

الجمهورية الجزائرية الديمقراطية الشعبية  
REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE  
وزارة التعليم العالي والبحث العلمي  
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR  
ET DE LA RECHERCHE SCIENTIFIQUE



جامعة محمد بوضياف - المسيلة  
UNIVERSITE MOHAMED BOUDIAF - M'SILA  
كلية التكنولوجيا  
FACULTE DE TECHNOLOGIE  
قسم الإلكترونيك  
DEPARTEMENT D'ELECTRONIQUE

## **MEMOIRE DE MASTER**

DOMAINE : SCIENCES ET TECHNOLOGIE  
FILIERE : GENIE ELECTRIQUE  
OPTION : CONTROLE INDUSTRIEL

---

**Intitulé : *Etude de la classification supervisée des données  
environnementales à l'aide de réseaux de neurones de  
fonctions à base radiales***

---

*Présenté par :*  
**Mr. Ali DJAIDJA**

*Proposé et dirigé par :*  
**Dr. Mohamed LADJAL**  
**Mr. Mohamed DJERIOUI**

N° d'ordre : 2016/CI13/89/385

**PROMOTION: JUIN 2016**

# DEDICACE

## Je dédie ce travail de master :

- ❖ Aux deux personnes, les plus chères pour moi dans ce monde, qui m'ont soutenu jusqu'au bout : ma mère qui m'a transmis toute sa volonté, sa persévérance et sa grande patience, et mon père qui a su m'imprégner de son savoir et de sa sagesse. Que Dieu les préserve et les bénisse.
- ❖ mon frère : *CHAWKI*.
- ❖ à toute ma famille.
- ❖ à tous mes amis surtout : *NIZAR, RIYADH, ABD EL FETAH, ALI*

# REMERCIEMENTS

*A Dieu, le tout puissant, nous rendons grâce pour nous avoir donné santé, patience, volonté et surtout raison.*

*En premier lieu, je tiens à remercier mon cher encadreur Dr. LADJAL Mohamed qui m'a aidé et conseillé durant ce travail.*

*Je remercie ensuite mon cher Co-encadreur Mr. DJERIOUI Mohamed pour m'avoir honoré par son encadrement, ses conseils précieux, sa patience et ses nobles valeurs humaines.*

*Mes remerciements vont également aux membres de jury pour m'avoir honoré par leur évaluation de ce travail.*

*Je remercie également tous les enseignants du département de l'électronique de l'université Mohamed Boudiaf - M'SILA pour leur aide et encouragement.*

*Enfin, je remercie tous ceux qui m'ont soutenu, encouragé et donné l'envie de mener à terme ce travail.*



# **Table des matières**

ABREVIATION ET NOTATION.....	VI
LISTE DES FIGURES.....	VIII
LISTE DES TABLEAUX.....	X
INTRODUCTION GENERALE.....	1

## **CHAPITRE I**

# **SURVEILLANCE DES PROCEDES DE PRODUCTION D’EAU POTABLE**

Introduction .....	4
1. Composition chimique de l’eau.....	4
2. Production d'eau potable .....	5
3. Définition de l’eau potable .....	5
4. Origine des eaux potables.....	6
4.1. Eaux de source, eaux minérales .....	6
4.2. Eaux brutes .....	6
4.2.1. Catégorie des eaux brutes.....	7

5. Contrôle sanitaire de l'eau.....	7
6. Critères de potabilité.....	8
7. Importance de l'analyse et du traitement.....	8
8. Processus de traitement des eaux .....	9
8.1. Prétraitement.....	10
8.2 Préoxydation .....	11
8.3 Clarification .....	12
8.4 Oxydation-Désinfection.....	13
8.5 Affinage .....	13
9 Surveillance des eaux potables.....	14
9.1 Mesure Les paramètres usuels .....	14
9.2 Mesure des paramètres spécifiques.....	14
9.2.1 Capteur physique.....	14
9.3 Qualité des capteurs .....	16
9.3.1 Sensibilité, précision, gamme de mesure .....	16
9.3.2 Fiabilité, environnement et dérive.....	16
9.4 Les méthodes de surveillance des eaux potables .....	17
9.4.1 Méthode classique : essais de traitabilité en laboratoire .....	17
9.4.2 Surveillance moderne .....	18
9.5 Méthode de surveillance .....	20
9.5.1 Méthodes de surveillance avec modèles .....	21
9.5.1.1 Redondances physiques et analytique.....	21
9.5.1.2 Méthodes d'estimation paramétrique.....	22
9.5.2 Méthode de surveillance sans modèles.....	22
9.5.2.1 Surveillance avec outils statistiques.....	22
9.5.2.2 Surveillance par reconnaissance de formes.....	23
Conclusion.....	25

## CHAPITRE II

# LES RESEAUX DE NEURONES ARTIFICIELS

Introduction .....	26
1. Historique des RNA .....	27
2. Modèle Biologique .....	28
2.1. Neurone.....	29
2.2. Structure de neurone .....	29
2.2.1. Le corps cellulaire (soma) .....	29
2.2.2. Axone .....	29
2.2.3. Synapse.....	29
2.3. Neurone formel .....	29
3. Réseaux de neurones artificiels .....	30
3.1. Définition .....	30
3.2. Domaine d'application des réseaux de neurones (RNA) .....	31
3.2.1 L'application au traitement du signal.....	31
3.2.2 L'application à la planification : .....	32
3.2.3 L'application au contrôle .....	32
3.2.4 L'application au diagnostic.....	32
3.3. Fonctionnement .....	33
4. Modélisation générale .....	33
5. Architecture des réseaux.....	34
6. Les types d'apprentissage des réseaux de neurones .....	37
6.1. L'apprentissage supervisé.....	38
6.2. L'apprentissage non supervisé.....	39

7. Les réseaux de neurones les plus utilisés .....	39
7.1. Perceptron à une seule couche .....	40
7.1.1 La convergence du perceptron .....	40
7.2. Perceptron Multi Couches (PMC) .....	41
8. Algorithme de retro-propagation du gradient (generalized delta rule).....	42
9. Formalisation d'apprentissage.....	42
9.1 Adaptation des poids.....	43
9.2 Algorithme de la rétro-propagation .....	46
10 Réseau de neurones de type RBF (Radial Basis Functions).....	47
10.1 Architecture .....	47
10.2. Algorithme d'apprentissage du réseau RBF.....	48
11. Les différentes applications de RNAs .....	49
11.1 Reconnaissance de formes .....	49
11.2 Traitement de la parole .....	50
11.3 Détection d'anomalies .....	50
11.4 Traitements dépendant du temps .....	51
Conclusion.....	51

## **CHAPITRE III**

# **SIMULATION ET RESULTATS**

Introduction .....	52
1. Problématique.....	52
1.1 Architecture du système de contrôle et de surveillance.....	52
1.2 Approche utilisée dans la surveillance.....	53
1.3 Application à la station de production d'eau potable Tilesdit.....	54
1.3.1 Le site TILES DIT (zone d'étude).....	54

1.3.2 La station de production d'eau potable TILES DIT.....	56
2. Description des données d'entrée .....	56
2.1. Données d'entrée.....	57
3. La technique de contrôle et de surveillance .....	59
3.2 Développement des réseaux de neurones .....	60
4 Simulation.....	60
4.1 Apprentissage et généralisation .....	60
4.1.1 Résultats d'apprentissage .....	61
4.1.2 Résultats de test.....	62
Conclusion.....	64
CONCLUSION GENERALE.....	65
BIBLIOGRAPHIE.....	67



**Abréviation et**

**IA** : Intelligence Artificielle

**RNA** : Réseaux de Neurones Artificiels

**RBF** : Radial Basis Function

**SVM** : Support Vector Machines.

**PMC** : Perceptron MultiCouche

**MLP** : Multi-layer Perceptron

**pH** : potentiel Hydrogène

**T°** : Température

**C** : Conductivité

**TU** : Turbidité

**OD** : Oxygène Dissous

**CO<sub>2</sub>** : dioxyde de carbone

**CaCO<sub>3</sub>** : Carbonate de calcium

**CaO<sub>2</sub><sup>+</sup>** : Peroxyde de calcium

**HCO<sub>3</sub><sup>-</sup>** : Bicarbonates

**θ**: La seuil.

**f** : fonction d'activation

**L** : indice de la couche cachée.

**W** : poids de la connexion<sup>i<sup>ème</sup></sup> entrée.

$\theta_j^L$  : biais ou la valeur de seuil interne du neurone

$K$  : numéro de neurone de la couche de sortie.

$Y_K$  : la sortie désirée

$R_K$  : la sortie actuelle du réseau

$T$  : longueur de l'ensemble d'apprentissage

$N$  : numéro de l'itération

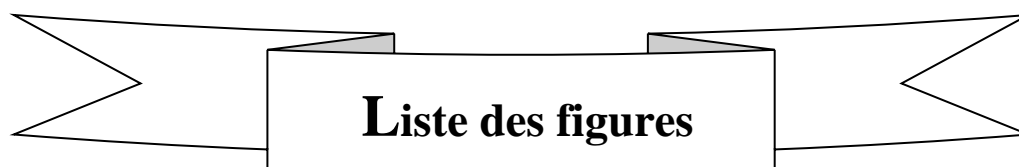
$Mg$  : Magnésium

$B$  : Bicarbonate

$DP$  : Dureté Permanente

$TAC$  : Titre Alcalin Complet

$NNCC$  : Le nombre de neurones dans la couche cachée



## Liste des figures

<b>Figure 1.1</b> Chaîne de traitement d'eau potable .....	10
<b>Figure 1.2</b> exemple d'une boucle de supervision d'une usine d'eau.....	19
<b>Figure 1.3</b> classification des méthodologies de surveillance.....	20
<b>Figure 1.4</b> sensibilité de la méthode à franchissement de seuils aux fausses alarmes.....	23
<b>Figure 1.5</b> schéma général du système de surveillance par reconnaissance de formes.....	24
<b>Figure 2.1</b> Neurone biologique .....	28
<b>Figure 2.2</b> Modèle de base d'un neurone formel.....	30
<b>Figure 2.3</b> Correspondance entre neurones biologiques et neurones artificiels.....	31
<b>Figure 2.4</b> Différents types de fonctions d'activation pour le neurone formel.....	34
<b>Figure 2.5</b> Réseau monocouche.....	35
<b>Figure 2.6</b> Réseau multicouche.....	35
<b>Figure 2.7</b> Réseau à connexion complète.....	36
<b>Figure 2.8</b> Réseau à connexions locales.....	36
<b>Figure 2.9</b> Réseau de neurones boucle.....	37
<b>Figure 2.10</b> Exemple d'apprentissage.....	38
<b>Figure 2.11</b> Réseaux de neurones les plus utilisés.....	39
<b>Figure 2.12</b> Structure d'une couche de neurones.....	40
<b>Figure 2.13</b> Exemple de réseau de type perceptron Multi-Couche.....	41
<b>Figure 2.14</b> Présentation schématique d'un réseau RBF.....	48
<b>Figure 3.1</b> Architecture du système de contrôle et de surveillance.....	53
<b>Figure 3.2</b> Carte géographique situant le barrage « Tilesdit ».....	54

**Figure 3.3**Image représentant le site du barrage « Tilesdit ».....54

**Figure 3.4**Image représentant le site de la station « Tilesdit ».....54

**Figure 3.5** Evolution temporelle des paramètres descripteurs de la qualité d’eau brute  
(Station Tilesdit) .....58



**Liste des tableaux**

<b>Tableau 2.1</b> Comparaison entre neurone Biologique et artificiel.....	31
<b>Tableau 2.2</b> Correspondance RNA - domaines d'application.....	32
<b>Tableau 3.1</b> Statistique descriptives des paramètres de la qualité d'eau brute (station tilesdit).....	58
<b>Tableau 3.2</b> Résultats d'apprentissage de RNA et neuronal-RBF. ....	62
<b>Tableau 3.3</b> Résultats du test de RNA et neuronal-RBF.....	62

## **INTRODUCTION GENERALE**

L'eau est une des ressources les plus importantes de la planète. Par contre, l'alimentation en eau peut être limitée en quantité et/ou en qualité. L'importance de la limitation du point de vue qualitatif est illustrée par le fait que 97 % de l'eau de la Terre est contenue dans des océans et a une teneur en sels d'approximativement 35 000 mg/l ce qui limite énormément l'utilisation de cette immense quantité d'eau comme source d'eau potable. Il est en effet très coûteux de produire de l'eau potable à partir d'eau salée. Il y a 2 % de l'eau de la Terre qui se trouve sous forme de glace et dans les glaciers, tandis que 0,3% se trouvent dans l'atmosphère. Les rivières et les lacs correspondent à seulement 0,1 % de la quantité d'eau totale [1]. Les nappes d'eau souterraine représentent seulement 0,6 % de l'eau de la Terre. De plus, environ la moitié de l'eau souterraine est localisée à des profondeurs plus grandes que 800m.

L'eau est un élément essentiel au fonctionnement des établissements de santé, mais elle peut constituer une source d'infections graves, en cas de contamination, particulièrement pour les patients les plus fragiles. Les principaux risques sanitaires liés à l'utilisation de l'eau dans les établissements de santé doivent donc être identifiés et évalués afin de déterminer les moyens à mettre en œuvre pour maîtriser ces risques [2].

Pour 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. La qualité de l'eau est garantie par le contrôle et la surveillance permanente des services qui s'occupent du traitement et de la distribution. Aujourd'hui, l'eau potable représente l'un des produits alimentaires les plus surveillés dans le monde [3].

La surveillance de potabilité de l'eau peut être mise en pratique selon deux types de méthodes classiques et modernes. Les méthodes classiques sont déterminées par une analyse chimique effectuée au laboratoire, cette méthode nécessite un temps d'analyses relativement important et peut donc être difficilement intégrée dans un système de surveillance et de diagnostic en temps réel de l'unité. En outre les méthodes modernes qui ont l'avantage de pouvoir effectuer un contrôle automatique permanent en temps réel, sont mieux placées pour être une alternative plus efficace. Une surveillance permanente des divers procédés traitement et des relatifs à la qualité de l'eau est devenue nécessaire, ou des systèmes de contrôle automatique infaillibles sont impératifs [4].

Durant ces dernières décennies, des efforts importants ont été réalisés dans le développement des méthodes de contrôle et de surveillance automatique des eaux potables; Ces méthodes dites de haut niveau disposent d'outils qui sont plutôt orientés vers la communication avec un opérateur expert. Celles-ci représentent les techniques de l'intelligence artificielle (IA) qui servent comme outil de base pour l'aide à la décision. Leur réponse est plus élaborée et peut être obtenue soit à partir de données brutes venant directement des variables de surveillance ; soit, à partir de données traitées venant des sorties de traitements de bas niveau [4].

Il est logique de supposer que le problème de contrôle et de surveillance de l'eau brute peut être vu comme un problème de reconnaissance de formes, représentant l'ensemble des observations ou mesure liées aux caractéristiques de celle-ci. Parmi les techniques d'IA connues dans ce domaine, figurent les réseaux de neurones artificiels. Ces techniques se démarquent des autres outils par leur capacité d'apprentissage et de généralisation. Les différents paramètres physico-chimiques exploités dans le traitement de l'eau, tels que le pH, la température, l'oxygène dissous, la conductivité, et les matières en suspension, sont transformés en signaux électriques à partir d'une fusion des données multi-sensorielle et transmis vers une station de contrôle qui assure l'acquisition et le traitement des données. La technique devant être utilisée au niveau du système de décision doit pouvoir effectuer un contrôle quasi permanent de cette ressource précieuse. De plus, le contrôle et la surveillance d'eaux (classification de données), permettant de déterminer les corrélations existantes entre les variables caractéristiques de l'eau brute puis de ne conserver que les caractéristiques apportant réellement une information pertinente. C'est qu'on peut utiliser ces sorties comme des variables d'entrée dans un autre système de surveillance [3].

Le travail présenté dans le cadre de cette thèse a pour objectif l'étude et la mise en œuvre de deux modèles d'apprentissage statistique Neuronal (RNA et neuronal-RBF) appliqué dans le domaine de reconnaissance de formes en classification. La station Tilesdit (Bouira-Algérie) est sélectionnée comme étant une zone d'étude où les données sont collectées. Une étude en simulation est effectuée pour valider, évaluer et comparer les performances de ces modèles dans un but de choix décisif adapté au problème de contrôle et de surveillance de la qualité des eaux potables.

Le mémoire est structuré en trois chapitres suivis d'une conclusion générale.

Le premier chapitre présente la généralisation de l'eau, définition et origine de l'eau potable, nous allons parler du processus de traitement des eaux. Suivent les différentes étapes de la chaîne de traitement des eaux potables. Nous terminerons par les paramètres descripteurs de l'eau ainsi leurs capteurs correspondants : les différentes méthodes et technique surveillance classique et modernes sont décrites.

Dans le deuxième chapitre, nous allons la notion de modèle biologique et neurone formel, nous décrivons définition réseaux de neurones artificiels, modélisation générale, son architecture des réseaux, Les types d'apprentissage des réseaux de neurones et rappelons l'architecture des réseaux de neurones artificiels les plus utilisés (Perceptron Multi Couches (PMC) et radial basis network (neuronal-RBF)), ainsi qu'algorithme de retro-propagation du gradient et formalisation d'apprentissage. Enfin nous décrivons les différentes applications de RNAs.

Le troisième et dernier chapitre est consacré à la simulation et vise l'application des techniques citées précédemment comme étant une solution de surveillance de la qualité de l'eau par reconnaissance de formes. L'objectif est de valider et d'évaluer les performances de chacune des méthodes présentées. Afin de mener une étude comparative permettant un choix décisif de la méthode la mieux adaptée à l'application proposée, on évaluera pour ces méthodes les paramètres liés au taux de reconnaissance et le temps d'apprentissage. Une discussion des résultats conclura cette étude de simulation pour choisir la technique la mieux adaptée. Une conclusion générale en fin de cette thèse est prévue, elle retrace les différentes étapes réalisées et souligne les perspectives envisagées.

## CHAPITRE I

# SURVEILLANCE DES PROCÉDES DE PRODUCTION D'EAU POTABLE

### Introduction

L'eau est un élément naturel d'une importance primordiale, indispensable à toute forme de vie, l'eau est une richesse nécessaire à toutes activités humaines, c'est un facteur de production déterminant dans le développement durable, elle devient de plus en plus au centre des intérêts stratégiques.

Ce premier chapitre est consacré au processus de contrôle de traitement de l'eau brute pour la rendre potable. Il s'agit de donner quelques généralités à propos de cette ressource naturelle qu'est l'eau, ainsi que les outils et moyens mis en œuvre pour son traitement. La chaîne de traitement d'eau potable la plus courante est de ce fait présentée. Les différents paramètres descripteurs de l'eau ainsi que leurs capteurs correspondants sont décrits. Les différentes méthodes et techniques de surveillance classiques et modernes sont citées. Notre application est enfin présentée en fin de chapitre, et le principe de l'approche adoptée dans le procédé de surveillance proposé.

### 1. Composition chimique de l'eau

L'eau n'est pas seulement un ensemble de molécules de  $H_2O$  (deux atomes d'hydrogène et un atome d'oxygène).

Elle contient en réalité naturellement une très grande variété de matières dissoutes, inertes ou vivantes : des gaz, des substances minérales ou organiques, des microorganismes (bactéries, virus, plancton), ainsi que des particules en suspension (fines particules d'argiles, limons et déchets végétaux). En effet, l'eau est un excellent solvant qui se charge en composés solides ou gazeux tout au long de son cycle, suivant les milieux (rivières, zones humides, roches, atmosphère, etc.) dans lesquels elle circule ou séjourne :

- quand elle tombe en pluie, elle se charge des poussières atmosphériques,
- quand elle ruisselle sur les sols (lessivage),
- quand elle s'infiltré dans le sous-sol, elle se charge des produits d'altération des roches.

La composition chimique de l'eau est ainsi complètement liée aux caractéristiques du bassin versant dans lequel elle opère son cycle (la nature du sol et du sous-sol, les espèces végétales et animales, mais également les activités humaines [5]).

## **2. Production d'eau potable**

La surface de notre planète bleue est constituée de deux tiers d'eau, qui est toutefois inégalement répartie sur la terre. Sur les 1.4 milliards de km<sup>3</sup> d'eau disponibles dans le monde, les 97 % sont de l'eau salée. Sur les 3 % d'eau douce restante, la plus grande partie est absorbée par la calotte glaciaire, les glaciers ainsi que par l'atmosphère et par le sol. Seul un 0.3 % de l'ensemble des provisions d'eau douce est disponible en tant qu'eau potable[5].

## **3. Définition de l'eau potable**

Plusieurs spécialistes ont définie l'eau par différents manières, mais en général l'eau potable est très malaisée. C'est en effet un terme générique qui ne peut s'appuyer sur un type unique, car toute eau que l'on peut consommer sans danger peut être considérée comme potable. A cette notion de danger potentiel peut se superposer une notion d'agrément vis-à-vis du gout et même de confort (aspect, température). Pour cela, plusieurs spécialistes ont défini l'eau comme suit :

- Une eau potable est une eau devant satisfaire à un certain nombre de caractéristiques la rendant propre à la consommation humaine.
- L'eau qui est fournie par le réseau de distribution doit être conforme aux normes de potabilités (limites), de qualité fixée par la réglementation. Lorsque la limite de qualité est dépassée, l'eau est déclarée non potable.
- On dit qu'une eau est potable lorsque sa consommation n'a pas de danger pour la santé humaine.
- Eau propre à la consommation, signifiant qu'elle ne contient pas de micro-organisme ou autres substances nocives.
- une eau potable est une eau que l'on peut boire sans risque pour la santé. Pout être consommable, l'eau doit être traitée afin d'éliminer les substances inertes ou vivantes qui peuvent être nocives pour l'organisme. Des normes sont d'ailleurs établies afin de fixer les teneurs limites [4].

## 4. Origine des eaux potables

### 4.1. Eaux de source, eaux minérales

Les eaux dites «de source » sont des eaux naturellement propres à la consommation humaine ; on parle alors d'eau potable. En Europe, les seuls traitements qu'il est permis de leur appliquer sont l'aération, la décantation et la filtration. Aux États-Unis, les traitements sont acceptés. Les eaux naturellement gazeuses, qui contiennent du dioxyde de carbone dissous, peuvent également être regazéifiées avant d'être embouteillées.

Les eaux minérales naturelles, possèdent des propriétés particulières : elles ont des teneurs en minéraux et en oligo-éléments qui peuvent leur donner des vertus thérapeutiques. Comme les eaux de source, elles ne peuvent être traitées. Une eau ne peut être qualifiée en France de minérale que si elle a été reconnue comme bénéfique pour la santé par l'Académie de médecine et le Ministère de la Santé,. En ce sens, c'est un aliment. Elle est d'origine profonde et de composition physico-chimique constante dans le temps, et satisfait à des exigences microbiologiques plus strictes que les autres eaux destinées à la consommation humaine (à l'émergence).

Certaines eaux très fortement minéralisées dépassent les critères législatifs de minéralisation tolérées pour les eaux brutes. Par extension, et certains composants tels que le fluor et les sulfates n'étant bénéfiques pour la santé qu'à faibles doses, il arrive qu'elles soient qualifiées de «non-potable ».[6]

### 4.2. Eaux brutes

De la qualité des eaux brutes (ressource en eau avant tout traitement de potabilisation) va dépendre :

- l'autorisation de prélever cette eau pour la rendre potable
- la filière de potabilisation appropriée.

En application du droit communautaire, les eaux superficielles destinées à la production alimentaire doivent répondre à des exigences de qualité très précises et fixés par des organismes spécialisés [6].

#### **4.2.1. Catégorie des eaux brutes**

Les eaux brutes sont classées en trois catégories, selon l'intensité du procédé :

- Bonne qualité : traitement physique simple et désinfection (dans une « chambre de contact »)
- qualité moyenne : traitement normal physique, chimique et désinfection
- qualité médiocre : traitement physique, chimique poussé, affinage et désinfection.[7]

### **5. Contrôle sanitaire de l'eau**

La qualité de l'eau potable est soumise à deux types de contrôle, que son distributeur soit public ou privé :

Un contrôle officiel, ponctuel, qui relève de la compétence des pouvoirs publics (agence régionale de santé). Il s'agit là du contrôle règlementaire fondamental ; il porte sur l'ensemble du système de distribution : points de prélèvement, stations de traitement, réservoirs et réseaux de distribution ainsi que les centres aquatique (piscines...), les campings et les restaurants.

Les échantillons d'eau prélevés aux différents points de contrôle sont analysés par des laboratoires agréés par le ministère chargé de la Santé. La nature et la fréquence de ces analyses sont fixées par le ministère de la santé. Elles dépendent notamment de la taille de la collectivité desservie : plus celle-ci est grande, plus les contrôles sont fréquents. Suivant le nombre de châteaux d'eau ou stations d'approvisionnement d'eau potable. Cela influence sur la nature et la fréquence des analyses. Une auto-surveillance permanente par les exploitants de leurs services de distribution (régies municipales ou sociétés déléguées). Bien que peu de municipalités le pratiquent en dehors des prélèvements obligatoires, qui incombent financièrement aux communes. La qualité du réseau influe aussi celle de l'eau selon son intégrité : Quand un réseau fuit, en cas de dépression, c'est l'eau de la nappe superficielle, éventuellement polluée ou chargée de microbes indésirables qui peut pénétrer dans le réseau. [7]

### **6. Critères de potabilité**

Pour être consommée, l'eau doit répondre à des critères de qualité très stricts. Fixés par le ministère de la santé, les critères d'une eau "propre à la consommation" sont au nombre de 63, ils portent sur :

- **La qualité microbiologique** : L'eau ne doit contenir ni parasite, ni virus, ni bactériopathogène.
- **La qualité chimique** : Les substances chimiques autres que les sels minéraux font l'objet de normes très sévères. Ces substances sont dites "indésirables " ou "toxiques"
  - ✓ **Les substances indésirables** : Leur présence est tolérée tant qu'elle reste inférieure à un certain seuil (le fluor et les nitrates par exemple).
  - ✓ **Les substances aux effets toxiques** : Le plomb et le chrome en font partie. Les teneurs tolérées sont extrêmement faibles.
- **La qualité physique et gustative**: L'eau doit être limpide, claire, aérée et ne doit présenter ni saveur ni odeur désagréable Les eaux adoucies ou déminéralisées : Les eaux traitées par un adoucisseur d'eau doivent contenir une teneur minimale en calcium ou en magnésium, de même qu'en carbonate ou en bicarbonate. [7]

## 7. Importance de l'analyse et du traitement

Une analyse régulière de l'eau est importante pour les raisons suivantes :

- Elle permet de définir les problèmes existants.
- Elle garantit une eau qui convient à l'utilisation prévue.
- Elle garantit une eau potable sûre.
- Elle permet de vérifier l'efficacité des procédés dans un système de production.

La qualité d'une réserve d'eau peut changer au fil du temps et même subitement. Si l'apparence, l'odeur et le goût de l'eau restent les mêmes, le changement de qualité risque de passer inaperçu. La seule façon de connaître la salubrité de l'eau potable, est de la faire analyser. Comme les bactéries, les parasites et les virus nuisibles sont invisibles à l'œil nu, une eau au goût et à l'apparence agréables n'est pas forcément potable. Ces microbes, qui vivent parfois dans l'eau souterraine et de surface, risquent de causer rapidement des maladies chez les humains qui consomment l'eau sans la traiter adéquatement. Certains contaminants chimiques que l'on retrouve dans les réserves d'eau peuvent causer des problèmes de santé à long terme, qui n'apparaissent que des années après la consommation. Une analyse fréquente de l'eau permet de déterminer le niveau de salubrité de l'eau et de vérifier si le système de traitement a atteint un degré de purification satisfaisant. Plusieurs analyses disponibles sont utiles pour déterminer la salubrité et la sûreté des réserves d'eau. L'analyse de base de l'eau potable comprend

plusieurs aspects d'analyse tels que celui des bactéries coliformes, des nitrates, du pH, du sodium, du chlorure, du fluorure, des sulfates, du fer, du manganèse, des matières totales dissoutes et celui de la dureté [8].

Si on soupçonne la présence d'un contaminant particulier dans l'eau, on peut procéder à d'autres analyses. On analyse parfois l'eau souterraine afin d'y détecter la présence d'arsenic, de sélénium ou d'uranium, par exemple. On peut aussi évaluer la contamination de l'eau de surface ou souterraine par les pesticides. Les réserves d'eau domestique doivent faire l'objet d'une analyse au moins une fois par an. L'eau potable provenant de puits peu profonds ou de réserves de surface, plus sujette à la contamination que l'eau souterraine ; doit être analysée plus souvent (chaque saison). Il est important d'analyser l'eau potable au robinet et à la source. Ces deux analyses permettent de vérifier l'efficacité du système de traitement et de détecter tout changement dans la qualité de l'eau à la source [8]. Il est important de souligner que l'eau avant qu'elle parvienne au consommateur, subi des traitements plus ou moins poussés, elle est stockée, acheminée, puis distribuée. L'eau potable est donc une denrée rare et précieuse qui a un coût, qu'il ne faut pas gaspiller. Par ailleurs, il faut garder à l'esprit qu'elle est produite à partir de ressources naturelles qu'il convient de protéger.

## **8. Processus de traitement des eaux**

Le traitement d'une eau brute dépend de sa qualité. La potabilisation d'une eau de surface nécessite de faire appel à un ensemble de procédés de traitement extrêmement divers qu'il faut assembler dans un ordre déterminé. L'exploitant devra d'une part, respecter certains principes élémentaires pour assurer le contrôle du processus de traitement et le contrôle de l'eau traitée, et d'autre part disposer d'un certain nombre de moyens techniques et humains. Ces processus peuvent être physiques comme la décantation, chimique comme la désinfection ou la coagulation, ou biologique tel que la filtration. Le traitement classique et complet d'une eau s'effectue en plusieurs étapes, il concerne plus spécialement les eaux de surface, dont certaines eaux (souterraines) ne sont pas nécessaires mais doivent également être traitées. Suivant les circonstances, ces deux types de traitement sont différents, mais de toute façon, ils présentent des points communs. Le facteur commun de tous les processus de traitement de l'eau est d'enlever les contaminants existants dans l'eau, afin d'améliorer la qualité de l'eau avant usage. La figure 1.1 représente la chaîne habituelle la plus complète et la plus courante dans une usine de potabilisation appliquée à une eau de surface qui comporte 5 grandes étapes qui

sont : prétraitement, pré-oxydation, clarification (coagulation, floculation-décantation ou flottation, filtration), oxydation-désinfection, traitement d'affinage [9, 10, 11, 12].

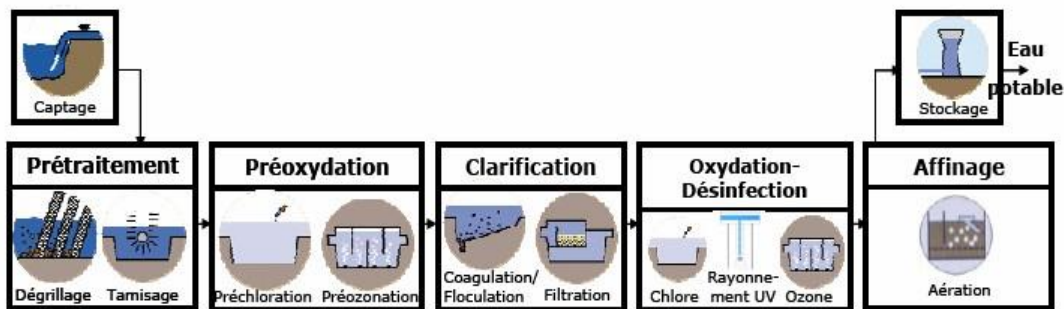


Figure. 1.1. Chaîne de traitement d'eau potable

### 8.1. 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 prétraitements de l'eau de surface[13]. Dans le cas d'une eau potable, les prétraitements sont principalement de deux types :

- le dégrillage,
- le tamisage.

Le dégrillage, premier poste de traitement, permet de protéger les ouvrages aval de l'arrivée de gros objets susceptibles de provoquer des bouchages dans les différentes unités de traitement. Ceci permet également 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. Le dégrillage est avant tout destiné à l'élimination de gros objets : morceaux de bois, etc. Le tamisage, quant à lui, permet d'éliminer des objets plus fins que ceux éliminés par le dégrillage. Il s'agit de feuilles ou de morceaux de plastique par exemple.

### 8.2 Préoxydation

L'oxydation est une opération essentielle à tout traitement des eaux. Elle est toujours incluse en fin de filière au niveau de la désinfection.

A l'issue du prétraitement, on a une eau relativement propre mais qui contient encore des particules colloïdales en suspension. Celles-ci n'ont en elles-mêmes rien de dangereux. Il nous arrive souvent de consommer de l'eau en contenant : le thé, le café, le

vin ou le lait qui sont chargés en matières organiques, mais qui s'oxydent spontanément en présence d'air. On va les détruire dans la mesure du possible par une oxydation. Celle-ci peut être faite de trois façons différentes :

- ajout de Chlore (préchloration)
- ajout de dioxyde de chlore
- ajout d'ozone (préozonation)

La préchloration est effectuée avant le procédé de clarification. Le chlore est le plus réactif et le plus économique, mais il a comme inconvénient de former avec certains micropolluants des composés organochlorés du type chloroforme ou des composés complexes avec les phénols du type chlorophénol dont le goût et l'odeur sont désagréables[13].

On préfère utiliser le dioxyde de chlore qui coûte plus cher mais qui n'a pas les inconvénients de l'oxydation par le chlore cités ci-dessus. Ce type de traitement est cependant réservé à des cas spécifiques. En effet, l'utilisation du dioxyde de chlore présente, lui aussi, des inconvénients non négligeables comme sa décomposition à la lumière, ce qui entraîne une augmentation du taux de traitement à appliquer en période d'ensoleillement. En conclusion, le dioxyde de chlore est un oxydant plus puissant que le chlore qui représente une alternative intéressante à l'utilisation du chlore lorsque celui-ci entraîne des problèmes de qualité d'eau.

Enfin, depuis quinze à vingt ans, on utilise comme oxydant l'ozone, qui non seulement a l'avantage de détruire les matières organiques en cassant les chaînes moléculaires existantes, mais également a une propriété virulicide très intéressante, propriété que n'a pas le chlore. Généralement utilisée en désinfection finale, cette technique peut être mise en œuvre en oxydation. Elle peut aussi être employée pour l'amélioration de la clarification. L'un des avantages d'une préozonation est l'oxydation des matières organiques, et une élimination plus importante de la couleur. Un autre avantage est la diminution du taux de traitement (taux de coagulant) dans le procédé de clarification. En somme, la préozonation est une solution de substitution à la préchloration. On évite ainsi les problèmes liés aux sous-produits de la chloration. Néanmoins, ce procédé ne résout pas tous les problèmes car certaines algues résistent à l'ozone. De plus, son coût reste beaucoup plus élevé que celui au chlore.

### 8.3 Clarification

La clarification est l'ensemble des opérations permettant d'éliminer les matières en suspension *MES* (minérales et organiques) d'une eau brute ainsi que des matières organiques dissoutes [14]. Suivant les concentrations de l'un et de l'autre des différents polluants, on peut être amené à pratiquer des opérations de plus en plus complexes qui vont de la simple filtration avec ou sans réactif jusqu'à la coagulation – floculation – décantation ou flottation – filtration.

La clarification comprend les opérations suivantes :

- Coagulation
- Floculation
- Filtration

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.

Un mauvais contrôle de ce procédé 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 (diminution de la turbidité), le contrôle du coagulant résiduel en sortie (réglementation de plus en plus stricte de la présence de coagulant résiduel dans l'eau traitée) et la diminution des contraintes et des coûts de fonctionnement (coûts des réactifs et des interventions humaines).

### 8.4 Oxydation-Désinfection

La désinfection est l'étape ultime du traitement de l'eau de consommation avant distribution. Elle permet d'éliminer tous les micro-organismes pathogènes présents dans l'eau [14]. Il peut cependant subsister dans l'eau quelques germes banals, car la désinfection n'est pas une stérilisation.

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 résiduelle en désinfectant. Une bonne désinfection via les réactifs oxydants demande la combinaison d'une concentration  $C$  avec un temps de contact  $T$  ; c'est le facteur  $C \cdot T$  (mg.min/L). Cette valeur varie avec les microorganismes concernés, le type de désinfectant et la température.

Les quatre principaux désinfectants utilisés en production d'eau potable sont les suivants :

- Le chlore
- Le dioxyde de chlore
- L'ozone
- Le rayonnement UV

La concentration en oxydant est pratiquement le seul paramètre sur lequel l'opérateur peut intervenir. Il faut retenir que l'efficacité de la désinfection dépend, en partie, du suivi de la concentration en oxydant. L'évolution de la concentration en oxydant est liée à la demande en oxydant de l'eau. Cette demande dépend de la qualité de l'eau, du pH, des températures (différentes entre été et hiver), des matières organiques, et de la concentration en ammoniacale. Dans la section 1.5, nous nous focaliserons davantage sur l'aspect physico-chimique de la désinfection, et en particulier sur la chloration de l'eau.

## 8.5 Affinage

Le traitement final traite de la mise à l'équilibre calco-carbonique. L'eau suit un cycle naturel dans lequel les éléments chimiques qu'elle contient évoluent[14]. L'eau de pluie contient naturellement du dioxyde de carbone ( $\text{CO}_2$ ). Quand celle-ci traverse les couches d'humus, riches en acides, elle peut s'enrichir fortement en  $\text{CO}_2$ . Lors de sa pénétration dans un sol calcaire, c'est-à-dire riche en carbonate de calcium ( $\text{CaCO}_3$ ), elle se charge en calcium  $\text{CaO}_2^+$  et en ions bicarbonates  $\text{HCO}_3^-$ . En fait, le calcium est dissous par l'eau chargée en  $\text{CO}_2$ . On dit qu'elle est entartrant ou incrustante. En revanche, quand l'eau de pluie traverse une roche pauvre en calcium (région granitique), elle reste très chargée en  $\text{CO}_2$  dissous. Cette eau est, en générale, acide. On dit qu'elle est agressive.

Il y a typiquement deux problèmes distincts : corriger une eau agressive et corriger une eau incrustante. La correction d'une eau agressive peut s'effectuer de plusieurs façons. Premièrement, on peut éliminer le  $\text{CO}_2$ . Par aération. Du fait de l'élimination du  $\text{CO}_2$ , le pH augmente et se rapproche du pH d'équilibre. Deuxièmement, on peut ajouter une base à

l'eau. L'ajout de base permet d'augmenter le pH et d'atteindre le pH d'équilibre. La correction d'une eau incrustante peut se faire soit par traitement direct soit en réduisant le potentiel d'entartrage par décarbonatation. Le traitement direct correspond à un ajout d'acide.

## 9 Surveillance des eaux potables

Le surveillance permanent des processus de traitement implique la mesure d'un certain nombre de paramètres à l'aide de capteurs en continu. Ceux-ci peuvent se classer en deux grandes familles : les paramètres usuels et les paramètres spécifiques de l'eau [4].

### 9.1. Mesure Les paramètres usuels

Les paramètres usuels sont principalement les débits, les niveaux de liquides ou de solides, les pressions, les températures. Dans toute installation de traitement d'eau la connaissance du débit est impérative.

### 9.2 Mesure des paramètres spécifiques

Dans les appareils utilisés pour la mesure des paramètres spécifiques de l'eau, les différentes méthodes d'analyse sont mises en œuvre de façon automatique, en particulier : la néphélométrie (mesure de turbidité), la mesure de résistivité ou de conductivité, la potentiométrie (mesure de pH), l'ampérométrie (mesure de concentration en agent oxydant, chlore, ozone), la photolorimétrie et la titrimétrie (mesure de la concentration de certaines substances dissoutes dans l'eau) [4]. On peut classer ces différents appareils en deux grandes catégories : celle des capteurs physiques et celle des analyseurs chimiques qui réalisent, préalablement à toute mesure, une ou plusieurs réactions chimiques.

#### 9.2.1 Capteur physique

**-Mesure de la turbidité :** Le turbidimètre mesure la quantité de lumière diffusée par un échantillon d'eau brute du fait de la présence de particules dans l'eau. Cette valeur est directement proportionnelle à la turbidité de l'échantillon mesuré. Un faisceau lumineux vient toucher la surface sous une incidence telle que ni lui-même ni le faisceau réfléchi ne peuvent impressionner une cellule photorésistante placée sensiblement perpendiculairement au faisceau incident. Par contre, la lumière diffusée par les particules en suspension vient modifier d'autant plus l'éclairement de la cellule que leur nombre est élevé, ce qui permet d'obtenir la mesure de la turbidité de cette eau. Selon l'appareil de

mesure, les gammes de mesure s'étendent de 0 à 5 000 unités néphéométriques de la turbidité (NTU).

- **Mesure de la conductivité** : Le principe mis en œuvre pour la mesure de la conductivité, et de son inverse la résistivité, est simple puisqu'il consiste à mesurer l'intensité du courant électrique recueilli aux bornes de deux électrodes de géométries connues, plongées dans l'eau et soumises à une différence de potentiel alternatif, dont la fréquence doit être d'autant plus élevée que la concentration en acides, sels ou bases dissous est grande, pour éviter les phénomènes de polarisation. La résistivité d'une eau étant fonction du degré de dissociation des molécules dissoutes, la plupart des appareils comportent une compensation automatique de température pour ramener la valeur de la mesure à une température de référence donnée.

- **Mesure de pH** : Industriellement, la mesure du pH se fait toujours par potentiométrie à l'aide de deux électrodes : une électrode de référence et une électrode de mesure. L'électrode de référence est plongée dans une solution de concentration constante en ions hydrogène. Une cloison, laissant passer le courant électrique, sépare la solution de référence de celle dont on veut mesurer le pH et dans laquelle est plongée l'électrode de mesure. Une tension, fonction linéaire de la concentration en ions hydrogène de la solution, apparaît alors aux bornes des électrodes. Il suffit donc de relier ces bornes à un voltmètre pour connaître la valeur du pH. En pratique, les électrodes sont réunies pour former une sonde.

- **Mesure d'Oxygène dissous** : L'ampèremètre est utilisé industriellement en traitement des eaux pour la mesure en continu de la concentration en agents oxydants et met en œuvre une méthode simplifiée d'analyse par ampérométrie. La cellule de mesure, qui est alimentée à débit constant en eau à analyser, comporte une cathode inattaquable, par exemple en platine, et une anode qui peut-être en cuivre, en cadmium, en argent, etc. En l'absence d'agent oxydant, la pile ainsi formée est polarisée et n'est traversée que par un courant très faible. Sa dépolarisation et, par conséquent, l'intensité du courant qu'elle débite sont sensiblement proportionnelles à la concentration de l'agent oxydant qui vient se réduire à la cathode. On mesure ainsi la concentration en chlore, ozone, oxygène d'une eau. L'inconvénient de ces appareils réside dans le fait qu'ils mesurent la somme des agents oxydants et qu'ils ne peuvent être vraiment utilisés que dans le cas où un seul corps se trouve en solution à concentration variable. L'effet d'un autre corps, éventuellement présent à concentration constante, peut être annulé par action sur le zéro de l'appareil [15].

### 9.3 Qualité des capteurs

Pour que le fonctionnement de l'ensemble du système de mesure correct, il est essentiel de s'assurer de la compatibilité de chacun des instruments mis en place en particulier les capteurs. L'information ainsi délivrée, surtout si elle est utilisée dans un système de surveillance, doit être la plus représentative possible de la valeur vraie du paramètre mesuré et être très fiable.

#### 9.3.1 Sensibilité, précision, gamme de mesure

De nombreux facteurs conditionnent l'écart entre la valeur du paramètre mesuré et l'information délivrée. Le premier facteur est la précision du capteur. Celle-ci, exprimée en pourcentage, est le quotient de l'incertitude de la valeur obtenue par l'étendue de mesure pour des conditions de mesure données. La précision du capteur est fonction du processus de mesure mais aussi des corrections annexes qui y sont apportées. Une bonne précision finale dépend d'une bonne corrélation entre une caractéristique ou un phénomène annexe, pouvant être mesuré avec une bonne précision, et le phénomène à mesurer. Un autre facteur peut être l'existence d'erreurs systématiques dues à un étalonnage incorrect ou trop peu fréquent du capteur. Les erreurs accidentelles peuvent également être causées par des signaux parasites, ou des absences de correction de température, de pression, etc. La sensibilité initiale d'un appareil de mesure est un autre facteur à prendre en compte. Celle-ci est la valeur minimum du paramètre à mesurer en dessous duquel l'appareil ne réagit pas. La sensibilité en fonctionnement est la plus petite variation du paramètre mesuré décelable par la mesure. Elle n'est pas nécessairement constante dans toute la gamme de mesure. Il faut enfin tenir compte de la gamme de mesure du capteur, qui correspond aux valeurs de seuils au delà desquels la précision et la sensibilité du capteur se dégradent [15].

#### 9.3.2 Fiabilité, environnement et dérive

La fiabilité est définie comme la capacité du capteur à fonctionner correctement, c'est-à-dire à fournir des données avec la précision annoncée. Elle dépend naturellement de la qualité de conception du matériel qui doit être robuste et avoir été entretenu conformément aux spécifications du constructeur. Mais elle dépend également de son adaptation à l'environnement dans lequel il va se trouver. Les contraintes des capteurs concernant la gestion de l'eau sont principalement l'humidité et la nature de l'eau. L'humidité peut provoquer de la condensation dans les boîtiers du matériel. Ceux-ci doivent être étanches, des submersions étant toujours possibles, et doivent comporter des

dispositifs éliminant la condensation. Cette atmosphère humide peut également provoquer des courts circuits au niveau des câbles de jonction ou d'alimentation. La nature de l'eau, notamment celle des rivières, peut perturber les capteurs immergés avec des dépôts en modifiant les réactions. C'est en particulier le cas de nombreuses sondes dont le nettoyage doit être effectué très régulièrement car ces dépôts provoquent une dérive du capteur. C'est le principal défaut de ce type de capteur dont la surveillance doit être constante, les dispositifs de nettoyage automatique sous forme de brosses ou de rétro-lavage de la sonde n'étant pas toujours efficaces.

En conclusion, pour tirer pleinement parti des avantages des capteurs de mesure et de l'instrumentation associée, il est indispensable d'accepter certaines contraintes telles que le nettoyage des sondes de mesures, l'étalonnage régulier, etc. Malgré ces précautions, certains facteurs peuvent encore perturber l'information délivrée par les capteurs [15].

#### **9.4 Les méthodes de surveillance des eaux potables**

Quand on parle de surveillance des eaux potables, il s'agit en fait de connaître l'état de l'eau en continue( à chaque instant) à partir des différents paramètres ayant trait à sa qualité.

##### **9.4.1 Méthode classique : essais de traitabilité en laboratoire**

Cette technique a pour but de connaître les différents paramètres de l'eau brute pour décider après sur son état propre, et par suite chercher les techniques et méthodes pour la rendre potable. ces méthodes sont traditionnelles, déterminées à l'aide d'un essai expérimental appelé (jar-test)[4]. On procède généralement à un certain nombre de mesures utiles pour le test de qualité tels que: le contrôle bactériologique, le contrôle de désinfection, et de contrôle physico-chimique ( pH, T°, Turbidité, conductivité, oxygène dissous,...). La dose optimale recherchée est déterminée en fonction de la qualité des différentes eaux comparées. la fréquence de ces jar-test est souvent irrégulière. En général dans les usines importantes, un seul essai est effectué par jour [16]. L'opérateur fera un nouvel essai entre temps pour changer la dose du coagulant uniquement si la qualité de l'eau traitée se dégrade. L'inconvénient de cette technique est qu'elle nécessite de façon non-stop des interventions et des déplacements sur site de l'opérateur. Ce type d'approche a également le désavantage d'avoir un temps de réponse relativement long. En effet, on ne modifie la dose du coagulant qu'une fois l'événement apparu, vérifié puis analysé. De plus, elle ne permet pas de suivre finement l'évolution de la qualité de l'eau brute. Par exemple,

si l'eau brute devient plus (facile à traiter) l'opérateur ne le verra pas forcément et donc ne modifiera pas la dose du coagulant, d'où un coût d'exploitation plus élevé que nécessaire et une économie non réalisée. en voici tout l'intérêt de disposer de contrôle automatique de ce procédé pour une meilleure efficacité de traitement et une réduction des coûts d'exploitation. La régulation de l'eau brute au niveau des usines de traitement doit se faire de façon immédiate en se basant sur une surveillance continue des paramètres descripteurs de la qualité de cette eau.

#### **9.4.2 Surveillance moderne**

La fonction surveillance moderne de l'exploitation d'un tel processus à travers des données quantifiables et qualifiables permet aussi de détecter les états anormaux de l'objet à surveiller et prendre ainsi les décisions pour un meilleur état. L'importance de la mesure en continu des paramètres physiques physico-chimiques à l'aide de capteurs dans un système de surveillance monté sur place vient du fait que [4] :

- ✓ Ces paramètres ne sont pas conservatifs et changent instantanément.
- ✓ Les mesures sont relativement simples, rapides et peu coûteuses.

Ces mesures permettent de détecter immédiatement des anomalies de la composition de l'eau ( élévation de pH par exemple ) ce qui permet une intervention immédiate.

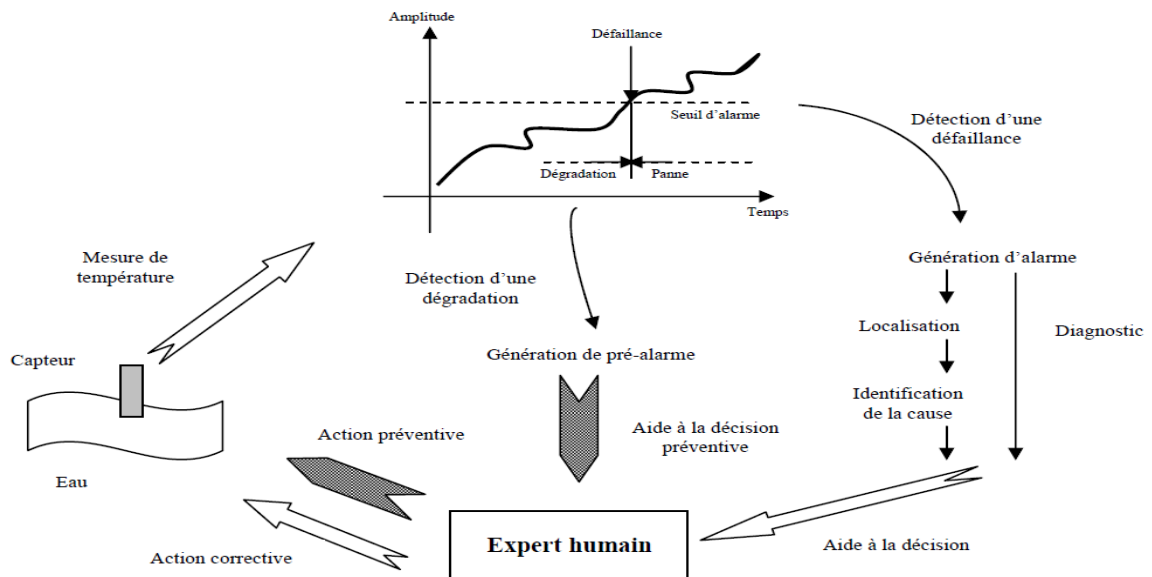
##### **- Les avantages des systèmes de surveillance**

La surveillance est un dispositif passif, informationnel qui analyse l'état du système. Elle consiste notamment à détecter et classer les anomalies en observant l'évolution du système puis à les diagnostiquer en localisant les éléments défaillants et en identifiant les causes premières, et prendre les décisions nécessaires et finales sur l'état de l'objet surveillé. La surveillance se compose de deux fonctions principales, qui sont la détection et le diagnostic. Pour détecter toute anomalie du système ou au niveau de l'objet à surveiller, il faut être capable de classer les situations observables comme étant normales ou anormales. Cette classification n'est pas triviale, étant donné le manque d'information qui caractérise généralement les situations anormales. Une simplification communément adoptée consiste à considérer comme anormale toute situation qui n'est pas normale. Quant à l'objectif de la fonction diagnostic, c'est de rechercher les causes et de localiser les organes qui ont entraîné une observation particulière. Cette fonction se compose de deux

fonctions élémentaires: localisation et identification. La localisation permet de déterminer le sous-ensembles fonctionnel défaillant. Alors que l'identification consiste à déterminer les causes qui ont mené à une situation anormale. Les avantages des systèmes de surveillance basés sur des méthodes modernes sont [4]:

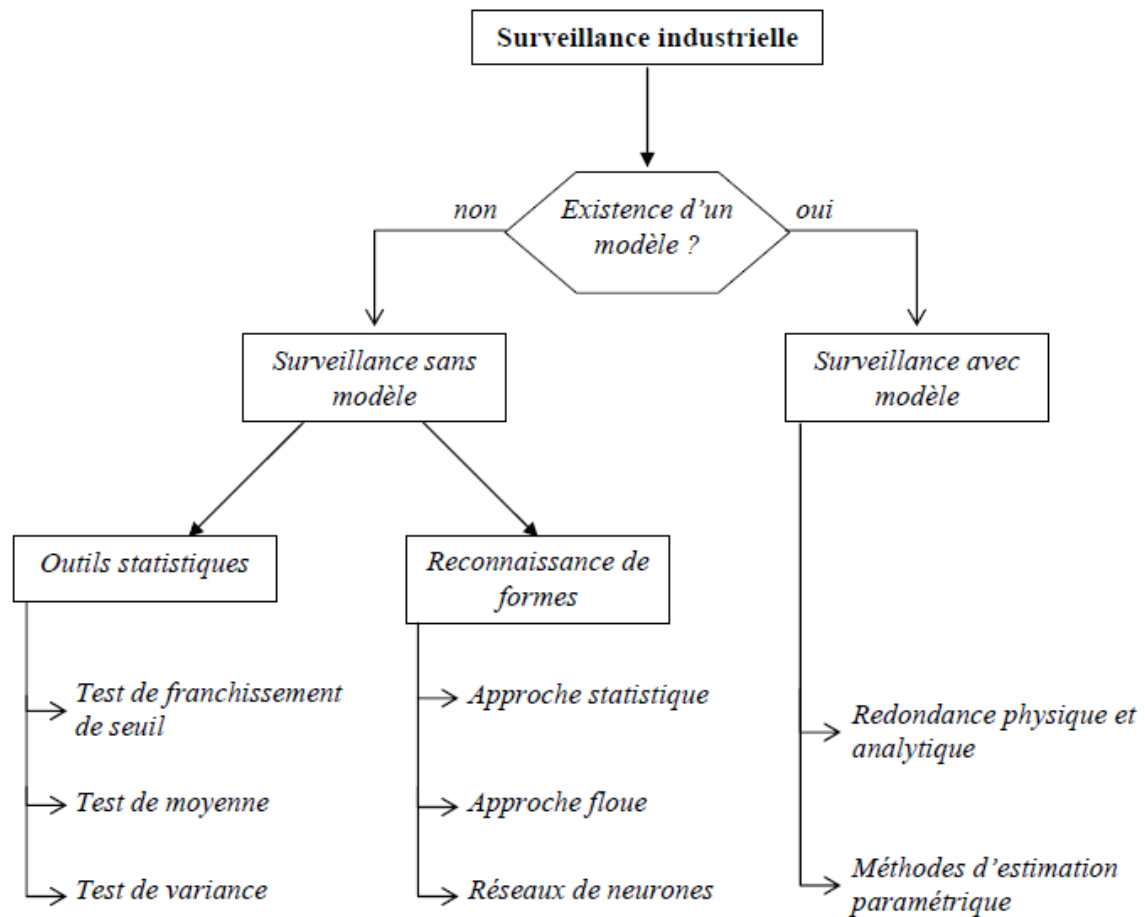
- Amélioration des conditions d'exploitation et des performances d'une installation.
- Augmentation de la productivité.
- Fonctions temps réel et différé.
- Aide à la décision et à la maintenance.

La figure 1.2 montre l'exemple d'une boucle de supervision (surveillance + action) dans une usine moderne.



**Figure 1.2** exemple d'une boucle de supervision d'une usine d'eau.

## 9.5 Méthode de surveillance



**Figure 1.3** classification des méthodologies de surveillance.

Les méthodes de surveillance industrielle telle qu'elles sont présentées dans ce paragraphe sont illustrées sur la Figure 1.3. L'existence d'un modèle formel ou mathématique de l'équipement détermine la méthode de surveillance utilisée. La surveillance avec modèle se compose essentiellement de deux techniques : méthodes de redondance physique et analytique et méthodes d'estimation paramétrique. D'un autre côté, les méthodes qui ne se basent pas sur l'existence du modèle se divisent en deux catégories : méthodes utilisant des outils statistiques et méthodes de reconnaissance des formes. Les outils statistiques établissent des tests sur les signaux d'acquisition. Ces tests ne sont capables d'assurer que la fonction détection de défaillances. Par contre, les techniques de surveillance par reconnaissance des formes sont plus élaborées par rapport aux simples tests statistiques et sont capables de détecter et de diagnostiquer les défaillances [16].

### 9.5.1 Méthodes de surveillance avec modèles

Les méthodes de surveillance avec modèle ont pour principe de comparer les mesures effectuées sur le système aux informations fournies par le modèle. Tout écart est alors synonyme d'une défaillance. Les outils de la théorie de la décision sont ensuite utilisés pour déterminer si cet écart est dû à des aléas normaux comme, par exemple, le bruit de mesure ou s'il traduit une défaillance du système. Ces méthodes peuvent être séparées en deux techniques : techniques de redondance physique et analytique et techniques d'estimation paramétrique. Ces deux techniques seront présentées brièvement.

#### 9.5.1.1 Redondances physiques et analytiques

##### a) Redondances physiques

Afin de fiabiliser la détection des défaillances à partir des signaux mesurés, il faut un moyen pour distinguer les défaillances capteurs des défaillances système. La méthode la plus simple consiste à utiliser la redondance physique. Il s'agit de doubler ou tripler des composantes de mesure du système. Si ces composantes identiques placées dans le même environnement émettent des signaux identiques, on considère que ces composants sont dans un état de fonctionnement nominal et, dans le cas contraire, on considère qu'une défaillance capteur s'est produite dans au moins une des composantes. Cette méthode par redondance physique a l'avantage d'être conceptuellement simple mais est coûteuse à être mise en œuvre et conduit à des installations encombrantes. Elle est, par conséquent, utilisée uniquement pour la surveillance des sous-ensembles critiques d'un système. Un autre inconvénient est que les composantes identiques fabriquées dans la même série peuvent se dégrader de la même façon et tomber en panne en même temps. Pour pallier ce dernier inconvénient, on peut utiliser des composantes différentes qui remplissent la même fonction.

##### b) Redondances analytiques

Les méthodes de redondance analytique nécessitent un modèle du système à surveiller. Ce modèle comprend un certain nombre de paramètres dont les valeurs sont supposées connues lors du fonctionnement nominal. Dans la mesure où la surveillance est établie à partir des mesures échantillonnées des grandeurs observables du système, la modélisation de ce dernier sous forme discrète semble être raisonnable. Le but des méthodes de redondance analytique est d'estimer l'état du système afin de le comparer à

son état réel. L'estimation de l'état du système peut être réalisée soit à l'aide de techniques d'estimation d'état, soit par obtention de relations de redondance analytique [16].

### 9.5.1.2 Méthodes d'estimation paramétrique

Les méthodes d'estimation paramétrique supposent l'existence d'un modèle paramétrique décrivant le comportement du système et que les valeurs de ces paramètres en fonctionnement nominal soient connues. Elles consistent alors à identifier les paramètres caractérisant le fonctionnement réel, à partir de mesures des entrées et des sorties du système. On dispose ainsi d'une estimation des paramètres du modèle, effectuée à partir des mesures prises sur le système et de leurs valeurs théoriques. Pour détecter l'apparition de défaillances dans le système, il faut effectuer la comparaison entre les paramètres estimés et les paramètres théoriques. Comme pour les méthodes de redondance analytique, la théorie de la décision sert alors à déterminer si l'écart observé est dû à des aléas normaux du fonctionnement ou à des défaillances. La différence entre les méthodes de redondance analytique et les méthodes d'estimation paramétrique est qu'on effectue, pour les premières, la comparaison entre l'état estimé et l'état théorique du système, alors que pour les secondes, on compare les paramètres estimés aux paramètres théoriques du système. Les méthodes d'estimation paramétrique requièrent donc l'élaboration d'un modèle dynamique précis du système à surveiller. Ceci restreint leur utilisation à des procédés bien définis [16].

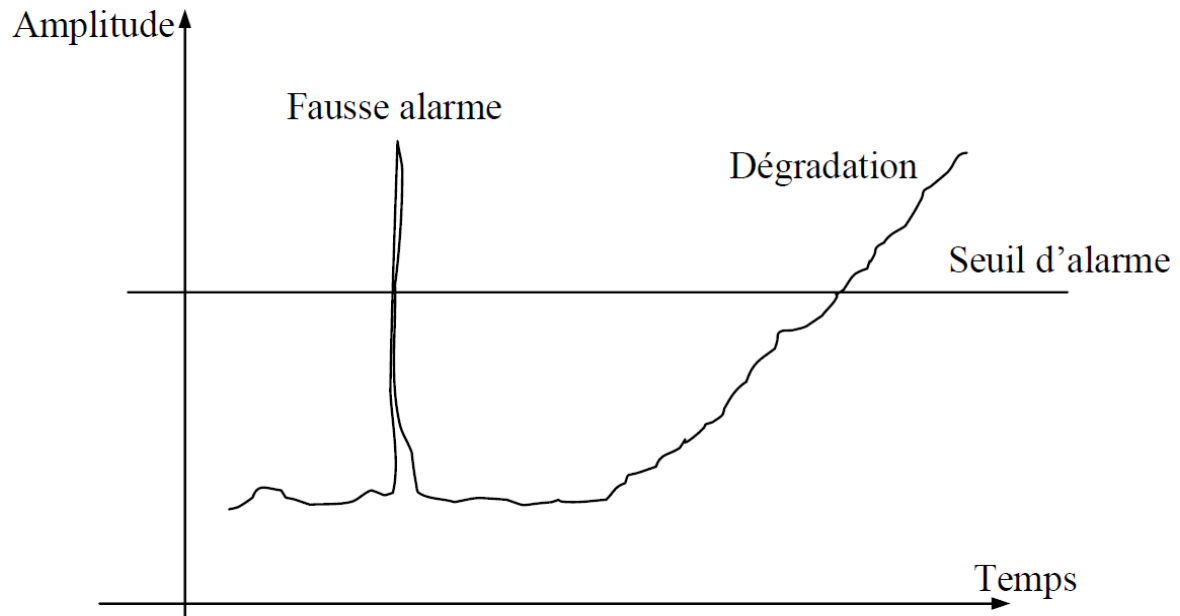
### 9.5.2 Méthode de surveillance sans modèles

Nombreuses sont les applications industrielles dont le modèle est difficile, voire impossible à obtenir suite à une complexité accrue ou à de nombreuses reconfigurations intervenants durant le processus de production. Pour ce type d'applications industrielles, les seules méthodes de surveillance opérationnelles sont celles sans modèle. Deux solutions existent dans ce cas : surveillance avec des tests statistiques et surveillance par reconnaissance des formes. La première technique est moins élaborée que la deuxième dans le sens où elle ne remplit qu'une partie de la surveillance.

#### 9.5.2.1 Surveillance avec outils statistiques

Les outils statistiques consistent à supposer que les signaux fournis par les capteurs possèdent certaines propriétés statistiques. On effectue alors quelques tests qui permettent de vérifier si ces propriétés sont présentes dans un échantillon des signaux mesurés [17].

a) **Test de franchissement de seuils** : Le test le plus simple est de comparer ponctuellement les signaux avec des seuils préétablis. Le franchissement de ce seuil par un des signaux capteurs génère une alarme. Ce type de méthode est très simple à mettre en œuvre mais ne permet pas d'établir un diagnostic des défaillances. Cette méthode est aussi très sensible aux fausses alarmes (Figure 1.4).



**Figure 1.4** sensibilité de la méthode à franchissement de seuils aux fausses alarmes

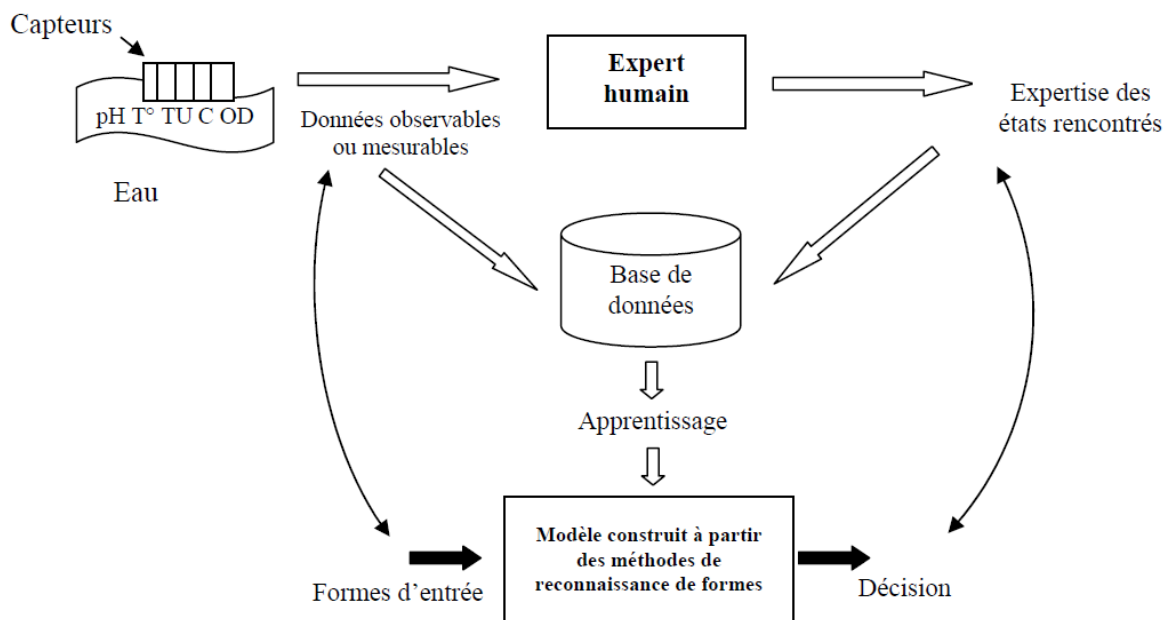
b) **Test de moyenne** : Contrairement à la méthode précédente, le test de comparaison est effectué sur la moyenne du signal contenu dans une fenêtre de  $n$  valeurs plutôt que sur une valeur ponctuelle.

c) **Test de variance** : On peut également calculer la variance d'un signal. Tant que cette variance se situe dans une bande située autour de sa valeur nominale, l'évolution du système est supposée normale.

### 9.5.2.2 Surveillance par reconnaissance de formes

L'approche de surveillance par reconnaissance de formes (ou par analyse des données), permet d'associer un ensemble de mesures effectuées sur le système à des états de fonctionnement connus. Cette fonction permet d'avoir une relation d'un espace de caractéristiques vers un espace de décision de façon à minimiser le risque de mauvaise classification. Les techniques de reconnaissance de formes sont nombreuses et pour certaines, les résultats sont globalement satisfaisants. Ces méthodes supposent qu'aucun

modèle n'est disponible. Cependant, elles effectuent les opérations de perception, de classification et de prédiction. Ces méthodes sont basées sur les théories des probabilités et de l'Intelligence Artificielle. La première peut se montrer insuffisante car elle suppose une connaissance a priori de tous les états de fonctionnement et ne prend pas en compte l'évolution du système. La seconde a l'avantage de se baser plutôt sur une phase dite d'apprentissage. Parmi les techniques proposées, on trouve en particulier les systèmes experts, les outils statistiques de reconnaissance de formes, le raisonnement à partir de cas (RàPC), l'ACP, les réseaux de neurones artificiels (RNA), SVM, logique floue et réseaux neuro-flous. La figure I.5 montre l'architecture générale qu'on peut imaginer d'une application de surveillance par reconnaissance de formes. L'expert humain joue un rôle primordial dans ce type d'applications. Toute la phase d'apprentissage supervisée dépend de son analyse des états du système où chaque état est caractérisé par un ensemble de données (formes d'entrée) recueillies sur le système. Après cette phase d'apprentissage, l'algorithme utilisé associera les classes correspondantes représentant ainsi les sorties du système aux formes d'entrée présentées. Le principal avantage de ces techniques est leur capacité d'apprentissage et de généralisation des données reçues en entrée. Ces techniques s'apprêtent parfaitement pour une éventuelle mise au point de systèmes de surveillance capables de s'adapter à d'éventuelles extensions et reconfigurations multiples [3].



**Figure 1.5** schéma général du système de surveillance par reconnaissance de formes

**Conclusion**

Ce premier chapitre a servi d'introduction au domaine de contrôle et de surveillance des eaux potables. Les différentes étapes d'une chaîne de traitement sont présentées. Les paramètres ainsi que les capteurs physico-chimiques utilisés comme source d'information ayant trait à la qualité de l'eau ont été particulièrement décrits. De même, les différentes techniques existant actuellement dans le domaine du contrôle et de surveillance des eaux potables ont été aussi évoquées.

## CHAPITRE II

# LES RESEAUX DE NEURONES ARTIFICIELS

### INTRODUCTION

L'évolution technologique durant les dernières années a permis aux scientifiques d'élaborer et de perfectionner des méthodes pour différents domaines. L'évolution des ordinateurs en particulier et la capacité d'intégration de composants formidable atteintes à nos jours ont permis une grande vitesse de calcul et une grande capacité mémoire. Parmi ces méthodes, il existe une méthode qui est utilisée dans plusieurs domaines de recherches et de différentes manières, ainsi elle peut être utilisée d'une manière complètement soft en utilisant uniquement l'ordinateur ou d'une manière hard en utilisant les circuits intégrés. Cette méthode est celle des réseaux de neurones artificiels (RNA). Les réseaux de neurones artificiels sont des outils puissant capables d'être utilisés dans prés que tous les domaines technologiques, et on peut citer : le traitement du signal, vision, parole, prévision, modélisation, aide à la décision, robotique, évaluation des écosystèmes, identification des bactéries, commande des processus, modélisation des systèmes physiques reconnaissance des formes, mesure, instrumentation,...[18].

Dans le cadre de notre travail, la solution adoptée rappelons-le une fois de plus, est une surveillance et un contrôle de la qualité de l'eau par les technique de reconnaissance de formes. Dans ce chapitre, nous allons donc pouvoir passer en revue des méthode d'apprentissage statistique appliquées à la classification et la régression.

Après une brève introduction, où nous allons la notion de modèle biologique et neurone formel, nous décrivons définition réseaux de neurones artificiels, modélisation générale, son architecture des réseaux, Les types d'apprentissage des réseaux de neurones et rappelons Les réseaux de neurones les plus utilisés (Perceptron Multi Couches (PMC)), ainsi que algorithme

de retro-propagation du gradient et formalisation d'apprentissage. Enfin nous décrivons les différentes applications de RNAs.

### 1. Historique des RNA

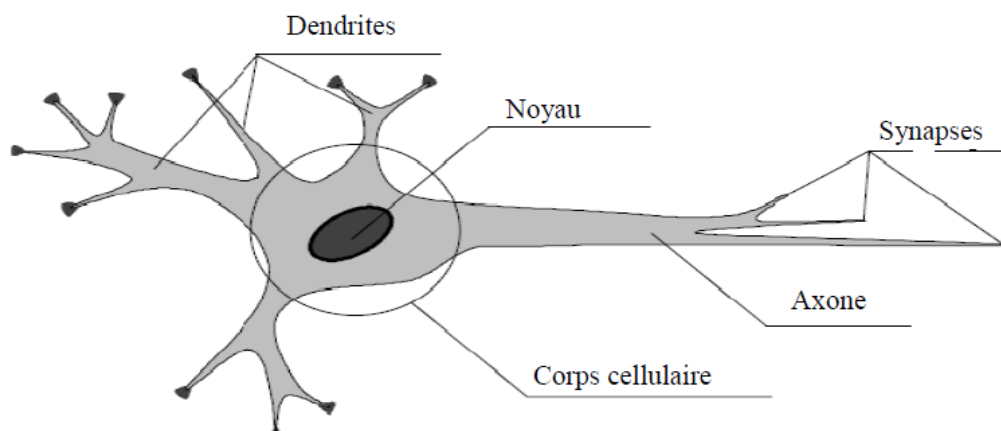
L'origine de l'inspiration des réseaux de neurones artificiels remonte à 1890 où W. James, célèbre psychologue américain, introduit le concept de mémoire associative. Il propose ce qui deviendra une loi de fonctionnement pour l'apprentissage des réseaux de neurones, connue plus tard sous le nom de loi de Hebb. Quelques années plus tard, en 1949, J. McCulloch et W. Pitts [19] donnent leurs noms à une modélisation du neurone biologique (un neurone automatique comportement binaire). Ce sont les premiers à montrer que des réseaux de neurones formels simples peuvent réaliser des fonctions logiques, arithmétiques et symboliques complexes. C'est ensuite que D. Hebb, physiologiste américain, présente en 1949 les propriétés des neurones par le conditionnement chez l'animal. 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. Les premiers succès de cette discipline remontent à 1957, lorsque F. Rosenblatt développe le modèle du Perceptron. Il construit le premier neuro-ordinateur basé sur ce modèle et l'applique au domaine de la reconnaissance des formes. Notons qu'à cette époque les moyens à sa disposition étaient limités et c'était une prouesse technologique que de réussir à faire fonctionner correctement cette machine plus de quelques minutes. C'est alors qu'en 1960, l'automaticien Widrow développe le modèle Adaline (Adaptative LinearElement). Dans sa structure, le modèle ressemble au Perceptron, cependant la loi d'apprentissage est différente. Celle-ci est à l'origine de l'algorithme de rétro-propagation de gradient très utilisé aujourd'hui 1969 : M.Minsky et S.Papert publient un ouvrage qui met en exergue les limitations théoriques avec les Perceptrons Multi Couches. M.Minsky et S. Papert publient ensuite en 1969 un ouvrage qui met en évidence les limitations théoriques du Perceptron. Ces limitations concernent l'impossibilité de traiter des problèmes non linéaires en utilisant ce modèle. Quelques années d'ombre se sont ensuite succédé de 1967 à 1982.

Le renouveau de cette discipline reprend en 1982 grâce à J. J. Hopfield, un physicien reconnu. Il présente une théorie du fonctionnement et des possibilités des réseaux de neurones. Il faut remarquer la présentation anticonformiste de son article. Alors que les

auteurs s'acharnent jusqu'alors à proposer une structure et une loi d'apprentissage, puis à étudier les propriétés émergentes, J. J. Hopfield fixe préalablement le comportement à atteindre par son modèle et construit, à partir de là la structure et la loi d'apprentissage correspondant au résultat escompté. Ce modèle est aujourd'hui encore très utilisé pour des problèmes d'optimisation. On peut citer encore la machine de Boltzmann en 1983 qui était le premier modèle connu, apte à traiter de manière satisfaisante les limitations recensées dans le cas du Perceptron. Mais l'utilisation pratique s'avère difficile, la convergence de l'algorithme étant extrêmement longue (les temps de calcul sont considérables). C'est ensuite qu'en 1985 la rétro-propagation de gradient apparaît. C'est un algorithme d'apprentissage adapté au Perceptron Multi Couches. Sa découverte est réalisée par trois groupes de chercheurs indépendants. Dès cette découverte, nous avons la possibilité de réaliser une fonction non linéaire d'entrée/sortie sur un réseau, en décomposant cette fonction en une suite d'étapes linéairement séparables. Enfin, en 1989 Moody et Darken exploitent quelques résultats de l'interpolation multi variables pour proposer le Réseau à Fonctions de base Radiales (*RFR*), connu sous l'appellation anglophone *Radial Basis Function network (RBF)*. Ce type de réseau se distingue des autres types de réseaux de neurones par sa représentation locale [16].

## 2. Modèle Biologique

Les neurones, au nombre d'une centaine de milliards, sont les cellules de base du système nerveux central. Chaque neurone reçoit des influx nerveux à travers ses dendrites (récepteurs), les intègre pour en former un nouvel influx nerveux qu'il transmet à un neurone voisin par le biais de son axone (émetteur) [20] comme le montre la Figure 2.1 :



**Figure 2.1** Neurone biologique

A partir de ces concepts on va donner une description des fondamentales du réseau de neurones.

## **2.1.Neurone**

Est une cellule vivante, qui peut prendre des formes variables (pyramidale, sphérique ou étoilée) sa forme est définie par une membrane qui sépare l'intérieur du neurone de l'extérieur. Ces neurones sont considérés comme l'élément de base qui constitue les unités élémentaires de traitement dans le cerveau [21].

## **2.2.Structure de neurone**

### **2.2.1.Le corps cellulaire (soma)**

Il est composé d'un noyau qui effectue les transformations biochimiques essentielles à la vie de neurone et se ramifie pour former les dendrites, qui établissent la liaison avec d'autres cellules [22].

### **2.2.2.Axone**

C'est les prolongements unique, qui diffuse le signal du neurone vers d'autres cellules donc son rôle est représenté dans la communication avec d'autre neurone. Il peut diviser à son extrémité pour entrer en contact avec un grand nombre d'autre cellules [22].

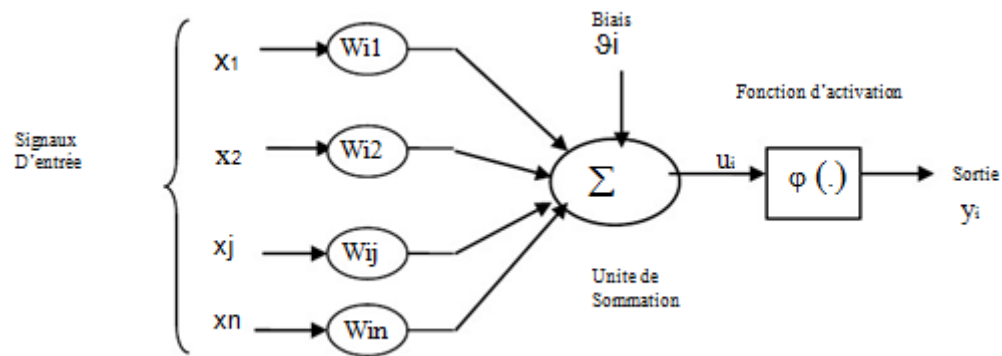
### **2.2.3.Synapse**

C'est un élément de jonction qui assure le contact du cytoplasme (membrane) d'un neurone et les membranes de ses voisins, il joue un rôle essentiel dans la transmission des signaux [22].

## **2.3.Neurone formel**

Est processeur très simple imitant grossièrement la structure et le fonctionnement d'un neurone biologique, la première version du neurone formel est celle de McCulloch et Pitts qui est largement utilisée [23].

Un neurone formel est un opérateur effectuant la somme pondérée de ses entrées externes puis elle s'active suivant la valeur de cette sommation pondérée puis modulée par une fonction d'activation Figure 2.2 [22].



**Figure. 2.2.**Modèle de base d'un neurone formel.

Le modèle de la figure I.2 est composé [19] :

- ✓ Des entrées du neurone formel  $x_i=1,2,\dots, n$  ;
- ✓ Des paramètres de pondération  $w_{ij}$  ;
- ✓ De la fonction d'activation ou de seuillage (non linéaire, sigmoïde, etc....) ;

Et d'une sortie du neurone formel.

### 3. Réseaux de neurones artificiels

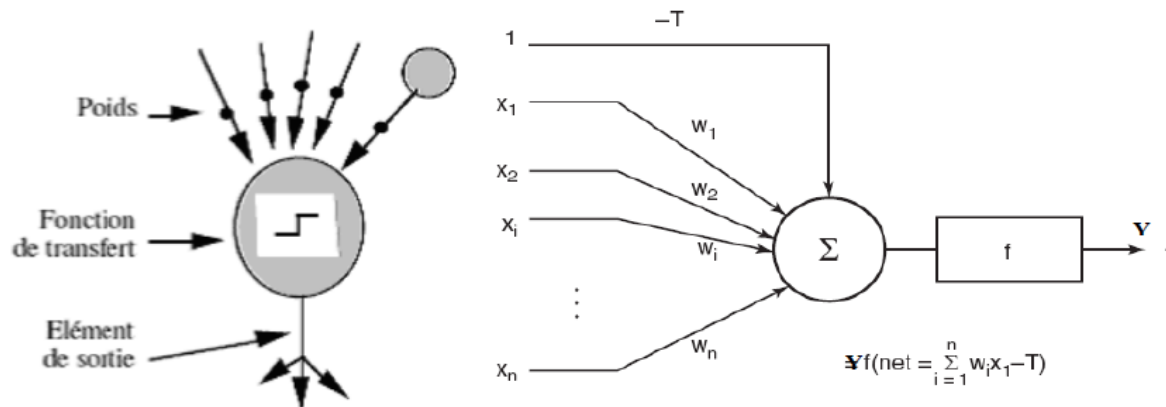
#### 3.1. Définition

Un réseau de neurones (RN) est un système d'opérateurs non linéaires interconnectés, recevant des signaux de l'extérieur par ses entrées, et délivrant des signaux de sortie, ces (RN) sont une métaphore des structures cérébrales et de traitement parallèle et distribué d'information et comportent plusieurs élément de traitement appelé neurone.

Chaque neurone fonctionne indépendamment des autres de telle sorte que l'ensemble est un système parallèle fortement interconnecté. L'information détenue par le réseau de neurone est distribuée à travers l'ensemble des constituants et non localisée dans une partie de mémoire sous la forme d'un symbole.

Le réseau de neurone ne se programme jamais pour réaliser une ou telle tâche, il est entraîné sur des données acquises, grâce à mécanisme d'apprentissage qui agit sur les constituants du réseau afin de réaliser au mieux la tâche souhaitée [22].

On peut comparer la correspondance entre les propriétés respectives de neurones biologiques et neurones artificiels comme le montre la Figure 2.3 [20].



**Figure 2.3:** Correspondance entre neurones biologiques et neurones artificiels

Cette comparaison est montrée dans Tableau 2.1. [24]

système nerveux	réseau neurone Artificiel
Neurone	Traitant élément, noeud, neurone artificiel, neurone abstrait,
Le corps cellulaire (soma)	Niveau de l'activation, fonction de l'activation, fonction du transfert, la fonction de la sortie
Axone	la communication avec d'autre neurone
Synapse	poids multiplicatifs

**Tableau2.1 :** Comparaison entre neurone Biologique et artificiel

### 3.2. Domaine d'application des réseaux de neurones (RNA)

Les (RNA) aujourd'hui ont des applications dans des domaines variés parmi lesquelles on cite :

#### 3.2.1 L'application au traitement du signal

Dans ce domaine les applications réalisées traitent essentiellement de la reconnaissance de signatures radar ou sonner.

La société NESTOR a développé un réseau de neurones qui identifie une cible à coup sur (100% de réussite) et reconnaît du bruit avec un taux de réussite de 95%.

Cette application à été construite à partir de signaux sonars et des traits caractéristiques de ces signaux déterminés par les experts comme utiles pour identifier une cible en environnement bruité. Le réseau agit dans ces comme un filtre pour éliminer le bruit [25].

### 3.2.2 L'application à la planification :

L'optimisation est souvent sollicitée pour la planification des actions ou des tâches. A cet effet, l'utilisation du modèle de Hop Field et du modèle de Kohonen paraît très utile pour la résolution des problèmes d'optimisation [25].

### 3.2.3 L'application au contrôle

Grâce à leur propriétés (parallélisme de traitement, capacité d'adaptation, et de généralisation, etc... ). Les réseaux de neurones sont appliqués au contrôle intelligent. Par exemple plusieurs tentatives ont été faites pour appliquer les réseaux de neurones ; Ces applications peuvent être classifiées en plusieurs méthodes, telles que : le contrôle supervisé. Le contrôle inverse et le contrôle neuronal adaptatif [25].

### 3.2.4 L'application au diagnostic

Les réseaux de neurones sont bien adaptés à la résolution des problèmes de diagnostic, utilisant la classification automatique des signaux et des formes. Dans ce contexte on distingue plusieurs applications des réseaux de neurones pour le diagnostic des défaillances et en particulier, pour le diagnostic des pannes des machines électriques.

Exemples d'applications de chaque modèle : [25]

Caractéristiques fonctionnelles	Type de RNA
Reconnaissance de formes	MLP, Hopfield, Kohonen, PNN
Mémoires associatives	Hopfield, MLP récurrents, Kohonen
Optimisation	Hopfield, ART, CNN
Approximation de fonctions	MLP, RBF
Modélisation et control	MLP, MLP récurrent, FLN
Traitement d'images	CNN, Hopfield
Classification et clustering	MLP, Kohonen, RBF, ART, PNN

**Tableau 2.2.** : Correspondance RNA - domaines d'application

### 3.3. Fonctionnement

Avant de pouvoir utiliser les capacités de classification et d'approximation de fonctions d'un réseau de neurones, il faut le construire, ceci se déroule en quatre temps : [26]

1. La construction de la structure du réseau (généralement empirique).
2. La constitution d'une base de données de vecteurs représentant au mieux le domaine à modéliser. Celle-ci est scindée en deux parties : une partie servant à l'apprentissage du réseau (on parle de base d'apprentissage) et une autre partie aux tests de cet apprentissage (on parle de base de test).
3. Le paramétrage du réseau par apprentissage.

Au cours de l'apprentissage, les vecteurs de données de la base d'apprentissage sont présentés séquentiellement et plusieurs fois au réseau. Un algorithme d'apprentissage ajuste le poids du réseau afin que les vecteurs soient correctement appris. L'apprentissage se termine lorsque l'algorithme atteint un état stable.

4. La phase de reconnaissance qui consiste à présenter au réseau chacun des vecteurs de la base de test. La sortie correspondante est calculée en propageant les vecteurs à travers le réseau, La réponse du réseau est lue directement sur les unités de sortie et comparée à la réponse attendue. Une fois que le réseau présente des performances acceptables, il peut être utilisé pour répondre au besoin qui a été à l'origine de sa construction.

### 4. Modélisation générale

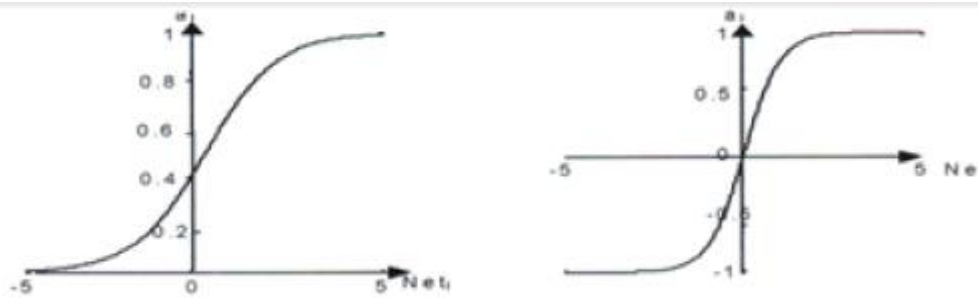
On peut modéliser un réseau de neurone par des élémentaires qu'il s'agit de : [27]

**\*La nature de ses entrées :** qu'ils peuvent être binaire (0 ou 1), (-1,1) ou réelles appartenant souvent à intervalle bornée [a, b].

**\*La fonction des entrées :** que sa signifie qu'elle peut définir le pré traitement effectuée sur les entrées.

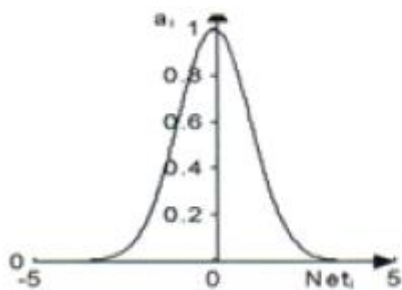
**\*Fonction de sortie :** Cette fonction calcule la sortie du neurone en fonction de son état d'activation.

**\*Fonction d'activation ou de seuillage :** Il existe de nombreuses formes possibles pour la fonction d'activation. Les plus courantes sont présentées sur la figure. 2.4, la plupart des fonctions d'activations sont continues, offrant une infinité de valeurs possibles comprises dans l'intervalle [0, +1] (ou [-1, +1]).

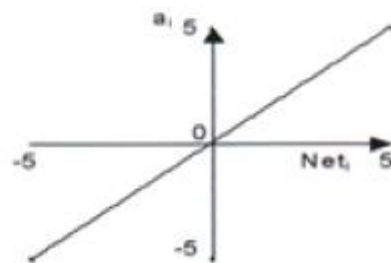


Fonction stochastique ( $T=1$ )

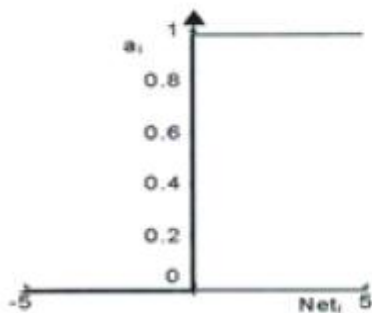
Fonction sigmoïde



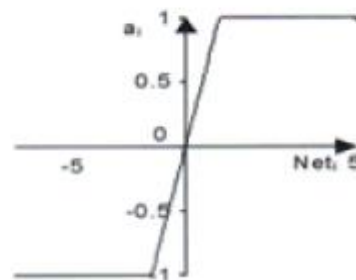
Fonction



Fonction linéaire ( $\lambda = 1$ )



Fonction seuil



Fonction linéaire

**Figure. 2.4.** Différents types de fonctions d'activation pour le neurone formel.

Toutes les fonctions d'activation utilisées doivent être différentiables, car l'architecture des réseaux de neurones l'impose pour que l'apprentissage soit possible.

## 5. Architecture des réseaux

### a- Réseau monocouche

La structure d'un réseau monocouche est telle que des neurones organisés en entrée

soient entièrement connectés à d'autres neurones organisés en sortie par une couche modifiable de poids (figure. 2.5).

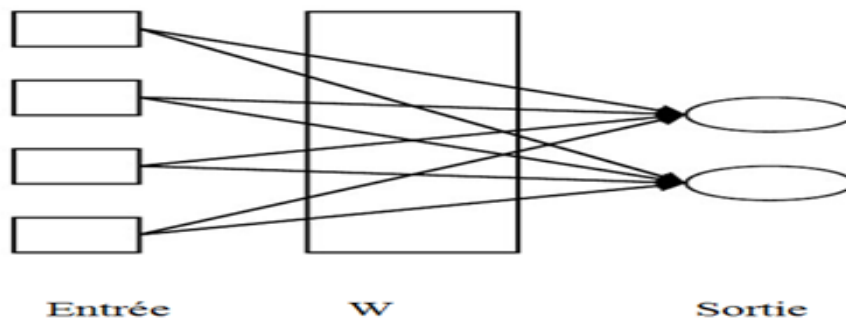


Figure. 2.5. Réseau monocouche

### b- Réseau multicouche

Les neurones sont arrangés par couche, les entrées des neurones de la deuxième couche sont en fait les sorties des neurones de la couche amont, les neurones de la première couche sont reliés au monde extérieur et reçoivent le vecteur d'entrée. Il peut y avoir une ou plusieurs sorties à un réseau de neurone.

Dans un réseau multicouche, il n'y a pas connexion entre neurone d'une même couche et les connexions ne se font qu'avec les neurones de la couche aval, et tous les neurones de la couche amont sont connectés à tous les neurones de la couche aval. On appelle [28] :

**Couche entrée** : contient l'ensemble des neurones d'entrées, cette couche est une couche passive, ses neurones n'effectuent aucun traitement

**Couche de sorties** : contient l'ensemble des neurones de sorties

**Couches cachées** : les couches intermédiaires n'ayant aucun contact avec l'extérieur.

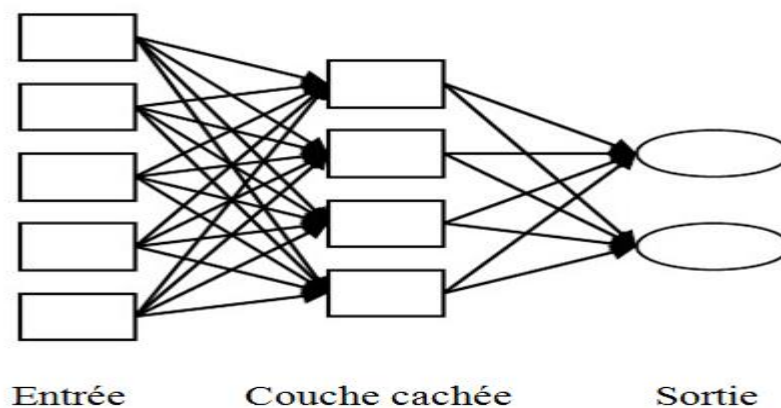
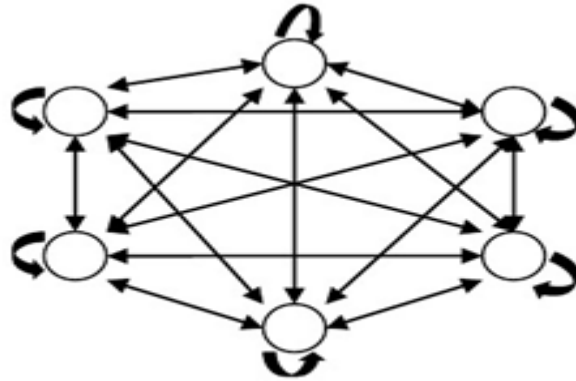


Figure. 2.6. Réseau multicouche.

**c- Réseau à connexion complète**

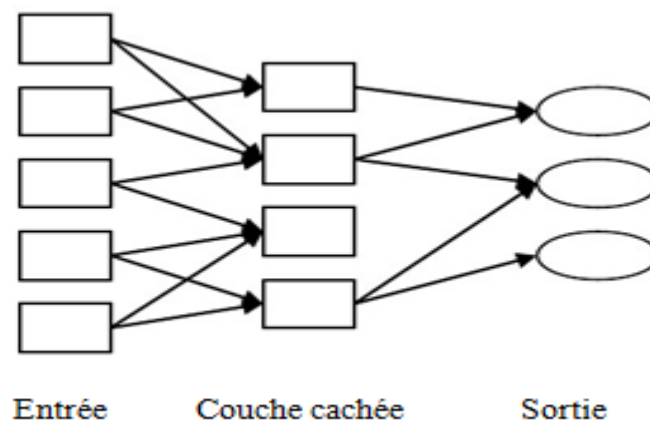
C'est la structure d'interconnexion la plus générale. Chaque neurone est connecté à tous les neurones du réseau (et à lui-même) (figure. 2.7) [28].



**Figure. 2.7.** Réseau à connexion complète.

**d- Réseau à connexions locales**

Il s'agit d'une structure multicouche, mais Chaque neurone entretient des relations avec un nombre réduit et localisé de neurones de la couche avale. Les connexions sont donc moins nombreuses que dans le cas d'un réseau multicouche classique (figure. 2.8) [28].



**Figure. 2.8.** Réseau à connexions locales.

### e- Les réseaux de neurones boucles (récurrents)

Un réseau de neurone boucle a temps discret réalise une ou plusieurs équations aux différences non linéaires, par composition des fonctions réalisées par chacun des neurones et des retards associés à chacune des connexions [29].

Ces réseaux caractérisent par la présence d'au moins une boucle de rétroaction au niveau des neurones ou entre les couches, et la prise en compte de l'aspect temporel du phénomène (figure 2.9). Mais ce sont des modèles plus durs à mettre en œuvre.

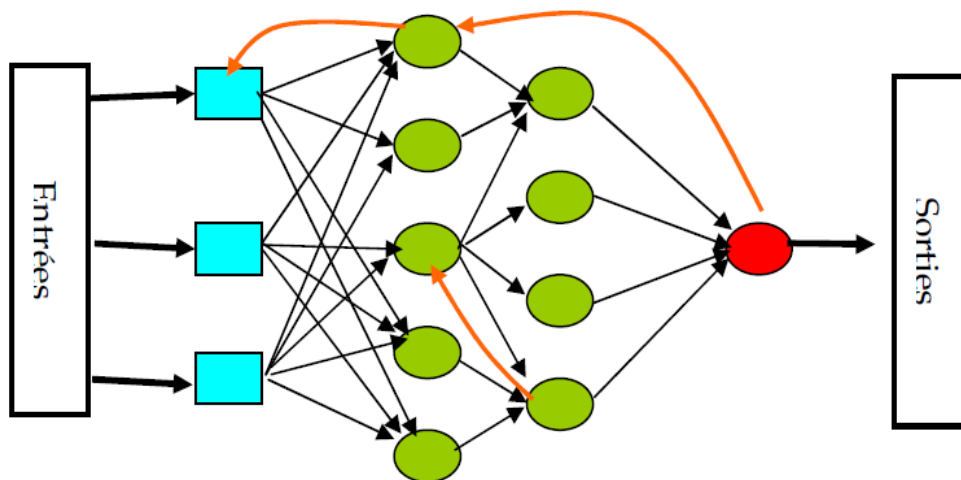


Figure 2.9. Réseau de neurones boucle.

## 6. Les types d'apprentissage des réseaux de neurones

L'apprentissage est vraisemblablement la propriété la plus intéressante des réseaux Neuronaux. Elle ne concerne cependant pas tous les modèles, mais les plus utilisés. L'apprentissage est une phase du développement du réseau de neurones durant laquelle on calcule les poids des neurones de telle manière que les sorties du réseau soient aussi proche que possible des sorties désirées [30].

L'apprentissage RNA est une phase qui permet de déterminer ou de modifier les paramètres du réseau, afin d'adopter un comportement désire.

Les procédures d'apprentissage peuvent se subdiviser, en trois grandes catégories: apprentissage supervise, non supervise.

Cette distinction repose sur la forme des exemples d'apprentissage. Dans le cas de l'apprentissage supervisé, les exemples sont des couples (entrées, sorties associées) alors que l'on ne dispose que de valeurs (entrées) pour l'apprentissage non supervisé.

L'exemple de la figure (2.10) est une partie de l'application de reconnaissance du caractère manuscrit. Il nous donne une idée sur la forme générale d'un apprentissage de réseau de neurones.

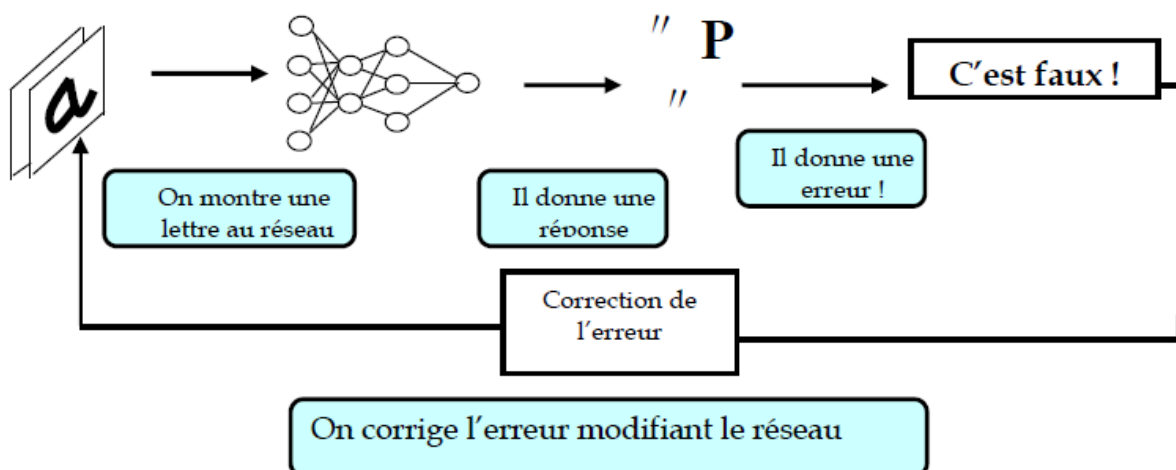


Figure 2.10. Exemple d'apprentissage

Les algorithmes d'apprentissages donnent de meilleurs résultats lorsqu'on leur fournit des exemples multiples et variés ; ainsi le réseau peut assimiler toutes les connaissances. Ils existent différentes règles d'apprentissages parmi lesquelles on peut distinguer [31] :

- La règle de Widrow-Hoff,
- La règle de Hbb,
- La règle de Perceptron,
- La règle de Grossbergen, ect...

### 6.1. L'apprentissage supervisé

L'apprentissage supervisé implique l'existence d'un professeur qui a pour rôle d'évaluer le succès (ou l'échec) du réseau quand il lui est présenté un exemple appartenant à la base d'apprentissage. Cette supervision consiste à renvoyer au réseau une information lui permettant de faire évoluer ses connexions (parfois aussi son propre architecture) afin de faire diminuer son taux d'échec. L'information peut être explicite, sous la forme d'une mesure de l'erreur commise, par exemple, ou globalement sur l'ensemble des exemples de la base. C'est

ce qui se passe dans le cas du Perceptron. Mais elle peut être plus implicite (apprentissage par renforcement), sous forme d'une simple appréciation (bon ou mauvais, punition ou récompense), sans mesure d'erreur, et même être globale, sur l'ensemble des tâches que le réseau doit exécuter. Dans ce dernier cas, il est facile de comprendre que l'apprentissage est le plus difficile, la difficulté majeure consistant pour le réseau à identifier les étapes du processus qui sont responsables de l'échec ou du succès. [29][32].

## 6.2. L'apprentissage non supervisé

L'apprentissage non supervisé implique la fourniture à un réseau autonome d'une quantité suffisante d'exemples contenant des répétitions (autrement dit de la redondance), telles que celui-ci en dégage les régularités automatiquement. Ces réseaux sont souvent appelés auto-organiseurs, ou encore à apprentissage compétitif. Dans l'apprentissage non supervisé, les données ne contiennent pas d'informations sur une sortie désirée, il n'y a pas de superviseur ou expert humain. Il s'agit de déterminer les paramètres du réseau de neurones suivant un critère à définir. Dans ce cas, les exemples présentés à l'entrée provoquent une auto adaptation du réseau afin de produire des valeurs de sortie qui soient proches en réponse pour des valeurs d'entrées similaires. Ce type d'apprentissage possède souvent moins de complexité dans le calcul par rapport à l'apprentissage supervisé. Bien entendu, l'architecture du réseau, préalablement définie par son utilisateur, est une forme de supervision. [29][32].

## 7. Les réseaux de neurones les plus utilisés

Les réseaux de neurones les plus utilisés dans la littérature sont représentés sur la figure (2.11) ci-dessous. Cependant, dans ce qui suit, nous allons faire une description détaillée que des MLP compte tenu que ce sont les réseaux qui vont être appliqués dans ce travail [33].

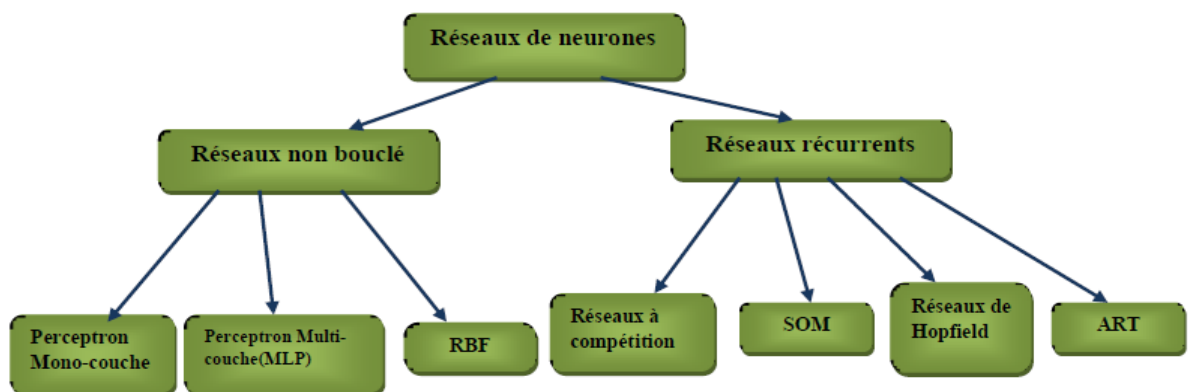


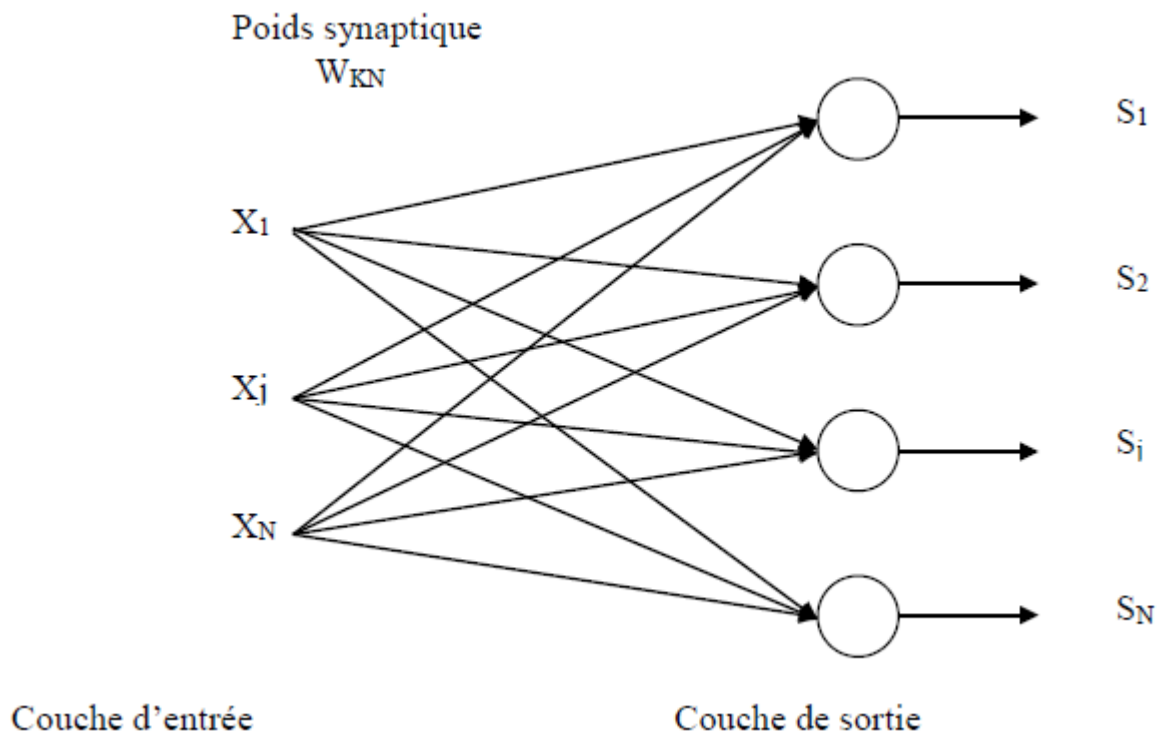
Figure 2.11. Réseaux de neurones les plus utilisés.

### 7.1. Perceptron à une seule couche

C'est le premier modèle qui fut présenté par F.Rosenblatt, dans ce type de réseaux, il y a une seule couche cachée qui lie les couches d'entrée aux couches de sortie. Seule la couche cachée est modifiable. [21]

L'application avec un tel modèles sont très restreintes dans une même structure ayant en commun les mêmes entrées, chaque neurone agit indépendamment des autres et en particulier ne reçoit aucune connexion en provenance des neurones de cette couche et traite des problèmes pour lesquels la sortie attendue n'est pas une seule valeur scalaire mais un vecteur scalaire.

L'exemple de réseau à une seule couche est donné par la Figure 2.12.



**Figure 2.12** : Structure d'une couche de neurones

#### 7.1.1 La convergence du perceptron

Au début de la convergence du Perceptron consiste à initialiser les petites. Valeurs aléatoires non nulle du poids synaptique  $W_i$  et le seuil  $\theta$  si l'erreur est minimale et e réseau converge la sortie est acceptable, si non il fait varie le vecteur d'entrée.

La sortie est donne par [25] :

$$Y = f \left( \sum_{i=0}^{n-1} w_i x_i - \theta \right) \quad (2.1)$$

Avec :

$\theta$ : La seuil.

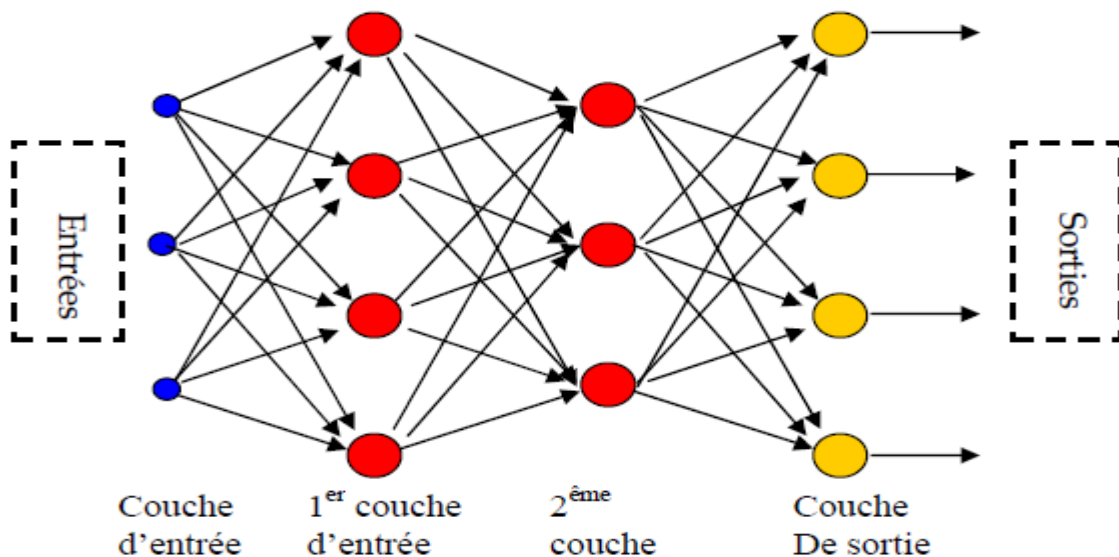
$f$  : fonction d'activation

## 7.2. Perceptron Multi Couches (PMC)

Le perceptron Multi-Couche est un réseau orienté de neurones artificiels organise en couches et ou l'information voyage dans un seul sens, de la couche d'entrée vers la couche de sortie.

La figure 2.13 donne l'exemple d'un réseau contenant une couche d'entrée, deux couches cachées et une couche de sortie. La couche d'entrée représente toujours une couche virtuelle associée aux entrées du système. Elle ne contient aucun neurone. Les couches suivantes sont des couches de neurones. Dans l'exemple illustre, il y a 3 entrées, 4 neurones sur la première couche cachée, trois neurones sur les deuxièmes et quatre neurones sur la couche de sortie. Les sorties des neurones de la dernière couche correspondent toujours aux sorties du système.

Dans le cas général, un perceptron Multi-Couche peut posséder un nombre de couches quelconque et un nombre de neurones (ou d'entrées) par couche également quelconque.



**Figure 2.13.** Exemple de réseau de type perceptron Multi-Couche.

La création d'un perceptron Multicouche pour résoudre un problème donne passe donc par l'inférence de la meilleure application possible telle que définie par un ensemble de données d'apprentissage constituées de paires de vecteurs d'entrées et de sorties désirées. Cette inférence peut se faire, entre autre, par l'algorithme dit de retro propagation. [34].

### 8. Algorithme de retro-propagation du gradient (generalized delta rule)

C'est un algorithme qui permet de stabiliser un réseau à plusieurs couches a été trouvé de façon indépendante par plusieurs équipes de chercheurs (Rumelhart, Parker et Le cun). La topologie d'un tel réseau est donc formée de plusieurs couches de neurones sans communication à l'intérieur d'une même couche.

L'apprentissage y supervisé, c'est a dire que l'on présente air réseau, en même temps, une forme et son modèle.

Le but est encore de minimiser l'erreur quadratique entre les sorties obtenues et celle souhaitées qui correspond de nouveau à une descente du gradient mais alors que dans un réseau à une seule couche on était assuré de la convergence de la méthode, ce n'est plus le cas dans un réseau multicouche, et on pourrait éventuellement rester bloqué sur un minimum local et non global, il s'avère qu'en fait ce la ne se produira pas dans la majorité des cas (mais cela reste expérimental).

Le principe est redistribuer sur toutes les couches, y compris les couches cachées, une partie de l'erreur de manière réursive, en partant de la couche de sortie et en remontant vers la couche de entrée [25].

### 9. Formalisation d'apprentissage

Pour un réseau multicouche à « m » entrées et « n » sorties, composé de L couches (couches cachées et couche de sortie), la somme de la  $j^{ème}$  unité cachée est : [36] [21]

$$S_j^L = \sum_{i=1}^n w_{ji}^L x_i + \theta_j^L \quad (2.2)$$

L : indice de la couche cachée.

W : poids de la connexion  $i^{ème}$  entrée.

$\theta_j^L$  : biais ou la valeur de seuil interne du neurone

f : est la fonction d'activation de ce neurone j tel que

$$I_j = f^L(S_j^L) = \sum_{j=1}^n w_{ji}^L x_1 + \theta_j^L \quad (2.3)$$

Les équations de la couche de sortie sont :

$$S_K^R = \sum_{j=1}^n w_{Kji}^R I_j + \theta_K^R \quad (2.4)$$

$$R_K = f_K^R(S_K^R) \quad (2.5)$$

$K$  : numéro de neurone de la couche de sortie.

L'objectif de rétro-propagation est l'adaptation des paramètres  $W_{ij}$  de telle façon qu'on minimise la fonction de coût donnée par :

$$E_P = \frac{1}{2} \sum_{K=1}^T E = \frac{1}{2} \sum_{K=1}^T (\delta)^2 \quad (2.6)$$

Avec :

$$\delta_K = Y_K - R_K \quad (2.7)$$

Où :

$Y_K$  : la sortie désirée

$R_K$  : la sortie actuelle du réseau

$T$  : longueur de l'ensemble d'apprentissage

### 9.1 Adaptation des poids

Après avoir calculé la sortie  $R_K$  et l'erreur  $E$  correspondante à l'ensemble des entrées à partir des équations (2.5) et (2.6), les poids du réseau sont alors ajustés par la méthode du gradient :

$$w_{ji}^L(n+1) = w_{ji}^L(n) + \Delta w_{ji}^L(n) \quad (2.8)$$

$$\Delta w_{ji}^L(n) = -\mu \frac{\partial E}{\partial w_{ji}^L(n)} \quad (2.9)$$

Avec :

$N$  : numéro de l'itération

$M$  : pas d'apprentissage représentant la vitesse de convergence, sa valeur est généralement choisie expérimentalement ( $0 < \mu < 1$ )

-Si  $\mu$  est trop petit, la convergence est lente mais la direction de la descente est optimale.

-Si  $\mu$  est trop grand, la convergence est rapide mais la précision est médiocre.

L'application des poids du réseau est faite tout d'abord pour la couche de sortie puis pour les couches cachées on a :

**a) pour la couche de sortie**

D'après (II-6)

$$E_P = \frac{1}{2} \sum_{K=1}^T (Y_K - R_K)^2 \quad (2.10)$$

La dérive de l'erreur  $E_P$  par rapport au poids synaptique  $W_k^R$  nous donne :

$$\frac{\partial E_P}{\partial w_{ji}^L} = -(Y_K - R_K) \frac{\partial R_k^L}{\partial S_k^L} \frac{\partial S_k^L}{\partial w_{Kj}^L} \quad (2.11)$$

Le dernier facteur de (II.11) est :

$$\frac{\partial S_k^R}{\partial w_{Kji}^R} = \frac{\partial}{\partial w_{Kj}^L} [w_{Kj}^R I_j + \theta_k^L] = I_j \quad (2.12)$$

En combinant (II-10) et (II-11) ; on a pour l'opposé du gradient :

$$\frac{-\partial E_P}{\partial w_{Kj}^R} = (Y_K - R_K) f_k^R(S_k^R) I_j \quad (2.13)$$

L'amplitude de la variation du poids étant proportionnelle a l'opposé du gradient, les poids de couche de sortie son renouvelés selon :

$$w_{Kj}^R(n+1) = w_{Kj}^R(n) + \Delta w_{Kj}^R(n) \quad (2.14)$$

$$\Delta w_{Kj}^R = \mu (Y_K - R_K) f_k^R(S_k^R) I_j \quad (2.15)$$

Il y a deux formes de fonctions de sorties qui nous intéressent ici :

- $f_k^R(S_k^R) = S_k^R$
- $f_k^R(S_k^R) = [1 + \exp(-S_k^R)]^{-1}$

La première est une fonction linéaire. La deuxième appelée sigmoïde. Pour la fonction linéaire la dérivée est :  $f_k^R = 1$

L'équation (II.14) devient :

$$w_{Kj}^R(n+1) = w_{Kj}^R(n) + \mu(Y_K - R_K)f_k^R(S_k^R)I_j \quad (2.16)$$

Pour une fonction d'activation type sigmoïde, la dérivée est :

$$f_k^R = f_k^R(1 - f_k^R) = R_K(1 - R_K) \quad (2.17)$$

L'équation (II.14) devient :

$$w_{Kj}^R(n+1) = w_{Kj}^R(n) + \mu(Y_K - R_K)(1 - R_K)R_K I_j \quad (2.18)$$

En définissant la quantité :

$$S_k^R = (Y_K - R_K)f_k^R(S_k^R) \quad (2.19)$$

Nous pouvons alors écrire l'équation d'adaptation des poids sous une forme indépendante de la fonction de sortie :

$$w_{Kj}^R(n+1) = w_{Kj}^R(n) + \mu \delta_k^R I_j \quad (2.20)$$

### b) pour les couches cachées

Nous devons répéter le même calcul que pour la couche de sortie. L'erreur total doit être rapporté d'une façon ou d'une autre aux sorties de la couche cachée :

$$E_P = \frac{1}{2} \sum_{K=1}^T (Y_K - R_K)^2 \quad (2.21)$$

$$= \frac{1}{2} \sum_{K=1}^T [Y_K - f_k^R(S_k^R)]^2 \quad (2.22)$$

$$= \frac{1}{2} \sum_{K=1}^T (Y_K - f_k^R(\sum w_{kj} I_j + \theta_k^R))^2 \quad (2.23)$$

Dans les équations (2-2) et (2-3), on constate que  $I_j$  dépend des poids de la couche cachée, nous pouvons exploiter ce fait pour calculer le gradient de  $E_P$  par rapport aux poids de la couche cachée :

$$\frac{\partial E_P}{\partial w_{ji}^L} = \frac{1}{2} \sum_k \frac{\partial}{\partial w_{ji}^k} (Y_K - R_K)^2 \quad (2.24)$$

$$= - \sum_K (Y_K - R_K) \frac{\partial R_K}{\partial S_K^R} \frac{\partial S_K^R}{\partial I_j} \frac{\partial I_j}{\partial S_j^L} \frac{\partial S_j^L}{\partial w_{ji}^L} \quad (2.25)$$

Les facteurs de l'équation (2-20) peuvent être calculés à partir des équations précédentes pour donner :

$$\frac{\partial E_P}{\partial w_{ji}^L} = - \sum_K (Y_K - R_K) f_k^R(S_k^R) W_{Kj}^R f_j^L(S_j^L) X_j \quad (2.26)$$

Les poids des couches cachées sont adaptés proportionnellement à l'opposé de l'équation (2.26).

$$\Delta W_{Kj}^R = \mu f_j^L(S_j^L) X_j \sum_K (Y_K - R_K) f_k^R(\delta_k^R) W_{Kj}^R \quad (2.27)$$

Si on utilise la définition de, l'équation (2-27) devient :

$$\Delta W_{Kj}^R = \mu f_j^L(S_j^L) X_j \sum_K (\delta_k^R) W_{Kj}^R \quad (2.28) = \mu \delta_j^L X_j$$

Avec :

$$\delta_j^L = f_j^L(S_j^L) X_j \sum_K \delta_k^R W_{Kj}^R \quad (2.29)$$

On consulte que chaque adaptation des poids dans la couche cachée dépend de l'erreur totale de la couche de sortie ce qui conduit à la notion de rétro propagation.

L'équation d'adaptation des poids dans ce cas est :

$$w_{ji}^L(n+1) = w_{ji}^L(n) + \mu \delta_j^L X_j \quad (2.30)$$

## 9.2 Algorithme de la rétro-propagation

**Etape 1 :** Initialiser les poids  $W_{ij}$  et les seuils internes des neurones à des petites valeurs aléatoires.

**Etape 2 :** Présenter le vecteur d'entrée et de sortie désirée.

**Etape 3 :** Calculer :

- La somme des entrées des neurones de la couche cachée en utilisant l'expression (2.1) - Les sorties de neurones de la couche cachée en utilisant l'expression (2.2)
- La somme des entrées des neurones de la couche de sortie en utilisant l'expression (2.3)
- Les sorties des réseaux en utilisant l'expression (2.4).

**Etape4 :** Calculer l'erreur pour les neurones de la couche de sortie en utilisant l'expression (2.19).

**Etape5 :** Réinjecter l'erreur de sortie en utilisant l'expression (2.29)

**Etape6 :** Ajuster :

- Les poids de la couche de sortie en utilisant l'expression (2.20)
- Les poids de la couche cachée en utilisant l'expression (2.30)

**Etape 7 :** Calculer E l'erreur en utilisant l'expression (2.10)

**Etape8 :** Si la condition sur l'erreur  $E - E_p < \epsilon$  est atteinte, aller a l'étape 9 sinon aller à L'étape 6 et refaire le calcul pour un autre époque.

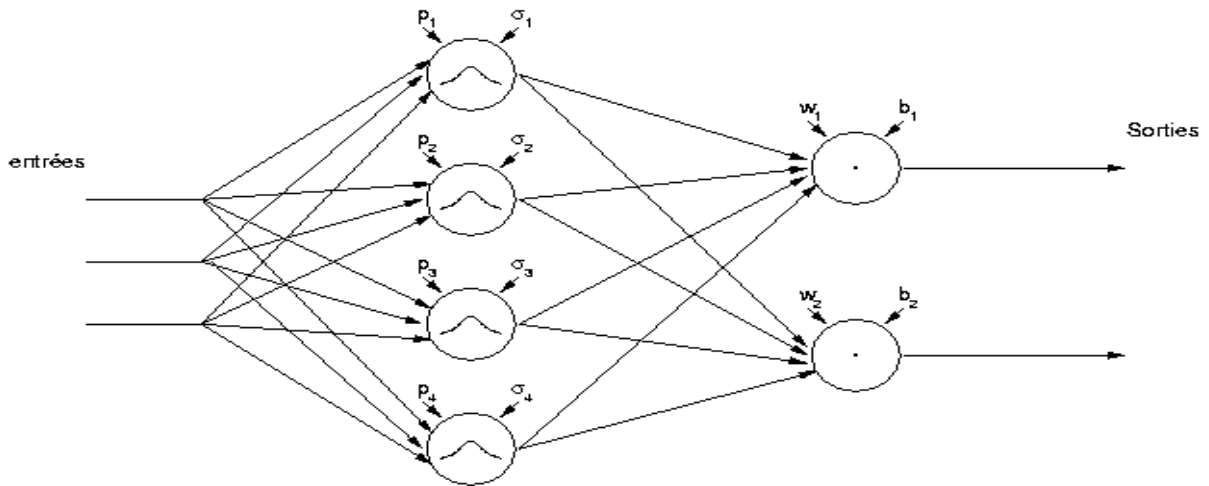
**Etape 9 :** FIN.

## 10 Réseau de neurones de type RBF(Radial Basis Functions)

Les réseaux à fonction radiales de base (RBF) sont des modèles connexionnistes simples à mettre en œuvre et assez intelligible, et sont très utilisés pour la classification. Leur propriétés théoriques ont été étudiées en détail depuis la fin des années 80 ; il s'agit certainement, avec le perceptron multicouche, du modèle connexionniste le mieux connu [3].

### 10.1 Architecture

Introduit par Powell et Broomhead, le réseau RBF (Radial Basis Functions) fait partie des réseaux de neurones supervisés. Il est constitué de trois couches (figure II.14): une couche d'entrée qui retransmet les entrées sans distorsion, une seule couche cachée qui contient les neurones RBF qui sont généralement des gaussiennes et une couche de sortie dont les neurones sont généralement animés par une fonction d'activation linéaire. Chaque couche est complètement connectée à la suivante et il n'y a pas de connexions à l'intérieur d'une même couche [3].



**Figure 2.14** Présentation schématique d'un réseau RBF

Ce réseau est constitué de N neurones d'entrée, M neurones cachés et J neurones de sortie. La sortie du m<sup>ième</sup> neurone de la couche cachée est donnée par [3] :

$$y_m^{(q)} = \exp \left[ - \frac{\|x^{(q)} - v_m\|^2}{(2\sigma_m^2)} \right] \quad (2.31)$$

$v_m$  est le centre du m<sup>ième</sup> neurone de la couche cachée où du m<sup>ième</sup> neurone gaussien et  $\sigma_m$  est la largeur du m<sup>ième</sup> gaussienne. La sortie du j<sup>ième</sup> neurone de la couche de sortie est donnée par:

$$z_j^{(q)} = \left( \frac{1}{M} \right) \left[ \sum_{(m=1, M)} w_{mj} y_m^{(q)} \right] \quad (2.32)$$

$m = 1, \dots, M$  et  $j = 1, \dots, J$ .

$w_{mj}$  sont les poids reliant la couche cachée à celle de la sortie.

### II.10. 2. Algorithme d'apprentissage du réseau RBF

L'apprentissage du réseau RBF a été présenté la première fois par Moody et Darken. Il consiste à régler quatre paramètres principaux: le nombre de neurones dans l'unique couche cachée ou le nombre des gaussiennes, la position des centres de ces gaussiennes, la largeur de ces gaussiennes et les poids de connexions entre les neurones cachés et le(s) neurone(s) de sortie. Le réseau RBF consiste à minimiser l'erreur quadratique totale E calculée entre les sorties obtenues du réseau et celles désirées [37] :

$$E = \sum_{q=1}^Q \sum_{j=1}^J t_j^{(q)} z_j^{(q)2} \quad (2.33)$$

Pour le réseau RBF, l'ajustement des poids  $w_{mj}$  reliant la couche cachée à celle de la sortie est réalisé par la règle de Widrow-Hoff. Il se fait comme suit :

$w_{mj}^{(i+1)} = w_{mj}^{(i)} + \eta(t_j - z_j)y_m$  (2.34)  $t_j$  est la sortie du  $j^{\text{ième}}$  neurone désirée,  $z_j$  est la sortie du  $j^{\text{ième}}$  neurone calculée,  $y_m$  est la sortie du  $m^{\text{ième}}$  neurone de la couche cachée et  $\eta$  est le pas d'apprentissage dont sa valeur est comprise entre 0 et 1.

En conclusion, mentionnons que la principale difficulté des réseaux RBF concerne la question du nombre de neurones radiaux à utiliser pour une application donnée. A priori, il n'existe pas de méthode pour fixer leur nombre, et cette architecture souffre de façon particulièrement aigüe de ce qu'on appelle la «malédiction de la dimension», à savoir l'augmentation exponentielle du nombre de neurones cachés requis en fonction de la dimension  $R$  de l'espace d'entrée. Lorsque  $R$  est grand, une façon d'atténuer ce problème consiste à remplacer les hyper-sphères qui résultent de l'imposition d'une variance fixe par des hyper-ellipses où la matrice de covariance n'est plus contrainte. On peut ainsi réduire le nombre de neurones à positionner au détriment du nombre de paramètres à estimer.

## 11. Les différentes applications de RNAs

Les réseaux de neurones sont utilisés dans de plus en plus de domaines, citons la classification, la reconnaissance de formes, l'identification et la commande de processus. Le choix d'utiliser tel ou tel type de réseau de neurones dépend de l'application mais aussi des capacités de traitement du processeur sur lequel ils s'exécutent.

### 11.1 Reconnaissance de formes

C'est un domaine privilégié d'application pour les RNAs et c'est lui qui a marqué leur début. Le terme de reconnaissance est un terme général qui désigne les processus traitant des données pour en extraire des informations afin de parvenir à leur classification. Une partie non négligeable des applications neuronales actuelles appartient à cette catégorie, parmi les quelles nous pouvons citer [36] :

- Reconnaissance de caractères Latins, Chinois, Coréens, Japonais, Russes ;
- Reconnaissance de visages ;
- Classification d'images et de documents ;
- Diagnostic des maladies ;

· Diagnostic des pannes dans les installations électriques, nucléaires, mécanique, ainsi que dans les systèmes de télécommunications.

## 11.2 Traitement de la parole

Le traitement des signaux stochastiques (filtrage, classification) provenant des systèmes non linéaires est une tâche difficile avec les méthodes classiques de traitement.

En effet, la résolution de ce type de problèmes implique la capacité de modéliser d'une façon adéquate, le système inconnu afin d'estimer sa sortie, ce qui n'est possible que pour des cas simples. La capacité d'apprentissage des réseaux de neurones leur permet d'extraire les propriétés statistiques des systèmes non linéaires, ce qui les rend aptes à traiter ce type de problèmes. Parmi les applications réalisées dans ce domaine nous pouvons citer [36] :

- Filtrage et suppression de bruits : amélioration du rapport signal/bruit.
- Analyse de signaux;
- Diagnostic de pannes;
- Démodulation Multi-Utilisateurs;
- Classification des cibles radars dans les avions ;
- Classification des bruits;

## 11.3 Détection d'anomalies

Ceci est une dérivation des reconnaissances de formes. On apprend à un réseau une image du fonctionnement (normal) d'un système et celui-ci sera ainsi capable d'indiquer tout état de dysfonctionnement quand certains paramètres engendrent une image (anormale). On trouve ce genre de système pour la surveillance de disjoncteurs à très haute tension (dans ce cas précis on écoute les bruits émis par le disjoncteur) ou encore dans la détection de colis piégés (examen du rayonnement gamma induit dans divers matériaux) [34].

## 11.4 Traitements dépendant du temps

Dans ce cas-ci on se retrouve devant des problèmes de prédiction et, d'identification et de commande de processus.

- *Prédiction* : Ici on va surtout utiliser des réseaux récurrents. Beaucoup de travaux sont consacrés à la prédiction de données financières et boursières et quelques résultats semblent devoir entretenir un optimisme mesuré.

- *Identification et commande de processus.* : On retrouve ce genre d'applications dans l'aide au pilotage de systèmes spécifiques tel que les réacteurs chimiques, une automobile sans pilote etc. Ces applications demandent peu de ressources car les temps caractéristiques sont longs.

Nous pouvons citer également les applications bancaires et financières, le traitement de la parole et la robotique [34].

### **Conclusion**

Les réseaux de neurones artificiels, sont le cœur de plusieurs applications réelles, un outil indispensable qui a fait ses preuves dans la pratique par leurs caractéristiques de généralisation et de robustesses face au bruit qui font rêver les chercheurs qui veulent avoir le parfait. Alors ces réseaux-là ne sont en fait qu'un :

- Ensemble de neurones formels inspiré du neurone biologique.
- Relié entre eux par des synapses contenant des poids.
- Utilisant un nombre d'algorithmes d'apprentissage afin de réaliser une tâche généralement la classification, prédiction, approximation.

Nous nous sommes concentrés principalement dans ce chapitre sur une architecture neuronale qui est concernée par notre travail de recherche et qui est généralement la plus utilisée pour la classification l'eau potable: le perceptron multi-couches.

## *CHAPITRE III*

# **SIMULATION**

### **Introduction**

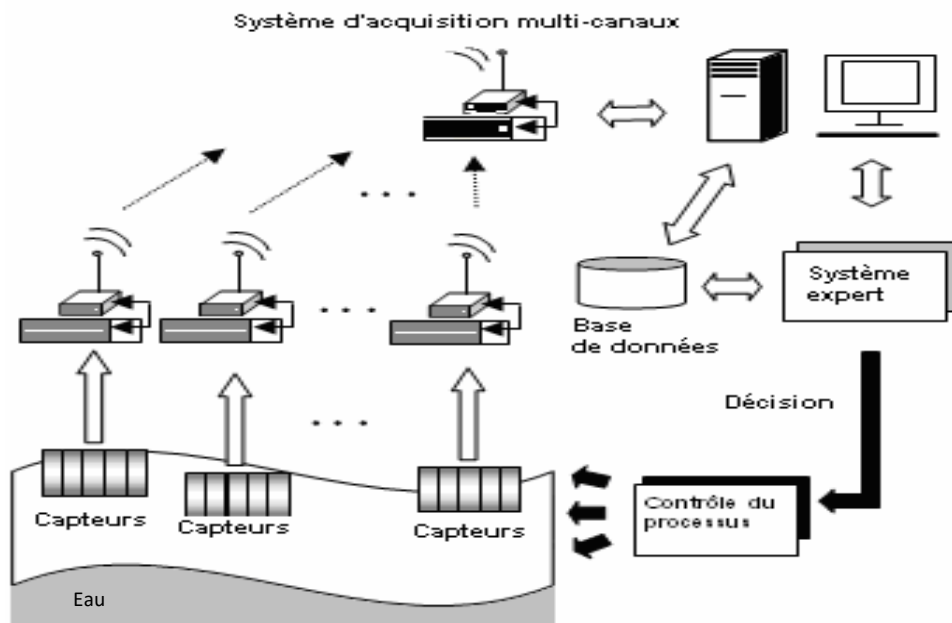
Ce dernier chapitre est consacré à la simulation et vise l'application des techniques étudiées comme étant une solution dans la surveillance de la qualité de l'eau par reconnaissance de formes. L'objectif est de valider et d'évaluer les performances de chacune des méthodes présentées à savoir les réseaux de neurones artificiels (RNA) et à base radiale (neuronal-RBF). Les exigences principales d'efficacité sont formulées sur deux points essentiels à savoir, les tests de spécification qui vérifient que le programme réalise bien la tâche pour laquelle il a été conçu, et les tests de performances qui vont servir à mesurer l'efficacité avec laquelle cette tâche est remplie. Afin de mener une étude comparative permettant un choix décisif de la méthode la mieux adaptée à l'application indiquée, on évaluera pour les méthodes exposées les paramètres liés au taux de reconnaissance et temps d'apprentissage. Une discussion des résultats conclura cette étude de simulation pour choisir la technique la mieux adaptée.

### **1. Problématique**

#### **1.1 Architecture du système de contrôle et de surveillance**

Il s'agit dans cette partie de travail d'évaluer les performances des deux techniques étudiées précédemment qui sont issues, rappelons-le, du domaine de l'intelligence artificielle à savoir, les réseaux neuronal RNA et les réseaux neuronal-RBF. Des techniques servant comme outils de base pour l'aide à la décision et présentant une réponse plus élaborée par rapport aux autres techniques se basant sur des données brutes, venant directement des variables de surveillance, ou à partir de données traitées venant des sorties de traitements de bas niveau. Le choix effectué sur la base des résultats obtenus, conduira à l'intégration de la technique sélectionnée au niveau d'un système de surveillance assurant un contrôle permanent de la qualité de l'eau. L'architecture de ce système imaginé est basée sur une approche multi- sensorielle et présentée dans la figure 3.1. Le processus de

contrôle est vu comme un problème de reconnaissance de formes, où les classes correspondent aux différents états de l'eau, et les formes représentent l'ensemble des observations ou mesures des paramètres liés à ses caractéristiques.



**Figure.3.1.** Architecture du système de contrôle et de surveillance.

Au niveau du système, on peut supposer que les différents paramètres physico-chimiques utilisées, tels que le pH, la température ( $T^\circ$ ), la conductivité (C), la turbidité (TU), l'oxygène dissous (OD), etc. sont transformés en signaux électriques à partir des capteurs, et transmis vers une station de contrôle qui assure l'acquisition, le traitement et l'analyse. La technique de surveillance utilisée effectue après chaque acquisition, la classification et la séparation des données en plusieurs classes bien différentes. Une suite d'acquisitions pourrait être envisagée plusieurs fois par jour, sous des conditions prédéfinies. Un module d'apprentissage supervisé par un expert, permet de collecter de manière continue les paramètres relatifs aux différents états de l'eau pour la mise en œuvre d'une base de connaissance complète.

## 1. 2 Approche utilisée dans la surveillance

La solution devant être adoptée par les techniques citées ci-dessus au problème de reconnaissance de formes posé, ne s'applique en fait que si on se trouve dans le cas d'un apprentissage supervisé. Nous procédons donc lors d'une étape préliminaire d'apprentissage, à paramétrer le classificateur pour la reconnaissance. L'étape de test ou de

reconnaissance proprement dite, s'effectue une fois le modèle statistique établi. Il y a ici tout l'intérêt pour dire que cette approche se caractérise par sa souplesse et sa généricité. A souligner toutefois que les méthodes de reconnaissance de formes à base d'apprentissage statistique sont les plus utilisées dans les systèmes de classification à fusion multi sensorielle. En général l'apprentissage est une étape assez longue, et nécessite plus de temps de calcul. Les techniques partagent ce point commun mais diffèrent sur un certain nombre d'autres points. L'étude comparative effectuée dans les paragraphes suivants en fera la différence. Ce critère (temps d'apprentissage) aussi important dans le choix du modèle de reconnaissance, évoque un traitement hors ligne devant être effectué par le système de surveillance. Le déroulement de cette opération en permanence contribue sans doute à enrichir une base de connaissance qu'on veut qu'elle soit la plus complète possible pour le modèle de surveillance implanté. Le système de contrôle doit donc pouvoir marier à la fois une surveillance directe de l'eau et un apprentissage en arrière-plan (en différé). Un opérateur (ou système) expert supervisant cet apprentissage permet de collecter de manière continue les paramètres relatifs aux différents états de l'eau.

### 1.3. Application à la station de production d'eau potable Tilesdit

#### 1.3.1. Le site TILES-DIT (Zone d'étude)

Le barrage « Tilesdit » est situé géographiquement dans la commune de Bechloul à 20 km du Sud-Est de la wilaya de Bouira, Algérie. Ce barrage est situé entre les coordonnées cartographiques et les coordonnées Lambert suivantes (figure 3.2) :

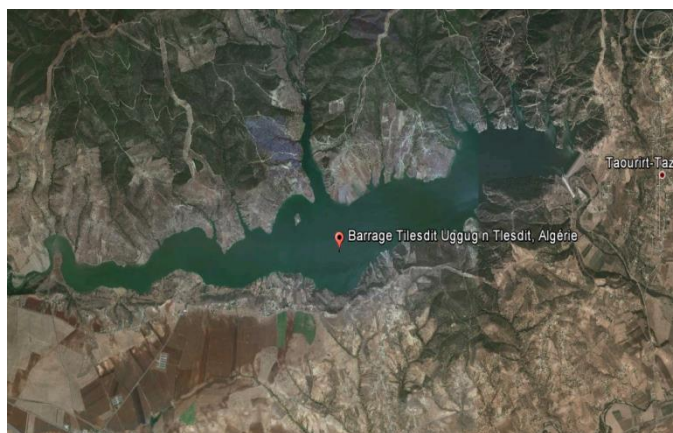
- Latitude : 35° 13' 22'' Nord.

- Longitude : 4° 14' 23'' Est.



**Figure 3.2.** Carte géographique situant le barrage « Tilesdit » [Google Maps].

Ce barrage disposant d'un volume de 167 million de mètres cubes d'eau, est conçu de façon à juguler la tension qui persiste dans la distribution d'eau au niveau de 12 communes (figure 3.3). De part son implantation dans la région de Bouira, le barrage Tilesdit dont la mise en eau à été effectuée vers la fin de l'année 2005, garantira de l'eau potable aux communes relevant de l'Est et du Sud-Est de la wilaya, c.-à-d: El-Asnam, Bechloul, El-Adjiba, Ahl-Ksour, Oud El-Berdi, Mesdour, Taguedit, Ahnif, Bordj O'khris, Ouled rached, Ath Mansour, Bouira et la zone industrielle de Sidi Khaled. Le transfert d'eau dont le lancement des travaux de réalisation à été prévu au début 2011, est destinée à l'alimentation de cinq autres communes de la daïra de Mansourah, dans la wilaya de Bordj Bou Arréridj. Des travaux sont en cours en vue de raccorder les communes de Takdit et Ait Laaziz, ainsi que d'autres communes rurales du Sud-Est de la wilaya de Bouira au réseau d'alimentation de ce barrage, qui devrait s'étendre jusqu'à Hammam k'sana. Parallèlement, la même direction a prévu 12 opérations portant sur la réalisation de réseaux AEP, en plus de 12 réservoirs d'une capacité globale de 74000 m<sup>3</sup> d'eau en vue de l'amélioration de l'approvisionnement en eau des communes alimentées par ce barrage notamment avec un débit de 72650 m<sup>3</sup>/jour. Il est également important de signaler qu'à l'horizon 2020, ce barrage garantira l'eau en faveur de 307200 habitants, selon les prévisions du secteur de l'hydraulique. Les travaux menés ont porté à 6 le nombre de stations de pompage pour un débit de 41 à 981 l/s. Une station de traitement d'une capacité de 74000 m<sup>3</sup>/jour, un réservoir d'eau traitée de deux compartiments d'un volume total de 13000 m<sup>3</sup> et de 6 réservoirs de capacité allant de 200 à 5000 m<sup>3</sup>.



**Figure 3.3.** Image représentant le site du barrage « Tilesdit » [Google Earth].

### 1.3.2. La station de production d'eau potable TILESDDIT

L'eau prélevée dans le barrage est pompée jusqu'à la station de traitement. Celle-ci se trouvant au même lieu, est mise en service depuis 2009 (figure 3.4). Elle effectue le processus d'épuration à travers les cinq étapes de traitement indiquées au premier chapitre à savoir : le prétraitement, la pré-oxydation, la clarification, la désinfection et l'affinage. L'étape de clarification est assurée par le procédé de coagulation-floculation, décantation et filtration, grâce à un décanteur et un étage de filtration sur sable.



**Figure 3.4.** Image représentant le site de la station « Tilesdit » [Google Earth].

## 2. Description de la base de données

Nous cherchons à décider sur la qualité de l'eau à travers ses paramètres descripteurs. Nous n'avons en fait aucune connaissance a priori sur un type de modèle représentant parfaitement ce procédé, par contre nous pouvons porter notre jugement sur la qualité de cette eau à partir de quelques données descriptives. Il y a huit (08) paramètres physico-chimiques qui sont souvent utilisés dans plusieurs travaux, et qui renseignent sur les dangers majeurs qu'il faut surveiller. Ces paramètres sont collectés à partir de la station Tilesdit et qui sont résumés comme suit : pH, Conductivité électrique (EC), Température (T°), Turbidité (TU), Magnésium (Mg), Bicarbonate (B), Dureté Permanente (DP), Titre Alcalin Complet (TAC).

L'objectif qui se trouve derrière la collecte des données relatives à ces paramètres est de trouver un modèle de classification permettant de distinguer deux états bien distincts de l'eau (état potable et état non potable) dans un premier temps, et trois états différents par la suite après sélection de la technique adoptée. La qualité de cette eau reflétée par sa potabilité repose en fait sur une corrélation qui ne peut être identifiée que statistiquement. Des données descriptives expérimentales recueillies sur une longue période (plusieurs années) pourraient atteindre cet objectif. A noter que le pH, la Conductivité, la

Température, la Turbidité, le Magnésium sont fortement dépendants des phénomènes saisonniers. Il y a donc intérêt de disposer d'au moins une année pour archiver des données afin de déterminer une base de connaissance assez complète capable de fonctionner normalement. D'où la nécessité d'une base de connaissance riche en informations exigeant d'abord une collecte des données sur une longue période, et la présence d'un expert et qui est notre cas dans la zone d'étude.

Dans un but de validation, une base de données réelle est utilisée. Les différents paramètres descripteurs de l'état de l'eau sont organisés selon les normes recommandées par les pouvoirs publics. Une classification multi-classe comportant trois classes de potabilité différentes (classe I, II et III) est appliquée sur les données réelles collectées à partir de la station de notre zone d'étude.

### **2.1. Données d'entrée**

Nous cherchons à appliquer notre approche de surveillance aux paramètres descripteurs de la qualité de l'eau brute fournis par les capteurs de mesure de la station. Notre connaissance du processus de traitement est limitée aux données enregistrées de la station durant les trois années écoulées 2009-2011. Ces mesures proviennent des différentes étapes de traitement, on y trouve :

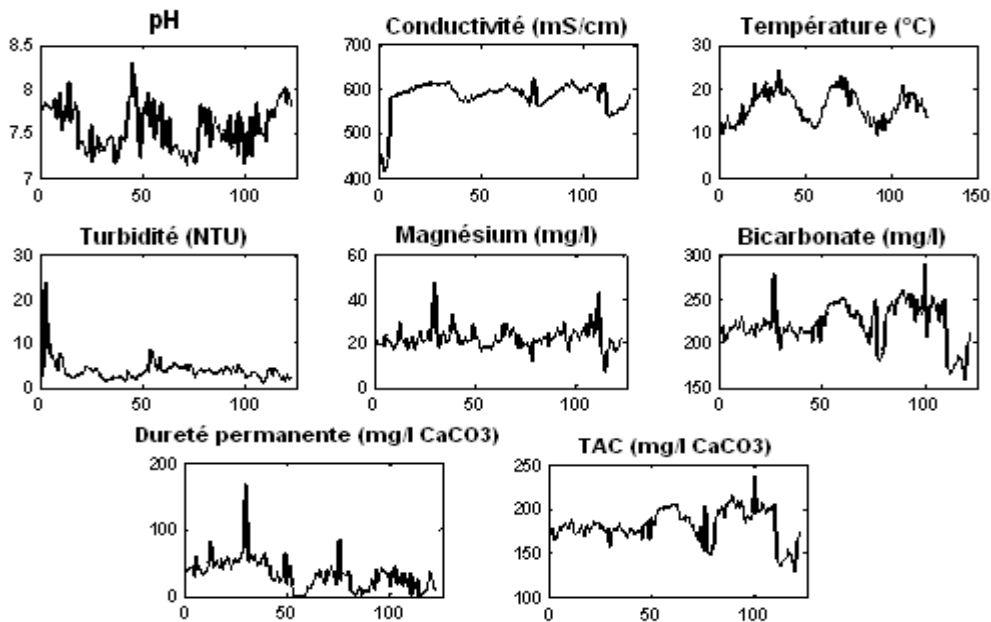
- des mesures en continu issues de capteurs physico-chimiques,
- des analyses faites au laboratoire.

Plusieurs paramètres descripteurs de la qualité de l'eau brute mesurés en ligne quotidiennement à raison de 3 fois/jour, en plus des essais effectués au laboratoire qui sont réalisés chaque semaine. Quatre paramètres descripteurs principaux sont mesurés directement des capteurs vers la station sont : Température, pH, Conductivité et Turbidité. Ces paramètres sont mesurés en continu et à tout niveau du processus de traitement (Eau brute, Eau décantée, Eau filtrée et Eau traitée). D'autres sont aussi mesurés quotidiennement tels que l'Ammonium et le Nitrite. Les paramètres tels que : Calcium, Magnésium, Chlorure, Sulfate, Bicarbonate, Dureté Total TH, Dureté Permanente, Titre Alcalin et Titre Alcalin Complet, sont mesurés une fois par semaine. La Couleur est par contre mesurée une fois/jour à tous les niveaux de traitement. Le Chlore Résiduel libre est ainsi mesuré chaque jour au niveau des étapes de décantation-filtration et à la sortie de la station (Eau traitée).

La base de données d'entrée constituée est composée d'un ensemble de 8 paramètres descripteurs de la qualité de l'eau brute qui sont, le pH, la Conductivité, la Température, la Turbidité, le Magnésium, le Bicarbonate, la Dureté et le Titre Alcalin Complet (TAC). Un ensemble de 122 échantillons correspondant à ces 8 paramètres est constitué pour mener cette étude. Des statistiques descriptives simples de ces données d'entrée sont illustrées dans le tableau 3.1. La figure 3.5 présente l'évolution temporelle de ces paramètres descripteurs.

**Tableau. 3.1.** Statistiques descriptives des paramètres descripteurs de la qualité d'eau brute (Station Tilesdit).

Paramètre descripteurs	Minimum	Maximum	Moyenne	Ecart-type
<b>pH</b>	7,150	8,300	7,567	0,250
<b>Conductivité à 25 °C</b>	414,000	624,000	585,393	36,278
<b>Température</b>	9,700	24,200	16,130	3,483
<b>Turbidité</b>	1,320	23,810	3,835	2,392
<b>Magnésium</b>	7,290	47,628	22,268	4,931
<b>Bicarbonate</b>	158,620	289,140	222,497	23,213
<b>Dureté Permanente</b>	0,000	168,000	32,287	23,029
<b>Titre Alcalin Complet (TAC)</b>	130,000	237,000	181,845	18,703



**Figure 3.5.** Evolution temporelle des paramètres descripteurs de la qualité d'eau brute (Station Tilesdit).

Nous cherchons à décider sur la qualité de l'eau à travers ses paramètres descripteurs. Nous n'avons en fait aucune connaissance a priori sur un type de modèle représentant parfaitement ce procédé, par contre nous pouvons porter notre jugement sur la qualité de cette eau à partir de quelques données descriptives. Il y a quatre paramètres physico-chimiques qui sont souvent utilisés dans plusieurs travaux, et qui renseignent sur les dangers majeurs qu'il faut surveiller. Ces paramètres sont résumés comme suit : Conductivité (C), pH, Température ( $T^\circ$ ) et Turbidité (TU).

L'objectif qui se trouve derrière la collecte des données relatives à ces paramètres est de trouver un modèle de classification permettant de distinguer trois états bien distincts de l'eau (état très bonne, état moyenne et état médiocre) par la suite après sélection de la technique adoptée. La qualité de cette eau reflétée par sa potabilité repose en fait sur une corrélation qui ne peut être identifiée que statistiquement. Des données descriptives expérimentales recueillies sur une longue période (plusieurs années) pourraient atteindre cet objectif. A noter que la turbidité, le pH, et l'oxygène dissous sont fortement dépendants des phénomènes saisonniers. Il y a donc intérêt de disposer d'au moins une année pour archiver des données afin de déterminer une base de connaissance assez complète capable de fonctionner normalement. D'où la nécessité d'une base de connaissance riche en informations exigeant d'abord une collecte des données sur une longue période, et la présence d'un expert.

Dans un but de validation, une base de données contient des différents paramètres descripteurs de l'état de l'eau sont selon les normes recommandées par les pouvoirs publics. Un cas d'étude est alors à envisager. Elle est une extension multi-classe représentant un cas de classification plus élaboré, comportant trois classes de potabilité différentes (classe I, II et III) et appliqué sur les données.

### **3 La technique de contrôle et de surveillance**

Les méthodes de reconnaissance de formes telle que le RNA et neuronal-RBF appliquées à la classification des données, présente l'avantage de couvrir un grand nombre d'applications. Elles sont utilisées pour les systèmes de décision de hauts niveaux, et fondées sur l'analyse de données expérimentales.

### 3.1 Développement des réseaux de neurones

Les réseaux de neurones artificiels (RNAs), sont des outils de l'intelligence artificielle, capables d'effectuer des opérations de classification. Le processus d'apprentissage est donc une phase très importante pour la réussite d'une telle opération. Une des qualités de ce type de technique, est leur adéquation pour la mise au point de systèmes de surveillance modernes, capables de s'adapter à d'éventuelles extensions et reconfigurations multiples.

Les différents types des RNAs sont des réseaux de processeurs élémentaires fortement connectés fonctionnant en parallèle. Chaque processeur calcule une sortie unique sur la base des informations qu'il reçoit. A chaque liaison neurone-neurone a associé un poids. Le développement d'un réseau de neurones se fait en plusieurs étapes : le choix de l'architecture puis l'identification des paramètres de cette architecture. Cette dernière phase connue sous le nom d'apprentissage est n'est autre que l'identification des paramètres d'un modèle non-linéaire, par une méthode itérative de modification des poids des connections. Dans notre cas, nous disposons de données journalières (la quantité des paramètres physico-chimiques). Nous avons donc opté pour un apprentissage supervisé. Dans notre application on va utiliser deux types d'architectures de RNA, le réseau de neurones multicouche PMC (ou Multi-layer Perceptron MLP) ainsi que les réseaux de neurones à fonctions radiales de base neuronal-RBF (Radial Basis network RBF), qu'on doit adopter dans notre travail.

## 4 Simulation

### 4.1 Apprentissage et généralisation

L'apprentissage est le processus d'ajustement des poids pour une classification optimale. Pour déterminer le nombre de couches cachées et le nombre de neurones correspondant à chacune d'entre elles, on a augmenté progressivement le nombre de couches et les nombres de neurones correspondants jusqu'à atteindre la précision voulue. Nous avons Utilisés pour cette phase une base de données réelle complète constituée de 800 échantillons pour chaque échantillon on possède les résultats de mesure réelle mais aussi d'analyses chimiques et physiques effectuées hors-ligne qui constituent un ensemble de 4 descripteurs de la qualité de l'eau : pH, C, T°, TU. Le nombre de neurones d'entrées est donc de 4.

L'algorithme d'apprentissage le plus couramment utilisé est celui de la rétro propagation de l'erreur. Dans notre travail, deux architectures sont abordées, les réseaux de neurones artificiels RNA et les réseaux de neurones de type neuronal-RBF. L'architecture du réseau de neurones de deux types choisis (RNA et neuronal-RBF) a les caractéristiques :

➤ **Caractéristiques de neuronal-RBF**

- Couche d'entrée : 4 neurones (4 paramètres).
- Couche de sortie : un neurone, classification multi-classes (03 classes).
- Les couches cachées : une seule couche
- La fonction d'activation : RBF

➤ **Caractéristiques de RNA**

- Couche d'entrée : 4 neurones (4 paramètres).
- Couche de sortie : un neurone, classification multi-classes (03 classes).
- Les couches cachées : multicouches
- La fonction d'activation : sigmoïde et linéaire.

Il s'agit de valider les deux techniques RNA et neuronal-RBF dans une étude comparative appliquée au contrôle de qualité de l'eau potable dans un but de classification multi-classes. Les classes sont : classes 1 : qualité très bonne, classes 2 : qualité moyenne, et classes 3 : qualité médiocre. Faut-il souligner dans ce que la base d'apprentissage représente l'information la plus importante et la délicate à constituer. Il s'agit bien d'utiliser une base d'entraînement constituée de données relatives aux différents états qualitatifs de l'eau suivant les normes recommandées. Dans un but de simulation, des bases de données réelles constituées des 4 paramètres physico-chimiques (pH, T°, C et TU), sont utilisées.

#### **4.1.1 Résultats d'apprentissage**

Les réseaux RNA et neuronal-RBF sont testés et validés en utilisant l'algorithme d'apprentissage de Levenberg-Marquardt et fonction d'activation (sigmoïde et linéaire) pour les réseaux de neurones artificiels de type PMC. Les résultats de l'apprentissage dans la base de données sont de 400 vecteurs de dimension 4 sont présentés dans le tableau 3.2. Les paramètres tels que les différents nombres de neurones dans la couche cachée

(NNCC), nombre d'itération, le temps d'apprentissage et le taux de reconnaissance pour les trois classes auparavant.

Base de données	Réseau	NNCC	Nombre d'itération	Temps d'apprentissage (sec)	Taux de reconnaissance
(4 paramètres) (C, pH, T°, TU)	RNA	8-1	1000	19,718	99.75 %
		15-12-1	1000	44,538	100 %
		16-12-8-1	1000	32,105	<b><u>100 %</u></b>
	Neuronal-RBF	290	290	11,263	<b><u>100 %</u></b>

**Tableau 3.2** Résultats d'apprentissage de RNA et neuronal-RBF.

On remarque une amélioration positive de taux de reconnaissance quand y a une association des neurones ou des couches cachées donne un meilleur résultat en termes de taux de reconnaissance. Le réseau de RBF présenté des meilleurs résultats en termes de taux de reconnaissance et de temps d'apprentissage. Donc, nous concluons que les réseaux des neurones de type RBF meilleur et plus efficace dans cette application.

#### 4.1.2 Résultats de test

Dans un but de validation de réseau choisis, une base de données destinée aux tests, constituée de 400 vecteurs de dimension 4, est utilisée pour les réseaux choisis. Une évaluation des performances de ces réseaux en matière de taux de reconnaissance est donc effectuée. Le tableau 3.3 montre ces résultats de test obtenus.

Base de données	Réseau	NNCC	nombre d'itération	temps de test (sec)	taux de reconnaissance
(4 paramètres) (C, pH, T°, TU)	RNA	8-1	1000	$31,2 \cdot 10^{-3}$	91,5 %
		15-12-1	1000	$46,8 \cdot 10^{-3}$	95,75 %
		16-12-8-1	1000	$31,21 \cdot 10^{-3}$	<b><u>96 %</u></b>
	Neuronal-RBF	290	290	0,0468	<b><u>95,75 %</u></b>

**Tableau 3.3** Résultats du test de RNA et neuronal-RBF.

## Comparaison

A travers les résultats de simulation obtenus dans les tests de validation des deux techniques RNA et neuronal-RBF, on déduit les résultats suivants :

- Le temps d'apprentissage de la méthode RNA augmente avec l'association des neurones et couches cachées. Le temps d'apprentissage de réseau neuronal-RBF nécessite moins de temps que la méthode RNA.
- Le taux de reconnaissance du réseau neuronal-RBF est meilleur et par fois similaire que celle du réseau RNA.

## Evaluation

Il apparaît clairement que les deux modèles étudiés RNA et neuronal-RBF présentent en général de très bons résultats, avec des taux de reconnaissance acceptables sur le plan décisionnel et plus de 90% et ils présentent une bonne adéquation de deux algorithmes testés pour ce type d'application. Par conséquent une diminution en matière de temps d'apprentissage dont l'utilisation du réseau neuronal-RBF est bien évidente, ce qui se voit clairement en comparant les tableaux 3.2 et 3.3. Le modèle neuronal-RBF est plutôt mieux placé du point de vue temps d'apprentissage et même de taux de reconnaissance, ce qui lui confère l'avantage d'une intégration dans un système de contrôle dynamique. L'enrichissement de la base de données peut faire une amélioration en matière de taux de reconnaissance, ce qui laisse envisager son intégration dans un système de surveillance en continu permettant la collecte en permanence des données supervisées par un expert. Un contrôle de potabilité étendu à une grande échelle, peut être pris en charge de façon dynamique par le système de la figure 3.1, puisque le temps de calcul réalisé est très faible. Le modèle RNA ne dispose pas en fait de règles bien présentes pour fixer le nombre de neurones cachés (et donc la complexité du modèle), ce qui présente un souci majeur pour l'obtention d'une architecture optimale. Le temps de calcul imparti à la phase d'apprentissage est très court, ce qui lui confère l'avantage d'une intégration dans un système de surveillance dynamique. Il faut toutefois souligner le fait que le principal souci pour l'application de ce type de modèle, est l'obtention d'un réseau « optimal ».

Etant donné que le modèle neuronal-RBF est retenue comme étant le meilleur choix, elle est appliquée dans ce cadre pour effectuer une classification multi-classe. L'impact de ce résultat est important sur les plans, aussi bien technique (temps d'apprentissage plus faible), qu'économique (nombre réduit de capteurs). Les caractéristiques affichées dans les

tableaux 3.2 et 3.3 soulignent l'intérêt théorique et pratique de deux modèles pour ce type d'application.

### **Conclusion**

Ce troisième et dernier chapitre a fait l'objet d'une étude en simulation concernant la mise en œuvre de deux techniques d'apprentissage statistique RNA et neuronal-RBF appliquées dans le domaine du contrôle et de surveillance des eaux potables. Cette étude a permis la validation et l'évaluation des performances de chacune de ces méthodes présentées. Une étude comparative dans le but d'un choix décisif de la méthode la mieux adaptée à l'application a été effectuée. Les paramètres liés au taux de reconnaissance et le temps d'apprentissage sont les deux critères de comparaison. La discussion des résultats obtenus, a permis d'opter pour la technique neuronal-RBF retenue pour ses qualités et avantages adaptés au problème posé.

## **CONCLUSION GENERALE**

Le travail de recherche présenté dans ce mémoire porte sur l'étude des réseaux de neurones artificiels et à base radiale : RNA et neuronal-RBF respectivement, appliquée à la reconnaissance de formes dans le domaine du contrôle et de surveillance des eaux potables. Dans l'objectif de découvrir laquelle des techniques est la plus efficace, dans ce domaine à partir d'une étude comparative. A cet effet, notre modeste travail peut être considéré comme une contribution aux solutions proposées, pour résoudre des problèmes d'intérêt stratégique à préoccupation nationale, utilisant des outils modernes à base de techniques avancées.

Cette étude a été structurée autour de trois chapitres essentiels. Le premier consacré à une introduction au domaine de l'eau, il a permis de présenter des généralités ainsi que les différentes méthodes de contrôle et de surveillance employées. Dans le second, Nous avons parlé de la théorie de ces méthodes citées auparavant (RNA et neuronal-RBF) de la classification multi-classe de données à apprentissage statistique supervisé. Nous avons traité de la définition et les caractéristiques de ces méthodes et leur mode de fonctionnement. Enfin le troisième et dernier chapitre, a fait l'objet d'une étude en simulation concernant la mise en œuvre de ces deux modèles d'apprentissage statistique appliqués dans le domaine du contrôle et de surveillance des eaux potables. Cette étude a permis la validation et l'évaluation des performances de chacune des méthodes présentées. Une étude comparative dans le but d'une sélection de la méthode la mieux adaptée à l'application proposée a été effectuée. Les paramètres liés au taux de reconnaissance et temps d'apprentissage, ont été les facteurs pertinents qui ont permis d'évaluer les méthodes étudiées.

Pour se faire, nous avons pris un cas réel d'une base de données à partir de la station Tilesdit, avec des données et paramètres réels. Les réseaux de neurones artificiels et à fonctions à base radiales sont les réseaux choisis pour cette application. La classification supervisée obtenue en combinant les deux techniques de reconnaissance de formes

appliquées au domaine de surveillance des eaux. Les deux méthodes exposées lors de cette étude présentent de bonnes performances en matière de taux de reconnaissance. Un intérêt d'usage et d'application de ces techniques dans ce domaine est donc bien justifié. Donc pour cela on trouve une validation et une bonne adéquation de la classification multi-classes pour un modèle intelligent telle que l'algorithme d'apprentissage et de généralisation. Le réseau neuronal-RBF est le mieux placé pour cette application. Le modèle RNA ne dispose pas en fait de règles bien présentes pour fixer le nombre de neurones cachés (et donc la complexité du modèle), ce qui présente un souci majeur pour l'obtention d'une architecture optimale. Le temps de calcul imparti à la phase d'apprentissage est très court, ce qui lui confère l'avantage d'une intégration dans un système de surveillance dynamique. Il faut toutefois souligner le fait que le principal souci pour l'application de ce type de modèle, est l'obtention d'un réseau « optimal ».

Plusieurs perspectives peuvent être envisagées, d'une part concernant la méthode de classification, qui peut être complétement avec l'utilisation d'autres algorithmes en parallèle, d'autres part concernant l'entrée du système en ajoutant de nouveaux capteurs en entrée pour des paramètres non mesurable en continu; des capteurs logiciels entre autres.

## Bibliographie

- [1]- **Mallevalle J., Odendaal P.E., Wiesner M.R.** « *The emergence of membranes in water and waste water treatment* » In: Water Treatment Membrane Process, Chapitre 1, McGraw-Hill, 1996 10 p.
- [2]- L'eau dans les établissements de santé : guide technique du ministère des solidarités, de la sante et de et de la famille en France.
- [3]- **M.LADJAL** « *Contribution au développement de systèmes de surveillance innovants dédiés au contrôle de la qualité des eaux potables* » thèse de doctorat, université de m'sila 2013.
- [4]- **A.SLIM** « *contribution de l'analyse multivariée a l'étude de régression et la classification supervisée des données environnementales* » mémoire de master, université de m'sila 2012.
- [5]-Traitement d'eau Par Catherine Bossard. Hubert METAYER association des techniciens de dialyse.
- [6] **K.SARDI** , « *Contrôle de la qualité de l'eau de la station d'hémodialyse De l'EHU 1<sup>er</sup> Novembre* » mémoire de master, université Mohamed Boudiaf de oran, 2014.
- [7] fiche ressources N3-les différentes étapes de traitement de l'eau : «[http://colleges.planetetp.com/IMG/pdf/fiche\\_ressource\\_n03\\_cle889a33.pdf](http://colleges.planetetp.com/IMG/pdf/fiche_ressource_n03_cle889a33.pdf)»
- [8]- **S. Harley**, « *L'analyse de la qualité de l'eau, Administration du rétablissement agricole* », TRE- 116-2002-03, Canada, Mars 2000.
- [9]- **N. Valentin, T. Denoex, F. Fotoohi**, « *Modelling of coagulant dosage in a water treatment plant* », Proceedings of EANN'99, September 1999, Warsaw, Poland, pp. 165-170
- [10]- **H. R. Hernández De León**, « *Supervision et diagnostic des procédés de production d'eau potable* », Thèse de Doctorat, Institut National des Sciences Appliquées de Toulouse, CNRS, N° 06584, France, 2006.
- [11]- **Igor Blîndu**, « *Outil d'aide au diagnostic du réseau d'eau potable pour la ville de Chisinau par analyse spatiale et temporelle des dysfonctionnements hydrauliques* », Thèse de Doctorat, Ecole nationale supérieure des Mines de Saint-Etienne, Université Jean Monnet - Saint-Etienne, N° d'ordre : 336ID, France, 2004.
- [12]- **Chauvin Arnoux**,« *La production d'eau potable pilotée plus efficacement* », Informatique industrielle, Solutions, Mesures, Décembre 2011, Vol. 840.

- [13]-CIDF Centre International De Formation. Principes généraux de traitement des eaux, Lyonnaise des Eaux, 2000.
- [14]- **DEGREMONT**, Mémento technique de l'eau :Tome 2. Lavoisier SAS – Lexique technique de l'eau, Paris, dixième édition, 2005.
- [15]- **N. VALENTIN** «*Construction d'un capteur logiciel pour le contrôle automatique du procédé de coagulation en traitement d'eau potable*» thèse de doctorat 2000.
- [16]- **M. ZEMOURI** «*Contribution à la surveillance des systèmes de production à l'aide des réseaux de neurones dynamiques: Application à la e-maintenance*» thèse de doctorat, université de France-comté, 2003.
- [17]- nouveaux risques sanitaires et nouveaux enjeux pour le contrôle de la qualité des eaux potables article de synthèse, les technologies de laboratoire N°3 Mars-Avril 2007 université paris Sud.
- [18]- **C. TOUZET**« *les réseaux de neurones artificiels, introduction au connexionnisme*» cours, exercice et travaux pratiques, paris, juillet 1992.
- [19]-**MC Culloch, W., Pitis, W.** - « *A Logical Calculus for the Ideas Immanent in NervousActivity* » - Bulletin Mathematics and Biophysics - 1943, N°. 5 - p. 115-133.
- [20]- **Guillaume B.**« *Contrôle sensori-moteur par réseaux neuromimétiques modulaires - Approche pour le pilotage réactif en atelier flexible-*» Thèse de Doctorat, Institut national des sciences appliquées de Lyon ,1995.
- [21]- **François B. et Michel V.** « *Les réseaux de neurones artificiels*» Presses Universitaire, Paris, 1996.
- [22]- **Bernard G.** « *Application de réseaux de neurones artificiels a la reconnaissance au de caractères manuscrits*» Thèse de Doctorat, Faculté polytechnique de Mons ,1996.
- [23]- **Gérard D.** « *Reducing the complexity of neural nets for industrial applications and biological models* » Neuronal Information Processing - From Biological Data to Modelling and Applications (World Scientific, 1998).
- [24]- **Poramate M.** « *Neural Preprocessing and Control of Reactive Walking Machines*» Springer- Verlag Berlin, 2007.
- [25]- **F. MEDJILI** « *Modélisation par Réseaux de Neurones Artificiels (RNA) et commande Prédictive non linéaire d'une station de production d'eau froide* » mémoire de magister en automatique, université de annaba 2010.
- [26]- **Katrina B.** « *Fuzzy neural predictive control of highly nonlinear ph process*» Computers chem. Engng, vol21, pp613-618, 1997.
- [27]- **Jagannathan S.**« *Neural Network Control of Nonlinear Discrete-Time Systems* » Taylor & Francis Group, New York, 2006.
- [28]- **F.Z. DAIKH.** « *contribution des approches de l'intelligence artificielle pour la stabilisation robuste des système non linéaires* » thèse de doctorat, université de oran 2015.

- [29]-**G. Dreyfus, J. Martinez, M. Samuelides, M. Gordon, F. Badran, S. thiria et L.Herault**, « *réseaux de neurones, méthodologies et application* »; EditionEyrolles.2002.
- [30]- **G. Zweingelstein**, «*diagnostic des défaillances; théories et pratique pour les systèmes Industriels* », col. traite des nouvelles technologies, séries diagnostic et maintenance, hermès .1995.
- [31]- **M. AMMAR**, « *mise en œuvre de réseaux de neurones pour la modélisation de cinétiques réactionnelles en vue de la transposition batch/continu* » thèse de doctorat, paris, juillet 2007.
- [32]- **Antoine Cornuéjols - Laurent Miclet** Avec la participation d'**Yves Kodratoff**; « *Apprentissage artificiel Concepts et algorithmes* » ; Eyrolles 2003.
- [33]- **A. Belgacem**, « *Classification des signaux EGC avec un système-multi-agent neuronale.*» mémoire de magister en informatique, université de Tlemcen 2012.
- [34]- **R. MAHADOU** « *diagnostic industriel par neuro-floue application a un système de production* » Thèse Ingénieur d'état en Informatique, Université de Batna 2007.
- [35]- **FABRICE. R.** « *Second Differentials in Arbitrary Feed-Forward Neural Networks*» ICNN, 96.
- [36]- **A. ASSOUM**, « *étude de la tolérance aux aléas logiques des réseaux de neurones Artificiels* », thèse de doctorat, institut national polytechnique de Grenoble, France.1997.
- [37]-**Sylvain Tertois**, « *Réduction des effets des non linéarités dans une modulation à l'aide de réseaux de neurones* », Thèse de Doctorat, Université de Rennes 1, France, N° d'ordre : 2924, 2003.

## RESUME

L'objectif des travaux de développement des outils de supervision/diagnostic d'une station de production d'eau potable dans son ensemble est la maîtrise des risques dans le domaine des métiers de l'eau est bien sûr en premier lieu la maîtrise de la qualité sanitaire de l'eau potable distribuée, objet d'une réglementation très stricte des pouvoirs publics. Ce mémoire aborde le problème de la surveillance des eaux potables en fonction des caractéristiques physicochimiques de l'eau brute à l'aide d'un modèle basé sur les techniques d'apprentissage statistiques telles que les réseaux de neurones artificiels et à base radiales. L'utilisation de ces techniques augmente dans l'industrie du traitement de l'eau potable puisqu'ils permettant le développement de robustes modèles non-linéaires d'unités de procédés complexes.

L'objectif de ce travail consiste à établir les différentes relations existant entre les paramètres descripteurs de l'eau et la décision adoptée sur la qualité des eaux propres. On cherche une évaluation complète exprimant les variables indicatrices de la qualité des eaux potables pour la décision sur les états différents de l'eau. Cette décision est basée sur l'application d'une technique de classification multi-classe et sur l'interprétation des informations obtenues sur tout l'ensemble de la base de données réelles pour les différentes classes des eaux considérées (plus de 03 classes).

### Mots clés :

Eau potable, Surveillance, classification multi-classe, Evaluation des performances, RNA, neural-RBF, Tilesdit, Simulation.

## ABSTRACT

The objective of the development work of monitoring tools / diagnostics of a drinking water production station as a whole is risk management in the area of the water sector is of course primarily the mastery of sanitary quality of drinking water distributed, subject to strict government regulation. This brief addresses the issue of the monitoring of drinking water based on the physicochemical characteristics of the raw water using a model based on statistical learning techniques such as artificial neural networks and radial basis. The use of these techniques increases in drinking water treatment industry as they enable the development of robust nonlinear complex process units models.

The objective of this work is to establish the relationship between the different water parameters descriptors and the decision adopted on the quality of clean water. We seek a full evaluation expressing the dummies of the quality of drinking water for the decision on water conditions. This decision is based on the application of a multi-class classification technique and the interpretation of the information obtained over the whole of the actual data for the different classes of the considered waters (over 03 classes).

### keywords :

Drinking water, monitoring, multi-class classification, Performance Evaluation, RNA, neural-RBF, Tilesdit, Simulation.

## ملخص

الهدف من هذا العمل هو تطوي ر أدوات الرصد / التشخيص محطة إنتاج مياه الشرب ككل لإدارة المخاطر في مجال قطاع المياه هو بالطبع في المقام الأول على إتقان الجودة الصحية من مياه الشرب الموزعة، تخضع لتنظيم الحكومة الصارم. هذه عناوين مختصرة قضية رصد مياه الشرب على أساس الخصائص الفيزيائية والكيميائية للمياه الخام باستخدام نموذج يقوم على تقنيات التعلم الإحصائية مثل الشبكات العصبية الاصطناعية وأساس شعاعي. استخدام هذه التقنيات زيادة في الشرب صناعة معالجة المياه لأنها تمكن من وضع قوة عملية معقدة وحدات النماذج غير الخطية.

والهدف من هذا العمل هو تحديد العلاقة بين مختلف واصفات المعلمات المياه والقرار الذي اتخذ على نوعية المياه النظيفة. إننا نسعى إلى تقييم كامل تعبر عن متغيرات مؤشرات نوعية مياه الشرب لاتخاذ قرار بشأن الحالات مختلفة من المياه. ويستند هذا القرار على تطبيق تقنية تصنيف متعدد الطبقات و تفسير المعلومات التي تم الحصول عليها على كامل البيانات الفعلية لفئات مختلفة من المياه يعتبر (أكثر من 03 فئات).

## RESUME

L'objectif des travaux de développement des outils de supervision/diagnostic d'une station de production d'eau potable dans son ensemble est la maîtrise des risques dans le domaine des métiers de l'eau est bien sûr en premier lieu la maîtrise de la qualité sanitaire de l'eau potable distribuée, objet d'une réglementation très stricte des pouvoirs publics. Ce mémoire aborde le problème de la surveillance des eaux potables en fonction des caractéristiques physicochimiques de l'eau brute à l'aide d'un modèle basé sur les techniques d'apprentissage statistiques telles que les réseaux de neurones artificiels et à base radiales. L'utilisation de ces techniques augmente dans l'industrie du traitement de l'eau potable puisqu'ils permettant le développement de robustes modèles non-linéaires d'unités de procédés complexes.

L'objectif de ce travail consiste à établir les différentes relations existant entre les paramètres descripteurs de l'eau et la décision adoptée sur la qualité des eaux propres. On cherche une évaluation complète exprimant les variables indicatrices de la qualité des eaux potables pour la décision sur les états différents de l'eau. Cette décision est basée sur l'application d'une technique de classification multi-classe et sur l'interprétation des informations obtenues sur tout l'ensemble de la base de données réelles pour les différentes classes des eaux considérées (plus de 03 classes).

### **Mots clés :**

Eau potable, Surveillance, classification multi-classe, Evaluation des performances, RNA, neural-RBF, Tilesdit, Simulation.

## ABSTRACT

The objective of the development work of monitoring tools / diagnostics of a drinking water production station as a whole is risk management in the area of the water sector is of course primarily the mastery of sanitary quality of drinking water distributed, subject to strict government regulation. This brief addresses the issue of the monitoring of drinking water based on the physicochemical characteristics of the raw water using a model based on statistical learning techniques such as artificial neural networks and radial basis. The use of these techniques increases in drinking water treatment industry as they enable the development of robust nonlinear complex process units models.

The objective of this work is to establish the relationship between the different water parameters descriptors and the decision adopted on the quality of clean water. We seek a full evaluation express the dummies of the quality of drinking water for the decision on water conditions. This decision is based on the application of a multi-class classification technique and the interpretation of the information obtained over the whole of the actual data for the different classes of the considered waters (over 03 classes).

### **keywords :**

Drinking water, monitoring, multi-class classification, Performance Evaluation, RNA, neural-RBF, Tilesdit, Simulation.