \sigma'(x_k)}{1 - \sigma(x_k)}$$

$$= \frac{2 \sigma(x_k) \sigma'(x_k)}{1 - \sigma(x_k)} = \delta_k(n)$$

$$\frac{\partial E(n)}{\partial \sigma(x_k)} = \frac{\delta_k(n) \sigma'(x_k)}{1 - \sigma(x_k)}$$

$$\frac{\partial E(n)}{\partial \sigma(x_k)} = \frac{\delta_k(n) \sigma'(x_k)}{1 - \sigma(x_k)}$$

$$\left[ \begin{array}{c} \delta_k(n) \\ \delta_k(n) \end{array} \right] = \frac{\delta_k(n)}{1 - \sigma(x_k)} \left[ \begin{array}{c} \sigma'(x_k) \\ -\sigma'(x_k) \end{array} \right]$$

$$\Rightarrow \delta_k(n) \left[ \begin{array}{c} \sigma'(x_k) \\ -\sigma'(x_k) \end{array} \right] = \delta_k(n) \sigma'(x_k)$$

$$\Rightarrow \delta_k(n) = \delta_k(n) \sigma'(x_k)$$

En conclusion :

$$\frac{\partial E(n)}{\partial W_{kj}} = \frac{\delta_k(n) \sigma'(x_k) \sum_j W_{ji}}{\sigma'(x_k)}$$

$$= \frac{\delta_k(n) \sum_j W_{ji}}{\sigma'(x_k)}$$

$$\delta_k(n)$$

En conclusion : on a :

$$\frac{\partial E(n)}{\partial W_{kj}} = -\delta_k(n) \sigma'(x_k) [1 - \sigma(x_k)] \sigma'(x_k) \quad (173-2)$$

On a aussi le dérivé de la fonction de transfert de la couche cachée :

$$\frac{\partial \sigma(x_k)}{\partial x_k} = \frac{\partial \sigma(x_k)}{\partial x_k} \frac{\partial x_k}{\partial x_k} = \frac{\partial \sigma(x_k)}{\partial x_k} \frac{\partial \sum_j W_{kj} x_j}{\partial x_k} = \frac{\partial \sigma(x_k)}{\partial x_k} \sum_j W_{kj} \quad (173-3)$$

avec :

$$\delta_k(n) = \delta_k(n) \sigma'(x_k) [1 - \sigma(x_k)] \quad (173-4)$$

correspondant au gradient local.

## 2. Cas de la couche cachée :

Considérons maintenant le cas des neurones de la dernière couche cachée (le cas des autres couches est semblable).

- La variable n désignera toujours la donnée d'entraînement c'est-à-dire un couple de vecteur d'entrée et de sortie désirée.
- L'objectif sera toujours d'adapter les poids de la couche courante en minimisant la somme des erreurs sur les neurones de la couche de sortie.
- Les indices i et j désignent respectivement un neurone sur la couche précédente et un neurone sur la couche courante.
- L'indice k servira à désigner un neurone sur la couche suivante.

Reprenons l'expression de la dérivée partielle de l'erreur totale E (n) par rapport à Wji,

mais en ne dérivant plus par rapport à l'erreur  $e_j(n)$  car celle-ci est inconnue :

$$\frac{\partial E(n)}{\partial W_j(n)} = \frac{\partial E(n)}{\partial y_j(n)} \cdot \frac{\partial y_j(n)}{\partial v_j(n)} \cdot \frac{\partial v_j(n)}{\partial W_j(n)} \quad (IV-11)$$

Par rapport aux résultats obtenus pour la couche de sortie, les deux derniers termes de cette équation restent inchangés, seul le premier terme requiert d'être évalué :

$$\frac{\partial E(n)}{\partial W_j(n)} = \frac{\partial \left[ \sum_{k=0}^K e_k(n) \right]}{\partial y_j(n)} \quad (IV-12)$$

Notre problème ici, contrairement au cas des neurones de la couche de sortie, est que tous les  $e_k(n)$  dans la somme ci-dessus dépendent de  $y_j(n)$ . On ne peut donc pas se débarrasser de cette somme ; Néanmoins, nous pouvons écrire :

$$\begin{aligned} \frac{\partial E(n)}{\partial W_j(n)} &= \sum_{k=0}^K \left[ e_k(n) \cdot \frac{\partial e_k(n)}{\partial y_j(n)} \right] \\ &= \sum_{k=0}^K \left[ e_k(n) \cdot \frac{\partial e_k(n)}{\partial v_k(n)} \cdot \frac{\partial v_k(n)}{\partial y_j(n)} \right] \\ &= \sum_{k=0}^K \left[ e_k(n) \cdot \frac{\partial |d_k(n) - \sigma(v_k(n))|}{\partial v_k(n)} \cdot \frac{\partial \sum_{i=0}^K W_{ki}(n) y_i(n)}{\partial y_j(n)} \right] \\ &= \sum_{k=0}^K [e_k(n) (-y_k(n) [-y_k(n)] W_{kj}(n))] \end{aligned}$$

Et en substituant l'équation (IV-12) on obtient :

$$\frac{\partial E(n)}{\partial W_j(n)} = \sum_{k=0}^K \delta_k(n) W_{kj}(n) \quad (IV-13)$$

En substituant l'équation (IV-13) dans l'équation (IV-11), on obtient :

$$\frac{\partial E(n)}{\partial W_j(n)} = Y_j(n) [1 - Y_j(n)] \sum_{k=0}^K \delta_k(n) W_{kj}(n) \quad (IV-14)$$

P :

$$\Delta W_{pj}(n) = -\tau \frac{\partial E(n)}{\partial W_{pj}(n)} = \eta \delta_j(n) Y_j(n) \quad (IV-15)$$

Ax : :

$$\delta_j(n) = Y_j(n) [1 - Y_j(n)] \sum_{k=0}^K \delta_k(n) W_{kj}(n) \quad (IV-16)$$



Les équations (IV-15) et (IV-16) sont valables pour toutes les couches cachées.

Notons cependant, que dans le cas de la première couche cachée du réseau, puisqu'il n'y a pas de couche précédente de neurones, il faut substituer la variable  $Y_i(n)$  par l'entrée  $X_i(n)$ .

### 3. Sommaire de l'algorithme :

L'algorithme de rétro propagation se résume donc à la série des étapes suivantes :

1. Initialiser tous les poids à des valeurs aléatoires dans l'intervalle [-0.5, 0.5].
2. Normaliser les données d'entraînement.
3. Présenter aléatoirement les données d'entraînement.
4. Pour chaque donnée d'entraînement  $n$  :
  - Calculer les sorties observées en propageant les entrées vers l'avant ;
  - Ajuster les poids en rétro propageant l'erreur observée :

$W_{ji}(n) = W_{ji}(n-1) + \Delta W_{ji}(n) = W_{ji}(n-1) + \eta \delta_j(n) Y_i(n)$	(IV-17)
--	---------

Où "le gradient local" est défini par :

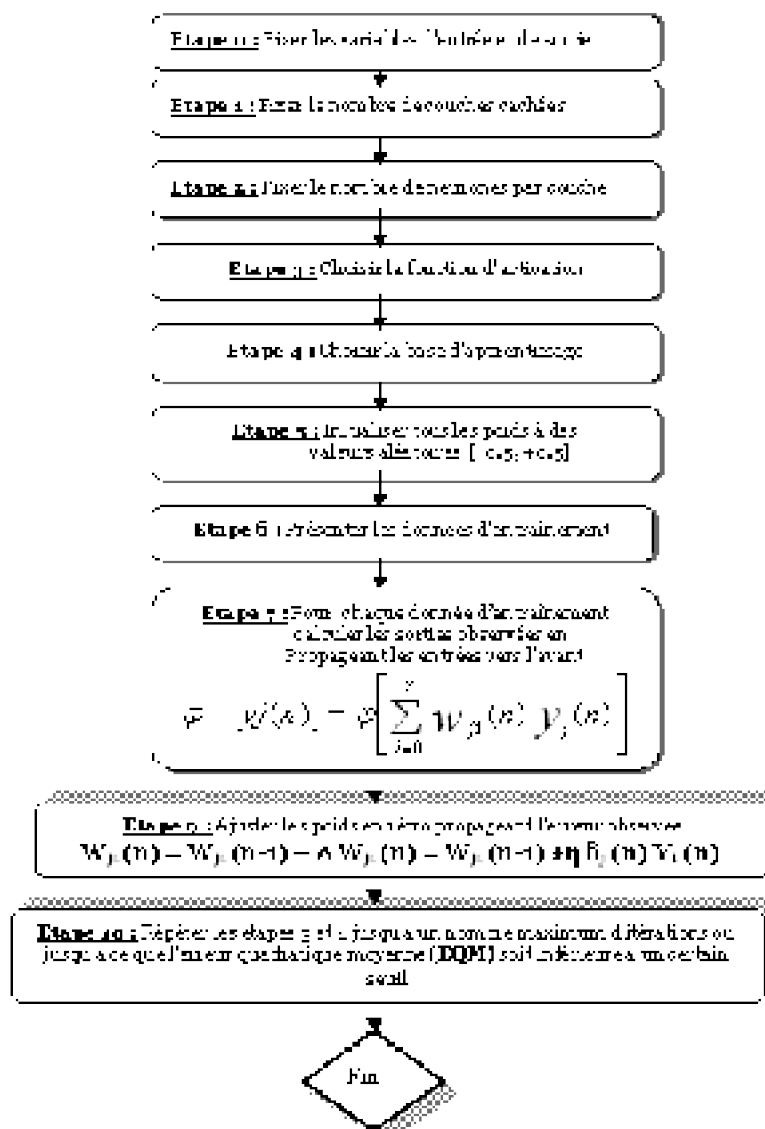
$$\delta_j(n) = e_j(n) Y_j(n) [1 - Y_j(n)] \text{ si } j \text{ appartient à une couche de sortie}$$

$W_{ji}(n) = W_{ji}(n-1) + \Delta W_{ji}(n) = W_{ji}(n-1) + \eta \delta_j(n) Y_i(n)$	(IV-18)
--	---------

$$\delta_i(n) = Y_i(n) [1 - Y_i(n)] \sum_k \delta_k(n) W_{ij}(n) \text{ si } j \text{ appartient à une couche cachée.}$$

Avec  $0 \leq \eta \leq 1$  représentant le taux d'apprentissage et  $Y_i(n)$  représentant soit la sortie du neurone  $i$  sur la couche précédente, si celui-ci existe, soit l'entrée  $i$  autrement.

Répéter les étapes 3 et 4 jusqu'à un nombre maximum d'itérations ou jusqu'à ce que l'erreur quadratique moyenne soit inférieure à un certain seuil.



4. Organigramme de rétropropagation de l'erreur :

## IV.7.2. Critères de validation.

Une fois le modèle construit se pose alors la question s'il rend fidèlement compte de l'échantillon. Pour cela, la modélisation nécessite des critères de validation sur la base desquels, le concepteur juge les performances finales du modèle qui sera retenu. Certains sont numériques et d'autres graphiques.

### IV.7.2.1. Le coefficient de détermination ( $R^2$ )

Celui-ci évalue le degré d'association entre deux variables et juge de la qualité de l'ajustement des valeurs observées et calculées :

$$R^2 = \left[ \frac{\frac{1}{N} \sum_{i=1}^N (Y_{iobs} - \bar{Y}_{obs})(Y_{ical} - \bar{Y}_{ical})}{\sigma_{obs} \cdot \sigma_{ical}} \right]^2$$

Avec :

$Y_{iobs}$  : Valeur observée.

$Y_{ical}$  : Valeur calculée par le modèle.

$\bar{Y}_{obs}$  : Moyenne de la variable observée.

$\bar{Y}_{ical}$  : Moyenne de la variable calculée.

Un coefficient de détermination proche de l'unité signifie que le modèle élaboré est très efficace.

#### IV.7.2.2. Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Y_{iobs} - Y_{ical})^2}{N}}$$

Un RMSE proche de zéro signifie que le modèle est très bon.

#### IV.7.2.3. Les critères graphiques

L'analyse graphique est indispensable et primordiale, cela est obtenu en portant sur un graphique les valeurs observées et celles calculées.

### IV.7.3. Choix d'initialisation des paramètres

Le choix initial des poids d'un réseau influence significativement le processus d'entraînement. Les poids doivent être choisis au hasard, mais il faut s'assurer que les sigmoïdes sont activées dans la région linéaire. Dans le cas où les poids ont des valeurs numériques élevées, les sigmoïdes vont saturer et engendrer des gradients faibles qui ralentiront l'apprentissage initial. Si les valeurs sont trop petites, l'apprentissage sera aussi très lent. Des valeurs de poids moyennes, amenant la sigmoïde dans la région

légèrement non linéaire, entraînent les gradients à être suffisamment grands pour permettre l'apprentissage, et assurer que le réseau apprenne les parties linéaires avant les parties non linéaires. Si la fonction d'activation est une sigmoïde et qu'elle est activée dans la région linéaire, les valeurs de la dérivée se trouveront dans la zone de valeurs les plus élevées et par conséquent, le gradient sera élevé. Accomplir une bonne initialisation nécessite la coordination de la normalisation de l'ensemble d'entraînement, du choix de la sigmoïde et du choix d'initialisation des poids. Cela est accompli au niveau de la couche d'entrée en normalisant tous les exemples d'entraînement.

Dans le cadre de ce mémoire nous avons adopté une couche d'entrée à 20 neurones maximum.

## IV.8. Applications

### IV.8.1. Base de données.

---

Nous avons élaboré une banque de données provenant des archives de la station de Boudouaou pour une période de 13 ans (1993-2005). Nous avons retenu six paramètres comme variables d'entrées.

Cependant la base de données telle qu'elle a été constituée n'est pas homogène.

Pour chaque enregistrement (qui sera par la suite défini comme un vecteur d'entrée pour les ACP et les réseaux de neurones); nous avons pris soin d'indiquer la date du jour de prélèvement correspondant ce qui nous a permis par la suite de classer la base de données selon un critère saisonnier.

Finalement les descripteurs retenus pour modéliser la dose du coagulant sont répartis comme suit :

- Données annuelles (six paramètres d'entrée : 1023 observations)
- Données saisonnières (six paramètres d'entrée) réparties comme suit :
  1. Hiver (Décembre, Janvier et Février).
  2. Printemps (Mars, Avril et Mai).
  3. Été (Juin, Juillet et Août).
  4. Automne (Septembre, Octobre et Novembre).

Nous disposons finalement de cinq fichiers de données (un annuel, et quatre saisonniers).

### IV.8.2. Statistique des paramètres étudiés

---

Nous avons procédé à l'analyse statistique des paramètres retenus dans le cadre de cette étude, les résultats obtenus sont résumés dans le tableau suivant :

Tableau IV.1 : Statistiques des paramètres retenus (fichier annuel : 1023 observations).

	Moyenne	Ecart type	Min	Max
Température (°C )	16.76	3.39	10	27.2
PH	7.77	0.268	7.11	8.92
Conductivité ( $\mu\text{s}/\text{cm}^{-1}$ )	1002.37	140.49	658	1867
Turbidité (NTU)	7.33	6.27	0.44	94
Oxygène dissous (mg/l)	4.56	2.88	0.13	13.2
Ultraviolet (NM)	0.11	0.069	0.013	1.1
Coagulant (mg/l)	22.06	14.07	3	100

### IV.8.3. Graphes des variations annuelles

Les évolutions des variations annuelles des différents paramètres étudiés sont représentées par les figures ci après :

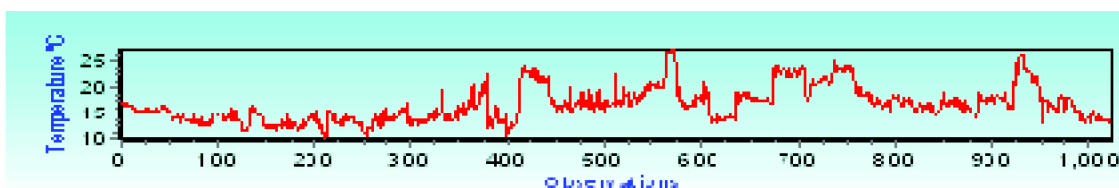


Figure IV.4 : Variation annuelle de la température-1023 observations

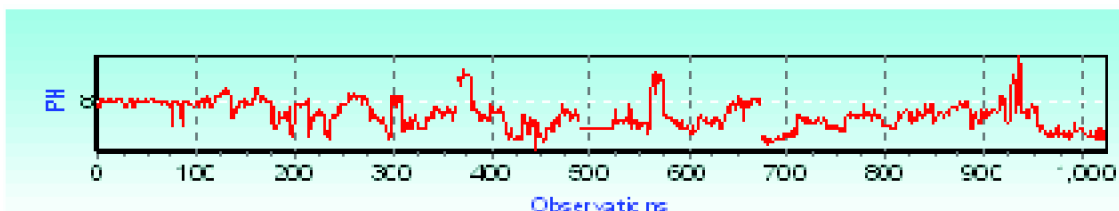


Figure IV.5 : Variation annuelle des PH-1023 observations

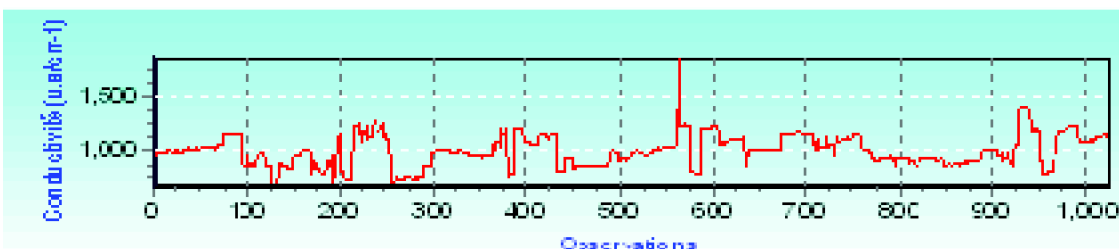


Figure IV.6 : Variation annuelle de la conductivité-1023 observations

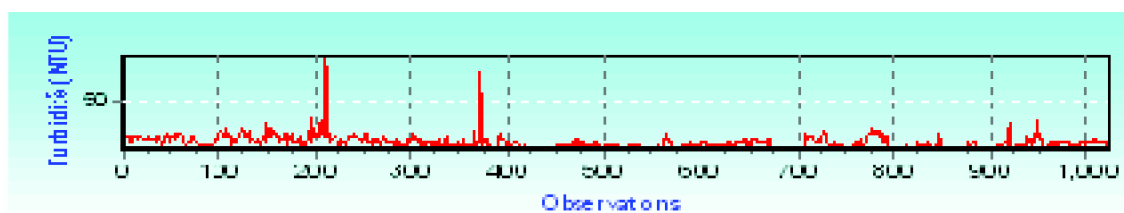


Figure IV.7 : Variation annuelle de la turbidité-1023 observations

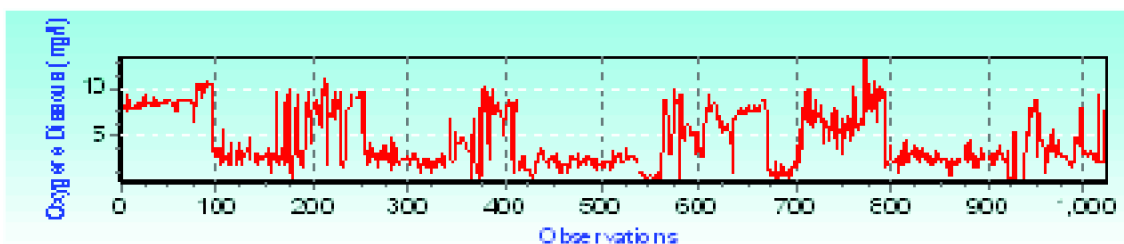


Figure IV.8: Variation annuelle de l'oxygène dissous-1023 observations

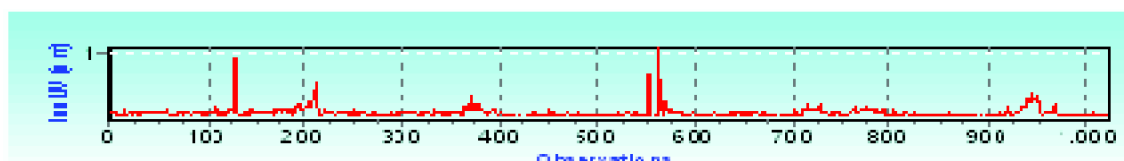


Figure IV.9 : Variation annuelle de l'ultraviolet-1023 observations

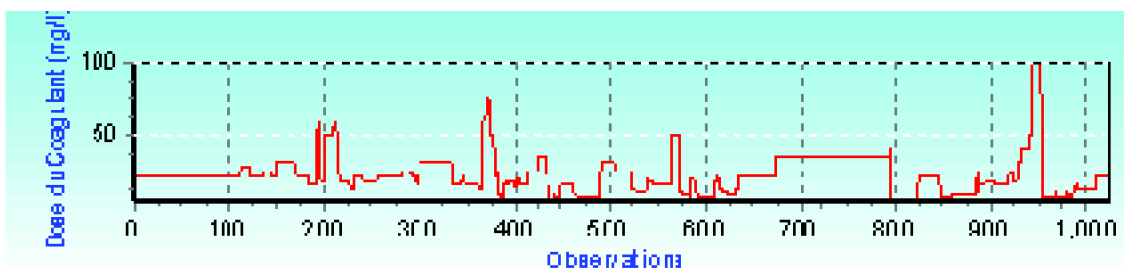


Figure IV.10 : Variation annuelle de la dose du coagulant-1023 observations

On constate des fluctuations notables pour les différents paramètres analysés, cela est dû surtout à la succession des saisons, contrairement à l'ultraviolet où l'on constate pratiquement de faibles variations, sauf pour quelques valeurs exceptionnelles.

Malgré les fluctuations des différents paramètres mesurés, nous constatons, que l'évolution de la dose du coagulant reste constante sur des intervalles plus ou moins longs.

#### IV.8.4. Application de l'ACP

Une approche basée sur l'analyse en composantes principales a été testée pour le fichier pris en considération (1023 observations) afin de déceler l'apport de chaque paramètre d'entrée (Température, PH, conductivité, turbidité, oxygène dissous, et ultraviolet) dans l'explication du phénomène étudié à savoir la dose du coagulant à injecter; afin



## Projection des variables sur un plan factoriel 1 x 2

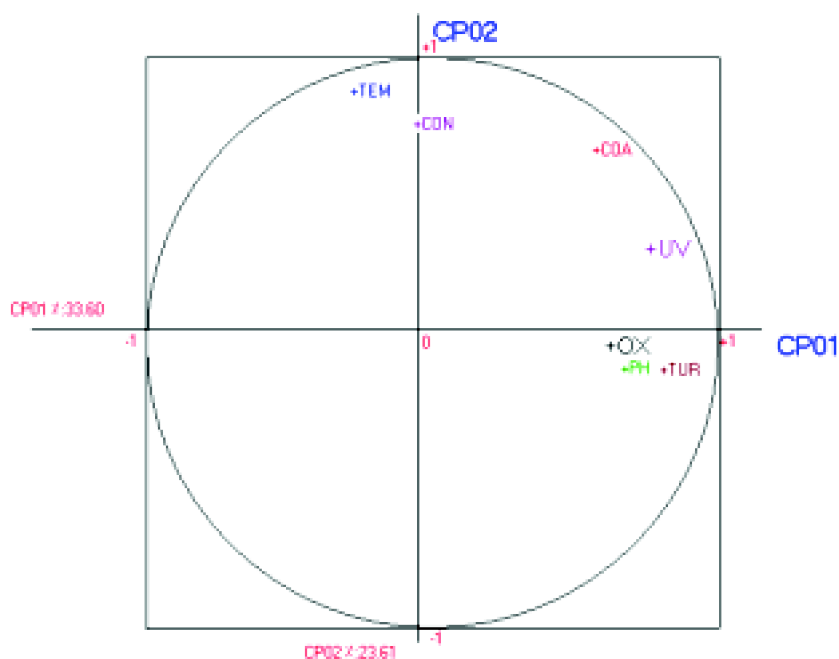


Figure IV.11: Cercle de corrélation -1023 observations.

On remarque la formation de deux groupes de variables, le premier formé par la turbidité, l'oxygène dissous, et le PH, les UV tandis que le deuxième n'est formé que par la conductivité et la température.

La dose du coagulant se trouve au milieu des deux groupes, ce qui explique qu'aucun groupe n'est dominant par rapport à l'autre.

On constate aussi que toutes les variables sont relativement proches de la circonférence du cercle de corrélation, cela démontre qu'elles ont toutes une influence sur le phénomène étudié.

A la fin de cette étape, nous nous sommes demandés si la qualité des données recueillies, n'avait pas de répercussion sur les résultats obtenus, en particulier le fait d'avoir le même degré d'influence des deux groupes de variables sur le phénomène étudié; nous nous sommes alors rapprochés des responsables de la SEAAL pour d'éventuelles explications, ceux-ci nous ont effectivement confirmés, que durant la période de sécheresse (2001-2002) qui a sévit, l'exploitation du barrage de Keddara s'en est ressentie, et que le pompage s'effectuait alors à partir du volume mort du barrage, et que par conséquent la dose du coagulant injectée à l'époque était très importante.

Devant cette situation nous avons été amené à revoir notre fichier de départ en éliminant toutes les valeurs enregistrées durant cette période.

Un nouveau fichier de données de 726 observations a été par conséquent retenu.

L'analyse statistique de ce dernier a donné les résultats résumés dans le tableau suivant :



	Moyenne	Ecart type	Min	Max
Température (°C )	16.53	3.49	10.2	26.2
PH	7.76	0.25	7.23	8.6
Conductivité ( $\mu$ .s/cm <sup>-1</sup> )	1009.48	122.27	668	1432
Turbidité (NTU)	7.58	4.54	0.44	32.4
Oxygène dissous (mg/l)	4.73	2.98	0.143	13.2
Ultraviolet (NM)	0.11	0.05	0.01	0.98
Coagulant (mg/l)	22.79	7.48	10	40

Les évolutions des variations annuelles des différents paramètres étudiés sont représentées par les figures ci après :

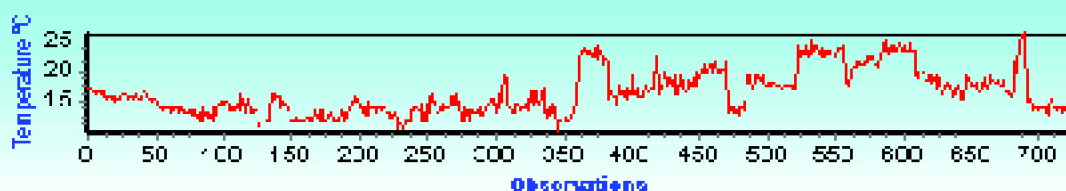


Figure IV.12 :Variation annuelle de la température-726 observations

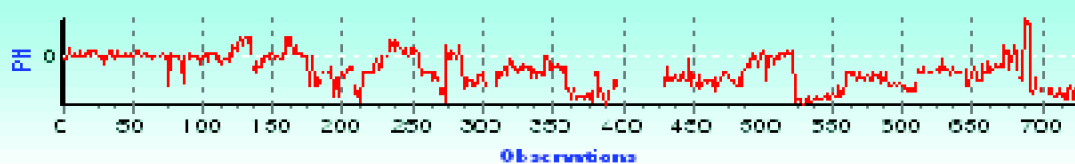


Figure IV.13 : Variation annuelle du PH-726 observations

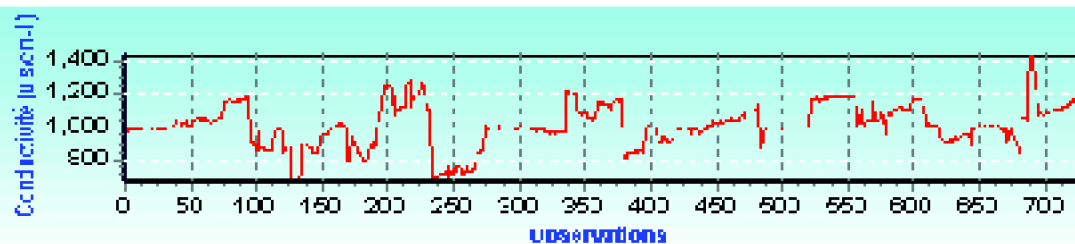


Figure IV.14 :Variation annuelle de la conductivité-726 observations

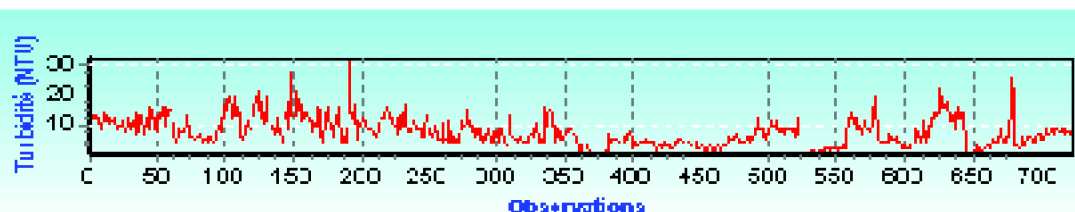


Figure IV.15 : Variation annuelle de la turbidité -726 observations

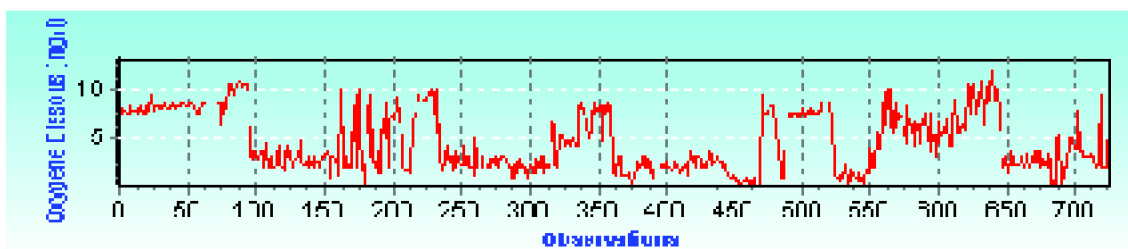


Figure IV.16 : Variation annuelle de l'oxygène dissous -726 observations

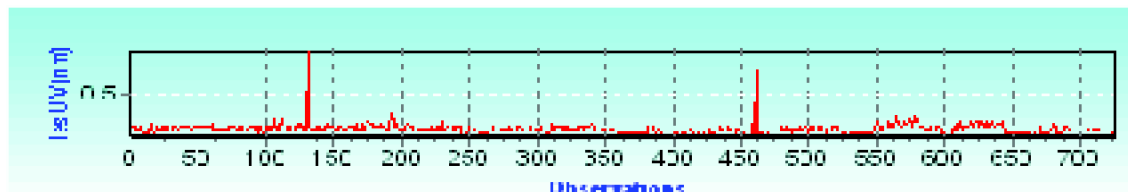


Figure IV.17 : Variation annuelle de l'ultraviolet-726 observations

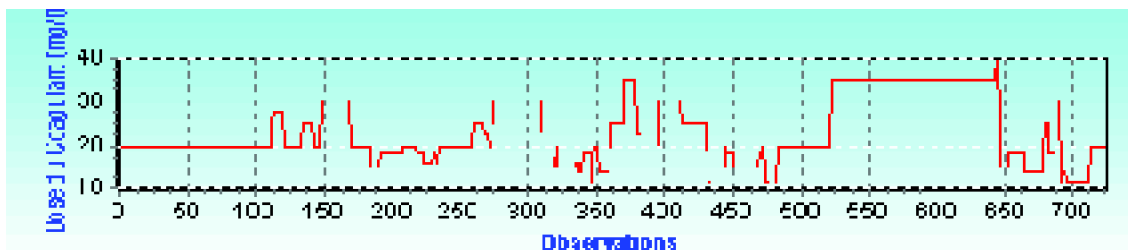


Figure IV.18 : Variation annuelle de la dose du coagulant-726 observations

Nous constatons que même après l'élimination des données de la période de sécheresse, la tendance de l'évolution de la dose du coagulant n'a pas subi de variations significatives, seules les valeurs maximales ont changé.

L'application de l'ACP au nouveau fichier constitué, et pour les sept paramètres étudiés a donné les résultats suivants :

Matrice des écart-types covariances							
	Coagulant	Température	PH	Conductivité	Turbidité	Oxygène	UV
Coagulant	35.21						
Température	12.90	1.00					
PH	1.57	0.33	0.065				
Conductivité	119.43	14.55	-0.40	158.03			
Turbidité	0.57	0.76	0.50	0.60	0.68		
Oxygène	4.0	-2.27	0.1	74.54	0.03	0.001	
UV	0.03	0.023	0.001	-0.50	0.00	0.001	0.001

Matrice des corrélations							
	Coagulant	Température	PH	Conductivité	Turbidité	Oxygène	UV
Coagulant	<b>1.00</b>						
Température	0.28	<b>1.00</b>					
PH	0.11	0.43	<b>1.00</b>				
Conductivité	0.12	0.55	-0.55	<b>1.00</b>			
Turbidité	0.06	0.40	0.51	0.20	<b>1.00</b>		
Oxygène	-0.03	-0.21	0.07	0.11	0.43	<b>1.00</b>	
UV	0.00	0.02	0.00	-0.11	0.00	0.00	<b>1.00</b>

Matrices des vecteurs propres en colonnes							
	U1	U2	U3	U4	U5	U6	U7
01	0.27	0.50	0.44	0.41	0.00	0.00	0.21
02	4.27	0.34	0.1	0.11	-0.57	-0.04	0.1
03	0.22	-0.011	0.002	0.03	-0.54	0.00	-0.17
04	0.24	0.29	0.63	0.1	0.10	0.00	0.07
05	0.03	0.24	0.10	0.00	0.49	0.15	0.02
06	0.02	0.40	0.49	0.10	0.00	0.43	0.02
07	0.01	0.00	0.02	-0.10	-0.28	0.02	-0.1

Matrice des valeurs propres							
U1	U2	U3	U4	U5	U6	U7	Somme
12.90	1.00	0.00	0.00	0.01	0.00	0.00	<b>13.91</b>

Contributions des 7 ACP							
U1 %	U2 %	U3 %	U4 %	U5 %	U6 %	U7 %	Somme
92.77	7.13	0.00	0.00	0.01	0.00	0.00	<b>100.00</b>

Cumul des contributions des CP							
U1 %	U2 %	U3 %	U4 %	U5 %	U6 %	U7 %	Somme
92.77	99.90	99.90	99.91	99.92	99.92	99.92	<b>100.00</b>

Matrice de corrélation CP-Variables							
	CP1	CP2	CP3	CP4	CP5	CP6	CP7
Coagulant	1.00	0.67	0.45	0.34	0.02	0.00	0.1
Température	0.27	1.00	0.17	0.06	-0.38	0.00	0.04
PH	0.11	0.17	1.00	0.00	0.00	0.00	0.02
Conductivité	0.12	0.17	-0.55	1.00	0.00	0.00	-0.1
Turbidité	0.06	0.40	0.51	0.20	1.00	0.00	0.00
Oxygène	0.02	-0.21	0.07	0.11	0.43	1.00	0.00
UV	0.00	0.02	0.00	-0.11	0.00	0.00	1.00

Tableau IV.4 : Résultats de l'analyse en composantes principales : fichier annuel (726 observations)

En effectuant une analyse en composantes principales sur le fichier étudié, nous avons noté qu'avec deux (02) CP seulement, nous pouvons expliquer 57.9 % de la variance totale.

## Projection des variables sur un plan factoriel 1 x 2

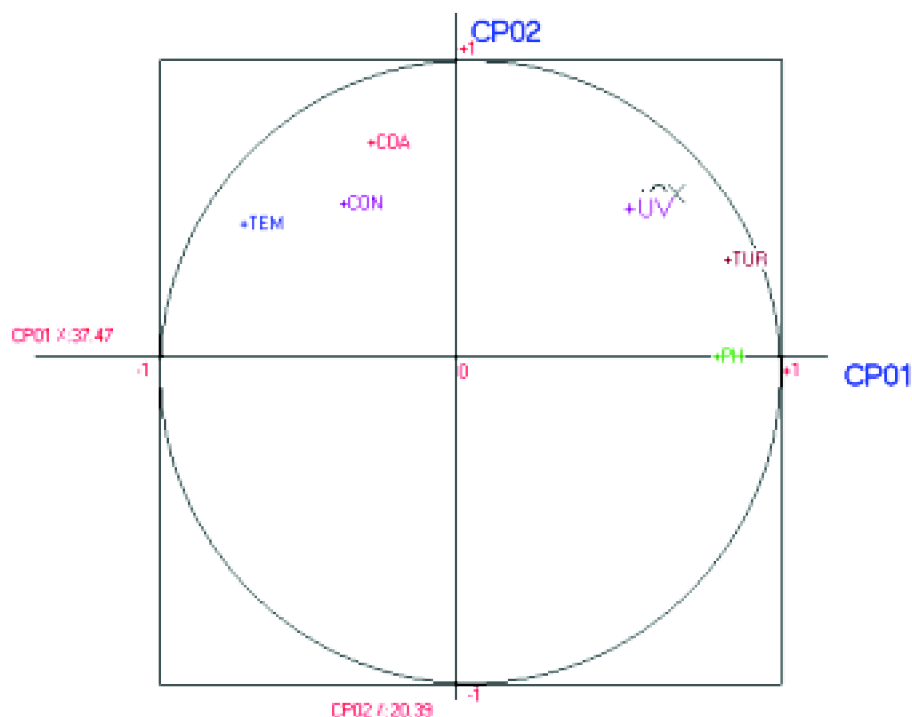


Figure IV.19 : Cercle de corrélation -726 observations.

D'après le cercle de corrélation, nous constatons comme précédemment le partage des paramètres en deux groupes, sauf que dans ce cas, la dose du coagulant ne se trouve plus au milieu mais fait partie du groupe formé par la température et la conductivité, ce qui nous a poussé à penser que ces deux paramètres étaient les plus prépondérants dans l'explication du phénomène étudié, et qu'il fallait les garder comme seule entrée dans le modèle réseau de neurones élaboré.

### IV.8.5 Application des réseaux de neurones

L'élaboration d'un modèle neuronale passe à la fois par le choix des paramètres d'entrée mais aussi par l'optimisation de l'architecture du réseau en lui-même. Ceci signifie qu'il faut trouver à la fois la meilleure répartition des données entre l'apprentissage et la validation, et le nombre de neurones constituant la couche d'entrée. L'optimisation à la fois du type de fonction non linéaire et de la géométrie du réseau de neurones a été faite après de nombreux essais empiriques. Nous avons fait varier le nombre de neurones au niveau de la couche d'entrée de 1 à 20. Chaque réseau a été entraîné avec la fonction sigmoïde. La couche de sortie n'est constituée que d'un seul neurone (dose de coagulant) à laquelle on a appliqué la fonction linéaire. Parmi les différentes configurations testées du réseau, nous avons retenu celle produisant la plus faible racine carrée de l'erreur quadratique moyenne (RMSE), ainsi que le meilleur coefficient de détermination  $R^2$  pour la partie validation.

Le fichier d'entrée (constitué de 726 observations) a été scindé en trois.

**Tableau IV.5 : Séparation du fichier de base**

Fichier	Observations
Total	<b>726</b>
Apprentissage	<b>601</b>
Validation	<b>100</b>
Modèle de prévision	<b>25</b>

L'application des réseaux de neurones au fichier retenu (constitué de la température et de la conductivité) a donné des résultats non satisfaisants en ce sens que les meilleures performances du modèle élaboré n'ont pas dépassé les 53 % pour le coefficient de détermination et 0.57 pour le RMSE pour la partie apprentissage, alors que pour la validation ces performances ont encore régressé pour atteindre les valeurs de 15 % pour  $R^2$  et 1.37 pour le RMSE.

Les tests graphiques sont représentés par les figures suivantes :

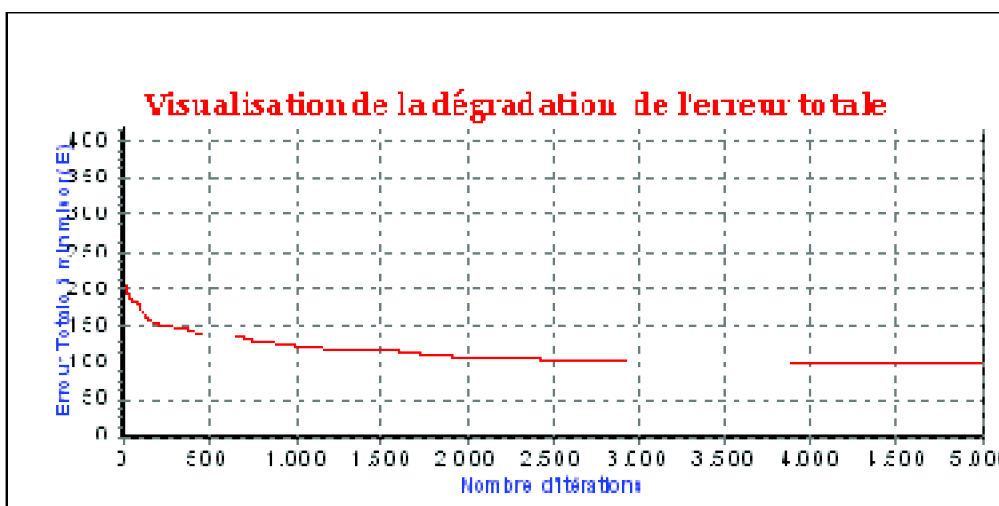


Figure IV.20 : Dégradation de l'erreur totale

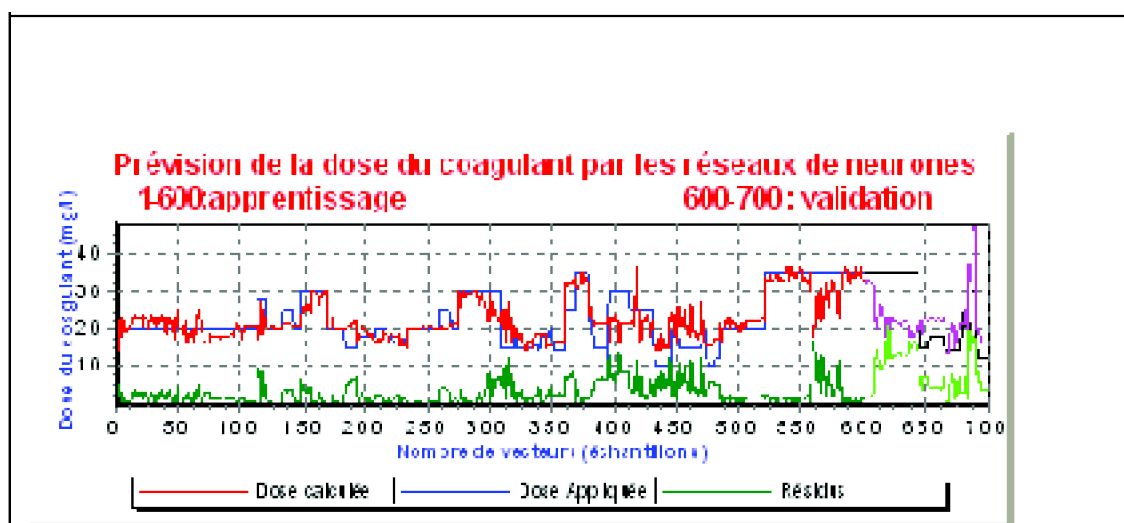


Figure IV.21 : Résultats obtenus par le réseau de neurones

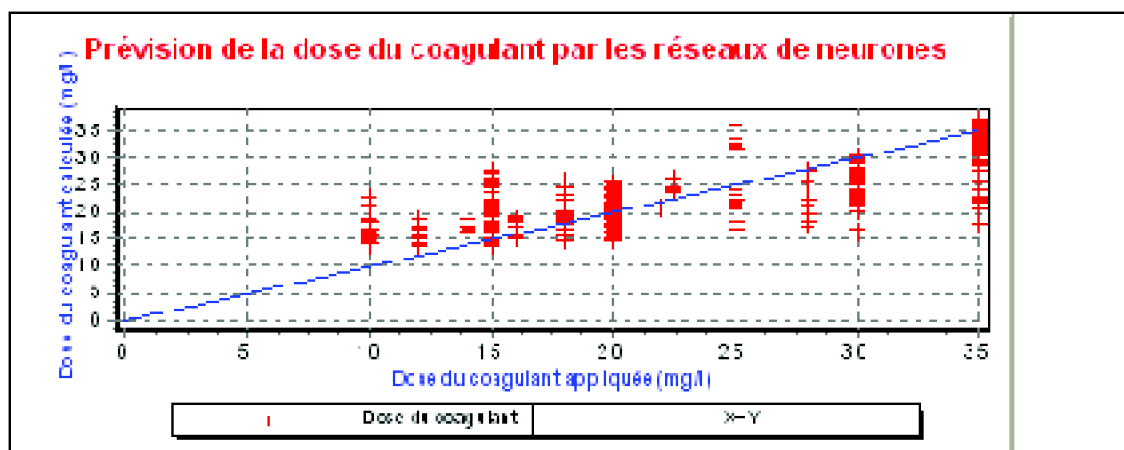


Figure IV.22 : Dose appliquée et dose prédite

Pour affiner le modèle, une approche basée sur des données saisonnières a été entreprise.

#### IV.8.6 Etude de l'effet saisonnier

Nous avons remarqué qu'en appliquant l'ACP aux quatre fichiers saisonniers les résultats ne se sont pas améliorés (Tableaux annexe A) et que la dose du coagulant dans les cercles de corrélation fait toujours partie du groupe formé par la température et la conductivité, sauf pour la saison hivernale où toutes les variables sont dispersées (voir figures ci-après).

Projection des variables sur un plan factoriel 1 x 2

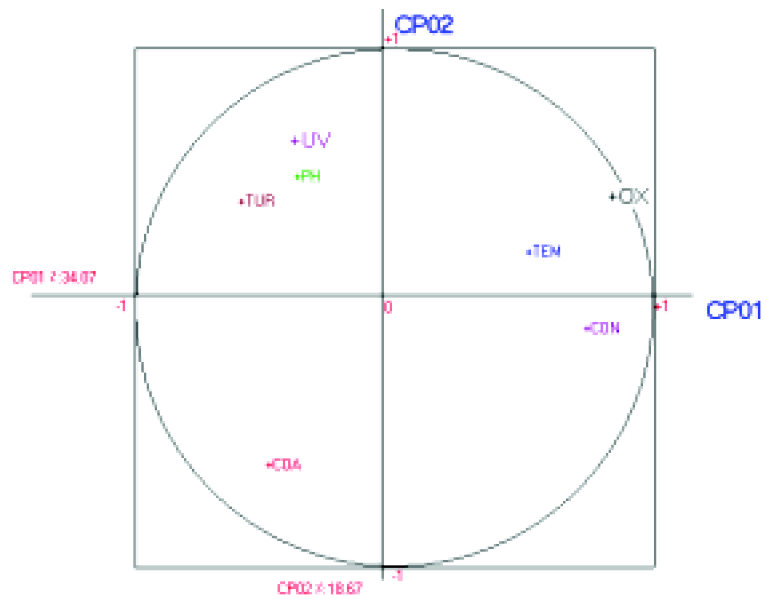


Figure IV.23: Cercle de corrélation –hiver

Projection des variables sur un plan factoriel 1 x 2

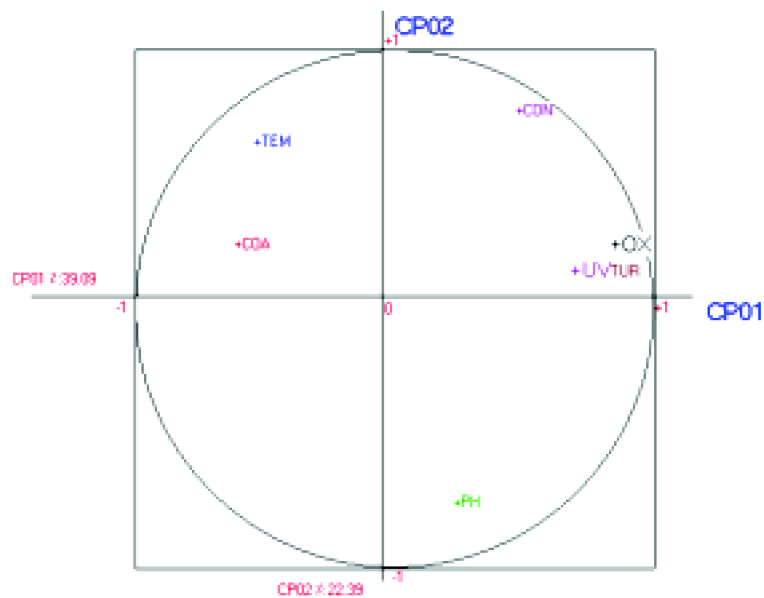


Figure IV.24 : Cercle de corrélation –printemps

### Projection des variables sur un plan factoriel 1 x 2

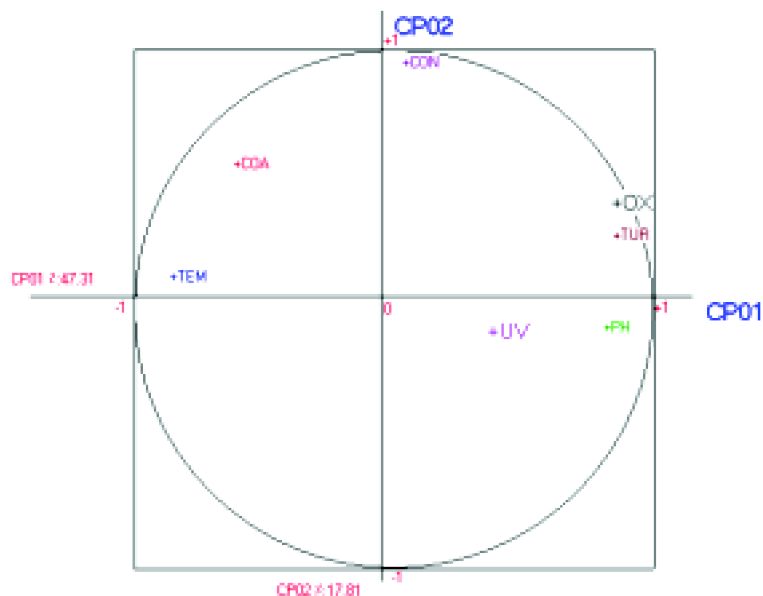


Figure IV.25: Cercle de corrélation –été

### Projection des variables sur un plan factoriel 1 x 2

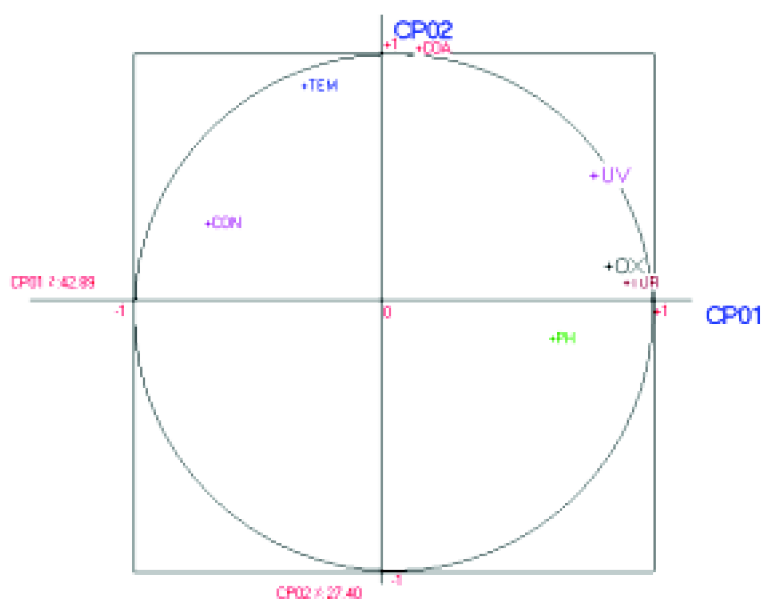


Figure IV.26 : Cercle de corrélation –automne.

Les performances obtenues par l'application des réseaux de neurones ne présentent aucun intérêt; en effet, le coefficient de détermination est de faible importance pour l'apprentissage tandis qu'il diverge complètement au niveau de la validation, pour cela nous avons décidé d'éliminer l'effet saisonnier de notre approche.



Saison	Nombre de neurones	R <sup>2</sup>		RMSE	
		apprentissage	validation	apprentissage	validation
Hiver	06	0.41	0.76	0.19	0.86
printemps	16	0.45	0.05	0.15	1.78
été	11	0.89	0.08	0.29	1.98
Automne	14	0.84	0.07	0.13	0.83

Tableau IV.7 :Performances des modèles saisonniers

## IV.9.Tentatives d'amélioration

Les résultats obtenus par l'application des réseaux de neurones n'étant pas satisfaisants, on s'est posé alors la question de savoir si l'introduction d'autres variables que celles obtenues par l'analyse en composantes principales ( température et conductivité) au fichier annuel 726 observations, ne permettraient pas l'amélioration des performances du modèle.

Des tentatives d'améliorations ont été réalisées en faisant intervenir à chaque fois à l'entrée du réseau, un nouveau paramètre, et on a suivi l'évolution des performances du modèle élaboré. Les variantes retenues sont les suivantes :

1. Trois entrées : température-PH-conductivité
2. Quatre entrées : température-PH-conductivité-turbidité
3. Cinq entrées : température- PH- conductivité- turbidité- oxygène dissous
4. Six entrées : température- PH- conductivité- turbidité- oxygène dissous- les ultraviolets.

Les résultats obtenus pour les différentes variantes retenues sont résumés dans le tableau suivant :

Variables d'entrée		Nombre de neurones	R <sup>2</sup>		RMSE	
			apprentissage	validation	apprentissage	validation
Deux entrées	Température	10	0.53	0.15	0.57	1.37
	Conductivité					
Trois entrées	Température	11	0.58	0.17	0.64	1.30
	Conductivité PH					
Quatre entrées	Température	09	0.64	0.00	0.49	1.04
	Conductivité					
	PH					
	Turbidité					
Cinq entrées	Température	11	0.72	0.62	0.38	1.15
	Conductivité					
	PH					
	Turbidité					
	Oxygène dissous					
Six entrées	Température	10	0.75	0.00	0.33	0.90
	Conductivité					
	PH					
	Turbidité					
	Oxygène dissous					
	Les UV					

Tableau IV.8 : Performances du modèle élaboré

Nous avons constaté que les performances s'amélioraient à chaque rajout d'un nouveau paramètre intervenant dans le phénomène étudié.

Les meilleurs résultats ont été obtenus en intégrant les six paramètres explicatifs, et cela est en concordance avec les constatations obtenues au niveau des cercles de corrélation ; en effet, aucun paramètre n'a de prépondérance par rapport aux autres mais que chacun d'eux contribue avec une certaine part à la détermination de la dose du coagulant nécessaire à l'amélioration de la qualité de l'eau.

Les meilleures performances du modèle sont : un coefficient de détermination R<sup>2</sup> de **75 %** pour l'apprentissage et **80 %** pour la validation.

Les performances graphiques pour chaque variante sont présentées dans les figures ci-après :

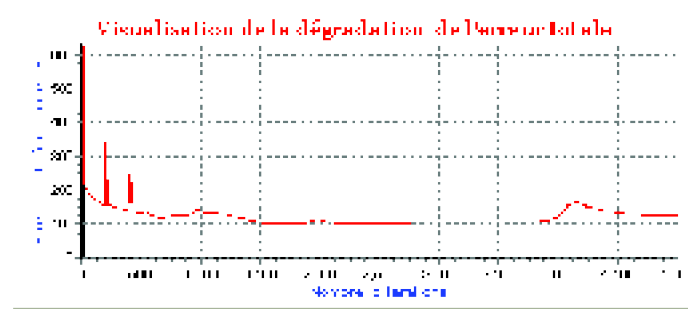


Figure IV.27: Résultats de l'ajustement de l'ensemble des données d'entrées

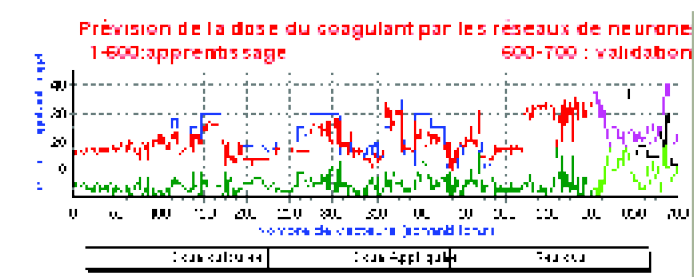


Figure IV.28: Résultats obtenus par le réseau de neurones à trois entrées

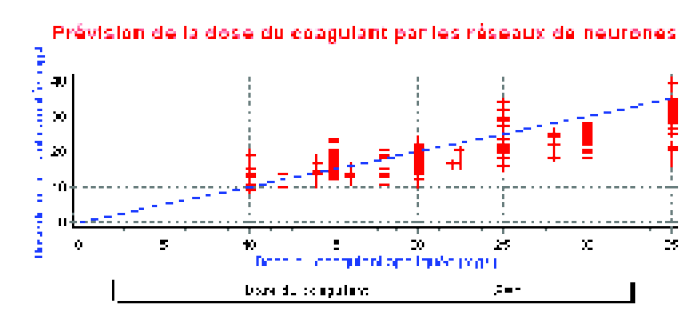


Figure IV.29: Dose appliquée en fonction des données d'entrées

*Performances du modèle à trois entrées*

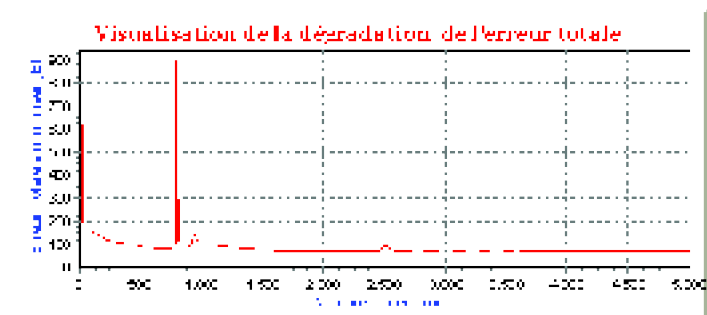


Figure IV.70 : Visualisation de la dégradation de l'erreur totale à quatre entrées

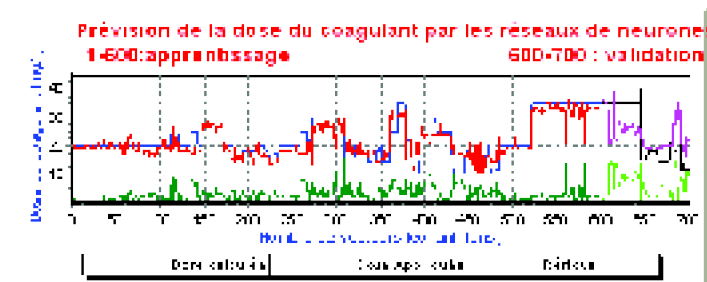


Figure IV.71 : Résultats obtenus par le réseau de neurones à quatre entrées

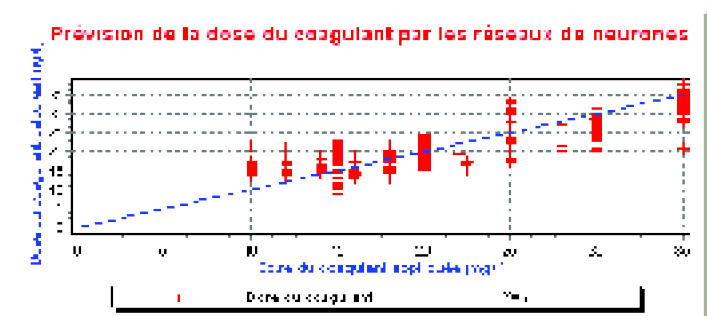


Figure IV.72 : Exemples de doses prédites à quatre entrées

*Performances du modèle à quatre entrées*

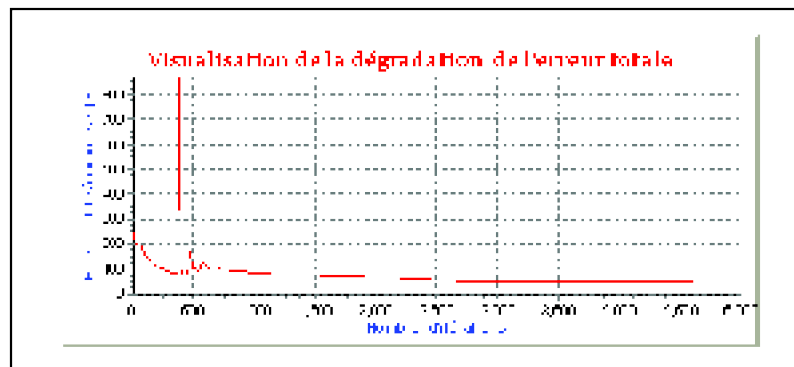


Figure IV.33: Dégradation de l'erreur avec la cinquième entrée

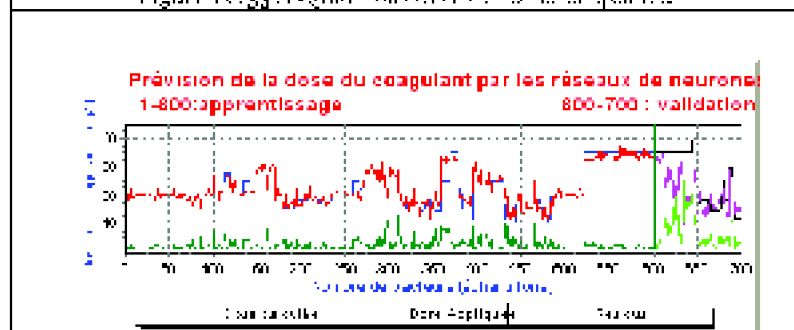


Figure IV.34: Résultats obtenus au réseau de neurones à cinq entrées

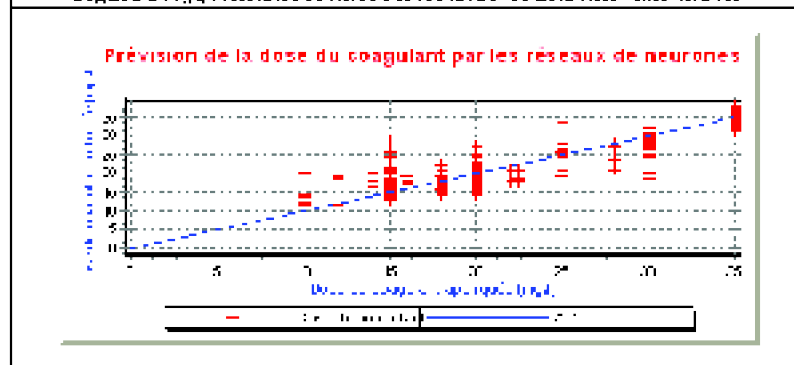


Figure IV.35: Données appliquées et doses prédites à cinq entrées

*Performances du modèle à cinq entrées*

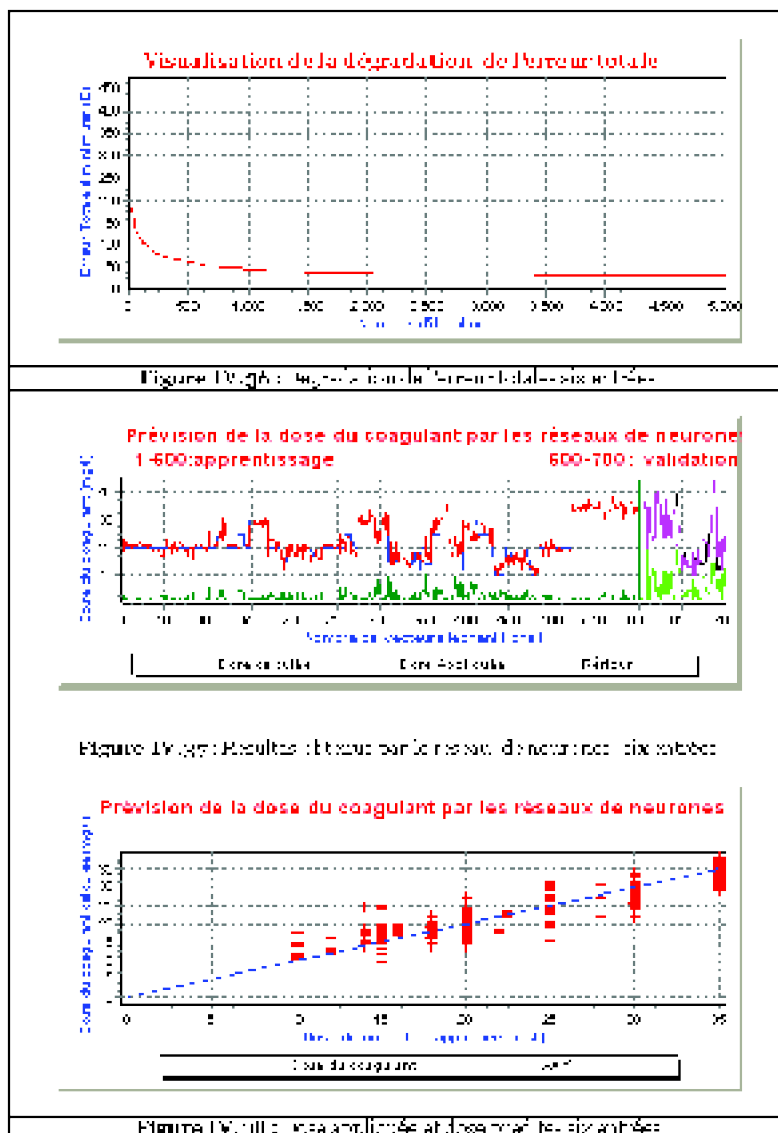


Figure IV.14: Application du réseau total à six entrées

Figure IV.15: Résultats obtenus par le réseau de neurone à six entrées

Figure IV.16: Régression appliquée au données pendant six années

*Performances du modèle à six entrées*

Les paramètres retenus par le modèle élaboré sont présentés dans le tableau suivant :

Neurone de sortie		Poids de la couche d'entrée							N° Neurone
biais	poids	biais	UV	Oxygène	Turbidité	Conductivité	PH	Température	
4.6%	1.77	-7.99	-2.3	-1.84	-0.89	-2.81	-3.64	-1.73	<b>1</b>
	2.16	-2.00	-0.79	-4.22	-0.80	3.80	-1.62	-0.05	<b>2</b>
	0.82	6.96	1.27	1.57	6.62	1.90	6.84	1.09	<b>3</b>
	0.02	-7.55	0.40	3.94	2.64	0.00	-5.64	6.00	<b>4</b>
	0.40	-6.50	2.15	0.30	-0.33	3.38	-5.11	8.43	<b>5</b>
	0.94	-6.94	6.73	-1.63	-1.63	2.33	2.06	5.82	<b>6</b>
	0.95	2.13	1.28	2.12	2.13	8.91	1.25	5.88	<b>7</b>
	0.56	-6.26	0.26	3.63	-0.07	0.56	-2.82	6.12	<b>8</b>
	0.06	-6.56	0.50	7.20	-0.49	-0.70	4.00	7.44	<b>9</b>
	1.87	3.07	0.74	7.73	2.19	-2.97	4.70	0.39	<b>10</b>
	1.75	0.59	-1.38	-0.78	0.22	-1.14	-1.99	-1.97	<b>11</b>
	0.91	3.28	3.00	2.01	1.1	5.60	4.04	2.08	<b>12</b>
0.10	-6.00	1.00	4.55	-1.20	0.71	-4.41	5.51	<b>13</b>	

Tableau IV.9 : Paramètres retenus

## IV.10. Prévision de la dose du coagulant par le modèle retenu

La prévision est une estimation préalable de l'état futur d'une grandeur bien déterminée à différentes échéances.

La démarche préconisée dans ce mémoire est la suivante ; une fois les paramètres du modèle optimisés, il suffit d'injecter les variables d'entrée (température, PH, conductivité, turbidité, oxygène dissous, et ultras violets) qui sont les paramètres descripteurs de l'eau brute, pour avoir en temps réel la dose à injecter.

Cette approche, permet à la fois de fournir un outil d'aide à la décision, pour le gestionnaire de la station de traitement, et de libérer l'opérateur de l'astreinte d'analyse de routine et d'optimiser les traitements en réduisant les coûts et les temps de réponse.

L'intérêt de la prévision donc est d'aboutir à une meilleure efficacité du traitement et une réduction des coûts d'exploitation.

Nous avons appliqué le modèle élaboré aux données non utilisées (25 observations) les résultats obtenus sont présentés dans le tableau suivant :

<b>Dose réelle (ppm)</b>	12	12	12	12	12	12	12	12
<b>Dose prévue (ppm)</b>	17.12	10.42	20.2	21.06	20.90	23.57	20.42	24.05
<b>Ecart (ppm)</b>	-5.12	-1.42	-8.2	-3.06	-8.98	-11.57	-8.42	-12.05
<b>Dose réelle (ppm)</b>	12	12	12	12	20	20	20	20
<b>Dose prévue (ppm)</b>	23.36	21.65	12.75	20.3	18.67	20.8	21.64	17.04
<b>Ecart (ppm)</b>	-11.36	-9.65	-9.75	-8.9	1.33	-0.8	-1.64	2.96
<b>Dose réelle (ppm)</b>	20	20	20	20	20	20	20	20
<b>Dose prévue (ppm)</b>	20.44	21.13	14.81	14.47	20.94	20.72	20.81	14.19
<b>Ecart (ppm)</b>	-0.44	-1.13	5.19	5.53	-0.94	-0.79	-0.81	5.91

Tableau IV.10 : Résultats de la prévision par le modèle retenu.

L'application du modèle retenu à la prévision de la dose du coagulant ; montre l'efficacité et la robustesse du modèle, en effet, l'écart maximal de la dose du coagulant prévue par rapport à la dose réelle ne dépasse pas **12** , tandis que l'écart minimal est de **0.13**, la moyenne des écarts est de l'ordre de **3.85**.

Ces résultats permettent en premier lieu aux gestionnaires de la station de traitement de Boudouaou d'avoir des ordres de grandeurs fort probables des doses de coagulants à injecter au niveau des essais « Jar Test » au lieu d'utiliser plusieurs doses, et d'autres part de limiter l'intervalle de recherche de la dose du coagulant à injecter permettant un traitement optimal de l'eau brute, et ceci conduira à la réduction des coûts des essais réalisés au niveau du laboratoire.



## Conclusion générale

Afin de répondre à la sévérité croissante des normes, aux volumes toujours plus grands d'eau à traiter, et à la pollution grandissante des réserves, d'importants efforts de recherche ont été consentis au cours des dernières décennies. Les études portent notamment sur les procédés de traitement des eaux ainsi que sur les techniques d'analyse et la mise au point d'instruments de surveillance automatique. Le traitement de l'eau avant usage constitue une étape indispensable et vitale pour les consommateurs, afin d'éviter tous les risques de maladies à transmission hydrique, et toute contamination.

Dans la chaîne de traitement qui est constituée de plusieurs étapes, nous nous sommes intéressé à la phase clarification, et plus particulièrement, à la modélisation de la dose du coagulant à injecter lors de l'opération (coagulation-floculation), en fonction des paramètres descripteurs de l'eau brute.

La station de traitement de Boudouaou qui alimente une grande partie de la wilaya d'Alger a été retenue comme station d'étude, pour le nombre de paramètres mesurés, et pour l'importance du point de vue débit traité.

Nous nous sommes attelés à rechercher une approche susceptible d'apporter une contribution au problème du traitement de l'eau brute et à améliorer par conséquent la gestion de la ressource en eau en Algérie qui accuse un déficit relatif dans les apports en eau, principalement dû aux fluctuations des conditions climatiques, et à l'irrégularité de la répartition spatiale de cette ressource.

En effet, la modélisation de la dose du coagulant à injecter en fonction des

paramètres descripteurs de l'eau brute, permet de mieux appréhender le problème du traitement et fournit un outil précieux pour les gestionnaires des stations de traitement.

Au début de ce travail, nous avons cherché à savoir si l'introduction des méthodes connexionnistes aux traitements des eaux, pouvait répondre aux attentes des gestionnaires des stations.

Un modèle basé sur les réseaux de neurones, est-il capable de satisfaire aux attentes de la modélisation de la dose du coagulant à travers la chaîne de traitement des eaux ? Ceci revient à se demander quelles sont les variables qui participent à la dite modélisation.

L'approche élaborée permet d'apporter des réponses satisfaisantes à certains problèmes du traitement de l'eau.

La démarche suivie dans le cadre de ce travail de recherche, est basée essentiellement sur les techniques de l'analyse en composantes principales et des réseaux de neurones.

Vu les limitations importantes rencontrées pendant l'exploitation de la base de données recueillies nous avons été obligé à réduire la taille de notre échantillon.

Nous avons essayé à travers l'application de l'analyse en composantes principales à déceler parmi l'ensemble des paramètres pris en considération, d'éventuelles variables prépondérantes à l'explication du phénomène étudié, et de réduire par conséquent leur nombre afin d'être utilisées comme seuls paramètres d'entrée dans le réseau de neurone.

Cependant, et malgré les corrélations significatives trouvées entre quelques variables explicatives (température et conductivité), et la variable expliquée (dose du coagulant), les résultats obtenus par les réseaux de neurones étaient de faible importance.

L'approche de la modélisation par la méthode des réseaux neuronaux, à la dose du coagulant à injecter, appliquée à la station de traitement de Boudouaou, a montré l'énorme potentiel de cette technique dans la modélisation de phénomènes complexes non linéaires, telle que le traitement des eaux. En effet, les résultats obtenus sont très satisfaisants que ce soit dans le domaine de la modélisation des paramètres descripteurs de l'eau brute, ou dans le domaine de la prévision.

L'avantage de l'approche préconisée dans ce mémoire c'est qu'elle a montré, qu'en premier lieu les résultats de la modélisation sont tributaires de la qualité des données d'entrées, ce qui confirme la théorie de la modélisation qui stipule que toute entrée fautive génère nécessairement une sortie fautive, et qu'en second lieu, la prise en compte de tous les paramètres descripteurs de l'eau brute est une condition principale dans la modélisation de la dose du coagulant à injecter.

## Références bibliographiques

- Adgar, A., Cox, C.S., Daniel, P.R., Billington, A.J, Lowdon, A.,(1995)** : Experiences in the application of the artificial neural networks to water treatment plant management, in Proceedings of the International COMDEM'95, Vol. 1, 33-38, Canada.
- Adrian, P.T., (2004)** : Étude comparée du colmatage en nanofiltration et en ultrafiltration. Thèse de maîtrise, Université Laval, 173 p.
- Albert, M. (1992)** : Les réseaux de neurones comme outils d'aide à la décision floue Rapport de D.E.A Ingénierie Informatique I.N.S.A de Lyon 64 p
- Alexander, L., Christoph, T., (1992)** : Combining Multiple Neural Network Paradigms and Applications using SESAME. In IJCNN, volume II, pages 528-534, Baltimore.
- Amirtharajah, A., O'Melia, C.R., (1990)** : *Coagulation Processes: Destabilization, Mixing, and Flocculation* , In: Water Quality and Treatment, fourth edition, New-York: McGraw-Hill, 1194 p.
- Anonym1, (2000)** : Centre International De Formation (C.I.D.F.), Principes généraux de traitement des eaux, Lyonnaise des Eaux.
- Anonym2, (1989)** : Degremont, Mémento technique de l'eau : Tome 1, Lavoisier - Technique et documentation, Paris.
- Azema-Barac, M., Hewetson, M., Recce, M., Taylor, J., P. Treleaven, and Vellasco, M., (1990)**: Pygmalion : Neural network programming environment. In Proc. Int.

Neural Network Conf., volume II, pages 709-712.

**Benoit, D., Philippe, E., Catherine, M., (1989):** Nacre: a neuron-oriented programming environment. In Proc. Neuro-Nimes, pages 183-200.

**Boisseau, J. M., Monier, B., Lefèvre, O., Denoeux, T. et Ding, X., (1995):** Préviation du trafic autoroutier par réseaux de neurones artificiels. In Les systèmes Intelligents dans les Entreprises, EC2 & Cie, Montpellier.

**Bouchard, C., Sérodes, J., (2002) :** *Production d'eau potable*, Notes de cours, Université Laval, 188 p.

**Bouroche, J.M., Saporta, G., (1980) :** L'Analyse des Données, Que sais-je, PUF.

**Boucheron, S., (1992) :** Théorie de l'apprentissage. Hermès, Paris.

**Brodart, E., Bordet J. P., Bernazeau, F., Mallevalle, J. et Fiessinger, F., (1989) :** Modélisation stochastique d'une usine de traitement d'eau potable. 2ème Rencontres Internationales Eau et Technologies Avancées, Montpellier.

**Canu, S., Sobral, R., Lengellé, R., (1990) :** Formel neural network as an adaptative model for water demand. Proceedings of INNC'90, Kluwer Academic Publishers, Vol. 1, pp. 131-136, Paris.

**Canu, S., Ding X., Granvalet, Y., (1997) :** Une application des réseaux de neurones pour la prévision à un pas de temps, S. Thiria, Y. Lechevallier, O. Gascuel et S. Canu éditeurs. « Statistiques et méthodes neuronales », Chapitre 7, pp.120-131, Dunod, Paris.

**Cox, C.S., Daniel, P., Adgar A., Lowdon, A., (1996):** Chemical dosing philosophies for a water treatment plant: results of some pilot plant experimentation, in Proceedings of the IEE Conference Publication, Vol. 427 (2), 1052-1057.

**Critchley, R.F., Smith, E.O., Pettit, P., (1990) :** Coagulation control at water-treatment plants in the north-west region of England, Journal IWEM, Vol. 4, 535-543.

**Ding, X., Canu, S., Denoeux, T., (1996) :** Neural network models for forecasting. In J. G. Taylor, editor, Neural Networks and their Applications, John Wiley and Sons, Chichester, pp. 243-252.

**Dreyfus, G., (2004) :** Réseauxde neurones ; Méthodologie et Application EYROLLES, pp 01-65.

**Dubuisson, B. (1990):** Diagnostic et reconnaissance de formes.Ed.Hèrmes.1990

**Edmond, M., Josef, S., (1992):** A Software Environment for Studying Computational Neural Systems. IEEE Trans. on Software Engineering, 8(7):575-589.

**Edzwald, J.K., Tobiason, J.E., (1999 ) :** *Enhanced coagulation : USA requirements and a broader view*, In: Removal of humic substances from water, Conference proceedings, Trondheim.

**Fahlman, S. E., (1988):** An empirical study of learning speed. Technical Report CMU-CS-88-162, School of Computer Science, Carnegie Mellon University - Pittsburgh, PA.

**Fuentes, L., Aldana, J.F., Troya, J.M. (1993) :** Urano : an object-oriented artificial neural network simulation tool. In Proc. Int. Workshop on Artificial Neural Networks, volume 686, pages 364-369. Springer-Verlag.

- 
- Gagnon, C., Grandjean, B.P.A., Thibault, J. (1997)** : Modelling of coagulant dosage in a water treatment plant, *Artificial Intelligence in Engineering*, Vol. 11, 401-404.
- Gayner, R. J. et Downs, T., (1994)** : On the properties of error functions that affect the speed of back propagation learning. In *Marinaro M. and Morasso P. G., editors, ICANN'94 -Proceeding of the International Conference on Artificial Neural Networks ; 26-29 May 1994, pp. 557-560, Sorrento, Italy.*
- Gorman R. P., Sejnowski T. J., (1988)** : Analysis of hidden units in a layered network trained to classify sonar targets. *Neural Networks*, Vol. 1, pp. 75-89.
- Hassibi, B., Stork, D. G., (1993)** : Second order derivatives for network pruning: Optimal brain surgeon. In *S. J. Hanson, J. D. Cawnan, and C. L. Giles, editors, Advances in Neural Information Processing Systems 5, pp. 164-171, Morgan Kaufmann, San Mateo, CA.*
- HEBB, D., (1949)**: *The organisation of the Behavior.* WILEY .NEW YORK.
- Hinton et Sejnowski, (1986)** : Learning and relearning in Boltzman machines In *Parallel Distributed Processing 1, chap .7*
- Hinton , G.E., (1989)**: Deterministic Boltzmann machine learning performs Steepest descent in weight-space. *Neural Computation*, 1 143–150.
- Hopfield , j.j., (1982)** : Neural Network and physical systems with emergent Collective abilities. *Proceedings of the National Academy Of Science , USA 81 , PP.3088-3092,*
- Hotelling, H., (1933)** : Analysis of a complex statistical variable into principal component", *J. Edu. Psy. , Vol 24, 417-441 and 498-520.*
- Kohonen , T. (1984)**: *Self organization and associative memory*, 2nd edition, Springer Verlag.
- Lebart L., Morineau A. et Piron M. (1977)** : *Techniques de la description statistique, Méthodes et logiciels pour la description des grands tableaux.* Dunod, Paris.
- Lebart, L., Salem, A., (1994)**: *Statistique Exploratoire multidimensionnelle,* Dunod.1994.
- Le Cun, Y., (1985)** : Une procédure d'apprentissage pour réseau à seuil asymétrique. In *Cognitiva 85:A la Frontière de l'Intelligence Artificielle des Sciences de la connaissance des Neurosciences, Paris: CESTA, pp. 599-604, Paris.*
- Lind, C., (1995)** : Coagulant road map, *Public Works*, 36-38.
- Lind, C., (1995)** : Coagulation Control and Optimization: Part one, *Public Works*, 56-57, October, 1994. Coagulation Control and Optimization: Part two, *Public Works*, 32-33, November.
- Marc, P., (2004)**: *Réseaux de Neurones Université de l'AVAL, 118p*
- Matilainen, A., Niina L., Susanna K., Tuula T., (2002)** : Removal of NOM in the different stages of the water treatment process. *Environmental International* 28, 457 – 465
- McCulloch, W. S., Pitts W.,(1943)** : A logical calculus of the ideas immanent in Nervous activity. *Bull Math .Biophysics*, 5; 113 –115.
- Minsky, M., Papert, S., (1969)**: *Perceptrons.* CAMBRIDGE: MIT press.
- Mirsepasi, A., Cathers, B. et Dharmappa, H.B.,(1995)**: *Application of Artificial Neural*
-

- Networks to the Real Time Operation of Water Treatment Plants, in *Proceeding of the International Conference on Neural Networks*, Vol. 1, 516-521, Perth, Australia.
- Moller, M., (1993)** : Efficient training of feed-forward neural networks. PhD thesis, Computer Science department; Aarhus University, Denmark.
- OUKHELLOU, L., (1997)** : Para métrisation et classification de signaux en contrôle non destructif. Application à la Reconnaissance des Défauts de Rails par Courant de Foucault. Thèse de doctorat –université paris sud, pp211.
- Parker, D.B., (1985)**: Learning logic. Technical report. Center of Computational Research in Economics and Management Science. MIT, Cambridge, MA.
- Pim, H.W., Wim, J. M., Philipsen, and Spaandonk, J.V.(1991)**: PLANNET: A New Neural Net Simulator. In O. Simula T. Kohonen, K. M#akisara and J. Kangas, Editors, *Artificial Neural Networks*, pages 1481-1484. Elsevier Science B.V.
- Pearson, K., (1901)** : On lines and planes of closest fit to systems of points in space", *Phil.Mag.*,2, 11, 559-572.
- P eretto , P. (1984)**: Collective properties of neural networks: a statical physics Approach. *Biological Cybernetics* 50 , PP.51-62
- Rodriguez, M. J., Sérodes, J. B., (1996)**: Advanced chlorination control in drinking water systems using artifical neural networks, *Water Supply*, Vol. 15 (2), 159-168, Amsterdam.
- Romain, B., (2003)** : Modèles Impulsionnels de Réseaux de Neurones Biologiques Thèse Doctorat, Paris pp 03-10
- Rosenblatt, F., (1958)** : The perceptron: A probabilistic model for information storage and organisation in the brain. *Psychological Review*, Vol 65, pp. 386-408.
- R umelhart , D.E., (1986)** : Learning internal representations by error propagation. *Parallel Distributed Processing Explorations in the Microstructure of Cognition*. Eds. D. E. Rumelhart and J.L. McClelland, Cambridge, MA, MIT Press, Bradford Books, vol. 1, pp. 318-362.
- Rumelhart D. E., Hinton G. E., Williams R. J., (1988)** : Learning internal representations by error propagation. In Rumelhart D. E. and McClelland J., editors, *Parallel Distributed Processing*. MIT Press, Cambridge, MA.
- Sejnowski, T. J., Rosenberg C. M., (1987)** : Parallel networks that learn to pronounce English text. *Complex Systems*, Vol. 1, pp. 145-168.
- Spearman, C., (1904)** : General intelligence, objectively detemined and measured. *Amer. Journal of Psychology*, 15, pp. 201- 293.
- Valiron, F., (1988)** Gestion des Eaux : automatisation - informatisation - télégestion, Presses de l'école nationale des ponts et chaussées, Paris.
- Valentin, N., (2000)** : Construction d'un capteur logiciel pour le contrôle automatique du procédé de coagulation en traitement d'eau potable. Thèse de doctorat –Université Technologie Compiègne- France, 153p.
- Vapnik, V., (1982)** : Estimation of Dependences Based on Empirical Data. Springer-Verlag New York Heidelberg Berlin.
- Vapnik, V., (1995)** : The Nature of Statistical Learning Theory. Springer-Verlag New

York Heidelberg Berlin.

**Vapnik, V. N., Chervonenkis, A. Y., (1971)** : On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and its applications*, 16:264-271.

**VON Neumann, J. (1958)** : Probabilistic Logics and the synthesis of reliable Organisms from unreliable components. In *Automata Studies*, Eds.C.E.Shannon.

**Widrow, B., Hoff, M., (1960)**: In *IRE Western Electric Show and Convention Record*, pp. 96-104.

**Werbos, P., 1974. (1974)**: Beyond regression: new tools for prediction and analysis in the behavioral sciences. PhD thesis, Harvard University, Cambridge, MA.





## Annexe A Résultats des ACP (Fichiers saisonniers)

### Annexe A.1: Résultats d'application de l'analyse en composantes principales Saison hiver (décembre-janvier-février).

Tableau A.1 : Statistiques des paramètres retenus (évolution saisonnière- hiver).

	moyenne	Ecart type	Min	Max
Température (°C )	14.28	1.49	11.2	17.3
PH	8.012	0.123	7.5	8.33
Conductivité ( $\mu\text{s}/\text{cm}^{-1}$ )	972.12	107.9	668	1166
Turbidité (NTU)	10.97	3.919	4.14	27
Oxygène dissous (mg/ L)	6.10	3.004	0.15	10.9
Ultraviolet (NM)	0.12	0.069	0.05	0.98
Coagulant (mg/ L)	21.82	3.612	20	30

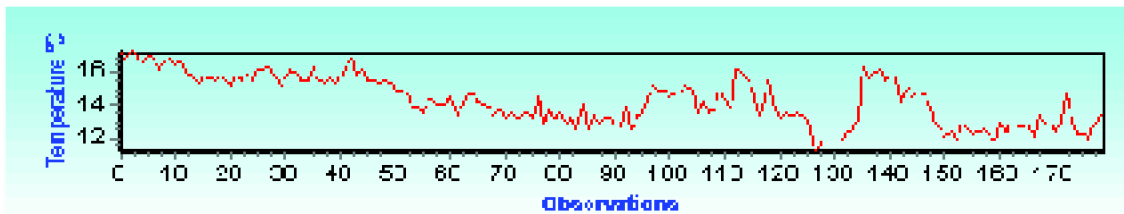


Figure A.1 : Variation saisonnière (hiver) de la température

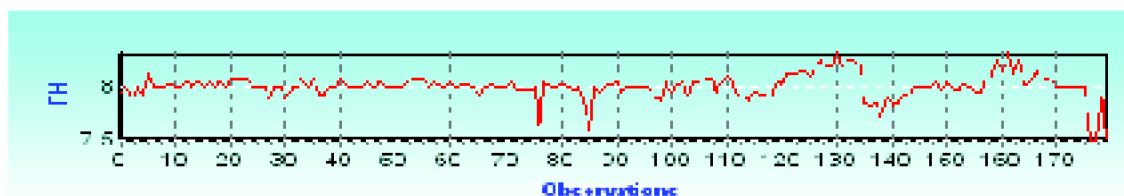


Figure A.2 : Variation saisonnière (hiver) du PH

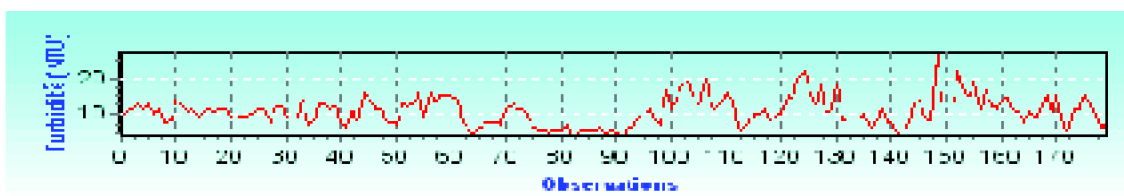


Figure A.3 : Variation saisonnière (hiver) de la conductivité

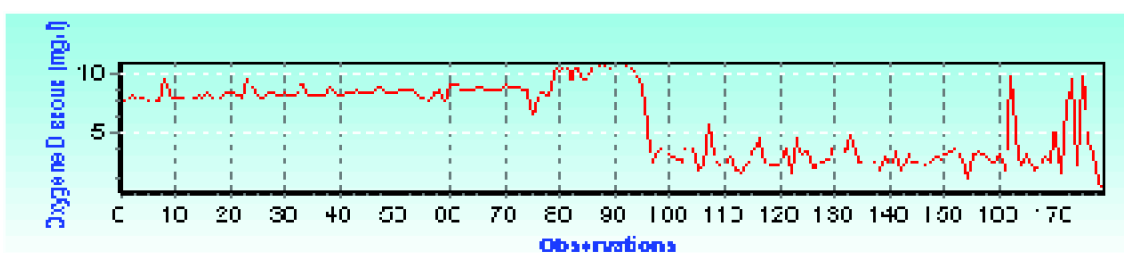


Figure A.4 : Variation saisonnière (hiver) de la turbidité

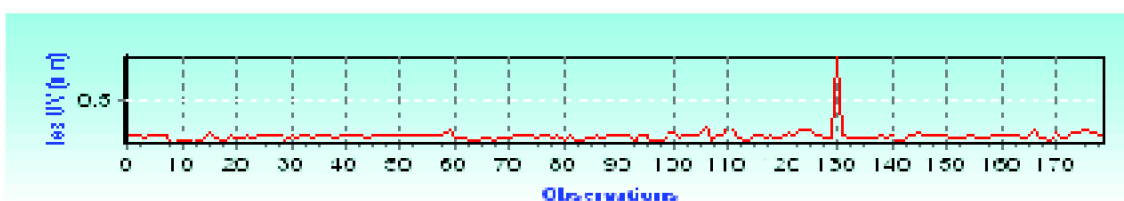


Figure A.5 : Variation saisonnière (hiver) de l'oxygène dissous

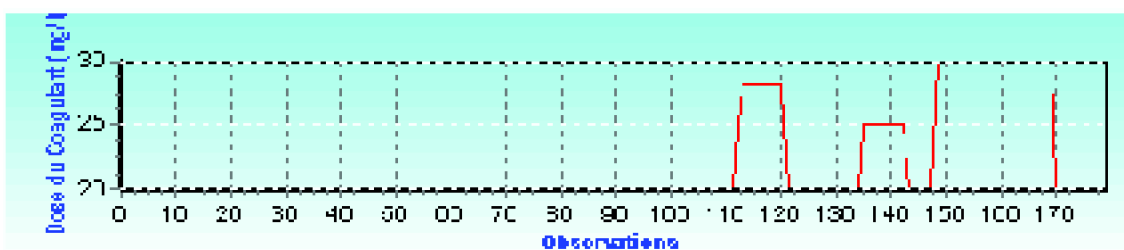


Figure A.6 : Variation saisonnière (hiver) de l'ultraviolet

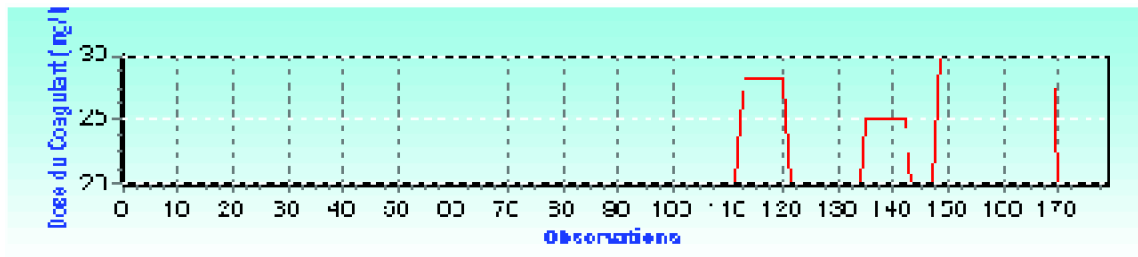


Figure A.7 : Variation saisonnière (hiver) de la dose du coagulant

Méthode de mesure standard						
Observation	Coagulant (mg/l)	PH	Conductivité (µS/cm)	Température (°C)	Alum (mg/l)	Fe (µg/l)
0	0.0	7.5	100	10.0	0.0	0.0
10	0.0	7.5	100	10.0	0.0	0.0
20	0.0	7.5	100	10.0	0.0	0.0
30	0.0	7.5	100	10.0	0.0	0.0
40	0.0	7.5	100	10.0	0.0	0.0
50	0.0	7.5	100	10.0	0.0	0.0
60	0.0	7.5	100	10.0	0.0	0.0
70	0.0	7.5	100	10.0	0.0	0.0
80	0.0	7.5	100	10.0	0.0	0.0
90	0.0	7.5	100	10.0	0.0	0.0
100	0.0	7.5	100	10.0	0.0	0.0
110	0.0	7.5	100	10.0	0.0	0.0
115	28.0	7.5	100	10.0	0.0	0.0
120	28.0	7.5	100	10.0	0.0	0.0
125	0.0	7.5	100	10.0	0.0	0.0
130	0.0	7.5	100	10.0	0.0	0.0
135	25.0	7.5	100	10.0	0.0	0.0
140	25.0	7.5	100	10.0	0.0	0.0
145	30.0	7.5	100	10.0	0.0	0.0
150	30.0	7.5	100	10.0	0.0	0.0
155	0.0	7.5	100	10.0	0.0	0.0
160	0.0	7.5	100	10.0	0.0	0.0
165	28.0	7.5	100	10.0	0.0	0.0
170	28.0	7.5	100	10.0	0.0	0.0

Tableau A.2 : Résultats de l'analyse en composantes principales – fichier hiver

## Annexe A.2 : Résultats d'application de l'analyse en

## composantes principales Saison: printemps (mars-avril-mai).

	moyenne	Ecart type	Min	Max
Température (°C )	13.79	1.39	10.6	19.5
PH	7.77	0.238	7.29	8.24
Conductivité ( $\mu\text{s}/\text{cm}^{-1}$ )	965.20	175.75	690	1278
Turbidité (NTU)	8.77	3.8	2.7	32.4
Oxygène dissous (mg/ l)	4.07	2.78	1.2	9.98
Ultraviolet (NM)	0.11	0.029	0.013	0.27
Coagulant (mg/ l)	22.06	5.35	15	30

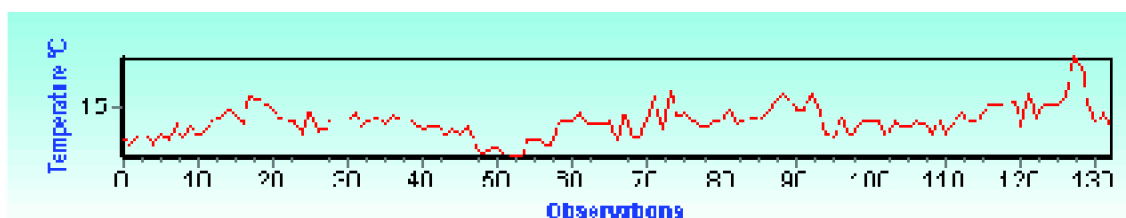


Figure A.8 : Variation saisonnière (printemps) de la température

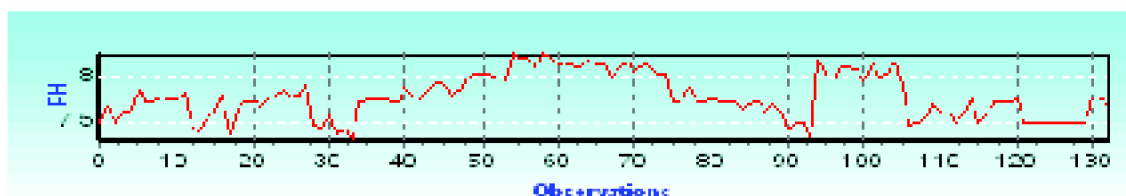


Figure A.9 : Variation saisonnière (printemps) du PH

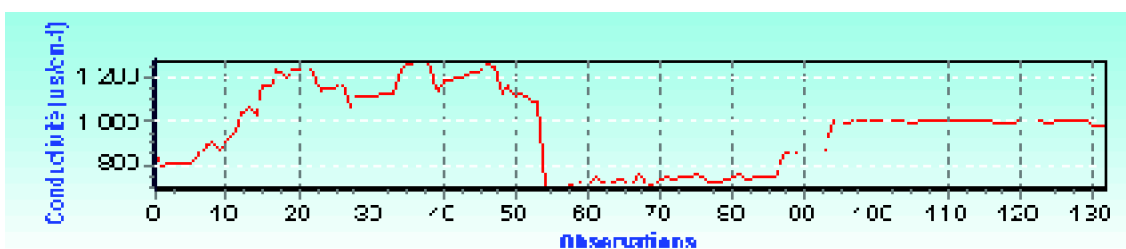


Figure A.10 : Variation saisonnière (printemps) de la conductivité

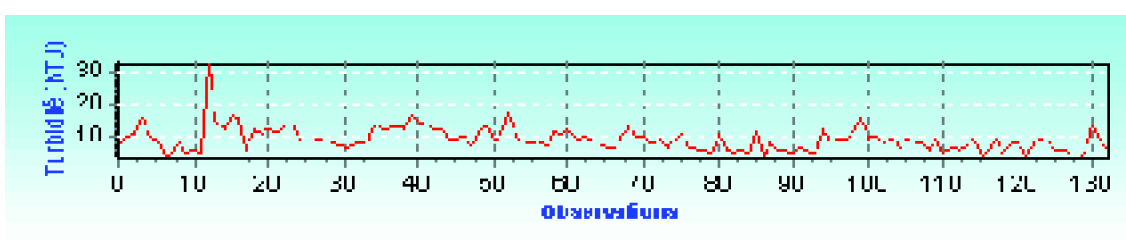


Figure A.11 : Variation saisonnière (printemps) de la turbidité

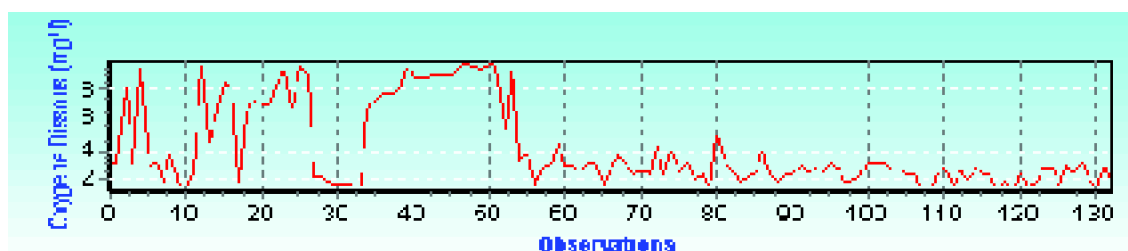


Figure A.12 : Variation saisonnière (printemps) de l'oxygène dissous

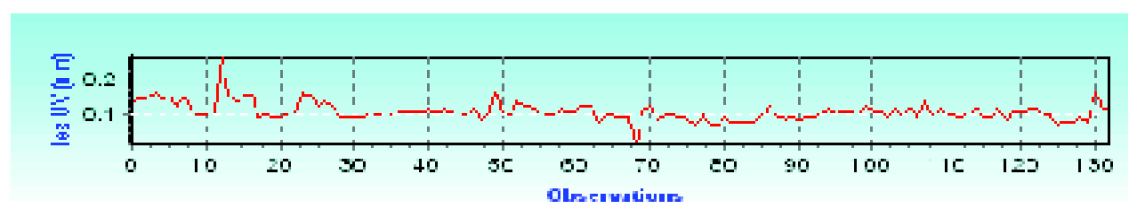


Figure A.13 : Variation saisonnière (printemps) de l'ultraviolet

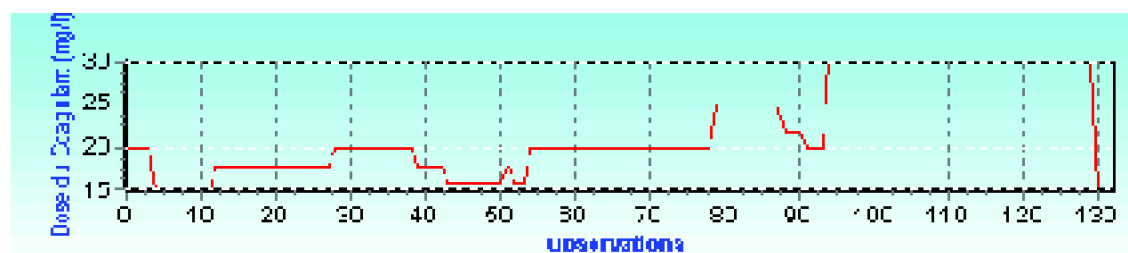


Figure A.14 : Variation saisonnière (printemps) de la dose du coagulant



	moyenne	Ecart type	Min	Max
Température (°C )	17.5	3.2	10.2	24.3
PH	7.618	0.17	7.27	8.03
Conductivité ( $\mu\text{s}/\text{cm}^{-1}$ )	1010.73	93.89	820	1221
Turbidité (NTU)	4.35	2.72	0.44	16.1
Oxygène dissous (mg/l)	3.08	2.30	0.4	8.8
Ultraviolet (NM)	0.097	0.058	0.023	0.74
Coagulant (mg/l)	19.63	7.00	10	35

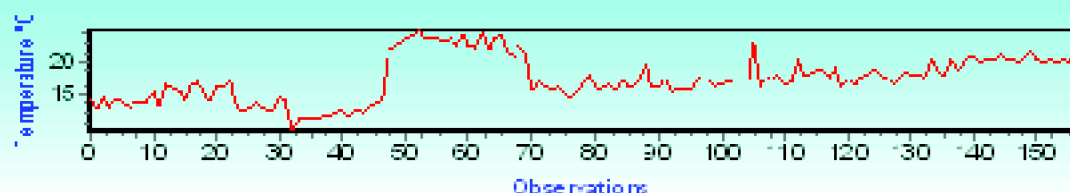


Figure A.15 : Variation saisonnière (été) de la température

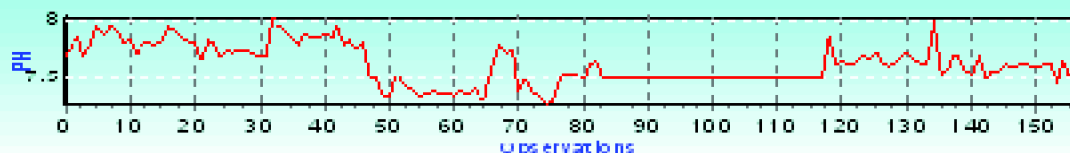


Figure A.16 : Variation saisonnière (été) du PH

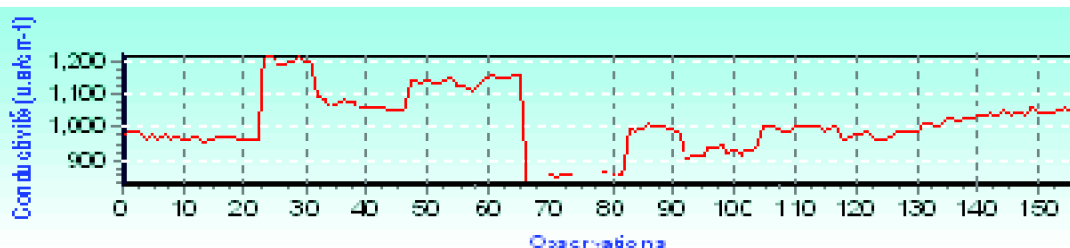


Figure A.17 : Variation saisonnière (été) de la conductivité

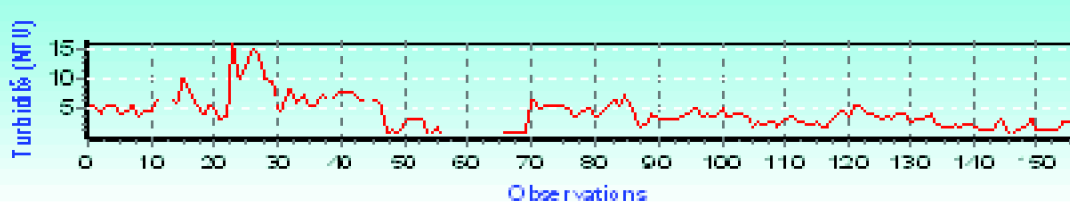


Figure A.18 : Variation saisonnière (été) de la turbidité



Figure A.19 : Variation saisonnière (été) de l'oxygène dissous

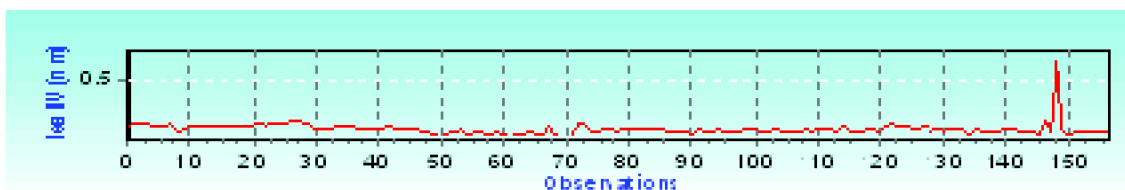


Figure A.20: Variation saisonnière (été) de l'ultraviolet

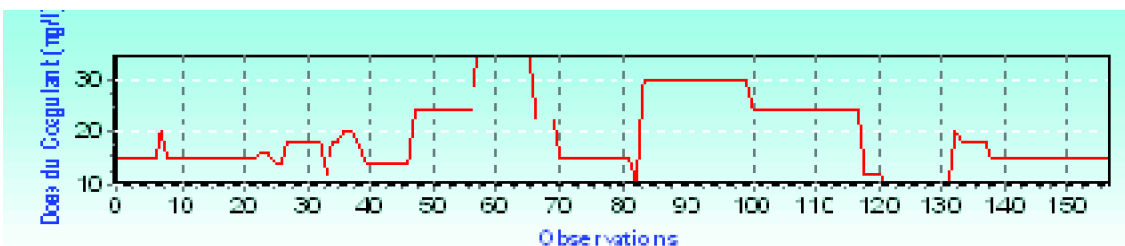


Figure A.21 : Variation saisonnière (été) de la dose du coagulant

Sécheresse - automne							
Coagulant	10/20	30/40	50/60	70/80	90/100	110/120	130/140
Coagulant	10/20	30/40	50/60	70/80	90/100	110/120	130/140
10/20	10/20	30/40	50/60	70/80	90/100	110/120	130/140
30/40	10/20	30/40	50/60	70/80	90/100	110/120	130/140
50/60	10/20	30/40	50/60	70/80	90/100	110/120	130/140
70/80	10/20	30/40	50/60	70/80	90/100	110/120	130/140
90/100	10/20	30/40	50/60	70/80	90/100	110/120	130/140
110/120	10/20	30/40	50/60	70/80	90/100	110/120	130/140
130/140	10/20	30/40	50/60	70/80	90/100	110/120	130/140
Sécheresse - hiver							
Coagulant	10/20	30/40	50/60	70/80	90/100	110/120	130/140
Coagulant	10/20	30/40	50/60	70/80	90/100	110/120	130/140
10/20	10/20	30/40	50/60	70/80	90/100	110/120	130/140
30/40	10/20	30/40	50/60	70/80	90/100	110/120	130/140
50/60	10/20	30/40	50/60	70/80	90/100	110/120	130/140
70/80	10/20	30/40	50/60	70/80	90/100	110/120	130/140
90/100	10/20	30/40	50/60	70/80	90/100	110/120	130/140
110/120	10/20	30/40	50/60	70/80	90/100	110/120	130/140
130/140	10/20	30/40	50/60	70/80	90/100	110/120	130/140
Sécheresse - printemps							
Coagulant	10/20	30/40	50/60	70/80	90/100	110/120	130/140
Coagulant	10/20	30/40	50/60	70/80	90/100	110/120	130/140
10/20	10/20	30/40	50/60	70/80	90/100	110/120	130/140
30/40	10/20	30/40	50/60	70/80	90/100	110/120	130/140
50/60	10/20	30/40	50/60	70/80	90/100	110/120	130/140
70/80	10/20	30/40	50/60	70/80	90/100	110/120	130/140
90/100	10/20	30/40	50/60	70/80	90/100	110/120	130/140
110/120	10/20	30/40	50/60	70/80	90/100	110/120	130/140
130/140	10/20	30/40	50/60	70/80	90/100	110/120	130/140
Sécheresse - été							
Coagulant	10/20	30/40	50/60	70/80	90/100	110/120	130/140
Coagulant	10/20	30/40	50/60	70/80	90/100	110/120	130/140
10/20	10/20	30/40	50/60	70/80	90/100	110/120	130/140
30/40	10/20	30/40	50/60	70/80	90/100	110/120	130/140
50/60	10/20	30/40	50/60	70/80	90/100	110/120	130/140
70/80	10/20	30/40	50/60	70/80	90/100	110/120	130/140
90/100	10/20	30/40	50/60	70/80	90/100	110/120	130/140
110/120	10/20	30/40	50/60	70/80	90/100	110/120	130/140
130/140	10/20	30/40	50/60	70/80	90/100	110/120	130/140



Tableau A.6 : Résultats de l'analyse en composantes principales –fichier été

## Annexe A.4 : Résultats d'application de l'analyse en composantes principales Saison automne (septembre-octobre-novembre).

	moyenne	Ecart type	Min	Max
Température (°C )	18.95	3.45	13	26.2
PH	7.67	0.24	7.23	8.6
Conductivité ( $\mu\text{s}/\text{cm}^{-1}$ )	1058.00	92.65	852	1432
Turbidité (NTU)	6.56	4.44	0.67	26
Oxygène dissous (mg/ l)	5.12	2.89	0.143	13.2
Ultraviolet (NM)	0.11	0.04	0.031	0.25
Coagulant (mg/ l)	25.80	9.42	10	40

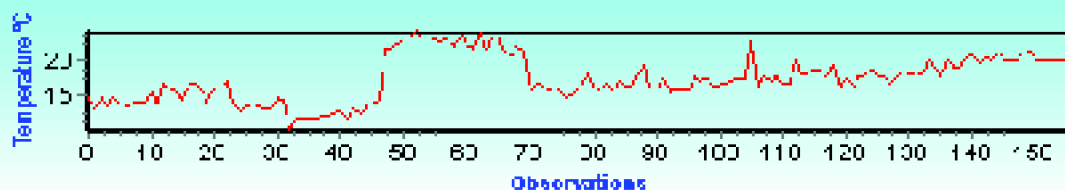


Figure A.22 : Variation saisonnière (automne) de la température

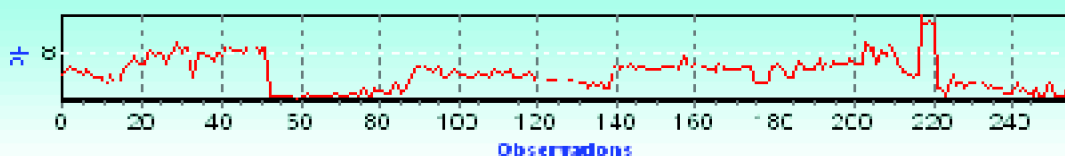


Figure A.23 : Variation saisonnière (automne) du PH

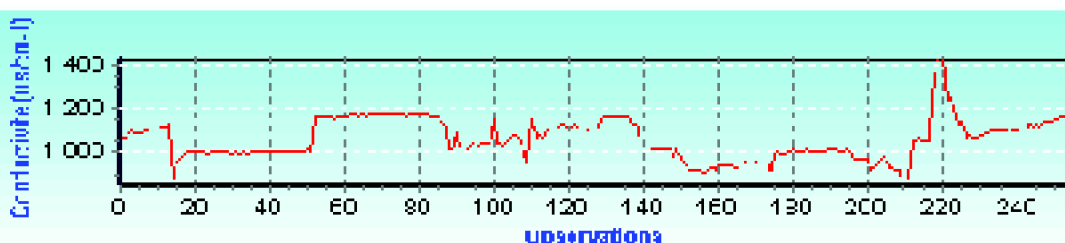


Figure A.24 : Variation saisonnière (automne) de la conductivité

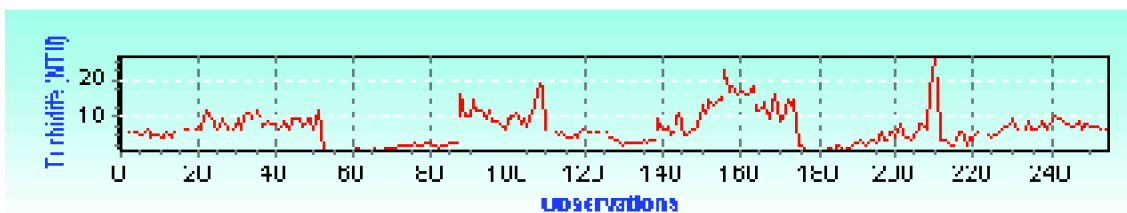


Figure A.25 :Variation saisonnière (automne) de la turbidité

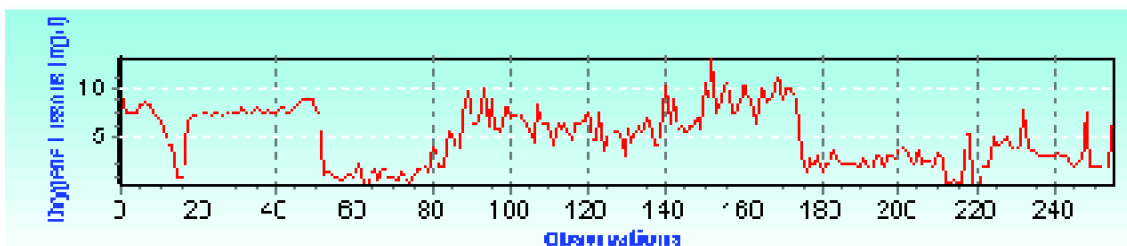


Figure A.26 :Variation saisonnière (automne) de l'oxygène dissous

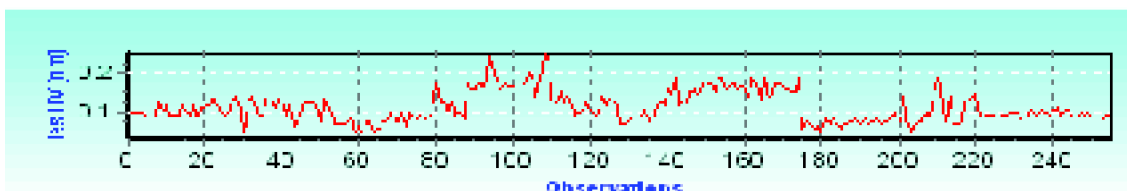


Figure A.27 :Variation saisonnière (automne) de l'ultraviolet

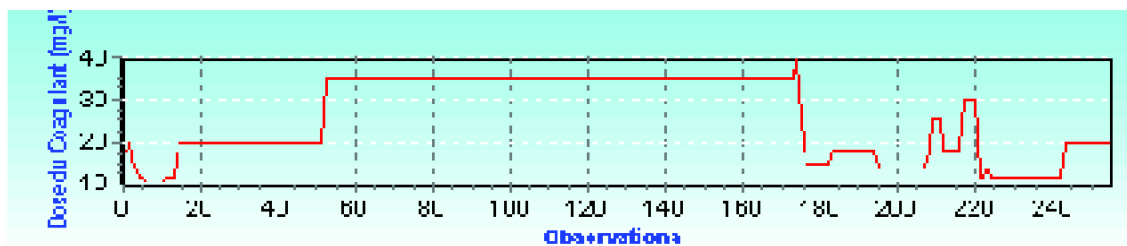


Figure A.28 : Variation saisonnière (automne) de la dose du coagulant

