UNIVERSITE KASDI MERBAH – OUARGLA -

FACULTE DES HYDROCARBURES ET ENERGIES RENOUVELABLES ET SCIENCES DE LA TERRE ET DE L'UNIVERS

Département des Sciences de la Terre et de l'Univers



THESE

En Vue De L'obtention Du Diplôme de Doctorat ès sciences en géologie

THEME

Etude et suivi de procédé d'épuration des eaux usées sous climat aride (simulation par l'application des méthodes d'intelligence artificielle) cas de la station d'épuration de Touggourt.

Soutenue publiquement par :

BEKKARI Naceureddine

Devant le jury :

| Dr. BELEKSIR Med Salah | M.C.A | Président | Université d'Ouargla |
|--------------------------|------------|--------------------|----------------------|
| Pr. ZEDDOURI Aziez | Professeur | Directeur de thèse | Université d'Ouargla |
| Dr. BENHADDYA Med Lamine | M.R.A | Examinateur | CRSTRA Touggourt |
| Dr. KHERIFI Wahida | M.R.A | Examinatrice | CRSTRA Biskra |
| Dr. SAGGAI Sofiane | M.C.A | Examinateur | Université d'Ouargla |
| Dr. BOUSELSAL Boualem | M.C.A | Examinateur | Université d'Ouargla |

ANNEE UNIVERSITAIRE 2019/2020

Remerciements

Je remercie tout d'abord <u>ALLAH</u> le Tout Puissant qui m'a guidé pour accomplir ce travail

Au terme de ce document, je tiens à remercier Mr. **BELEKSIR Med Salah** pour avoir bien voulu accepter de m'honorer de sa présence et de présider mon jury de thèse de Doctorat en science.

Mes sincères remerciements et ma gratitude s'adressent aussi à Mr. **ZEDDOURI Aziez** Professeur au Département de Sciences de la terre et de l'Univers, à l'université Kasdi Merbah-Ouargla pour avoir accepté de diriger ce travail, pour ses orientations, ses conseils précieux, ses critiques et sa compréhension.

Mes remerciements vont aussi à Monsieur **Benhaddya Mohammed Lamine** Maître de recherche **A** au Centre de Recherche Scientifique et Techniques sur les Régions Aride (CRSTRA) Touggourt d'accepter de juger ce travail.

Mes remerciements vont aussi à Madame **KHERIFI Wahida** Maître de recherche **A** au Centre de Recherche Scientifique et Techniques sur les Régions Aride (CRSTRA) Biskra, d'accepter de juger ce travail.

Mes remerciements vont aussi à Monsieur **SAGGAI Sofiane** Maître de conférences **A** au département de l'Hydraulique et Génie Civil à l'université Kasdi Merbah-Ouargla d'accepter de juger ce travail.

Mes remerciements vont aussi à Monsieur **Bouselsal Boualem** Maître de conférences **A** au département de Sciences de la terre et de l'Univers, à l'université Kasdi Merbah-Ouargla d'accepter de juger ce travail.

Je remercie également mes chers collègues **Dr Youcef Halis**, **Amiri Khaled**, Dr **Ben Aissa Mohammed Hocine**, **Dr Mihoub Adil** et **Dr Hadjoudj Moussa**.

Mes remerciements vont aussi à toute ma famille, tous mes amis, celles et ceux qui m'ont apporté leur aide pour la réalisation de ce travail.

عملية الحمأة المنشطة هي أكثر أنظمة معالجة المياه العادمة البيولوجية شيوعًا. يتأثر أداء عمليات التنقية عمومًا بعدة عوامل مثل التقلب في تركيزات وحجم مياه الصرف الصحي والتقلبات الكامنة في عملية المعالجة.

تعد نمذجة هذه العملية مهمة لتحسين كفاءة معالجتها وبالتالي جودة المخلفات السائلة التي يتم تصريفها في البيئة المستقبلة. يمكن أن تساعد النماذج المشغل بالتنبؤ بأداء محطة المعالجة من أجل اتخاذ تدابير تصحيحية فعالة من حيث التكلفة وفي الوقت المناسب لضمان كفاءة معالجة متسقة لتلبية معايير التصريف. ومع نظرًا للخصائص شديدة التعقيد وغير الخطية لهذا البيولوجي، فإن النمذجة التقليدية لعملية المعالجة هذه لا تزال تشكل تحديًا. الشبكة العصبية الاصطناعية متعددة الطبقات المدركة (RNA-PMC)هي أداة التنبؤ الأكثر استخدامًا في محطات معالجة المياه العادمة نظرا لأداءها المتميز.

في هذه الرسالة تم تقييم أداء محطة معالجة مياه الصرف الصحي المنزلي لمدينة تقرت. تضمنت البيانات التي تمت در استها وتحليلها إحصائيا الموصلية الكهربائية (CE) ، والأس الهيدروجيني (PH) (T) ، والأكسجين الذائب (DO) والطلب على الأكسجين الكيميائي الحيوي (BOD) ، والطلب على الأكسجين الكيميائي (CDD) (MES) نيتروجين الأمونيا (NH4) (NO₃) ، النتريت (NO₂)) والطلب على الأكسجين الكيميائي (PO4) (NO₃) إلى المواد الصلبة العالقة في حوض التهوية (MESBA) وإعادة تدوير الحمأة .(PC4) تم تنفيذ أداء تنقية هذه الم التحليلات الإحصائية للبيانات التي تم جمعها خلال فترة الدراسة المقدرة ب 10 أشهر. ثم تم تطوير النماذج العصبية (RNA-PMC) في هذه الدراسة للتنبؤ بتركيزات المعلمات الرئيسية للتلوث العضوي عند مخرج المحطة بما في ذلك (BO5₃) (DEO₅) . (MES)

تظهر النتائج أن كفاءة تنقية محطة معالجة مياه الصرف الصحى هذه مرضية للغاية لإز الة الملوثات وفقًا لمعايير التصريف، يمكن ملاحظة أن هناك تباينًا كبيرًا إلى حد ما في التلوث العضوي وفقًا للتغير ات في الظروف PO₄ NO₃ من هذه المناخية. من أجل تحديد البنية المناسبة لنماذج الشبكات العصبية للمعلمات الثلاثة التي تمت در استها، تطوير نوعين من النماذج، الأول هو النموذج البسيط الذي يستخدم معلمات المياه العادمة الداخلة الى المحطة كمدخلات للنموذج والثاني هو النموذج الشامل الذي يستخدم بالإضافة الى المعلمات الداخلة معلمات مخرج المحطة (المياه المصفاة) تم الحصول على افضل النتائج مع ابنية الشبكات [9-MES DCOs DBO₅s [1-30-7] [1-30-7] [1-45 بالنسبة للنموذج البسيط[11-35-10-1] -10-1] وظيفة التنشيط DCOs MES, DBO₅s 35-10], [1-10-30-13] الى خوارزمية التعلم ليفنبرغ ماركوارت (LM) .أظهرت النتائج أن نموذج RNA-PMC يمكن أن يتنبأ بالنتائج التجريبية مع مراحل التعلم والتحقق والاختبار للمعلمات الثلاثة المتوقعة. تشير النتائج بين 0.9 0,999 الإجمالية إلى أن نهج نمذجة RNA-PMC يمكن أن يوفر تقييمًا موثوقًا لأداء محطات معالجة مياه الصرف الصحي من خلال محاكاة والتغيرات الموسمية في درجة ك المحطة عبر مجموعة واسعة من التغييرات في نوعية المياه، يمكن أن يكون هذا أداة فعالة لمحاكاة ومراقبة وتوقع أداء محطة المعالجة.

: طرق تصفية المياه، المياه

RESUME

Le procédé d'épuration par boues activées (BA) est le système d'épuration biologique des eaux usées le plus couramment utilisé. Les performances des procédés d'épuration sont généralement affectées par plusieurs facteurs tels que la fluctuation dans les concentrations des éléments polluants, le volume des eaux usées et la variabilité inhérente du procédé de traitement.

La modélisation de ce processus est importante pour améliorer l'efficacité de son traitement et donc la qualité de l'effluent rejeté dans le milieu récepteur. En effet, les modèles peuvent aider l'exploitant à prédire les performances de la station d'épuration afin de prendre des mesures correctives rentables et opportunes permettant de garantir une efficacité constante du traitement afin de respecter les normes de rejet. Cependant, en raison des caractéristiques très complexes et non linéaires de ce système biologique, la modélisation traditionnelle de ce processus de traitement reste un défi. Les techniques de l'intelligence artificielle notamment le réseau de neurones artificiel perceptron multicouche (RNA-PMC) est l'outil de prédiction le plus couramment utilisé dans les stations d'épurations (STEP) à BA l'hors de leur pouvoir prédateur.

Dans cette thèse la performance de la STEP des eaux usées domestique de la ville de Touggourt a été évaluée. Les données étudiées et analysées statistiquement incluent la Conductivité Electrique(CE), le Potentiel hydrogène (pH), la température de l'eau (T), l'Oxygène Dissout (OD), la demande biochimique en oxygène (DBO₅), la demande chimique en oxygène (DCO), les matières en suspension (MES), l'azote ammoniacal (NH₄), le Nitrate (NO₃), le Nitrite (NO₂), les ortho phosphates (PO₄) à l'entrée et à la sortie de la station, en plus de la Matière en suspension dans le bassin d'aération (MESBA) et le recyclage des boues(RECY) . La performance épuratoire de cette station a été effectuée à travers les analyses statistiques des données collectées durant une période d'étude de 10 mois. Puis des modèles neuronaux (RNA-PMC) ont été développés dans cette étude afin de prédire les concentrations des paramètres clés de la pollution organique à la sortie de la station notamment la (DBO₅s), (DCOs) et (MESs) de la STEP de Touggourt.

Les résultats montrent que les rendements épuratoires de cette STEP sont très satisfaisants pour l'élimination des matières polluantes en accord avec les normes de rejet, sauf pour les NO₃ et les PO₄. D'après ces résultats on remarque qu'il y a une variation assez

importante en matière de pollution organique selon les variations des conditions climatiques. Afin de déterminer l'architecture appropriée des modèles de réseaux neuronaux pour les trois paramètres étudiés, deux types de modèle ont été élaborés, le premier est le modèle simple qui utilise les paramètres de l'influent de la STEP comme entrées du modèle et le deuxième c'est le modèle extensif qui utilise en plus des paramètres de l'influent les paramètres de l'effluent de la STEP ayant une corrélation significative avec le paramètre prédit. Les meilleurs résultats ont été obtenus avec des architectures des réseaux [9-45-1], [7-30-1] et [7-30-1] pour la DBO₅s, DCOs et MESs respectivement pour le modèle simple et [13-30-10-1], [10-35-10-1] et [11-35-10-1] pour la DBO₅s, DCOs et MESs respectivement pour le modèle extensif, en utilisant des fonctions d'activation sigmoïdes à tangente hyperbolique et un algorithme d'apprentissage de Levenberg-Marquardt (LM). Les résultats ont montré que le modèle RNA-PMC peut prédire les résultats expérimentaux avec un coefficient de corrélation élevé qui varie de 0,9 à 0,999 au cours des phases d'apprentissage, de validation et de test pour les trois paramètres prédits. Les résultats globaux indiquent que l'approche de modélisation RNA-PMC peut fournir une évaluation fiable de la performance des stations d'épuration des eaux usées (STEP) en simulant le comportement de la station sur une large fluctuation affectant la qualité des eaux, y compris les fluctuations de différentes durées et les variations saisonnières de température. Cela peut donc être un outil efficace pour simuler, contrôler et prévoir les performances de la station d'épuration.

Mots-clés : Procèdes d'épurations, eaux usées, l'intelligence artificielle, simulation, climat aride.

ABSTRACT

The Activated Sludge Process (ASP) is the most commonly used biological wastewater treatment system. The performance of purification processes is generally affected by several factors such as fluctuation in pollutant concentrations, volume of wastewater and the inherent variability of the treatment process.

The modeling of this process is important for improving its treatment efficiency and thus the quality of the effluent released into the receiving water body. This is because the models can help the operator to predict the performance of the plant in order to take cost-effective and timely remedial actions that would ensure consistent treatment efficiency and meeting discharge consents. However, due to the highly complex and non-linear characteristics of this biological system, traditional modeling of this treatment process has remained a challenge. The multilayer perceptron artificial neural network (ANN -MLP) is the prediction tool most commonly used in ASP out of their predatory power.

In this thesis the performance of wastewater treatment plant (WWTP) of domestic sewage of the Touggourt city was evaluated. The data studied and statistically analyzed included the Electrical Conductivity (CE), Hydrogen Potential (PH), Water Temperature (T), Dissolved Oxygen (DO), Biochemical Oxygen Demand (BOD₅), Chemical Oxygen Demand (COD), suspended solids (SS), ammonia nitrogen (NH₄), nitrate (NO₃), nitrite (NO₂), orthophosphates (PO₄) at the inlet and outlet of the WWTP, plus suspended solids in the aeration basin (MESBA) and sludge recycling (RECY). The purification performance of this WWTP was carried out through the statistical analyzes of the data collected during a study period of 10 months. Then neural models ANN -MLP were developed in this study to predict the concentrations of the key parameters of organic pollution including the effluent Biochemical Oxygen Demand (BOD₅eff), effluent Chemical Oxygen Demand (CODeff) and effluent suspended solids (SSeff) of Touggourt WWTP.

The results show that the treatment efficiencies of this WWTP are very satisfactory for the removal of pollutants in accordance with the discharge standards, except for NO_3 and PO_4 . From these results, it can be seen that there is a fairly significant variation in organic pollution according to the variations in climatic conditions. In order to determine the appropriate architecture of the neural network models for the three parameters studied, two types of

model have been developed, the first is the simple model which uses the parameters of the influent of STEP as inputs to the model and the second it is the extensive model which uses the effluent parameters in addition to the influent parameters, witch having a strong correlation with the predicted parameters. The best results have been obtained with network architectures [9-45-1], [7-30-1] and [7-30-1] for BOD₅s, CODs and SSs respectively for the simple model and [13- 30-10-1], [10-35-10-1] and [11-35-10-1] for BOD5s, CODs and SSs respectively for the extensive model, using sigmoid activation functions with hyperbolic tangent and Levenberg-Marquardt (LM) as learning algorithm. The results showed that the ANN -MLP model could predict experimental results with a high correlation coefficient varied from 0.9 to 0.999 during the learning, validation and testing phases for the three predicted parameters. The overall results indicated that the ANN -MLP modeling approach can provide reliable performance evaluation of wastewater treatment plants WWTP can be done by simulating the plant behavior over a wide range of influent disturbances, including fluctuation in different durations, seasonal temperature variations. So it can be an effective tool for simulating, controlling and predicting the performance of WWTP.

Keywords: Purification processes, wastewater, artificial intelligence, simulation, arid climate.

Remerciements Ι Π RESUME Ш ABSTRACT V Table de matière VII Liste des figures XII Liste des tableaux XV Listes d'abréviation XVI Introduction générale 1 **PARTIE 1 : SYNTHESE BIBLIOGRAPHIQUE** Chapitre l : Généralités sur les eaux usées et les procédés d'épuration 5 **I.1** Introduction 5 **I.2** Les eaux usées 5 **I.2.1** Définition 5 I.2.2 Origine des eaux usées 6 **I.2.2.1** Incidences des effluents sur la santé humaine 6 **I.2.2.2** Contamination de l'eau potable 7 **I.2.2.3** Dégradation de l'environnement 7 **I.2.2.4** Eutrophisation des eaux réceptrices 8 I.2.2.5 Toxicité directe 8 I.2.3 Procédés d'épuration des eaux usées 8 I.2.3.1 But d'épuration 8 I.2.3.2 Prétraitements 9 I.2.3.2.1 Dégrillage 9 I.2.3.2.2 Dessablage 10 I.2.3.2.3 Dégraissage déshuilage 10 I.2.3.3 Traitements Primaires 11 **I.2.3.4** Traitements secondaires (traitement biologique) 11 I.2.3.4.1 Traitement biologique intensifs : 12 I.2.3.4.1.1 Boues activées 12 I.2.3.4.2 Traitement biologique extensifs: 12 **I.2.3.4.2.1** Le lagunage 12

Table de matière

| I.2.3.4.2.2 La filtration /percolation | 13 |
|--|----|
| I.2.3.4.2.3 La phytoépuration ou filtres plantés de macrophytes | 14 |
| I.2.3.5 Traitements tertiaires | 14 |
| I.2.3.5.1 Traitement des boues | 14 |
| I.3 Conclusion | 15 |
| Chapitre II : Techniques de l'intelligence artificielle (IA) | |
| II.1 Introduction | 16 |
| II.2 Le réseau de neurone artificiel | 17 |
| II.2.1 Le neurone biologique | 17 |
| II.2.2 Le neurone formel (Artificiel) | 18 |
| II.2.3 Structure de réseau de neurones artificiels | 19 |
| II.2.4 Architecture des réseaux de neurones artificiels | 20 |
| II.2.4.1 Réseau de neurones artificiels direct (à alimentation avant) | 20 |
| II.2.4.2 Réseau de neurones artificiels récurrent | 21 |
| II.2.5 Fonction de transfert | 22 |
| II.2.6 Apprentissage des réseaux de neurones | 23 |
| II.2.6.1 Règle d'apprentissage du perceptron | 23 |
| II.2.6.2 Apprentissage supervisé | 24 |
| II.2.6.3 Apprentissage non supervisé | 24 |
| II.2.7 Algorithme d'apprentissage | 24 |
| II.2.7.1 Algorithme d'apprentissage de retro propagation (RP) | 24 |
| II.2.8 Généralisation | 26 |
| II.2.8.1 Sur-apprentissage et sous-apprentissage | 27 |
| II.2.8.2 Arrêt prématuré | 28 |
| II.2.9 Avantage des réseaux de neurones artificiels | 30 |
| II.3 Revue de littérature | 30 |
| II.4 Conclusion | 34 |
| PARTIE 2 : MATERIEL ET METHODES | |
| Chapitre III : Matériels et méthodes | 35 |
| III.1 Description de la station d'épuration de Touggourt | 35 |
| III.1.1 Données techniques | 36 |
| III.1.2 Etapes de traitement | 36 |
| III.1.3 Equipement de la station | 37 |

| III.2 Législation sur la protection de l'environnement régissant les usines de | 38 |
|---|----|
| traitement des eaux usées | |
| III.3 Description et collecte des données | 38 |
| III.3.1 Conductivité électrique (CE) | 39 |
| III.3.2 Potentiel hydrogène (PH) | 39 |
| III.3.3 Température (T) | 39 |
| III.3.4 Oxygène Dissout (DO) | 40 |
| III.3.5 Demande biochimique en oxygène (DBO) | 40 |
| III.3.6 Demande chimique en oxygène (DCO) | 41 |
| III.3.7 Matière en suspension (MES) | 41 |
| III.3.8 Composés azotes | 41 |
| III.3.9 Les ortho phosphore (PO ₄) | 42 |
| III.3.10 Matières en suspension dans le bassin d'aération (MESBA) | 42 |
| III.3.11 Boues recyclées (RECY) | 42 |
| III.3.12 Récapitulatifs des fréquences et des Méthodes d'analyse | 43 |
| III.4 Prétraitement des données | 43 |
| III.5 La boite à moustache (box and whisker plot) | 44 |
| III.6 L'analyse en composantes principales ACP | 44 |
| III.6.1 Généralités | 44 |
| III.6.2 Principes de l'ACP | 45 |
| III.7 Logiciels utilisés | 45 |
| III.8 Simulation par le réseau de neurone artificiel | 46 |
| III.8.1 Taille de la base des données | 46 |
| III.8.2 Développement du modèle | 46 |
| III.8.3 Division des données | 48 |
| III.8.4 Mesure des performances de prédiction | 49 |
| III.8.5 Normalisation des données | 49 |
| III.9 Conclusion | 51 |
| PARTIE 3 : RESULTATS ET DISCUSSIONS | |
| Chapitre IV : Analyse statistique et performance de la STEP de Touggourt | 52 |
| IV.1 Introduction | 52 |
| IV.2 Traitement des données par la boite a moustache : | 52 |
| IV.3 Statistique descriptive des paramètres de pollution : | 55 |

| IV.4 Performances épuratoires de la STEP de Touggourt | 56 |
|---|-----|
| IV.4.1 La température | 57 |
| IV.4.2 Le pH | 58 |
| IV.4.3 La conductivité électrique | 59 |
| IV.4.4 Matières en suspension (MES) | 60 |
| IV.4.5 Oxygène dissout | 61 |
| IV.4.6 La demande biochimique en oxygène(DBO ₅) | 62 |
| IV.4.7 Demande chimique en oxygène (DCO) | 63 |
| IV.4.8 Ortho phosphates (PO ₄) | 64 |
| IV.4.9 Matières azotées (NH ₄ , NO ₂ , NO ₃) | 65 |
| IV.4.10 Evaluation de la pollution organique des eaux brutes | 69 |
| IV.5 Conclusion | 70 |
| Chapitre V : Analyse en composantes principales (ACP) | 71 |
| V.1 Introduction | 71 |
| V.2 Matrice de corrélation | 71 |
| V.3 Analyse des facteurs | 74 |
| V.4 Analyse des variables et des individus | 76 |
| V.5 Conclusion | 81 |
| Chapitre VI : Simulation de la STEP de Touggourt par le RNA | 82 |
| VI.1 Introduction | 82 |
| VI.2 Le choix des variables d'entrées | 82 |
| VI.3 Prédiction de la DBO ₅ par le modèle PMC-RNA | 83 |
| VI.3.1 Prédiction de la DBO ₅ s par l'utilisation des paramètres d'entrée de la STEP | |
| (model simple RNA-PMCSDBO ₅ s) | 83 |
| VI.3.2 Prédiction de la DBO ₅ s par l'utilisation des paramètres d'entrée et de sortie de | 0.0 |
| la STEP (model extensif RNA-PMCEDBO ₅ s) | 89 |
| VI.4 Prédiction de la DCOs par le modèle PMC-RNA | 94 |
| VI.4.1 Prédiction de la DCOs par l'utilisation des paramètres d'entrée de la STEP | |
| (model simple RNA-PMCSDCOs) | 94 |
| VI.4.2 Prédiction de la DCOs par l'utilisation des paramètres d'entrée et de sortie de | |
| la STEP (model extensif RNA-PMCEDCOs) | 100 |
| VI.5 Prédiction de la MES par le modèle RNA-PMC | 105 |
| | 105 |

| (model simple RNA-PMCSMESs) | |
|--|-----|
| VI.5.2 Prédiction de la MESs par l'utilisation des paramètres d'entré et de sortie de la STER (model extensif RNA_RMCEMESs) | 110 |
| STEF (model extensil KIVA-FWCEWESS) | |
| VI.6 Conclusion | 115 |
| Conclusion générale | 116 |
| Perspectives | 119 |
| Références bibliographique | 120 |
| Annexe01 | 131 |
| Annexe02 | 133 |

| Liste | des | figures |
|-------|-----|---------|
|-------|-----|---------|

| Figure 1: Structure d'un neurone biologique | |
|--|-----|
| Figure2: Neurone formel (artificiel). | |
| Figure 3: Réseau de neurones multicouches (PMC) avec deux couches cachées | |
| Figure 4: Réseau de neurones artificiels directs (par alimentation avant) (Feed-Forward) | |
| Figure 5: Réseau de neurones récurrent | 21 |
| Figure 6 : Présentation graphique du phénomène de sur-apprentissage "Cas | |
| d'approximation d'une fonction" | 27 |
| Figure 7 : Exemple du phénomène de sous-apprentissage | 28 |
| Figure 8 : Illustration du critère de la validation croisée | 29 |
| Figure 9 : Localisation de la station d'épuration des eaux usées de Touggourt | 35 |
| Figure 10 : Schéma de procédé d'épuration des eaux usées de la STEP de Touggourt | 36 |
| Figure 11 : Equipement de la STEP de Touggourt | 37 |
| Figure 12 : la structure de réseau neuronale RNA-PMC utilisé dans ce travail | 48 |
| La figure 13 : Graphique en boîte à moustache des paramètres d'entrée et de sortie de la | |
| STEP de Touggourt. | 54 |
| Figure 14 : Représentation graphique de la variation de la T à l'entrée et à la sortie de | |
| STEP. | |
| Figure 15 : Représentation graphique de la variation de du PH à l'entrée et à la sortie de | |
| STEP. | 58 |
| Figure 16 : Représentation graphique de la variation de la CE à l'entrée et à la sortie de | |
| STEP. | 59 |
| Figure 17 : Représentation graphique de la variation de la MES à l'entrée et à la sortie de | |
| STEP. | 60 |
| Figure 18 : Représentation graphique de la variation de du OD à l'entrée et à la sortie de | - 1 |
| STEP. | 61 |
| Figure 19 : Représentation graphique de la variation de du DBO ₅ à l'entrée et à la sortie de | |
| STEP. | 63 |
| Figure 20 : Représentation graphique de la variation de du DCO à l'entrée et à la sortie de | |
| STEP. | 63 |
| Figure 21 : Représentation graphique de la variation de du PO ₄ à l'entrée et à la sortie de | - 1 |
| STEP. | 64 |
| Figure 22 : Représentation graphique de la variation de du NH ₄ à l'entrée et à la sortie de | 66 |

| STEP. | |
|---|------------|
| Figure 23 : Représentation graphique de la variation de du NO ₂ à l'entrée et à la sortie de | |
| STEP. | 67 |
| Figure 24 : Représentation graphique de la variation de du NO ₃ à l'entrée et à la sortie de | <i>c</i> 0 |
| STEP. | 68 |
| Figure 25 : Représentation graphique du rapport de biodégradabilité des eaux brutes. | 69 |
| Figure 26. Projection des variables et des individus sur le plan 1-2 (ACP) | 77 |
| Figure 27. Projection des variables et des individus sur le plan III- IV (ACP) | 78 |
| Figure 28. Projection des variables et des individus sur le plan 5-6 (ACP) | 80 |
| Figure 29 : Comparaison entre les meilleurs modèles simples RNA-PMCSDBO ₅ s de | |
| chaque scénario | 86 |
| Figure 30 : L'architecture du meilleur modèle simple RNA-PMCSDBO ₅ s | 87 |
| Figure 31: Comparaison entre les valeurs prédites et les valeurs mesurées pour le meilleur | ~ ~ |
| modèle simple RNA-PMCSDBO ₅ s | 87 |
| Figure 32: Diagramme de régression entre les valeurs prédites et mesurées pour le meilleur | |
| modèle simple RNA-PMCSDBO5s | 88 |
| Figure 33: Comparaison entre les meilleurs modèles extensifs RNA-PMCEDBO ₅ s de | |
| chaque scénario | 91 |
| Figure 34: L'architecture du meilleur modèle extensif RNA-PMCEDBO ₅ s | 92 |
| Figure 35: Comparaison entre les valeurs prédites et les valeurs mesurées pour le meilleur | |
| modèle extensif RNA-PMCEDBO5s | 92 |
| Figure 36: Diagramme de régression entre les valeurs prédites et mesurées pour le meilleur | |
| modèle extensif RNA-PMCEDBO5s | 94 |
| Figure 37: L'architecture du meilleur modèle simple RNA-PMCSDCO5s | 97 |
| Figure 38: Comparaison entre les meilleurs modèles simples RNA-PMCSDCOs de chaque | |
| scénario | 97 |
| Figure 39: Comparaison entre les valeurs prédites et les valeurs mesurées pour le meilleur | 0.0 |
| modèle simple RNA-PMCSDCOs | 98 |
| Figure 40 : Diagramme de régression entre les valeurs prédites et mesurées pour le | |
| meilleur modèle RNA-PMCDCOs | 99 |
| Figure 41: Comparaison entre les meilleurs modèles extensifs RNA-PMCEDCOs de | 102 |
| chaque scénario | 102 |
| Figure 42 : L'architecture du meilleur modèle extensif RNA-PMCEDCOs | 103 |

| Figure 43: Comparaison entre les valeurs prédites et les valeurs mesurées pour le meilleur | |
|---|-----|
| modèle extensif RNA-PMCEDCOs | |
| Figure 44: Diagramme de régression entre les valeurs prédites et mesurées pour le meilleur | |
| modèle extensif RNA-PMCEDCOs | 105 |
| Figure 45: Comparaison entre les meilleurs modèles simples RNA-PMCSMESs de chaque | |
| scénario | 107 |
| Figure 46: L'architecture du meilleur modèle simple RNA-PMCSMESs | 108 |
| Figure 47: Comparaison entre les valeurs prédites et les valeurs mesurées pour le meilleur | 100 |
| modèle simple RNA-PMCSMESs | 109 |
| Figure 48 : Diagramme de régression entre les valeurs prédites et mesurées pour le meilleur | |
| modèle simple RNA-PMCSMESs | 110 |
| Figure 49 : Comparaison entre les meilleurs modèles extensifs RNA-PMCEMESs de | |
| chaque scénario | 112 |
| Figure 50 : L'architecture du meilleur modèle extensif RNA-PMCEMESs | 112 |
| Figure 51: Comparaison entre les valeurs prédites et les valeurs mesurées pour le meilleur | |
| modèle extensif RNA-PMCEMESs | 113 |
| Figure 52: Diagramme de régression entre les valeurs prédites et mesurées pour le meilleur | |
| modèle extensif RNA-PMCSEMESs | 114 |

Tableau 1: Les principales fonctions de transfert 23 Tableau 2: Méthodes d'analyse et fréquence d'échantillonnage des paramètres 43 étudiés **Tableau 3**: Synthèse de l'analyse statistique des données à l'entrée de la STEP de 55 Touggourt Tableau 4: Synthèse de l'analyse statistique des données à la sortie de la STEP de 56 Touggourt Tableau 5: rapports exprimant la pollution organique des eaux brute de la STEP de 69 Touggourt Tableau 6 : Matrice de corrélation des paramètres physico-chimiques à l'entrée et à 72 la sortie de la STEP de Touggourt Tableau 7 : Matrice des composantes 75 Tableau 8 : Données d'entrée-sortie du modèle simple RNA-PMCSDBO5 pour la 83 prédiction de la DBO5 à la sortie de la STEP Tableau 9 : Performances de la prédiction de la DBO₅s par le modèle simple RNA-84 PMCSDBO₅ Tableau 10: paramètres d'entrées-sortie du modèle extensif RNA-PMCEDBO5s 89 pour la prédiction de la DBO5s a la sortie de la STEP Tableau 11: Performances de la prédiction de la DBOs par modèle extensif RNA-90 PMCEDBO5 Tableau 12: Données d'entrées-sortie de modèle simple RNA-PMCSDCO pour la 94 prédiction de la DCO a la sortie de la STEP Tableau 13: Performances de la prédiction de la DCOs par le modèle simple RNA-95 **PMCSDCOs** Tableau 14: paramètres d'entrée-sortie de modèle extensif RNA-PMCEDCOs pour 100 la prédiction de la DCOs à la sortie de la STEP Tableau 15: Performances de la prédiction de la DCOs par modèle extensif RNA-101 **PMCEDCOs** Tableau 16 : paramètres d'entrée-sortie du modèle simple RNA-PMCSMES pour la 105 prédiction de la MES à la sortie de la STEP Tableau 17: Performances de la prédiction de la MESs par le modèle simple RNA-106 **PMCSMES**

Liste des tableaux

| Tableau 18 : paramètres d'entrées-sortie du modèle extensif RNA-PMCEMESspour la prédiction du MESs à la sortie de la STEP | 110 |
|--|-----|
| Tableau 19 : Performances de la prédiction du MESs par modèle extensif RNA-PMCEDCO5 | |
| Annexe | |
| Tableau01: Normes Algériennes des rejets d'eaux usées traités | 131 |
| Tableau02: Les normes des eaux usées rejetées selon l'OMS (1971) | |

| RNA | Réseau de Neurones Artificiel |
|-------------------|--|
| РМС | Perceptron multicouches |
| BA | Boues Activées |
| RNA-PMC | Réseau de neurones artificiel perceptron multicouche |
| STEP | Stations d'épurations |
| СЕ | Conductivité électrique |
| РН | Potentiel hydrogène |
| Τ | Température de l'eau |
| OD | Oxygène dissout |
| DBO5 | La demande biochimique en oxygène |
| DCO | La demande chimique en oxygène |
| MES | Les matières en suspension |
| NH4 | L'azote ammoniacal |
| NO ₃ | Le nitrate |
| NO ₂ | Le nitrite |
| PO ₄ | Les ortho phosphates |
| MESBA | Matières En Suspension dans le Bassin d'aération |
| RECY | Recyclage des boues |
| ASM 1, ASM 2, ASM | Activated Sludge model |
| | Intelligence Artificialle |
| | |
| ACP | Analyse en composantes principales |
| K | Coefficient de correlation |
| REQM | Racine carrée de l'erreur quadratique moyenne |
| CO ₂ | Dioxyde de carbone |
| RNAR | Réseau de neurones artificiels récurrent |
| RP | Retro propagation |
| NTK | L'azote total kjeldahl |
| CRSTRA | Centre de Recherche Scientifique et Technique sur les Zones Arides |
| Norg | Azote organique |
| NH3 | Ammoniac |

Liste des abréviations

| АРНА | American Public Health Association |
|--------------------|--|
| AFNOR | Association Française de Normalisation |
| MATLAB | Matrix laboratory |
| LM | Levenberg-Marquardt, algorithme |
| HCO ₃ - | Carbonates |
| N2 | Azote gazeux |
| OMS | Organisation Mondiale de la Sante |

INTRODUCTION GENERALE

Introduction générale

La qualité de l'eau est une question très importante car c'est un facteur essentiel pour différents aspects de la vie, biotope et biocénose. Pour développer toute évaluation opérationnelle de la gestion et de la planification des ressources en eau, il est nécessaire de développer l'évaluation des paramètres de qualité de l'eau.

L'élimination des eaux usées a revêtu une importance croissante au début des années soixante-dix. La protection de milieu naturel et plus globalement notre cadre de vie, la pollution de l'eau inquiètent les dirigeants des pays développés qui commencent à prendre des mesures pour protéger l'environnement (Boukerroucha, 2011).

Les stations d'épuration des eaux usées sont considérées comme la première ligne dans la bataille de l'humanité pour éliminer la pollution de l'eau et restaurer la qualité de l'eau dans les cours d'eaux. Les stations d'épuration des eaux usées sont conçues et construites de manière à éliminer une quantité prédéterminée des polluants contenus dans les eaux usées brutes. La quantité de polluant à enlever répond aux critères standards de qualité de l'eau correspondant à la quantité de polluant qui peut être contenue dans l'eau et ne pas avoir un impact négatif surtout si on va la réutiliser. Une défaillance dans l'élimination de la quantité de polluants requis peut signifier non seulement que l'eau ne peut pas être utilisée comme prévu mais aussi des millions de dinars algériens perdus (Djeddou, 2014).

L'utilisation inappropriée de la station d'épuration peut entrainer des graves problèmes pour l'environnement et la santé publique. En effet, les effluents des stations d'épuration vers les cours d'eau récepteurs peuvent causer différentes maladies d'origine hydrique, non seulement aux êtres humains, mais également à la vie aquatique (Hamed *et al.*, 2004). Par conséquent, la STEP doit respecter les directives strictes en matière d'effluent avant son rejet dans les eaux réceptrices (Guclu *et al.*, 2010).

Dans nos jours, la technologie biologique la plus couramment utilisée pour les procédés de traitement des eaux usées est le procédé de boues activées qui se base sur l'utilisation de bactéries et d'autres microorganismes pour éliminer les contaminants en les assimilant.

L'une des principales caractéristiques qui ont affecté l'efficacité de la STEP est la fluctuation des influents par la perturbation du volume horaire et quotidien des charges polluantes des eaux usées (Metcalf et Eddy, 1978; Zeng *et al.*, 2004). Pour manipuler et contrôler ce phénomène, il est nécessaire d'utiliser des systèmes d'automatisation et des systèmes experts, tels que les systèmes de modélisation.

En fait, l'étude de la modélisation dynamique et du contrôle opérationnel des processus de traitement des eaux usées a fait l'objet d'une attention croissante dans les secteurs universitaire et industriel.

Plusieurs modèles de systèmes de traitement des eaux usées ont été développés et appliqués pour les stations de traitement des eaux usées, notamment l'ASP. Il existe deux groupes de modèles de traitement des eaux usées ; l'un est le modèle piloté par les données et le deuxième c'est le modèles mécaniste (ASM 1, ASM 2, ASM 2d et ASM 3) (Henze *et al.*, 1987). Cependant, ces modèles présentent des faiblesses en raison de la complication d'application pratique.

En raison de sa simplicité et de la précision de ses prévisions, l'utilisation de l'intelligence artificielle (IA), en particulier RNA, pour modéliser le processus de traitement des eaux usées, devient une alternative prometteuse réalisable avec le développement des capacités informatiques.

Les modèles de RNA utilisent un ensemble d'équations non linéaires pour déterminer des modèles complexes et des relations entre entrée et sortie. Ils peuvent donc être utilisés pour la prévision, l'estimation, la simulation et la classification (Suh *et al.*, 2009).

Plusieurs études ont montré les performances des RNA pour la prévision en ligne et la surveillance des processus dans un certain nombre des STEP des eaux usées. Pai (2008) on a donné un exemple dans lequel il utilise des réseaux de neurones pour prédire les effluents des stations de traitement des eaux usées en utilisant la qualité de l'influent. Güçlü *et al.* (2008) ont appliqué une prévision de RNA pour améliorer l'élimination du carbone dans le processus de traitement par BA. La capacité du modèle RNA pour la modélisation de la DBO₅ a été démontrée par Dogan *et al.* (2008). Mjalli et *al.* (2007) ont obtenu une bonne précision pour la prédiction dans le modèle RNA par l'annulation des effluents dans une STEP. Singh *et al.*

(2009) ont mis au point des RNA à rétro-propagation à trois couches pour la modélisation de la DBO₅ et de l'OD dans une STEP.

Giustolisi (2006) et Laucelli *et al.* (2009) ont déclaré que le RNA avait acquis de l'importance en raison de sa flexibilité et de la quantité considérable de données acquises par les réseaux de surveillance.

La station d'épuration de Touggourt, est une station d'épuration à boue activée traite les eaux usées domestiques. Elle est située au sud de la région d'Oued Righ (région aride au sud-est de l'Algérie). Elle s'étend sur une superficie de 5 hectares. Elle a été mise en service le 20/11/1993 et réhabilitée en 2004 et traite aujourd'hui une partie des rejets d'eaux usées déversées par la ville de Touggourt

Dans une station d'épuration, différentes variables explicatives clés peuvent être utilisées pour évaluer les performances de la station. Ces variables incluent la (CE), (PH), (T), (OD), (DBO₅), (DCO), (MES), (NH₄), (NO₃), (NO₂), (PO₄), (MESBA) et le (RECY).

L'objectif principal de cette étude est de suivre les performances épuratoires de la STEP de Touggourt (rendements en élimination de la pollution) par une série des données des analyses physico-chimique des eaux (brutes et traitées) et de développer une méthodologie basé sur la technique d'intelligence artificielle notamment l'application de réseau de neurone pour modéliser et contrôler cette STEP.

Pour atteindre les objectifs ci-dessus, l'étude est organisée en trois parties.

La première comprend la synthèse bibliographique, divisée en deux chapitres : le premier comprend des généralités sur les eaux usées et les différentes techniques d'épuration, et la deuxième traite la technique de l'intelligence artificielle notamment le réseau de neurone artificiel.

- La deuxième partie de ce travail s'intitule matériels et méthodes de travail ;
- La troisième partie met en exergue les résultats et les discussions avec trois chapitres :

- Traitement des données et étude de performances ;
- Application de l'analyse en composantes principales de (ACP) ;
- Application du modèle neuronale et interprétation des résultats.

Une conclusion générale décrit les principales conclusions et recommandations. Enfin, les références bibliographiques.

<u>PARTIE 01 :</u> SYNTHESE BIBLIOGRAPHIQUE

<u>Chapitre I</u>

I Généralités sur les eaux usées et les procédés d'épuration

I.1 Introduction

L'eau est indispensable. Tous les êtres vivants sur la Terre micro-organismes, plantes, animaux, humains et même notre cerveau sont principalement constitués de l'eau. En outre, l'eau est utilisée dans plusieurs domaines, tels que la consommation domestique, la production industrielle, l'irrigation, la production d'énergie ainsi que le nettoyage. Cependant, bien que plus de 70% de la surface de la terre soit recouverte de l'eau, seulement 0,5% de celle-ci convient à toutes les utilisations humaines (Gleick, 1996). Cette petite fraction diminue à mesure que l'agriculture, l'industrie et les besoins domestiques consomment de plus en plus de cette petite fraction, tandis que les déchets résultants sont des polluants qui dégradent la qualité de l'eau. En effet, les eaux usées contiennent une quantité considérable de matières organiques qui, si elles étaient rejetées en grande quantité dans les eaux réceptrices, entraîneraient une diminution des niveaux d'oxygène dissout et d'autres problèmes environnementaux. De plus, des composants toxiques peuvent être présents en raison de composants industriels (Metcalf et Eddy, 2003). Par conséquent, afin de protéger l'environnement et de préserver la vie, les eaux usées doivent être traitées de manière adéquate avant d'être rejetées. En particulier, le traitement biologique des eaux usées contribue à réduire la teneur en matières organiques des eaux usées. Bien qu'il existe plusieurs méthodes de traitement biologique des eaux usées, la plus souvent utilisée est le système de traitement biologique par boues activées (Spellman, 2003). Dans la section suivante, on discute plus de détails sur les eaux usées et leurs procédés d'épuration.

I.2 Les eaux usées

I.2.1 Définition

Selon Rejsek (2002), les eaux résiduaires urbaines (ERU), ou eaux usées, sont des eaux chargées de polluants, solubles ou non, provenant essentiellement de l'activité humaine. Une eau usée est généralement un mélange de matières polluantes répondant à ces catégories, dispersées ou dissoutes dans l'eau qui a servi aux besoins domestiques ou industriels (Grosclaude, 1999). Donc sous la terminologie d'eau résiduaire, on groupe des eaux d'origines très diverses qui ont perdu leurs puretés ; c'est-à-dire leurs propriétés

naturelles par l'effet des polluants après avoir été utilisées dans des activités humaines (domestiques, industrielles ou agricoles).

I.2.2 Origine des eaux usées

D'après Rodier (2005), On peut classer comme eaux usées, les eaux d'origine urbaines constituées par des eaux ménagères (lavage corporel et du linge, lavage des locaux, eaux de cuisine) et les eaux vannes chargées de fèces et d'urines ; toute cette masse d'effluents est plus ou moins diluée par les eaux de lavage de la voirie et les eaux pluviales. Peuvent s'y ajouter suivant les cas les eaux d'origine industrielle et agricole. L'eau, ainsi collectée dans un réseau d'égout, apparaît comme un liquide trouble, généralement grisâtre, contenant des matières en suspension d'origine minérale et organique à des teneurs extrêmement variables. En plus des eaux de pluies, les eaux résiduaires urbaines sont principalement d'origine domestique mais peuvent contenir des eaux résiduaires d'origine industrielle d'extrême diversité. Donc les ERU sont constituées par :

- Des eaux résiduaires ou eaux usées d'origine domestique, industrielle et/ou agricole
- Des eaux pluviales ou de ruissellement urbain.

I.2.2.1 Incidences des effluents sur la santé humaine

La mauvaise gestion des déchets ménagers est à l'origine du problème de la santé publique d'autant plus qu'ils constituent le facteur dominant de création de nids de production des vecteurs de menace de la santé comme les moustiques, mouches, cafards, souris...

Soumise à une urbanisation galopante et non planifiée, les villes des pays en développement apparaissent comme des espaces à risques potentiels sanitaires (Vilaginès, 2003).

En général, les déchets ménagers sont mal gérés à causes de l'absence d'infrastructures d'hygiène et d'assainissement de base, un manque de synergie d'action des acteurs... cela se traduit par une hygiène défectueuse qui offre des conditions bioécologiques favorables au développement de germes pathogènes (virus, bactéries, parasites) responsables de nombreuses maladies qui sévissent dans les quartiers les transformant de plus en plus en espace potentiellement `'épidémiogène" (un espace dont le fonctionnement génère des germes pathogènes qui provoquent des processus pathologiques et qui contribuent à faire apparaître et propager des phénomènes morbides au sein d'une population) (Vilaginès, 2003).

I.2.2.2 Contamination de l'eau potable

Étant donné que l'eau destinée à la consommation, est traitée et désinfectée, les éclosions fulgurantes de maladies d'origine hydrique sont rares. Mais des cas isolés de contamination microbienne de l'eau potable ayant pour origine des ERU, des eaux pluviales insuffisamment traitées ont été signalés. Des méthodes analytiques de plus en plus précises pour la détection des parasites et des virus ont donné naissance à une préoccupation à l'égard de l'innocuité d'eaux qui satisfont par ailleurs aux normes de qualité actuelles pour l'eau potable. Dans le cadre d'une étude épidémiologique portant sur le territoire de la Communauté urbaine de Montréal, Payment *et al* (2002) ont signalé que le risque de troubles gastro-intestinaux était plus élevé chez les personnes consommant de l'eau du robinet (incidence de 0,76) ayant pour origine des eaux de surface contaminées par des eaux usées, que celui déterminé pour les personnes ayant consommé la même eau, mais filtrée dans une unité domestique d'osmose inversée (incidence de 0,50) (Payment *et al.*, 2002).

I.2.2.3 Dégradation de l'environnement

Le rejet dans les eaux réceptrices d'ERU à charge de DBO₅ élevée peut provoquer une réduction immédiate de l'oxygène dissout dans la colonne d'eau de même que des effets à plus long terme (à l'échelle de mois ou d'années) découlant de l'accumulation de matériaux consommant l'oxygène dans les sédiments benthiques (demande d'oxygène des sédiments). Le manque d'oxygène dissout menace souvent les poissons et d'autres organismes en été car la solubilité de l'oxygène dans l'eau diminue avec l'augmentation de sa température. Mais sous les climats plus froids, lorsque les cours d'eau et les lacs sont recouverts de glace pendant plusieurs mois, ce manque d'oxygène dissout peut survenir en hiver, la couverture de glace prévenant toute ré aération (Payment *et al.*, 2002). La réduction de la concentration d'oxygène dissout peut avoir des incidences écologiques, comme un appauvrissement de la diversité biologique et la perte d'espèces. Ainsi que les concentrations élevées d'ammoniac pouvaient être à l'origine des hécatombes de poissons.

I.2.2.4 Eutrophisation des eaux réceptrices

Les eaux résiduaires urbaines apportent des substances nutritives (N et P) dans les plans d'eau récepteurs et favorisent ainsi l'eutrophisation. Comme les substances nutritives peuvent s'accumuler dans les sédiments benthiques et être libérées dans l'eau ultérieurement, la charge en substances nutritives a un effet cumulatif et un effet immédiat. Les incidences sur les écosystèmes aquatiques de l'ajout de substances nutritives sont sources d'importantes préoccupations car ces quantités supplémentaires peuvent favoriser la croissance des producteurs primaires (algues et plantes aquatiques à racines) à des niveaux nuisibles pour l'écosystème (p. ex., modification de la dynamique énergétique et de la structure du réseau trophique, modification de l'habitat et perte d'espèces). Ces changements écologiques peuvent, à leur tour, influer sur l'utilisation humaine des ressources aquatiques notamment en ce qui a trait aux activités récréatives, aux pêches et à la qualité de l'eau utilisée à des fins urbaines, industrielles et agricoles (Metcalf et Eddy, 2003).

I.2.2.5 Toxicité directe

La toxicité des effluents urbains est fonction de divers facteurs dont la taille et l'étendue des installations industrielles et urbaines, le type et l'efficacité des procédés de traitement et de désinfection et les caractéristiques physiques, chimiques et biologiques des eaux réceptrices. Dans le cas des ERU, la toxicité est généralement attribuée à l'ammoniac, au chlore résiduel total (effluents chlorés), au cyanure, aux sulfures, aux phénols, aux tensioactifs et à de nombreux métaux lourds (notamment le cuivre, le zinc, le chrome et le nickel) (Lynch *et al.*, 2002). D'autres facteurs, comme la température, le pH, la dureté, l'alcalinité et l'oxygène dissout, ont tendance à modifier la toxicité des constituants chimiques. En outre, les composés peuvent réagir entre eux et la toxicité résultante ne reflète pas nécessairement celle des composés individuels. Par conséquent, étant donné les nombreux facteurs et leurs interactions ainsi que la spécificité au site des effets dans le milieu récepteur, il est difficile de formuler des généralisations sur la toxicité des ERU (Bonjoch *et al.*, 2004).

I.2.3 Procédés d'épuration des eaux usées

I.2.3.1 But d'épuration

Le rôle d'une STEP des eaux usées est l'élimination de la pollution jusqu'à un niveau défini par la réglementation en vigueur pour assurer que le rejet des eaux traitées n'affecte pas

le milieu récepteur, et selon cette réglementation, les procédés de traitement sont mis en œuvre selon plusieurs niveaux de traitements.

Les niveaux de traitement d'une station sont définis selon la succession suivante :

Les prétraitements, le traitement primaire et le traitement secondaire. Lorsque l'eau traitée devra être rejetée en milieu particulièrement sensible, un traitement tertiaire est nécessaire.

Une STEP comporte généralement une phase de prétraitement, pendant laquelle les éléments les plus grossiers sont éliminés par dégrillage (pour les solides de grandes tailles), puis par flottaison/décantation (pour les sables et les graisses). Vient ensuite un traitement dit primaire, une décantation plus longue, pour éliminer une partie des MES. Des traitements physico-chimiques et/ou biologiques sont ensuite appliqués afin d'éliminer la matière organique. Ils sont généralement suivis d'une phase de clarification qui est encore une décantation. Enfin, un traitement des nitrates et des phosphates est exigé en fonction de la sensibilité du milieu récepteur. Il existe également des traitements dits extensifs, comme le lagunage, qui combinent des traitements biologiques, physiques et naturels.

I.2.3.2 Prétraitements

Les eaux usées brutes à leur arrivée à la station doivent généralement subir, un prétraitement qui composé d'un certain nombre d'opérations successives, uniquement physiques ou mécaniques. Il est destiné à extraire de l'eau usée, la plus grande quantité possible d'éléments dont la nature ou la dimension constitueront une gêne pour les traitements ultérieurs. Selon la nature des eaux à traiter et la conception des installations, le prétraitement peut comprendre les opérations : (le dégrillage), principalement pour les déchets volumineux, (le dessablage) pour les sables et graviers et (le dégraissage-déshuilage) pour les huiles et les graisses.

I.2.3.2.1 Dégrillage

Lors de l'opération de dégrillage, les eaux usées passent au travers d'une grille, dont l'espacement est déterminé de sorte qu'il puisse retenir matières grossières les plus volumineuses et flottantes charriées par l'eau brute, qui pourraient nuire, obstruer ou provoquer des bouchages dans conduites d'alimentation de l'installation, et nuire à l'efficacité de la station. Le dégrillage permet aussi de protéger la station contre l'arrivée intempestive des gros objets, les éléments retenus sont, ensuite, éliminés avec les ordures ménagères.

Cette opération est effectuée avant le poste de relevage afin de protéger les pompes ou les vis d'Archimède et de ne pas gêner leur fonctionnement.

I.2.3.2.2 Dessablage

Le dessablage a pour but d'extraire des eaux brutes les graviers, les sables, les verres brisés, les coquilles d'œufs, et les particules minérales plus ou moins fines ayant une vitesse de sédimentation sensiblement supérieure à la matière organique. Le dessablage est prévu pour protéger les équipements mécaniques à l'abrasion et à l'usure, de réduire la formation de dépôts dans les canalisations et les canaux, et de réduire la fréquence de nettoyage du digesteur qui est nécessaire en raison des particules accumulés.

Un but secondaire, mais cependant pas le moins extrêmement souhaitable du système d'élimination du sable est de séparer les grains de la matière organique dans les eaux usées. Cette séparation permet à la matière organique d'être traité dans les processus subséquents.

Dans un dessableur à flux horizontal, pour assurer l'élimination des grains et empêcher que la matière organique se dépose, trois conditions doivent être remplies (Steel *et* Mcghee, 1979) :

- La vitesse d'écoulement à la sortie du dessableur doit être égale à la vitesse de sédimentation des particules inertes ;
- La vitesse horizontale doit être inférieure à la vitesse d'érosion des particules inertes ;
- La vitesse horizontale doit être supérieure à la vitesse de décantation des particules organiques.

I.2.3.2.3 Dégraissage déshuilage

Le déshuilage-dégraissage se rapporte à l'extraction de toutes les matières flottantes d'une densité inférieure à celle de l'eau. Ces matières sont de natures très diverses et leurs quantités s'estime par la mesure des « matières extractibles par solvants ». La teneur des eaux usées en matières extractibles est de l'ordre de 30 à 75 mg/L. Néanmoins, certains rejets industriels (abattoirs, laiteries...) peuvent élever ces valeurs à 300- 350 mg/L.

Les huiles et graisses, lorsqu'elles ne sont pas émulsionnées, sont séparées sous forme de boues flottantes dans des ouvrages comportant une zone d'aération où les bulles d'air augmentent la vitesse de montée des particules grasses et une zone de tranquillisation où s'effectue la récupération.

Le temps de séjour dans ce type d'ouvrage est de 5 à 12 min. Le débit d'air insufflé est de l'ordre de 0,2 m3 par mètre cube d'eau et par heure.

Le plus souvent, les fonctions de dessablage et de déshuilage sont combinées dans un même ouvrage (Gaïd, 1993).

I.2.3.3 Traitements Primaires

Le traitement "primaire" fait appel à des procédés physiques naturels, filtration et décantation plus ou moins aboutie, éventuellement assortie de procédés physico-chimiques, tels que la coagulation- floculation.

I.2.3.4 Traitements secondaires (traitement biologique)

L'épuration biologique a pour but d'éliminer la matière polluante biodégradable contenue dans l'eau domestique (décantée ou non) en la transformant en matières en suspension : micro-organismes et leurs déchets, plus facilement récupérables (Dufournet, 1974 ; Gaïd, 1993).

La dégradation peut se réaliser par voie aérobie (en présence d'oxygène), pour les eaux usées et en anaérobie (en l'absence d'oxygène) pour le traitement des boues.

Le traitement biologique classique des eaux domestiques s'effectue par voie aérobie, qui consiste à dégrader les impuretés grâce à l'action d'une biomasse épuratrice, à laquelle doit être fourni l'oxygène nécessaire à son développement (Gaïd, 1993).

I.2.3.4.1 Traitement biologique intensifs :

Les techniques les plus développées au niveau des STEP urbaines sont des procédés biologiques intensifs. Le principe de ces procédés est de localiser sur des surfaces réduites et d'intensifier les phénomènes de transformation et de destruction des matières organiques que l'on peut observer dans le milieu naturel. Trois grands types de procédés sont utilisés :

- ✤ Les lits bactériens ;
- ✤ Les disques biologiques ;
- Les boues activées.

I.2.3.4.1.1 Boues activées

C'est le procédé le plus répandu actuellement pour l'épuration des eaux résiduaires urbaines des petites, moyennes ou grandes collectivités. Le procédé à boues activées est un système en continu dans lequel des micro–organismes sont mis en contact avec des eaux usées renfermant des matières biodégradables pendant un temps suffisant. Ces amas biologiques sont maintenus en agitation au sein de l'eau de façon à assurer un contact avec toute la partie de l'effluent. L'oxygénation est fournie en quantités suffisantes par des aérateurs (Gomella et Guerree, 1982). Ainsi, dans le bassin d'aération, en présence d'oxygène, les micro-organismes vont se développer et se reproduire aux dépens des matières biodégradables formant ainsi des flocons décantables, orientés par la suite vers un clarificateur. A la sortie une eau traitée et des boues seront produites, une partie de ces boues sera expédiée vers les organes de traitement de boues et l'autre partie réintroduite dans l'aérateur (Urios, 2005).

I.2.3.4.2 Traitement biologique extensifs :

I.2.3.4.2.1 Le lagunage

Un traitement par lagunage comprend en général trois types de bassins : un bassin anaérobie, un bassin facultatif et un bassin de maturation. Le bassin anaérobie permet de diminuer la charge en matière organique. L'anaérobie est obtenu en apportant un effluent très chargé en matière organique. Dans ces lagunes, une profondeur importante est en principe un élément favorable au processus (5 à 6 m, par exemple). Ce bassin n'est applicable que sur des effluents à forte concentration et, le plus souvent, à titre de prétraitement avant un deuxième stade d'épuration de type aérobie surtout dans les pays à climat chaud où le terrain est disponible à coût raisonnable. Le lagunage utilise des mécanismes naturels pour traiter les eaux usées. Il est fort développé dans les petites communes rurales, en raison de sa rusticité et de sa performance d'épuration honorable. Par contre, ces procédés conviennent moins bien aux communes plus grandes, vu les grandes surfaces de bassins nécessaires.

Le bassin facultatif permet le développement d'algues photosynthétiques qui vont produire de l'oxygène nécessaire au développement des bactéries aérobies. Cet apport peut être complété exceptionnellement par des aérateurs pour stimuler l'activité biologique et diminuer les surfaces. Il existe deux types de bassins facultatifs, selon les végétaux qu'ils comprennent :

- Les bassins à microphytes : ils contiennent des algues microscopiques (essentiellement les algues vertes ou bleues),
- Les bassins à macrophytes : ils contiennent des végétaux macroscopiques, sous formes libres (ex. lentilles d'eau) ou fixées (ex. roseaux).

I.2.3.4.2.2 La filtration /percolation

La filtration ou percolation consiste à traiter les eaux usées par l'intermédiaires d'un sol ou d'un massif filtrant (Vasel, 2007). On filtre les effluents à raison de quelques centaines de litres d'effluents par mètres carrées de massif filtrants et par jours. Deux mécanismes entrent en jeu :

- Filtration de MES : plus le sable est grossier, plus la fixation de MES se fera en profondeur et les MES finissent par colmater le filtre. Pour lutter contre le bouchage du massif filtrant, il faut alterner phase de filtration et phase de séchage, l'élimination de MES permet l'élimination des microorganismes qui y sont fixées ;
- L'adsorption des bactéries libres par les grains du sable de filtre il se forme alors un film biologique contaminé, surtout dans la partie supérieure, ce film va permettre une dégradation microbienne de la matière organique et des substances dissoutes dans l'effluent (PO₄, NO₃, etc). Cette dégradation consomme de l'OD et produit du CO₂, il faut donc aérer régulièrement le film pour éviter l'asphyxie du milieu.

Les techniques de filtrations/percolation permettent l'élimination des « gros » microorganismes (protozoaires et helminthes) par filtration/adsorption au début du massif filtrant. L'élimination des virus et des bactéries est fonction du milieu poreux.

I.2.3.4.2.3 La phytoépuration ou filtres plantés de macrophytes

L'usage des végétaux aquatiques dans le traitement des eaux usées (phytoépuration) provient de l'observation des rôles de zones humides dans la préservation de la qualité des milieux aquatiques d'où le nom de marais filtrant artificiel. Son utilisation pour le traitement des eaux usées remonte à une centaine d'années. Les marais filtrants artificiels utilisés pour le traitement des eaux usées domestiques sont aussi appelés filtres plantés de macrophytes, ils sont constitués d'un lit de sol ou d'un autre milieu, tel que du gravier ou du sable, implanté avec du macrophytes et qui est inondé ou maintenu en condition saturée (niveau d'eau près de la surface) (Cors, 2007). L'environnement étant ainsi propice à l'établissement de plantes adaptées aux conditions de sol saturé et produisant un important réseau de racines dans le milieu. Le traitement des eaux usées s'effectue au moyen d'une combinaison de processus physiques, chimiques et biologiques, incluant la sédimentation, la précipitation, l'adsorption sur les particules de sol, l'assimilation par les plantes et les transformations microbiologiques (Cors, 2007).

I.2.3.5 Traitements tertiaires

Les traitements tertiaires ou d'affinage visent principalement l'élimination de la pollution azotée et phosphatée ainsi que la pollution bactériologique des eaux usées domestiques, ayant déjà subi au préalable des traitements primaires et secondaires qui s'avèrent insuffisants pour répondre aux normes de rejet. Pour cela les traitements tertiaires s'imposent et deviennent plus que nécessaires, afin de garantir une meilleure protection des milieux récepteurs.

I.2.3.5.1 Traitement des boues

La réduction de volume est classiquement obtenue à travers des opérations de séparation de phases liquide/solide par décantation, filtration ou évaporation rencontrées dans les techniques d'épaississement, de déshydratation et de séchage thermique. La dégradation des matières organiques de la boue par des procédés biologiques (digestion, compostage) ou

thermiques (incinération à 850 °C, oxydation par voie humide de boues liquides épaissies sous 45 bar à 250 °C) conduira également à un volume final moindre.

La stabilisation sera concrètement obtenue en ralentissant, voire en supprimant, la biodégradation putride des matières organiques de la boue, à travers différentes voies, biologique, chimique ou physique avant (phase liquide) ou après (phase pâteuse) l'étape de déshydratation.

La déshydratation constitue souvent l'étape ultime de la filière de traitement des boues :

Une siccité minimale peut en effet être imposée contractuellement (généralement > 30 %) en vue de l'évacuation de la boue ou être requise en vue d'une incinération dans des conditions d'auto combustibilité.

Les principaux débouchés des boues produites par les STEP restent la possibilité d'une valorisation agricole (Gaïd, 1993).

I.3 Conclusion

Ce chapitre a examiné le contexte des eaux usées et les STEP. Il a présenté l'importance du traitement des eaux usées, son histoire et la structure des STEP classiques.

Grâce aux développements des procédés de traitements, la production des eaux usées issues de l'activité humaine, n'est plus un problème, mais une source alternative pour une récupération des eaux traitées afin de les réutiliser pour répondre aux besoins croissants dans l'agriculture et l'industrie.
CHAPITRE II

II Techniques de l'intelligence artificielle (IA)

II.1 Introduction

Les techniques d'intelligence artificielle concernent plusieurs domaines liés à la simulation de l'intelligence humaine dans une machine informatique. Le principal attrait des techniques d'IA réside dans le fait qu'elles sont capables de représenter des systèmes dotés de caractéristiques non linéaires, sans la tâche difficile de traiter avec des mathématiques déterministes non linéaires. Les stratégies existantes de modélisation de l'IA peuvent être divisées en trois catégories, à savoir « boîte blanche », « boîte noire » et « boîte grise » en fonction du type de connaissances utilisé pour l'élaboration du modèle. Dans les stratégies de modélisation de type "boîte blanche", également appelées modèles déterministes, le développement du modèle est principalement motivé par la connaissance des mécanismes et des équilibres pertinents. En d'autres termes, les équations du modèle sont élaborées à partir des équations générales de bilan appliquées à la masse et à d'autres quantités conservées, ce qui donne un ensemble d'équations différentielles. Un modèle de type "boîte noire" ou "entrées-sorties" est principalement basé sur les données mesurées obtenues à partir du processus. Cependant, le modèle de la boîte noire est aussi bon que les données qui ont été utilisées pour les calibrer (Gernaey et al., 2004). En effet, on ne pense pas que les modèles à « boîte noire » avaient des propriétés d'extrapolation ; par conséquent, il faut obtenir un grand nombre de données qui couvrent l'éventail des fluctuations des variables d'entrée pertinentes pour la modélisation de processus. Enfin, le modèle "boîte grise" peut être défini comme une combinaison appropriée des modèles "boîte noire" et "boîte blanche", de manière que le modèle soit développé à l'aide de techniques pilotées par les données et puisse en même temps extraire des données.

Parmi les outils et techniques d'IA couramment utilisés figurent les RNA (boîte noire), la logique floue (boîte grise), les systèmes experts (boîte blanche) et une grande variété de techniques de recherche telles que les algorithmes génétiques. Cependant, la suite de ce chapitre ne présente que le RNA, car il est au cœur de ce travail.

II.2 Le réseau de neurone artificiel

Un réseau de neurones est un système de traitement d'informations qui tente de simuler la façon dont un cerveau humain traite les signaux électriques.

Grace à leur traitement parallèle de l'information et à leurs mécanismes inspirés des cellules nerveuses (neurones), ils infèrent des propriétés émergentes permettant de solutionner des problèmes qualifiés de complexes. Ils ont été développés pour résoudre les problèmes de contrôle, de la reconnaissance des formes et des mots, de prise de décision, de mémorisation comme une alternative à l'IA. Les travaux dans le domaine des RNA ont été motivés par la fascination de l'humanité à la complexité du cerveau humain.

II.2.1 Le neurone biologique

Le neurone biologique est une cellule vivante spécialisée dans le traitement des signaux électriques. Les neurones sont reliés entre eux par des liaisons appelées axones. Ces axones vont eux-mêmes jouer un rôle important dans le comportement logique de l'ensemble. Ces axones conduisent les signaux électriques de la sortie d'un neurone vers l'entrée (synapse) d'un autre neurone. Les neurones font une sommation des signaux reçus en entrée et en fonction du résultat obtenu vont fournir un courant en sortie (El Kebir, 2009 ; Karima, 2006). (Figure1) La structure d'un neurone se compose de trois parties :

- La somma : ou cellule d'activité nerveuse, au centre du neurone.
- L'axone : attaché au somma qui est électriquement actif, ce dernier conduit l'impulsion conduite par le neurone.
- Dendrites : électriquement passives, elles reçoivent les impulsions d'autres neurones.



Figure 1 : Structure d'un neurone biologique (Graupe, 2007).

II.2.2 Le neurone formel (Artificiel)

Le neurone artificiel (ou cellule) est un processeur élémentaire. Il reçoit un nombre variable d'entrées en provenance de neurones appartenant à un niveau situé en amont (on parlera de neurones "amont"). A chacune des entrées est associé un poids w représentatif de la force de la connexion. Chaque processeur élémentaire est doté d'une sortie unique, qui se ramifie ensuite pour alimenter un nombre variable de neurones appartenant à un niveau situé en aval (on parlera de neurones "avals").

A chaque connexion est associé un poids (Figure2).



Figure2 : Neurone formel (artificiel).

Le neurone formel doit être capable de :

1. Recevoir en entrée différentes informations provenant des neurones

environnants;

- 2. Analyser ces informations, de manière à envoyer en sortie une réponse ;
- 3. Ajuster cette réponse avant de l'envoyer aux neurones suivants.

Un neurone est l'unité de traitement de base d'un réseau de neurones. La figure 2 illustre un neurone d'entrée avec les entrées x1 à xn constituent le vecteur d'entrée x. La fonction f est appelée fonction d'activation du neurone. Chaque entrée xi est pondérée par un poids wi, ce sont les poids des entrées qui vont permettre au neurone d'apprendre et de modifier sa sortie au fur et à mesure de l'apprentissage. La sortie du neurone est y = f (w.x), où w.x représente le produit scalaire du vecteur poids par le vecteur des entrées. Le biais b fournit une variable supplémentaire qui peut être ajusté pour obtenir la performance du réseau désiré. Cette somme est l'argument de la fonction de transfert f, qui prend l'argument et produit une sortie. La fonction de transfert qui limite la gamme d'amplitude admissible du signal de sortie à des valeurs finies dans l'intervalle [-1, 1]. Les poids wi et le biais b sont des paramètres ajustables qui permettent au réseau de neurones de présenter le comportement désiré (Haykin, 1999 ; Krenker *et al.*, 2011 ; Demuth *et al.*, 2013).

II.2.3 Structure de réseau de neurones artificiels

Haykin (1999), définit un réseau de neurones, comme un processeur distribué massivement parallèle qui est composé d'unités de traitement simples, ayant la capacité de stocker des connaissances empiriques et de les rendre disponible pour l'utilisation.

Un réseau de neurones artificiel (figure3) est en général composé d'une succession de couches dont chacune prend ses entrées sur les sorties de la précédente. Chaque couche (i) est composée de Ni neurones, prenant leurs entrées sur les Ni-1 neurones de la couche précédente. À chaque synapse est associé un poids synaptique, de sorte que les Ni-1 sont multipliés par ce poids, puis additionnés par les neurones de niveau i, ce qui est équivalent à multiplier le vecteur d'entrée par une matrice de transformation. Mettre l'une derrière l'autre (les différentes couches d'un réseau de neurones) reviendrait à mettre en cascade plusieurs matrices de transformation et pourrait se ramener à une seule matrice, produit des autres, s'il n'y avait à chaque couche, la fonction de sortie qui introduit une non linéarité à chaque étape (Touzet, 1992). Ceci montre l'importance du choix judicieux d'une bonne fonction de sortie : un réseau de neurones dont les sorties seraient linéaires n'aurait aucun intérêt.



Figure 3 : Réseau de neurones multicouches (PMC) avec deux couches cachées (Rustum, 2009).

II.2.4 Architecture des réseaux de neurones artificiels

En générale, Il y'a essentiellement deux types de topologies de RNA, à savoir réseau de neurones artificiel directe ou (par alimentation avant) (Feed- forward) et le réseau de neurones artificiels récurrent (RNAR) (Bishop, 1995 ; Demuth *et al.*, 2013).

II.2.4.1 Réseau de neurones artificiels direct (à alimentation avant)

Un réseau de neurones artificiels direct est un RNA où les connexions entre les unités ne forment pas un cycle dirigé. Ceci est différent de réseaux de neurones récurrents.

Le réseau neuronal direct a été le premier et le plus simple type des RNA imaginés. Dans ce réseau, l'information se déplace dans une seule direction, vers l'avant, à partir des nœuds d'entrée, par l'intermédiaire des nœuds cachés (le cas échéant) et de nœuds de sortie. Il n'y a pas de cycles ou des boucles dans le réseau.



Figure 4 : Réseau de neurones artificiels directs (par alimentation avant) (Feed-Forward) (Kriger, 2007).

II.2.4.2 Réseau de neurones artificiels récurrent

Un réseau de neurones artificiels récurrent (RNAR) est une classe de réseau de neurones, où les connexions entre les unités forment un cycle dirigé. Ceci crée un état interne du réseau qui lui permet un comportement temporel dynamique. Contrairement aux réseaux de neurones directs (feed-forward), les RNAR peuvent utiliser leur mémoire interne pour traiter les séquences arbitraires d'intrants. Cela les rend applicable à des tâches telles que la reconnaissance de l'écriture manuscrite connecté non segmenté, où ils ont obtenu les meilleurs résultats connus (Graves *et al.*, 2009).



Figure 5: Réseau de neurones récurrent (Kriger, 2007).

Trouver une configuration optimale du réseau est l'un des problèmes les plus importants dans la conception d'un RNA. Les paramètres du réseau de neurones suivants doivent être sélectionnés :

- Le nombre de couches cachées ;
- Le nombre de neurones dans chaque couche cachée ;
- ✤ Le type de fonction d'activation, qui peut varier dans la même couche.

Ces paramètres doivent être choisis avec soin particulier pour éviter les phénomènes de sur- apprentissage et sous-apprentissage et réaliser une phase d'apprentissage convergent (Khan et Ondrusek, 2000). En raison d'un manque d'informations sur les variables représentées dans la couche cachée, les RNA sont considérés par la plupart des chercheurs comme un modèle « boîte noire » (Hu, 1997 ; Moshiri et Cameron, 2000 ; Olden et Jackson, 2002).

II.2.5 Fonction de transfert

La sortie du neurone est déterminée par l'activation (ou transfert) en fonction du neurone. Le but de la fonction d'activation est d'introduire une non-linéarité dans le réseau. Cette non-linéarité associée aux interconnexions des neurones accomplit une cartographie des données propre à partir de signaux d'entrée pour les activités de sortie correspondants (Zupan et Gasteiger, 1993). Les principales fonctions de transfert sont présentées dans le tableau 1 ci-dessous.

Il existe d'autres fonctions d'activation moins populaires tels que la fonction satlins (linéaire saturée symétrique), la fonction poslin (linéaire positive), et la fonction Softmax (Burton, 2006). Le succès ou l'échec du processus d'apprentissage dépend du fait que la fonction d'activation est dérivable ou non sur la base de cette exigence, la fonction non-linéaire la plus populaire est la fonction sigmoïde pour la couche cachée en raison de son dérivé facilement calculé (Rustum, 2009). Parmi les fonctions non-linéaires les plus populaires, la fonction logistique et la fonction tangente hyperbolique (Demuth *et al.*, 2013).

| Nom de la fonction | Relation d'entrée/sortie | Icône |
|-----------------------------|---|--------------|
| seuil | $\begin{array}{cc} a=0 & \operatorname{si} n < 0 \\ a=1 & \operatorname{si} n \geq 0 \end{array}$ | |
| seuil symétrique | $\begin{array}{ccc} a=-1 & \mathrm{si}\; n<0\\ a=1 & \mathrm{si}\; n\geq 0 \end{array}$ | [] |
| linéaire | a = n | \mathbb{N} |
| linéaire saturée | a = 0 si n < 0 $a = n \text{si } 0 \le n \le 1$ a = 1 si n > 1 | |
| linéaire saturée symétrique | $ \begin{array}{ll} a=-1 & \sin n<-1 \\ a=n & \sin -1\leq n\leq 1 \\ a=1 & \sin n>1 \end{array} $ | \vdash |
| linéaire positive | a = 0 si $n < 0a = n si n \ge 0$ | \square |
| sigmoïde | $a = \frac{1}{1 + \exp^{-n}}$ | \sum |
| tangente hyperbolique | $a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$ | F |
| compétitive | a = 1 si <i>n</i> maximum a = 0 autrement | С |

Tableau 1 : Les principales fonctions de transfert

II.2.6 Apprentissage des réseaux de neurones

L'apprentissage est la propriété la plus intéressante des RNA. L'apprentissage est une phase du développement d'un réseau de neurones durant laquelle le comportement du réseau est modifié jusqu'à l'obtention du comportement désiré. Au niveau des algorithmes d'apprentissage, on peut trouver deux grandes classes : l'apprentissage supervisé et l'apprentissage non supervisé.

II.2.6.1 Règle d'apprentissage du perceptron

Le perceptron est un réseau oriente de neurones artificiels organisé en couches et ou l'information voyage dans un seul sens, de la couche d'entrée vers la couche de sortie. Faire apprendre un neurone, consiste à régler (estimation adéquate) ses poids de manière à ce que la sortie du neurone évolue dans le sens que l'on souhaite sans bien sûr changer les entrées. Il s'agit dans ce cas d'un apprentissage supervisé (on spécifie la sortie que l'on souhaiterait voir,

et le neurone va adapter ses poids pour essayer de s'approcher de la valeur de la sortie. On appelle cette sortie la cible du neurone). La règle d'apprentissage du perceptron va donc faire évoluer les poids du neurone vers une cible que l'on a spécifiée (Heaton, 2008 ; Prevost, 2012).

II.2.6.2 Apprentissage supervisé

Dans l'apprentissage supervisé, un ensemble d'entrées et sorties adéquates sont utilisées pour former le réseau. Le réseau produit ensuite ses propres sortis. Ces sorties sont comparées avec les sorties correctes et la différence (erreur) est utilisée pour modifier les poids. Ce type d'apprentissage est utilisé dans notre thèse en raison de la nature des données disponibles (Haykin, 1999 ; Krenker, 2011).

II.2.6.3 Apprentissage non supervisé

Apprentissage non supervisé ou auto-organisation est le cas où un réseau développe ses propres règles de classification par l'extraction d'informations à partir des entrées présentées au réseau.

II.2.7 Algorithme d'apprentissage

II.2.7.1 Algorithme d'apprentissage de retro propagation (**RP**)

La rétro-propagation est un exemple d'apprentissage supervisé. Dans de nombreuses applications des RNA sont essentiellement des réseaux fonction de cartographie. Le problème avec la cartographie est de découvrir la relation entre les entrées et les sorties du réseau cible. Cette relation est selon toute vraisemblance, probablement non linéaire, comme c'est le cas avec le procédé par boues activées. L'algorithme de rétro propagation a été effectivement utilisé dans de nombreuses applications nécessitant une cartographie d'entrée- sortie non linéaire. Avec la possibilité d'approcher une fonction non-linéaire continue, le réseau de RP a cartographie extraordinaire (prévision) capacités (Rui et El Keib, 1995).

Le terme RP se rapporte à la manière dont le gradient est calculé pour les réseaux multicouches non linéaires. La généralisation de la règle d'apprentissage Widrow- Hoff aux réseaux multicouches et des fonctions de transfert non linéaires différentiables créé l'algorithme de rétro propagation. L'algorithme est basé sur la plus grande pente (descente de gradient) techniques accordés à chacune des couches du réseau par la règle de la chaîne (Ng, 1997). La dérivée partielle de la fonction d'erreur par rapport aux poids est calculée en tant que :

$$\nabla E(\mathbf{k}) = \frac{\partial E(\mathbf{k})}{\partial W(\mathbf{k}-1)}$$
(1)

Avec:

W: Vecteur des poids.

E(k) : Fonction de K^{ème} erreur

La fonction d'erreur est définie comme :

Où Y_p est la sortie du réseau

Y_m est la sortie mesuré

La constante $\frac{1}{2}$ est choisie pour faciliter le calcul d'une dérivée de la fonction de coût qui est essentiel dans l'estimation des paramètres.

L'objectif est de minimiser la fonction d'erreur E(k) en prenant le gradient d'erreur par rapport au vecteur des poids, W doit être adapté Les pondérations sont mises à jour à l'aide:

$$w(k) = W(k-1) + \frac{\partial E(k)}{\partial W(k-1)}....(3)$$

Avec :

W est le poids

est le taux d'apprentissage et:

$$\frac{\partial \mathbf{E}(\mathbf{k})}{\partial \mathbf{W}(\mathbf{k}-1)} = -\mathbf{e}(\mathbf{k})\frac{\partial \mathbf{E}\mathbf{y}_{\mathrm{m}}(\mathbf{k})}{\partial \mathbf{W}(\mathbf{k}-1)} = -\mathbf{e}(\mathbf{k})\frac{\partial O_{0}(\mathbf{k})}{\partial \mathbf{W}(\mathbf{k}-1)}....(4)$$

Il devient évident que le succès de la technique de rétro-propagation dépend de trois facteurs : le taux d'apprentissage , la distance entre la production réelle et la production

prévue, et la fonction d'activation. Le taux d'apprentissage détermine la vitesse du réseau se règle en réponse à des erreurs. Si le taux d'apprentissage est trop élevé, les poids ajustés oscillent entre des points avec des hauteurs similaires sur la surface d'erreur (Park, 1996) et le réseau ne converge pas vers un minimum. Si le taux d'apprentissage est trop faible, l'entrainement est lent et le réseau est piégé dans un minimum local. Le taux d'apprentissage devrait se situer dans la gamme 0 1.

L'algorithme de rétro propagation pourrait se résumer comme suit :

- Initialiser les poids des liens entre les neurones. Souvent une valeur entre 0 et 1.
- Application d'un vecteur entrées-sorties à apprendre.
- Calcul des sorties du RNA à partir des entrées qui lui sont appliquées et calcul de l'erreur entre ces sorties et les sorties idéales à apprendre.

• Correction des poids des liens entre les neurones de la couche de sortie et de la première couche cachée selon l'erreur présente en sortie.

• Propagation de l'erreur sur la couche précédente et correction des poids des liens entre les neurones de la couche cachée et ceux en entrées.

• Boucler à la 2e étape avec un nouveau vecteur d'entrées-sorties tant que les performances du réseau de neurone (erreur sur les sorties) ne sont pas satisfaisantes.

Un seul passage des trois premières étapes ci-dessus est connu comme une itération. La formation dure habituellement dans certains cas, des milliers d'itérations que l'erreur du réseau est inférieure à un certain seuil. De cela, il est évident que la modélisation des réseaux de neurones est extrêmement chronophage (Haykin, 1999 ; Krenker, 2011).

II.2.8 Généralisation

Il est rarement utile d'avoir un simple RNA qui mémorise un ensemble de données (approximation de fonction), puisque de nombreux algorithmes pour consultation peuvent faire la mémorisation beaucoup plus efficacement.

En règle générale, il est nécessaire que le RNA soit capable d'effectuer avec précision une généralisation sur de nouvelles données. La généralisation peut être définie comme la capacité du réseau à faire une réponse hypothétique à une entrée qui n'a pas été vu avant, et de faire une réponse raisonnable (Park, 1996).

II.2.8.1 Sur-apprentissage et sous-apprentissage

Sur-apprentissage ou sous-apprentissage est essentiellement la façon dont le RNA réalise des prédictions pour les cas qui ne sont pas dans le jeu de données d'entrainement, comme d'autres méthodes d'estimation non linéaire.

En règle générale, le but de l'entrainement est de faire des prédictions pour données de sortie à venir dans lequel seules les entrées du réseau sont connues. Le résultat de la formation du réseau conventionnel est un ensemble unique de poids qui peuvent être utilisés pour faire de telles prédictions (Kriger, 2007).



Figure 6 : Présentation graphique du phénomène de sur-apprentissage "Cas d'approximation d'une fonction" (Parizeau, 2006).

Par exemple, la figure 6 illustre le problème du sur-apprentissage dans le contexte d'une tache d'approximation d'une fonction. La droite en pointillé montre une fonction linéaire que l'on voudrait approximer en ne connaissant que les points noirs. La courbe en trait plein montre ce qu'un réseau hypothétique pourrait apprendre. On constate que la courbe passe par tous les points d'entraînement et donc que l'erreur est nulle. De toute évidence, ce réseau ne généralisera pas bien si l'on échantillonne d'autres points sur la droite (Parizeau, 2006).

Le phénomène de sous-apprentissage est le contraire du sur-apprentissage. Cela se produit lorsque le modèle est incapable de capturer la variabilité des données. Le réseau de neurones artificiels n'aura aucun pouvoir prédictif, ni qu'il sera en mesure de cartographier correctement les données d'entraînement. Ceci est le résultat d'un sous-apprentissage, ou de tenter d'utiliser ce modèle qui est trop simple pour décrire un ensemble de données. La figure 7 présente graphiquement le phénomène de sous-apprentissage.

La compréhension de ces deux phénomènes permet d'aller dans l'espace entre les deux extrêmes. C'est dans cet intervalle lorsque le modèle a un pouvoir prédictif de l'ensemble des données de validation (Parizeau, 2006).



Figure 7 : Exemple du phénomène de sous-apprentissage (Parizeau, 2006).

Une solution à ce problème consiste à utiliser un autre critère d'arrêt et basée sur une technique dite de validation croisée (en anglais «cross-validation»). Cette technique consiste à utiliser deux ensembles indépendants de données pour entrainer le réseau : un pour l'apprentissage (l'ajustement des poids) et l'autre pour la validation, c'est-à-dire le calcul d'un indice de performance (Parizeau, 2006).

II.2.8.2 Arrêt prématuré

Le but de l'entrainement du réseau n'est pas d'apprendre la représentation exacte des données d'entrainement par lui-même, mais de construire un modèle du processus qui génère ces données, il est important que le réseau présente une bonne généralisation. L'arrêt prématuré est la technique la plus utilisée pour résoudre le problème du sur-apprentissage et de trouver le réseau ayant la meilleure performance sur de nouvelles données. Cet arrêt prématuré implique la séparation des données disponibles en 3 sous-ensembles : ensemble d'apprentissage, ensemble de validation et ensemble de test. Au cours de l'entrainement du

réseau, le message d'erreur sur l'ensemble de validation est contrôlé, ainsi que l'erreur sur l'ensemble d'apprentissage (Bishop, 1995 ; Demuth *et al.*, 2013).

Le critère d'arrêt consiste alors à stopper l'apprentissage lorsque l'indice de performance calculée sur les données de validation cesse de s'améliorer pendant plusieurs périodes d'entrainement. La figure 8 illustre le critère de la validation croisée dans le cas d'un indice de performance que l'on cherche à minimiser. La courbe en pointillées de ce graphique représente l'indice de performance d'un réseau Hypothétique calculée sur les données d'apprentissage, alors que la courbe en trait plein montre le même indice mais calculée sur les données de validation. On voit qu'il peut exister un moment au cours de l'apprentissage ou l'indice en validation se détériore alors que le même indice continue à s'améliorer pour les données d'entrainement. C'est alors le début du « sur-apprentissage » (Parizeau, 2006 ; Adeloye et De Munari, 2006).



Figure 8 : Illustration du critère de la validation croisée (Parizeau, 2006)

Il ya aussi beaucoup d'autres problèmes importants qui sont difficiles qu'un réseau de neurones est incapable de les apprendre sans avoir à mémoriser l'ensemble de la formation, tels que :

- Prévoir des nombres aléatoires ou pseudo-aléatoires.
- Factoriser de grands nombres entiers.
- Déterminer si un grand nombre entier est premier ou composite.

II.2.9 Avantage des réseaux de neurones artificiels

Les réseaux de neurones sont caractérisés par les capacités positives suivantes :

- Pour résoudre des tâches où il est avantageux d'utiliser une machine, mais où il est impossible d'avoir de réponses du programme à toutes les issues possibles.
- 2) Apprendre des exemples limités « les réseaux de neurones artificiels sont particulièrement adaptés à des problèmes où la solution est complexe et difficile à définir, mais fournit une abondance de données à partir de laquelle une réponse peut être apprise (Tarassenko, 1998).
- 3) Capacité à généraliser, à savoir la carte d'entrées semblables aux sorties similaires : les réseaux de neurones sont capables d'interpoler à partir d'une expérience d'apprentissage précédent. Avec une conception soignée, un RNA peut être formé pour donner la bonne réponse à des données qu'il n'a pas préalablement rencontrées (Tarassenko, 1998).
- Possibilité de cartographier les fonctions linéaires et non linéaires : la cartographie non linéaire donne souvent aux RNA l'avantage de traiter les problèmes complexes du monde réel.
- 5) Efficacité de calcul : la formation d'un RNA pour un calcul intensif peut être modeste (Tarassenko, 1998), mais les exigences de calcul d'un RNA entièrement formés peuvent être amélioré pour des problèmes plus importants, grâce à un traitement parallèle (Haykin, 1999).
- 6) Robuste en présence de bruit.
- 7) Capacités multi variables.

II.3 Revue de littérature

Un fonctionnement plus sûr et un contrôle adéquat d'une STEP peuvent être atteints par le développement d'un outil de modélisation pour prédire la performance de la station sur la base des observations passées de certains paramètres clés des eaux usées. Une STEP des eaux usées implique plusieurs procédés physiques, biologiques et chimiques complexes.

Souvent, ces procédés présentent des comportements non linéaires qui sont difficiles à décrire par des modèles mathématiques linéaires (Plazl *et al.*, 1999). En outre, la variabilité des caractéristiques influentes, en termes de taux, de composition, de débit, pourrait

influencer les paramètres du modèle, et du contrôle opérationnel de façon significative (Hamoda *et al.*, 1999). Par conséquent, la modélisation d'une STEP est une tâche difficile et la plupart des modèles disponibles ne sont que des approximations basées sur des hypothèses (Lee et Park, 1999 ; Cote *et al.*, 1995; Hamoda *et al.*, 1999; Plazl *et al.*, 1999 ; Bhat *et al.*, 1990).

Les techniques employant les RNA peuvent être utilisées pour modéliser ces processus de stations d'épuration. Il peut être utilisé pour une meilleure prédiction de la performance des processus en raison de leur grande précision, d'adéquation et d'applications très prometteurs en génie (Govindaraju, 2000 ; Maier et Dandy, 2000 ; Neelakantan *et al.*, 2001). Elles s'appuient normalement sur des données historiques représentatives du processus. Dans une STEP, il y'a certaines variables explicatives clés qui peuvent être utilisées pour évaluer le rendement de la station. Ces variables comprennent la DBO₅, DCO et MES.

La plupart des études disponibles sur l'application des RNA dans les STEP ont pour but la modélisation de ces variables. On constate que les modèles à base de RNA constituent un outil efficace et robuste dans la prédiction de la performance d'une station d'épuration.

Plusieurs auteurs ont appliqué les techniques des RNA pour simuler les performances des STEP, dans ce qui suit, on cite les travaux les plus significatifs :

Häck et Köhne (1996) dans leurs travaux pour la prédiction permettant le calcul approximatif ou l'estimation des paramètres du processus de traitement qui sont temporairement non disponible en cas de la rupture d'un analyseur en ligne dans la STEP des eaux usées municipales de Siegen (Allemagne). Les résultats de la prédiction présentent des coefficients de corrélation de 0,92 pour la prévision de la DCO et de 0,82 pour la prévision de NH₄.

Zhu *et al.* (1998) ont proposé un réseau de neurones à retard temporisé (TDNN) pour prédire l'efficacité d'un procédé de traitement biologique en termes de mesures de DBO₅. Les résultats de la simulation des auteurs à partir des données de processus réels ont montré que la performance du modèle réseau de neurones à retard temporisé (coefficient de corrélation) est améliorée à 0,845 contre 0,80 pour un réseau PMC initial.

Belanche *et al.* (1999) proposent deux modèle RNA à retard temporisé pour la prédiction de la DBO₅ et la DCO de la station de traitement des eaux usées de la Costa Brava (Espagne),

Les résultats de prédiction ont trouvé un indice de corrélation de 0,504 pour la prédiction de la DCO. Ils mentionnent qu'un modèle hybride RNA-Logique flou améliore considérablement la précision de la prédiction.

Choi et Park (2001) proposent un modèle RNA associé à une ACP pour extraire les données d'une STEP des eaux usées industrielles. Sur les onze paramètres de qualités des eaux usées industrielles sont réduits à seulement cinq principaux composants. Ces composants principaux sont devenus des entrées du modèle RNA.

Les dispositifs d'extractions sont alors employés pour prévoir l'azote total kjeldahl (NTK) des eaux usées industrielles à l'entrée de la station en utilisant un RNA. Le système hybride montre un perfectionnement en matière de prédiction et réduit le problème du sur-apprentissage.

Oliveira-Esquerre *et al.* (2002) ont obtenu des prédictions satisfaisantes de la concentration de la DBO₅ à la sortie d'une STEP des eaux usées industrielles au Brésil. Les résultats de la simulation des auteurs à partir des données de processus réels ont montré que la performance du modèle RNA associé à une analyse en composantes principales (coefficient de corrélation) c'est amélioré à 0,80 contre 0,74 pour un modèle RNA initiale.

Chen *et al.* (2003) développent un modèle RNA pour prédire la concentration de l'azote dans les effluents traités qui pourront être réutilisé pour la recharge des nappes souterraines. L'exactitude de ce modèle était supérieure à 90%. Par contre dans ce travail, aucune mention n'est faite sur la qualité des données.

Hamed *et al.* (2004) appliquent un modèle de RNA pour la prédiction des performances d'une STEP des eaux usées résiduaires avec un procédé de traitement conventionnelle importante dans la grande zone du **Caire (Egypte)**. Les résultats obtenus à partir de cette étude ont indiqué que les valeurs R^2 varient de 0,63 à 0,81 pour la DBO₅ et de 0,45 à 0,65 pour les MES. Les auteurs ont conclu que le modèle RNA a été gêné par la limitation des données, les données corrompues, et la restriction du modèle basée uniquement sur deux paramètres.

Mjalli *et al.* (2007) ont employé un modèle RNA pour prévoir la qualité des effluents traités et rejetés (DBO₅, DCO et MES) à la sortie de la STEP des eaux résiduaires de Doha, (Qatar).

Ils ont employé les données brutes des paramètres de pollutions qui caractérisaient la qualité d'eaux usées à savoir DBO₅, DCO et les MES. Les auteurs ont employé les données sur un an qui ont été prélevées tous les 5 jours. Ils ont obtenu un coefficient de corrélation supérieur à 0,7, mais ils n'ont pas distingué si ces résultats étaient obtenus de l'apprentissage, de la validation ou du test de l'ensemble des données.

Vyas *et al.* (2011) appliquent deux modèles de RNA pour la prédiction des performances d'une STEP des eaux usées industrielles dans la zone industrielle de Bhopal (Inde). La prédiction donne des résultats très satisfaisants pour les deux modèles. Pour le premier modèle, la valeur de R est de 0,90 ce qui montre une bonne corrélation entre DBO₅ réelles et prévues avec une précision de 90%. De même le deuxième montre de meilleurs résultats, la valeur de R est de 0,73 et une précision de 88%.

Nasr et al., (2012) proposent un modèle RNA pour prédire le rendement, construire une base de données pour effectuer un contrôle du fonctionnement du processus de traitement de la STEP des eaux usées d'El-Agamy (Egypt). L'étude indique que le model RNA peut prédire le rendement de la station entre les variables de sortie observées et prédites avec un coefficient de corrélation jusqu'à 0,90. En outre, le model RNA fournit une analyse efficace et peut être considéré comme un outil de diagnostic pour comprendre et simuler le comportement non linéaire de la station.

Bekkari et Zeddouri (2019) Appliquent un modèle neuronal pour prédire la station d'épuration des eaux usées de la ville de Touggourt en matière de la demande chimique en oxygène, où ils ont utilisé des différentes combinaisons entre les algorithmes d'apprentissage et les fonctions d'activations. Les résultats de cette étude ont montré les meilleurs résultats avec l'algorithme Levenberg-Marquardt et le couple des fonctions d'activations hyperbolique tangent sigmoïde dans la couche cachée et la fonction linière dans la couche de sortie avec un coefficient de corrélation qui ont dépassé 0,85.

II.4 Conclusion

Ce chapitre fournit une brève introduction théorique aux RNA et commence par examiner les différences et les similitudes entre le neurone biologique et artificiel. Les RNA_s possèdent une propriété remarquable qui est à l'origine de leur intérêt pratique dans des domaines très divers. Les RNA sont des structures mathématiques non linéaires qui sont capables de représenter le rapport fonctionnel non linéaire arbitraire et complexe entre les données d'entrée et les données de sortie de n'importe quel système. Les RNA ont été employés avec succès pour modéliser les rapports non linéaires complexes de séries chronologiques d'entrée-sortie dans une large variété de domaines. On s'est concentré principalement dans ce chapitre sur une architecture neuronale (Perceptron multicouche à rétro propagation) qui concerne notre travail de recherche et qui est généralement la plus utilisée pour les stations d'épuration à boues activées.

PARTIE 2 : MATERIEL ET METHODES

<u>Chapitre III</u>

III Matériel et méthodes

III.1 Description de la station d'épuration de Touggourt

La station d'épuration de Touggourt, est située au sud de la région d'Oued Righ (région aride au sud-est de l'Algérie), à environ 650 km d'Alger (Figure 9) Elle se trouve exactement dans la municipalité de Tebesbest, sur la route d'El Oued. Elle s'étend sur une superficie de 5 hectares. Elle a été mise en service le 20/11/1993 et réhabilitée en 2004 et traite aujourd'hui une partie des rejets d'eaux usées déversées par la ville de Touggourt. L'office National d'Assainissement est chargé de son exploitation.



Figure 9 : Localisation de la station d'épuration des eaux usées de Touggourt.

III.1.1 Données techniques

La capacité nominale de traitement et qualité requise des eaux usées de cette station avant traitement sont :

- \blacktriangleright Débit moyen : 9 360 m³/j ;
- \blacktriangleright Débit de pointe : 670 m³/h ;
- Demande biologique en oxygène(DBO) : 54g/habitant ;
- Charge moyenne de DBO/j : 3 375kg/j ;
- Charge moyenne de DCO/j : 5 625kg/j (source : S.T.E.P Touggourt, 2008).

III.1.2 Etapes de traitement

Comme le montre la figure 10, le plan de traitement consiste en une installation conventionnelle à boues activées dans laquelle les eaux usées entrant passent à travers le dégrillage et le dégraissage, puis entrent directement dans deux bassins d'aération. Après le traitement secondaire dans ces bassins d'aération, l'eau passe dans deux bassins de décantation et finalement dans une chambre de désinfection avant d'être rejetée dans les eaux réceptrices. (S.T.E.P Touggourt, 2008)



Figure 10 : Schéma de procédé d'épuration des eaux usées de la STEP de Touggourt.

III.1.3 Equipement de la station



Figure11 : Equipement de la STEP de Touggourt.

La figure 11 montre les différents équipements de la STEP de Touggourt, qui sont arrangés selon leurs ordres dans le procédé de boues activées :

- 1. Dégrillage mécanique, grilles mécaniques constituées des barreaux de large espacement de 20 mm ;
- Canaux de dessablage et déshuilage : avec portion d'extraction des sables, et centrale de production d'air pour aération des dessableurs ;

- Epuration biologiques : Deux bassins d'aérations fonctionnant en parallèle ; Equipement par bassin : capacité totale 7 200 m³ ; Deux aérateurs de 45 KWcapacité d'aération 80 kg O/h ;
- 4. Clarification : deux décanteurs circulaires à pont diamétral de 24 m muni d'un racleur. Nous devons signaler ici que l'étape de l'addition du chlore n'est pas toujours activée à cause des déversements des eaux dans le canal qui ne seront pas réutilisées;
- 5. Epaississeur de boues : un bassin circulaire : de 10 m de diamètre, muni d'un agitateur à pieux fixé sur le pont ;
- 6. Retour des boues : Deux vis d'Archimède à débit unitaire de 500 l/s ;
- Lits de séchage des boues: Il existe 16 lits de séchage d'une surface totale de 3 200 m².

III.2 Législation sur la protection de l'environnement régissant les stations d'épuration des eaux usées

Toutes les stations de traitement de l'eau du monde et d'Algérie détiennent une licence leur permettant de se conformer à la législation nationale / provinciale de protection de l'environnement en vigueur dans leur pays, car elles sont considérées comme des procédés à haut risque et dangereux, du point de vue de l'environnement, de l'écologie et de la santé publique. La licence de protection de l'environnement pour les STEP contient des limites législatives strictes et des spécifications légales concernés: (1) le débit maximum autorisé (2) type d'eaux usées autorisées à être traitées; (3) la concentration maximale admissible de polluants dans l'effluent; (4) les charges annuelles maximales admissibles pour chaque polluant rejeté de la STEP dans l'environnement; (5) échantillonnage, fréquence, techniques d'analyse et étalonnage des installations et du matériel permettant de mesurer le débit, la concentration et la charge de rejet des polluants; (6) le traitement minimum requis pour un influent dans la STEP; (7) notification des incidents et des accidents, enquêtes, actions correctives et préventives et comptes rendus.

III.3 Description et collecte des données

Une base de données journalière décrivant l'exploitation de la STEP de Touggourt pendant une période d'une année avec un total de 150 vecteurs de données a été obtenue. Les données ont été échantillonnées et analysées tous les 2 à 3 jours entre 8 heures et 10 heures lorsque la station a reçu le débit de pointe. Les échantillons sont prélevés à travers : l'influent de la station (qui se trouve dans la fosse principale de collecte des eaux usées avant le dégrillage), et les effluents de la station (qui se trouvait dans la fosse de mesure et du contrôle des eaux épurées) (voir Figure 10). Les paramètres étudiés ont été choisis parmi ceux couramment utilisés dans les procédés de traitement aérobie des eaux usées que l'on trouve dans la littérature. Ces données proviennent de différentes sources. Les paramètres effectués sur le terrain juste après le prélèvement des échantillons comprennent le débit, la température, la conductivité électrique et le pH. Les paramètres analysés au laboratoire du Centre de recherche scientifique et technique pour les zones arides (CRSTRA-Touggourt), incluent, CE, PH, T, OD, DBO5, DCO, MES, NH4, NO3, NO2, PO4, MESBA et le RECY. Les paramètres mesurés sont discutés dans les sous-sections suivantes.

III.3.1 Conductivité électrique (CE)

La conductivité est la propriété que possède une eau de favoriser le passage d'un courant électrique. Elle est due à la présence dans le milieu d'ions qui sont mobiles dans un champ électrique. Elle dépend de la nature de ces ions dissouts et de leurs concentrations (Rejsek, 2002). L'unité de la conductivité est le siemens par mètre (S/m).

III.3.2 Potentiel hydrogène (PH)

L'acidité, la neutralité ou l'alcalinité d'une solution aqueuse peut s'exprimer par la concentration en H_3O^+ (noté H^+ pour simplifier). De manière à faciliter cette expression ; on utilise le logarithme décimal de l'inverse de la concentration en ion H^+ : c'est le pH. (Mathieu *et al.*, 2003).

III.3.3 Température (T)

Il est important de connaître la température de l'eau avec une bonne précision. En effet, celle-ci joue un rôle dans la solubilité des sels et surtout des gaz, dans la dissociation des sels dissout donc sur la CE, dans la détermination du pH, pour la connaissance de l'origine de l'eau et des mélanges éventuels, etc. (Rodier, 2005).

III.3.4 Oxygène Dissout (DO)

L'oxygène dissout est l'une des mesures les plus importantes et les plus utiles dans les processus de boues activées et constitue également la base des tests de la DBO5 et taux d'absorption d'oxygène. L'OD est mesurée en ligne dans la station à l'aide de sondes OD. Les sondes ont également un capteur pour mesurer la température. L'aération est contrôlée en maintenant un point de consigne pour la concentration d'OD dans les bassins d'aération. Un apport suffisant en oxygène est important pour l'élimination des matières carbonées et essentielles au processus de nitrification. Une faible concentration en oxygène peut non seulement inhiber la nitrification, mais également dégrader la décantabilité des boues, détériorer la qualité des effluents et entraîner une prédominance des bactéries filamenteuses (Chen et al., 1993 ; Spellman, 2003). D'autre part, un apport excessif en oxygène entraîne des coûts opérationnels élevés. De plus, un apport élevé en oxygène peut conduire à une nitrification excessive et à une nouvelle décantation des boues. Les concentrations d'OD dans les bassins d'aération sont maintenues à environ 1,5 à 4 mg / 1 ; la valeur de 2 mg / 1 est utilisé comme point de consigne. Des valeurs supérieures à 4 mg /l n'améliorent pas les opérations de manière significative, mais augmentent considérablement les coûts d'aération. L'aération représente généralement plus de 50% des besoins énergétiques totaux de la station d'épuration (Gray, 2004; Spellman, 2003).

III.3.5 Demande biochimique en oxygène (DBO5)

Le test de DBO₅ cherche à mesurer la demande biochimique en oxygène consommé par l'échantillon sur une période déterminée. Il est donc apparu comme l'absorption d'oxygène dans la bouteille de DBO₅ et ne pas refléter nécessairement les besoins en oxygène du processus de traitement des eaux usées à traiter. On dit souvent que la DBO₅ représente entre 60% et 70% de la DBO₅ ultime (Orhon et Artan 1994 ; Spellman, 2003). La demande en oxygène est une mesure extrêmement importante de la qualité des eaux usées, car elle mesure le potentiel d'appauvrissement en oxygène de l'eau et constitue donc un indicateur important de la pollution organique. Il est nécessaire d'évaluer l'efficacité globale des processus de traitement, car celle-ci influe directement sur la qualité de l'effluent traité et sur l'économie du processus. Compte tenu du temps imparti à l'analyse, le test de DBO₅ n'est certainement pas adapté à l'exploitation et de contrôle. De plus, la DBO₅ n'est pas une valeur ponctuelle mais dépend du temps et n'est pas une mesure précise.

III.3.6 Demande chimique en oxygène (DCO)

La DCO est largement utilisée pour caractériser le potentiel organique des eaux usées. Le test de La DCO mesure la quantité d'oxygène nécessaire à l'oxydation chimique des matières organiques de l'échantillon en CO₂ et H₂O. Dans le test de la DCO, les matériaux biodégradables et non biodégradables sont oxydés ; cependant, on ne peut pas distinguer entre matière organique biodégradable et inerte. Il n'y a pas de relation uniforme entre la DCO et la DBO₅ des eaux usées, sauf que la valeur de la DCO doit être supérieure à la DBO₅ (Olsson et Newell, 1999). Une corrélation empirique entre la DCO et la DBO₅ pour une eau usée particulière peut être déterminée, ce qui est utile car cette méthode ne prend que quelques heures. Le rapport DCO / DBO₅ ou DBO₅ / DCO fournit une estimation de la proportion de matière organique biodégradable présente dans les eaux usées (Rustum *et al.*, 2008-a).

III.3.7 Matière en suspension (MES)

Selon Rejsek (2002), la pollution particulaire est due à la présence de particules de grande taille, supérieure à 10 μ m, en suspension dans l'eau, et que l'on peut assimiler aux MES. En fait, les matières en suspension ne sont pas des particules solides véritablement en suspension que dans des conditions moyenne d'écoulement des effluents correspondant à une vitesse minimale de 0,5 m/s. En fonction de la taille des particules, on distingue les matières grossières ou décantables (diamètre supérieur à 100 μ m) et les MES. On peut également prendre en compte une partie des matières colloïdales de dimension inférieure qui constituent la limite entre la phase solide et la phase dissoute (entre 1 et 10⁻² μ m).

III.3.8 Composés azotes

L'azote est présent dans les effluents sous différentes formes : azote organique (Norg), azote ammoniacal (ammoniac NH₃, NH₄), NO₃, NO₂. Plusieurs analyses sont possibles :

• Azote total : représente la somme de tous ces composés ;

• Azote Kjeldahl : généralement utilisée sur les sites urbains et industriels, cette mesure représente l'azote organique et l'azote ammoniacal ;

- Azote ammoniacal ;
- Nitrate ;

• Nitrites.

III.3.9 Les ortho phosphates (PO₄)

Le phosphore peut exister dans les eaux en solution ou en suspension, à l'état minéral ou organique. Les composés phosphorés qui, sans hydrolyse ou minéralisation, répondent au test spectrophotométrique sont considérés comme étant des ortho phosphates. L'hydrolyse en milieu acide fait apparaître le phosphore hydrolysable et minéralisation, le phosphore organique. Chaque fraction (phosphore en solution ou en suspension) peut être séparée analytiquement en ortho phosphates, phosphore hydrolysable et phosphore organique.

Suivant les cas, la teneur en phosphates peut être exprimée en mg/L de PO₄ ou de P₂O₅. $1mg/LPO_4 = 0.747 mg/L P_2O_5 = 0.326 mg/L P$ (Rodier ; 2005).

III.3.10 Matières en suspension dans le bassin d'aération (MESBA)

La concentration de matières en suspension dans le bassin d'aération est examinée quotidiennement. Un seul échantillon est collecté et analysé chaque jour. Le niveau de MESBA est maintenu à travers le recyclage des boues décantées à un niveau suffisant pour traiter la charge organique entrante. En règle générale, l'augmentation de la valeur MESBA produit une boue plus ancienne et plus dense, tandis que l'abaissement de la valeur MESBA produit une boue plus jeune et moins dense. La concentration en matière en suspension dans le bassin d'aération est contrôlée par un ajustement manuel des taux de recyclage pour atteindre une valeur optimale d'environ (2500 mg / l). Cela correspond à un âge de boue compris entre 3,3 et 7,5 jours.

III.3.11 Boues recyclées (RECY)

Le maintien d'un rapport de boue recyclée optimal a pour objectif de maintenir une concentration suffisante de boue activée dans le bassin d'aération, de manière à obtenir le degré de traitement requis dans le délai souhaité. Le taux de boue recyclée affecte le bilan de matière solide entre le clarificateur et le bassin d'aération. Dans la station d'épuration des eaux usées de Touggourt, le système de recyclage des boues contenait des pompes, une minuterie pour réguler le débit des boues. En règle générale, le débit de boue recyclée est lié au débit d'entrée de la station d'épuration par un facteur constant compris entre 0,4 et 1,5, de sorte que le débit des boues augmente avec l'augmentation du débit (Harremoës *et al.*, 1993).

III.3.12 Récapitulatifs des fréquences et des Méthodes d'analyse

Les échantillons ont été analysés à l'aide des méthodes standards pour l'analyse des eaux, de l'American Public Health Association (APHA), et Norme française (AFNOR) à travers l'analyse de l'eau Rodier 8 et 9ème Edition. Un bref résumé des paramètres mesurés, leur fréquence d'échantillonnage et leur méthode d'analyse sont présenté dans le tableau cidessous

| N° | Paramètres de processus | Fréquence d'analyse | Méthode d'analyse |
|-----|-------------------------|---------------------|-------------------------------------|
| 1 E | DBO ₅ | 2 à 3 jours | Méthodes respirométrique Rodier |
| | | | 9 ^e - E |
| 2 | DCO | 2 à 3 jours | ISO 15705 Rodier 9 ^e - E |
| 3 | MES | 2 à 3 jours | NF T 90-105 |
| 4 | OD | 2 à 3 jours | NF T 90-106 |
| 5 | NH ₄ | 2 à 3 jours | NF T 90-110 |
| 6 | PO ₄ | 2 à 3 jours | APHA 20th Edn 4500 -P-H |
| 7 | NO ₃ | 2 à 3 jours | NF T 90-012 |
| 8 | NO ₂ | 2 à 3 jours | NF T 90-012 |
| 9 | CE | 2 à 3 jours | NF T 90-031 |
| 10 | РН | 2 à 3 jours | NF T 90-008 |
| 11 | MESBA | chaque jour | NF T 90-105 |
| 12 | Température | 2 à 3 jours | NF T 90-100 |
| 13 | Recyclage de boue | chaque jour | - |

Tableau : 2 Méthodes d'analyse et fréquence d'échantillonnage des paramètres étudiés

III.4 Prétraitement des données

Le prétraitement des données est une étape essentielle, principalement en raison de la nature hétérogène des informations utilisées et obtenues à partir du contrôle opérationnel, et de l'échantillonnage manuel. Il s'agit de transformer des mesures en une quantité statistiquement significative (par exemple : convertir des impulsions électriques en débits de données. Le prétraitement des données comprend également la détermination si les données nécessitent un cryptage, une unification ou une conversion et aussi d'éliminer les valeurs extrêmes et les valeurs erronées. Quelle que soit la méthode de modélisation, la qualité des données de la sortie largement dépend de la qualité des données de l'entrée. Par conséquent, le traitement des données fournit des techniques à afin d'améliorer la qualité des données et

rendre les valeurs mesurées valides. Ceci est important pour obtenir des résultats de simulation fiables. Différentes méthodes d'exploration des données disponibles ont été appliquées. En règle générale, les statistiques descriptives notamment la moyenne, le maximum, le minimum, et l'écart type ont été calculées. De plus, des histogrammes et des courbes ont été tracés.

III.5 La boite à moustaches (box and whisker plot)

En raison de l'effet probable des valeurs aberrantes sur les performances d'un modèle RNA, les données expérimentales brutes ont été filtrées en excluant tous les points aberrants, points inhabituels. La filtration de ces données a été réalisée en supprimant tous qui se trouvent à au moins 1,5 fois l'étendue interquartile (Q3 - Q1) du bord de la boîte à moustache de la série des données. Les valeurs aberrantes sont déterminées par le traçage du diagramme à moustaches qui est une méthode la plus utilisée dans la recherche des valeurs aberrantes. Ce graphique composé d'un rectangle duquel deux droites sortent afin de représenter certains éléments des données. La valeur centrale du graphique est la médiane. Les bords du rectangle sont les quartiles (Pour le bord inférieur, un quart des observations ont des valeurs plus petites et trois quart ont des valeurs plus grandes, le bord supérieur suit le même raisonnement). Les extrémités des moustaches sont calculées en utilisant 1.5 fois l'espace interquartile (la distance entre le 1er et le 3ème quartile). Les graphiques illustrent l'étendue des valeurs aberrantes dans chaque variable, comme indiqué par les points qui dépassent les moustaches. En outre, il montre la plage de chaque variable et par conséquent, la caractérisation de la station d'épuration.

III.6 L'analyse en composantes principales ACP III.6.1 Généralités

L'analyse en composantes principales est une technique descriptive permettant d'étudier les relations qui existent entre des variables descriptives, sans tenir compte, à priori, d'une quelconque structure, ni des variables, ni des individus (Palm, 1998), les domaines d'application de cette méthode sont très variés, et de nombreux exemples sont proposés, notamment, par Jackson (1991).

L'ACP est une méthode d'analyse des données, lorsqu'un nombre très réduit de composantes suffit à concentrer le maximum d'informations, cette méthode, a le mérite de faciliter la visualisation et l'analyse préliminaire des données (Oukhellou ,1997). L'utilisation de l'ACP

pour l'exploitation des données remonte au début du siècle dernier. Elle est principalement issue des travaux de psychomètres américains (Pearson, 1901 ; Spearman, 1904 ; Hotelling, 1933 ; Thurstone, 1947).

III.6.2 Principes de I'ACP

L'ACP a deux objectifs majeurs (Saporta, 1990), le premier est de déterminer des représentations graphiques planes, appelées cartes, qui permettent d'analyser au mieux les proximités et les écarts entre les différentes variables. En particulier, elles doivent permettre d'isoler les individus "atypiques" et de regrouper les individus « semblables ». Le deuxième objectif de l'ACP est de fournir des "résumés linéaires" des variables, c'est à dire de remplacer les variables initiales par des combinaisons linéaires de celles-ci, ces nouvelles variables, appelées composantes principales ou facteurs, ont des propriétés qu'il est important de présenter pour faciliter l'interprétation de l'analyse. Chaque composante principale ou facteur est une combinaison linéaire des variables originales qui explique un maximum de la variabilité totale de données tout en étant non corrélé avec les autres composantes. Les coefficients de combinaisons linéaires appelées en anglais « loadings » représentent les degrés de corrélation entre les variables et le facteur. Les facteurs ne sont pas facilement interprétables parce qu'il est nécessaire que les facteurs subissent une rotation qui a pour effet de maximiser ou minimiser certains « loadings ». La rotation la plus utilisé dans des études environnementales est la rotation VARIMAX. L'ACP permet donc de réduire des tableaux de grande taille en un petit nombre de variables tout en conservant un maximum d'information.

III.7 Logiciels utilisés

Les modèles développés ont été mis en œuvre à l'aide du langage de programmation MATLAB 8.1 (Version–R2013a) (Math Works, Inc., USA) avec les boîtes à outils Neural Networks (Math Works, Inc., USA). Des analyses statistiques complémentaires ont été réalisées à l'aide de logicielle SPSS (v15.0) et l'EXCEL 2007.

Le langage de programmation MATLAB a été choisi pour le développement du modèle car RNA nécessite des calculs matriciels intensifs. Les boîtes à outils Réseau de neurones de MATLAB fournissent un support complet pour la conception, la mise en œuvre et la simulation rapide des modèles. Leur méthodologie cohérente et leur organisation

modulaire fournissent un cadre flexible pour l'expérimentation et simplifient la personnalisation.

III.8 Simulation par le réseau de neurone artificiel

L'objectif principal de cette recherche est de trouver des modèles RNA pouvant prédire les concentrations des principaux paramètres de contrôles des eaux usées traités dans la STEP de Touggourt (DBO₅, DCO et les MES).

L'un des défis est de trouver une topologie de réseau optimale en faisant varier les types d'entrées du réseau et ainsi de résoudre le problème particulier de la cartographie d'entrée- sortie. L'architecture du réseau approprié pour une application donnée est fortement dépendante du problème, de la diversité de l'ensemble d'apprentissage et de la complexité de la fonction sous-jacente (Garvey, 1997). Dans cette étude, les trois paramètres de la sortie de la STEP de Touggourt sont modélisés séparément par les paramètres de l'influent et de l'effluent de la STEP en considérant diverses variables uniques comme entrées de PMC-RNA afin d'examiner l'effet de chaque variable sur la variation des valeurs de chaque paramètre cible. De même, des modèles distincts ont été réalisés afin de montrer l'effet des variables d'entrée multiples sur la variation des valeurs des paramètres cibles. Ces entrées sont utilisées pour former les PMC-RNA en groupes de deux, trois, quatre, ...

III.8.1 Taille de la base des données

La taille de la base des données est déterminée par l'expérimentation. Dans le cas de notre étude, on a choisi d'utiliser l'ensemble des données des paramètres à savoir 150 données, pour les modèles des paramètres DBO₅, DCO et MES.

III.8.2 Développement du modèle

Dans cette étude, le développement du modèle de RNA se base sur trois étapes importantes : l'apprentissage, la validation et le test. L'efficacité de la prévision dépend fortement de la qualité des données d'apprentissage. L'ensemble des données d'apprentissage a été utilisé dans la construction du modèle RNA en ajustant les poids, tandis que L'ensemble des données de validation a été utilisé pour déterminer un point d'arrêt pour l'algorithme de RP afin d'éviter le sur apprentissage et également pour donner le nombre optimal d'unités cachées (détermination de l'architecture de réseau). Enfin, L'ensemble de données de test a été utilisé pour obtenir les caractéristiques de performance telles que la précision et le test de la généralisation de la construction du modèle. L'ensemble des données de test n'est généralement pas appliqué dans le processus exécuté par le modèle.

Dans cette étude, On a utilisé l'algorithme d'apprentissage le plus rapide et le plus couramment adoptés, L'algorithme de Levenberg-Marquardt, ou algorithme LM.

Afin de déterminer l'architecture appropriée des modèles de réseaux neuronaux pour les trois paramètres étudiés, deux types de modèle ont été élaborés, l'un c'est le modèle simple qui utilise que les paramètres de l'influent de la STEP comme entrées de modèle et le deuxième c'est le modèle extensif qui utilise en plus des paramètres de l'influent les paramètres de l'effluent de la STEP ayant une forte corrélation avec le paramètre prédit.

Dans ce travail, on a utilisé une seule couche cachée pour le modèle simple composée de plusieurs neurones (ils étaient ouverts au changement pour obtenir le meilleur modèle). En effet, l'augmentation du nombre de couches cachées augmente la charge de calcul sans gain de performance. Par contre, pour le modèle extensif et due au nombre important des paramètres d'entrée et afin d'améliorer le fonctionnement et la précision de modèle, on a utilisé deux couches cachées avec un nombre fixe de neurone dans la deuxième couche (10 neurones).

La fonction d'activation tangente hyperbolique convient la plupart des types de RNA, en particulier aux problèmes de prédiction. La fonction tangente hyperbolique a été utilisée dans ce travail. L'efficacité du modèle RNA a été estimée par deux paramètres. La figure12 récapitule la structure de réseau neuronale RNA-PMC utilisé dans ce travail.





III.8.3 Division des données

Dans cette étude, l'ensemble de données ont été divisées en trois partitions distinctes (sous-ensembles); (15%) des données originales ont été prises pour l'ensemble de validation, (15%) pour l'ensemble de tests et (70%) pour l'ensemble d'apprentissage. Afin de prendre en

compte les variations saisonnières possibles (effet de la saisonnalité) et d'inclure l'effet saisonnier individuel, on a divisé l'ensemble de données de sorte que les données de chaque saison sont incluses dans les trois phases de formation de modèle. Les données couvrent donc les quatre saisons : printemps, été, automne et hiver.

III.8.4 Mesure des performances de prédiction

La performance du modèle RNA a été estimée par deux paramètres : le coefficient de corrélation (R) et la racine carrée de l'erreur quadratique moyenne (REQM) calculée entre les données prédites et les données mesurées, comme indiqué dans l'équation 5 et 6 respectivement (Pendashteh *et al.*, 2011). Le modèle optimal a été obtenu lorsque la valeur REQM du réseau était minimale et la valeur de R été maximale.

Où N est le nombre des données de modèle, x_i et y_i correspondent aux valeurs mesurées et estimées du modèle respectivement. \overline{x} Et \overline{y} sont les valeurs moyennes calculée de x_i et y_i respectivement.

III.8.5 Normalisation des données

La normalisation des données d'entrée et de sortie est très critique (Baughman *et al.*, 1995 ; Choi *et* Park., 2001). Il a été utilisé pour réduire l'erreur dans les prédictions par le réseau neuronal. Si les données d'entrée et de sortie ne sont pas du même ordre de grandeur, certaines valeurs peuvent sembler plus significatives qu'elles ne le font réellement. En outre, une fonction d'activation typique telle que la fonction sigmoïde ou la fonction tangente hyperbolique ne permet pas de distinguer deux valeurs différentes de xi lorsque les deux sont très grandes, car les deux produisent des valeurs de sortie de seuil identiques de 1-0. Les variables doivent être adaptées à des plages uniformes correspondant aux limites des fonctions d'activation dans la couche en sortie, afin de

:

garantir que les valeurs des différentes variables sont dans la même plage et que toutes les variables reçoivent la même attention pendant la phase d'apprentissage (Choi *et* Park., 2001). En règle générale, les fonctions d'activations sigmoïdes, tangentes hyperboliques et Purelin sont utilisées pour lesquelles les sorties sont dans les intervalles [0 1], [-1 1] [-++], respectivement.

Les techniques statistiques fonctionnent bien, quand chacune des variables d'entrée a le même poids relatif. Cependant dans notre cas d'étude les différentes variables sont de nature physique différente, caractérisée par des unités différentes, ce qui nous a obligés à les normaliser. La normalisation vise à amener la plage d'évolution des valeurs entre 0 et 1.

Pour ce cas d'étude, nous avons normalisé les données en utilisant la formule suivante

$$x_{ni} = \frac{(x_i - x_{min})}{(x_{max} - x_{min})}.$$
(7)

Où x_i est la valeur initiale, x_{max} et x_{min} sont les maximums et minimums des valeurs initiales et x_{ni} est la valeur normalisée. Après la formation du modele, les données de sortie ont été adaptées aux valeurs réelles à l'aide de l'équation suivante :

$$x_i = x_{ni}(x_{max} - x_{min}) + x_{min}......(8)$$
III.9 Conclusion

Ce chapitre présente la méthodologie appliquée dans cette étude. Il commence par une description détaillée de la STEP étudiée dans ce travail. Ensuite, une description des paramètres utilisés et leurs caractéristiques sont présentées. Un aperçu des logiciels utilisés dans cette étude est également présenté brièvement. Le chapitre se termine par les critères des modèles neuronaux développés dans cette étude.

PARTIE 03 :

RESULTATS ET

DISCUSSIONS

Chapitre IV

IV Analyse statistique et performance de la STEP de TouggourtIV.1 Introduction

Les travaux expérimentaux ont permis de créer une base de données regroupant les résultats d'analyses menés pendant la période d'étude. Un traitement des données est nécessaire pour calculer les paramètres statistiques et la vérification de l'homogénéité de la série des données à travers le test de la boîte de moustache qui permit d'écarter les points aberrants afin d'assurer la précision de cette étude. Une présentation graphique s'effectuer dans ce chapitre pour estimer les performances de la station par rapport aux normes. Cette présentation a été faite sur les paramètres de qualité pour les eaux brutes et les eaux épurées afin de construire un aperçue globale sur cette station. Cette étude a été réalisée sur la base des données de 10 mois d'exploitation d'octobre 2016 jusqu'au juillet 2017 pour la CE, PH, T, OD, DBO₅, DCO, MES, NH₄, NO₃, NO₂ et le PO₄ à l'entrée et à la sortie de la station, en plus de la matière en suspension dans le bassin d'aération et le recyclage des boues. On a utilisé pour cette étude deux logiciels de calcul statistique notamment (EXCEL® et SPSS).

IV.2 Traitement des données par la boite à moustaches :

La figure 13 montre le graphique en boîte à moustaches des paramètres d'entrée et de sortie de la STEP de Touggourt.

À l'entrée de la station de Touggourt, la figure13 montre la présence de quelques points aberrants pour la DBO₅e et NH₄, cela peut indiquer des charges polluantes inhabituelles ou des erreurs de mesure, par contre pour la MESe aucun point aberrant n'est remarqué, tandis que pour la DCOe on distingue un seul point aberrant. On note aussi que 50% de la série est comprise entre 120-160 avec espace interquartile de 40mg/l, 216- 294 avec espace interquartile de 78mg/l, 41.47- 50.87 avec espace interquartile de 9.4mg/l et 155-322 avec espace interquartile de 167mg/l pour la DBO₅e, DCOe, NH₄ et MESe respectivement. Les points aberrants sont plus proches des autres points. On note que les moustaches de la boîte pour la MESe sont assez importantes cela est dû au charriage suite à des travaux de maintenance au niveau du réseau d'assainissement. La présentation pour les PO₄e indique une série assez variée, avec deux valeurs aberrantes, la moitié de cette série est compris entre 6.34 - 7.71 avec un espace interquartile de 1.37mg/l, et des moustaches assez longes qui impliquent une variation du PO₄e au niveau de réseaux d'assainissement qui a deux sources : l'un c'est le PO₄ lui-même rejeté dans le réseau d'assainissement et l'autre résulte de la transformation du phosphore organique. Pour le PHe, CEe et ODe le graphique montre des étendus très limités avec quelques valeurs aberrantes résultant de la qualité des eaux usées et leur dégradation au niveau du réseau d'assainissement. Les espaces interquartiles sont de 0.17, 0.6 ms/cm, et 0.08 mg/l pour le pHe, CEe et OD respectivement.

D'une part, La présence des ions NH₄, NO₂ et NO₃, dans une eau brute n'est pas forcément due aux rejets (domestiques et/ou industriels) mais leurs concentrations peuvent varier suite aux réactions chimiques d'oxydation le long des réseaux d'assainissement (l'oxydation du NH₄ donne des ions du NO₂ et l'oxydation de cette dernière donne des ions du NO₃). D'autre part, l'échantillonnage étant ponctuel, l'information recueillie à travers ces échantillons n'a pas permet de juger la validité des données. Toutefois, nous pouvons remarquer les fortes concentrations des ions NH₄ et NO₃ à l'entrée de la station.

À la sortie de la STEP, notons que la plupart des valeurs des différents paramètres sont inférieures aux normes algériennes de rejet. De ce fait, la présence de points aberrants n'affecte pas les performances de la STEP. Sauf les éléments azotés notamment les NO_3 et NO_2 qui ont un nombre des valeurs aberrantes importants cela est dû à la transformation de l'azote organique et ammoniacal, le rendement épuratoire et la quantité d'oxygène dans les eaux de la sortie de la STEP.

Selon les représentations graphiques en boîtes à moustaches des données, on note que dans le cas des paramètres de fonctionnement de la station notamment (RECY, MESBA), les espaces interquartiles sont assez larges 1255 m3/j et 664 mg/l respectivement ce qui indique la concentration de 50% des séries de données et aussi une longueur des moustaches importantes cela confirme la variabilité de la pollution de l'eau brute.



Figure 13 : Graphique en boîte à moustache des paramètres d'entrée et de sortie de la STEP de Touggourt.

IV.3 Statistique descriptive des paramètres de pollution :

Dans les tableaux 3 et 4 on résume les principales caractéristiques descriptives de la qualité des eaux à l'amont et à l'aval de la station durant la période d'étude de la STEP de Touggourt notamment le nombre d'observations, valeur maximum, valeur minimum, moyenne et l'écart-type pour chaque paramètre.

Pour les données d'entrée de la station, on constate que les valeurs de l'écart-type des paramètres de pollution sont assez variées due principalement à la variabilité moyenne de la charge polluante à l'entrée de la STEP.

L'analyse de l'écart-type annuelle de la DBO₅ à la sortie montre que la station de Touggourt est statistiquement stable pour la période d'étude. (Une station est statistiquement stable lorsque l'écart-type pour la DBO₅ des effluents est inférieur à 10 mg/l, et l'écart-type pour la MES des effluents est inférieur à 70 mg/l) (Niku *et al.*, 1982).

| Paramètre | Nomenclature | Unité | Min | Max | Moy | Ecart-type | |
|-----------------------|-------------------|-----------|---|--------|----------|------------|--|
| Demande biochimique | DBO ₅ | mg/l | 90.00 | 240.00 | 146,1745 | 33.26485 | |
| en oxygène | | 8/- | , | , | , | | |
| Demande chimique en | DCO | mg/l | 79.00 | 398.00 | 258 2584 | 63 68282 | |
| oxygène | Deo | 1115/1 | 79,00 | 570,00 | 250,2501 | 03,00202 | |
| Azote Ammoniacale | NH ₄ e | mg/l | 18,03 | 52,29 | 35,4451 | 5,56323 | |
| Oxygène dissout | O.D | mg/l | 0,08 | 1,15 | 0,1973 | 0,13892 | |
| Matière en suspension | MES | mg/l | 100,00 | 542,00 | 238,9866 | 103,95435 | |
| Ortho phosphates | PO ₄ | mg/l | 3,64 | 9,73 | 6,9023 | 1,01998 | |
| Température | Т | °C | 18,20 | 33,00 | 25,6554 | 3,47961 | |
| Nitrites | NO ₂ | mg/l | 0,21 | 3,62 | 0,7998 | 0,51637 | |
| Nitrates | NO ₃ | mg/l | 0,02 | 5,61 | 1,5280 | 1,40357 | |
| Potentiel hydrogène | PH | - | 6,94 | 7,96 | 7,3829 | 0,16683 | |
| Conductivité | CE | ms/cm | 4.68 | 8 78 | 6 2253 | 0.67466 | |
| électrique | CL | 1115/0111 | 4,00 | 0,70 | 0,2233 | 0,07400 | |

Tableau 3: Synthèse de l'analyse statistique des données à l'entrée de la STEP de Touggourt

| Paramètre | Nomenclature | Unité | Min | Max | Moy | Ecart-type | |
|---------------------|-------------------|-----------|---------|----------|-----------|------------|--|
| Demande | | | | | | | |
| biochimique en | DBO ₅ | mg/l | 7,00 | 28,00 | 14,3389 | 3,26233 | |
| oxygène | | | | | | | |
| Demande chimique | DCO | ma/l | 12.00 | 40.00 | 25.0738 | 6 30030 | |
| en oxygène | DCO | mg/1 | 12,00 | 40,00 | 25,0758 | 0,30030 | |
| Azote Ammoniacale | NH ₄ e | mg/l | 0.2 | 7.2 | 2.22 | 1.9 | |
| Oxygène dissout | O.D | mg/l | 1,21 | 6,93 | 3,8603 | 1,26666 | |
| Matière en | MES | ma/l | 8 56 | 52 12 | 23 1116 | 7 07703 | |
| suspension | IVIL25 | mg/1 | 8,50 | 52,42 | 23,1110 | 1,91103 | |
| Ortho phosphates | PO_4 | mg/l | 0,30 | 5,82 | 2,6606 | 1,36338 | |
| Température | Т | °C | 13,10 | 30,70 | 24,0805 | 3,90416 | |
| Nitrites | NO ₂ | mg/l | 0,02 | 2,49 | 0,5750 | 0,41637 | |
| Nitrates | NO ₃ | mg/l | 0,02 | 33,93 | 9,1900 | 10,03854 | |
| Potentiel hydrogène | PH | - | 7,08 | 7,81 | 7,4310 | 0,12629 | |
| Conductivité | CE | ms/cm | 5 13 | 7 17 | 6 1621 | 0 30281 | |
| électrique | CL | 1115/0111 | 5,15 | /,1/ | 0,1021 | 0,39281 | |
| Matière en | | | | | | | |
| suspension dans le | MESBA | mg/l | 1086,21 | 2768,17 | 1950,1892 | 435,17538 | |
| bassin d'aération | | | | | | | |
| Volume de recyclage | RECV | $m^{3/i}$ | 5950 00 | 11050.00 | 8854 4044 | 1130 73681 | |
| des boues | NEC I | III /J | 5750,00 | 11030,00 | 0004,4044 | 1130,73081 | |

Tableau 4: Synthèse de l'analyse statistique des données à la sortie de la STEP de Touggourt

IV.4 Performances épuratoires de la STEP de Touggourt

Les graphiques ci-dessous expriment les variations de la concentration des déférences paramètres physico-chimique des eaux usées à l'entrée et à la sortie de la STEP de Touggourt.

IV.4.1 La température



Figure 14 : Représentation graphique de la variation de la Température à l'entrée et à la sortie de STEP de Touggourt (2016-2017).

La température de l'eau est un facteur important dans l'environnement aquatique qui influe sur les réactions physico-chimiques et biologiques (Chapman et al., 1996). C'est un facteur clé de l'activité biologique ayant des répercussions écologiques (Leynaud, 1968). L'augmentation de la température implique une réduction de la viscosité, une élévation de la tension de vapeur saturante à la surface (évaporation), une diminution de la solubilité des gaz (oxygène). Quelques-uns de ces effets peuvent avoir une action bénéfique ; par exemple l'augmentation de la température favorise l'autoépuration et accroît la vitesse de sédimentation, ce qui peut présenter un intérêt dans les stations d'épuration. Dans l'eau de mer réchauffée, la production biologique est améliorée (Beauchamp, Ross, Withehouse). Les rejets d'eaux chaudes peuvent favoriser le développement d'espèces comme Naegleria fowleri qui peut être à l'origine de méningo-encéphalites amibiennes primitives à évolution souvent fatale. Ces amibes peuvent se développer à partir de 20 °C mais la température idéale pour leur prolifération se situe autour de 30°C surtout si cette température se maintient pendant un mois ou plus. (Rodier, 2005). Comme montré dans la figure 14, les valeurs de la température à l'entrée varient entre 18.2 et 33 °C, à la sortie, elles se situent entre 13.10 et 30 °C en dessous de 30 ° C Norme algérienne (tableau1 annexe1) (JORA ; 2006). On remarque que la plupart des valeurs sont comprises entre 20 et 30 même dans la période hivernale, cela est dû à la nature des eaux de consommation de la région qui dépassent 55°C (eau albienne).

IV.4.2 Le pH

La figure suivante représente la variation de pH à l'entrée et à la sortie de la STEP de Touggourt. Le pH des eaux naturelles est compris entre 6 et 8,5 (Chapman, 1996). Ceci est un élément important pour l'interprétation de la corrosion des tuyaux des installations de traitement et également un paramètre très important qui influence l'activité biologique de la microflore de l'eau. La grande majorité des micro-organismes se développe à un pH compris entre 4,5 et 8,0 et dans la plage optimale comprise entre 5,5 et 7,5 (Botton et al. ; 1990, Meinck et al. ; 1977). La plupart des organismes ne peuvent tolérer un pH supérieur à 9,5 ou inférieur à 4. Il est généralement recommandé de maintenir un pH compris entre 6,5 et 7,5 pour les STEP des eaux usées (Metcalf et Eddy, 1991). En effet, un pH faible favorise la croissance des champignons filamenteux et autres organismes à l'origine des boues flottantes (Arcand, 1989), alors que les bactéries nitrifiantes ont besoin d'un pH compris entre 7,4 et 9 pour les nitrosomonas, 8,5 et 9,1. Les bactéries Nitrobacter Acinetobacter déphosphatent se développent bien à pH acide. Les valeurs de pH enregistrées à l'entrée de la station varient de 6,94 à 7,96 et de 7,08 à 7,81 à la sortie, ce qui correspond aux normes algériennes (5,5 à 8,5) des eaux usées rejetées dans l'environnement (annexe1). On note qu'il y a une légère variabilité des valeurs du pH entre l'entrée et la sortie, cela est dû à l'effet de l'épuration et résultant de l'activité biologique.



Figure 15 : Représentation graphique de la variation de du PH à l'entrée et à la sortie de STEP (2016-2017).





Figure 16 : Représentation graphique de la variation de la CE à l'entrée et à la sortie de STEP (2016-2017).

La conductivité est la capacité de l'eau à conduire le courant entre deux électrodes. Elle sert également à déterminer la quantité de sel dissoute dans l'eau (Pescod, 1985 ; Rodier, 1984). La conductivité est également fonction de la température de l'eau : elle est plus importante lorsque la température augmente. À l'entrée de la STEP, les valeurs mesurées de la CE sont comprises entre 4.68 et 8.78 ms/cm, donc supérieurs à la norme de rejet. À la sortie de la STEP, elles se situent entre 5,13 et 7,17 au-dessus du maximum fixé par la norme algérienne : 3 ms/cm (annexe1). Les valeurs obtenues restent très élevées par rapport à la norme en raison de la salinité des eaux de consommation de la région d'étude (sud-est de l'Algérie). La figure 16 montre la variation de la CE dans le temps de la STEP de Touggourt.





Figure 17 : Représentation graphique de la variation de la MES à l'entrée et à la sortie de STEP (2016-2017).

Les matières en suspension sont toutes des particules organiques et inorganiques décantables et non décantables dans les eaux usées. Dans les cours d'eau, ces matériaux favorisent la réduction de la luminosité et de la production organique en raison de la diminution de la concentration en OD consécutif à la réduction du phénomène de photosynthèse. Les niveaux élevés de matières en suspension peuvent être considérés comme une forme de pollution. Une certaine augmentation de MES peut également provoquer un réchauffement de l'eau, ce qui aura pour effet de réduire la qualité des organismes vivant dans les eaux froides (Hebert *et al.*; 2000). D'après la figure 17 on observe que la concentration en MES à l'entrée de la STEP varie d'un mois à un autre où elle arrive jusqu'à 542 mg/l en raison du charriage des travaux de maintenance au niveau des réseaux d'assainissement et l'instabilité de la qualité des eaux à l'entrée. La concentration en MES est importante durant la période sèche à cause de l'évaporation de l'eau et le manque d'apports liquide. À la sortie, elle varie entre 8,56 et 52,42 mg/l mais reste inférieur à la norme algérienne de rejet (30 mg / 1; annexe 1) dans la plupart de la période d'étude

IV.4.5 Oxygène dissout

D'après la figure 18, l'oxygène dissout au cours de la période d'étude est très faible dans les eaux d'entrée, il varie de 0,08 à 1,15 mg/l et entre 1,21 mg / l et 6,93 mg/ l dans les eaux de sortie.



Figure 18 : Représentation graphique de la variation de l'OD à l'entrée et à la sortie de STEP(2016-2017).

Les valeurs minimales ont été observées au période hivernal, tandis que les valeurs maximales ont été observées au période estival dans les effluents traités de la STEP. L'OD dans les eaux usées revêt une grande importance pour la vie aquatique. Il est considéré comme le facteur qui reflète l'activité biologique aquatique et détermine les modifications biologiques provoquées par les organismes aérobies ou anaérobies. Dans la présente étude, on a obtenu une valeur presque nulle d'OD des eaux entrant à la STEP. Cela peut être dû au mélange des effluents industriels et au déversement des déchets solides municipaux dans les eaux usées. Les valeurs nulles d'OD peuvent également être dues aux conditions de stagnation de l'eau et à l'augmentation de la charge de déchets résultant de l'addition régulière d'aliments et de pesticides.

En effet, on note une augmentation des valeurs d'oxygène des eaux traitées par rapport aux eaux brutes. Cette augmentation peut être expliquée par la présence de l'aération artificielle et la diminution de la charge polluante dans ces eaux. La variabilité de l'OD dans les eaux traitées peut également être due au taux d'oxygène fournis par les aérateurs des bassins d'aération. Les concentrations en OD à la sortie de la STEP respectent les normes de l'OMS (3mg/l), sauf dans la période hivernale où on constate que les valeurs d'OD sont inférieures à la norme.

IV.4.6 La demande biochimique en oxygène(DBO5)

La demande biochimique en oxygène est définie comme la teneur en OD consommée par les micro-organismes au cours d'une période d'incubation à l'obscurité à 20 ° C pendant 5 jours afin de décomposer la matière organique dissoute ou en suspension contenue dans un litre d'eau. Il permet donc l'évaluation de la matière organique biodégradable. La DBO₅ est le paramètre le plus largement utilisé pour mesurer la qualité de l'eau. Il est également pris comme mesure de la concentration en matières organiques présente dans l'eau. Plus la quantité de matière biodégradable présente est importante, plus la demande en oxygène est importante et plus la DBO₅ est importante. La DBO₅ de l'eau naturelle est inférieure à 2 mg / 1. Les eaux usées ménagères ont des concentrations supérieures à 10 mg / l (Chapman, 1996). Au cours de la période d'étude, la demande biologique en oxygène varie de 90 mg / l à 240 mg / l dans les eaux brutes et de 7 mg / l à 28 mg / l dans les eaux traitées. La valeur minimale a été observée aux mois de l'hiver, tandis que la valeur maximale a été observée aux mois chauds dans les eaux brutes cela est dû à la dilution de ces eaux résultant de la quantité importante de l'eau rejetée dans les réseaux d'assainissement dans la période hivernale qui a conduit à la diminution de la charge polluante en terme de DBO₅. D'après les résultats on note que toutes les valeurs de DBO5 des eaux traitées respectent les normes algériennes de rejets (40 mg/l). En comparant entre les valeurs de DBO_5 des eaux brutes et celles traitées on remarque un bon rendement épuratoire de la STEP de cette pollution organique. La figure 19 montre la variation de la DBO₅ dans le temps dans la STEP de Touggourt.



Figure 19 : Représentation graphique de la variation de la DBO₅ à l'entrée et à la sortie de STEP (2016-2017).

IV.4.7 Demande chimique en oxygène (DCO)



Figure 20 : Représentation graphique de la variation de la DCO à l'entrée et à la sortie de STEP (2016-2017).

La demande chimique en oxygène (DCO) représente la consommation en oxygène pour l'oxydation non biologique de toutes les matières organiques de l'effluent, qu'elles soient biodégradables ou non. Il permet d'évaluer la concentration de matières organiques ou inorganiques, dissoutes ou en suspension dans l'eau à travers la quantité d'oxygène nécessaire à l'oxydation chimique totale (Rodier, 1996).

Les valeurs moyennes mensuelles de DCO (figure 20) des eaux brutes varient entre 79 et 398 mg / l, elles décroissent en hiver et s'augmentent en été suivant celles de MES et DBO₅. Il existe une corrélation très significative entre la DCO et la DBO₅. La DCO des eaux traitées de la STEP de Touggourt se situe entre 12 et 40 mg / l, en dessous de la norme algérienne des rejets (120 mg / l) annexe1. De plus, le rapport annuel moyen DCO / DBO₅ est de 1.76, ce qui permet de déduire que les eaux usées chargées de matière organique sont biodégradables avec certaines souches (Henze *et al.*; 1997). On note que l'augmentation de la DCO à l'entrée de la STEP est due à des rejets contenant des produits chimiques tels que les détergents, ou à cause de l'augmentation de la température qui implique l'augmentation de l'évaporation dans les mois chauds. Notant aussi que l'abattoir de Touggourt fait partie du réseau de la station dont les jours de l'abattage correspondent aux valeurs les plus importantes de la DCO.



IV.4.8 Ortho phosphates (PO₄)



L'origine des eaux usées contenant du phosphore est multiple (Villebrun, 1989). Il provient du métabolisme humain (urines et matières fécales), du détergent ménager, des déchets (non détergents) collectés avec les eaux de cuisine (vaisselle, déchets alimentaires liquides), eaux de lavage (lessive, sols, ...) et des déchets industriels. Les formes de phosphore des eaux usées sont solubles ou particulaires essentiellement se constituent de phosphore inorganique (principalement des polyphosphates et d'ortho phosphates) dont une partie provient de la dégradation de Phosphore organique : phospholipides, esters, polynucléotides, ATP, ADP, ...Les eaux usées domestiques ont une teneur en phosphore généralement comprise entre 10 et 25 mg / l. Le phosphate excrété dans l'urine représentait environ 25 à 50% des déchets ménagers, le phosphore étant principalement présent sous forme de phosphates inorganiques (Petry et al., 2002). En effet, de nombreux détergents contiennent des phosphates. Dans les pays méditerranéens où l'eau est plus dure, la consommation de détergents est plus importante (Fox et al., 2002). L'élimination des PO₄ des eaux usées nécessite des traitements coûteux et rarement utilisés. Les traitements primaires, secondaires et tertiaires ont des rendements d'élimination de P respectivement 20, 40 et 85%. Comme l'azote, l'élément phosphore est considéré comme polluant dans l'environnement à des teneurs supérieures à 0,5 mg / 1 (Rodier, 1996). Les rejets de phosphore dans les écosystèmes aquatiques sont l'un des problèmes environnementaux les plus graves car ils contribuent à accélérer l'eutrophisation de ces zones, en particulier dans les zones fermées ou semi-fermées telles que les lagunes, les estuaires ... (Rodier, 1984). D'après les résultats obtenus indiqués sur la figure 21, on constate que les eaux brutes à l'entrée de la STEP sont caractérisées par des valeurs moyennes de PO₄ comprises entre 3,64 mg/l et 9,73 mg/l. tandis que les concentrations en aval de la STEP varient entre 0,30 mg/l et 5,82 mg/l. Les valeurs des PO₄ à la sortie restent en dessus de la norme algérienne (2 mg/l) dans la plupart partie de la période d'étude. Ceci confirme le faible rendement épuratoire de la STEP en matière d'élimination de phosphore, ce qui nécessite un traitement tertiaire.

IV.4.9 Matières azotées (NH4, NO2, NO3)

L'azote dans les eaux usées se trouve par sa nature minérale ou organique. L'azote minéral, y compris NH₄, NO₂ et NO₃ constitue la majeure partie de l'azote total. L'azote organique est principalement un composant de protéines, de polypeptides, d'acides aminés et d'urées. Le rapport entre les formes organiques et minérales dépend de la longueur du réseau d'assainissement car il est la première étape de la transformation de l'azote organique,

nommée l'ammonification. L'ammonification se poursuit jusqu'à l'entrée de la station d'épuration où l'azote existe principalement sous la forme de NH₄. Le mécanisme de l'élimination biologique de l'azote s'effectue en deux étapes, à savoir la nitrification et la dénitrification. La nitrification est l'oxydation biologique de l'azote ammoniacal en nitrate. Cette transformation s'effectue en deux étapes : la nitritation, suivie de la nitratation. Elle est réalisée en présence de l'oxygène par des bactéries autotrophes qui utilisent NH₄) et les carbonates (HCO $_{3}$) comme source d'énergie. La nitritation est l'oxydation de l'azote ammoniacal en azote nitreux NO₂ par des bactéries autotrophes du genre Nitrosomonas, Nitrosococcus ou Nitrospira.

La réaction chimique de l'azote ammoniacal est exprimée par l'équation suivante :



$$2NH_4^+ + 3O_2 \longrightarrow 2NO_2^- + 2H_2O + 4H^+ \dots (9)$$



La nitratation se produit en présence des bactéries nitratantes (nitrobacters) qui transforment les nitrites en NO_3^- . La réaction chimique se présente par suit :

 $2NO_2 + O_2 \longrightarrow 2NO_3 \dots (10)$

Pour convertir 1 mg de NH4⁺ en NO3⁻ il faut 4,57 mg d'OD.

Les microorganismes responsables de la nitrification (Nitrosomonas et Nitrobacter) sont très fragiles. Ils ont besoin de températures constantes (jamais inférieures à 12 ° C), et d'un apport suffisant en oxygène. La dénitrification est une réduction du nitrate en azote gazeux (N2) par des bactéries hétérotrophes du genre Pseudomonas en l'absence d'oxygène dissout en présence de nitrates (Pelmont, 1993). La transformation de 1 mg de NO₃ en N₂ apporte 2,86 mg d'O2. La réaction chimique est comme suit :

 $NO_3^- + 6H^+ + 5e \longrightarrow 1/2N_2 + 3H_2O....(11)$

Les valeurs d'ammonium NH₄ enregistrés à l'entrée de la STEP (figure 22) varient entre 18,03 et 52,29 mg / l, tandis que les valeurs de sortie varient entre 0.26 et 9.27 mg /l. On remarque que plusieurs valeurs de NH₄ à la sortie dépassent la norme OMS (2 mg/l) notamment dans la période hivernale lorsque la teneur en oxygène est insuffisante et la température diminue pour assurer sa transformation.



Figure 23 : Représentation graphique de la variation de NO₂ à l'entrée et à la sortie de STEP (2016-2017).

D'après la figure 23 on note que les teneurs en NO_2 des eaux usées brutes et traitées sont faibles. Ainsi, la différence de ces valeurs entre l'entrée et la sortie renseigne une légère variation, dont les valeurs se situent entre 0,21 et 3,62 mg/l au niveau des eaux brutes. Ceci est en fonction de la qualité d'eau usée. Les valeurs mesurées après le traitement varient de 0,02 à 2,49 mg/l. Les valeurs de NO₂ dans l'eau épurée sont largement inférieures à la limite fixée par l'OMS (1mg/l) sauf dans quelque jour du mois de février où on peut les considérer comme des points aberrants.

Selon les résultats indiqués sur la figure 24, on observe que les valeurs de NO₃ obtenues d'après les analyses, varient entre 0,02 et 5,61 mg/l à l'entrée et de 0,02 et 33,93 mg/l à la sortie. On peut dire que Les valeurs de NO₃ sont élevées mais restent au-dessous de la norme de rejet fixée par l'OMS (50 mg/l).



Figure 24 : Représentation graphique de la variation de NO₃ à l'entrée et à la sortie de STEP(2016-2017).

La partie de l'azote libérée par la réaction de nitrification de l'ammonium combiné avec l'oxygène fourni par le processus de traitement est ajoutée à la partie de NO_2 et NO_3 de la sortie de la STEP confirment les équations 9 et 10. En présence d'une quantité très importante en oxygène à la sortie de la STEP, les teneurs de nitrate sont augmentés, ce que l'on observe à travers la figure 24, Ce qui nécessite un traitement tertiaire.

IV.4.10 Evaluation de la pollution organique des eaux brutes

Pour une meilleure appréciation de l'origine des eaux usées brutes entrantes à la STEP de Touggourt, le rapport DCO/DBO₅ donne un aperçu très important afin d'estimer le degré de pollution de ces eaux (Tableau 5).

| | DCO/DBO5 |
|------------|----------|
| Moyenne | 1,6790 |
| Ecart-type | 0,32039 |
| Minimum | 0,72 |
| Maximum | 2,22 |

Tableau 5: rapports exprimant la pollution organique des eaux brutede la STEP de Touggourt



Figure 25 : Représentation graphique du rapport de biodégradabilité des eaux brutes(2016-2017).

La figure 25 présente le rapport de biodégradabilité DCO/DBO₅ au cours de la période d'étude, on note que ce rapport est varié entre 0.72 et 2.81, ce qu'implique que les eaux brutes entrantes à la STEP sont biodégradables avec un rapport DCO/DBO₅ inférieur à 3, même s'ils contiennent des charges polluantes élevées et assez variées (Petry *et al.*, 2002).

IV.5 Conclusion

D'après les résultats obtenus depuis l'analyse statistique et le test de l'homogénéité de la série des données à travers la boîte à moustaches, on observe que les valeurs des paramètres sont très homogènes avec quelques points aberrants. La représentation graphique montre une grande différence entre les valeurs de pollution des eaux traitées et celles des eaux brutes, ceci dénote que le procédé de boue activée et très utile pour l'épuration des eaux usées domestique de Touggourt. À travers le rapport de la biodégradabilité, on conclut que les eaux brutes entrantes à la station sont des eaux biodégradables. On note que la station nécessite un traitement complémentaire en matière de traiter le NO₃ et le PO₄ à cause de leur teneur élevée à la sortie de la station. Les résultats montrent que les concentrations des paramètres de pollution augmentent dans la période estivale et diminuer dans la période hivernale. Les résultats montrent que la station d'épuration respecte les normes de rejet pour la plupart des paramètres.

Chapitre V

V Analyse en composantes principales (ACP)

V.1 Introduction

Afin de découvrir les relations existantes entre les paramètres physicochimiques, de mieux évaluer la STEP de la ville de Touggourt et aussi pour réduire le nombre important de ces paramètres analysés à l'entrée et à la sortie de la STEP et devant être considérés dans l'élaboration du modèle, une analyse statistique en composantes principales (ACP) a été appliquée à l'ensemble des paramètres pendant la période d'étude. Cette technique permet également de réduire les nombres des variables d'origine. Pour le traitement des données par analyse en composantes principales, on a utilisé les paramètres suivants : DBO₅, DCO, NH₄, OD, MES, PO₄, T, NO₂, NO₃, pH et CE à l'entrée et la sortie de la STEP, en plus le RECY et la MESBA.

V.2 Matrice de corrélation

Comme première mesure et compte tenu de la taille de la base de données mesurées ; on a établi la matrice des corrélations, elle est par hypothèse symétrique par rapport au diagonal, elle présente une première idée des relations entre les différentes variables. La matrice de corrélation de Pearson est obtenue par le calcul des valeurs de coefficient de corrélation (R). Les valeurs des coefficients R compris entre -1,0 et 1,0 et s'agissant d'un index sans dimension. Il convient d'affirmer qu'une valeur positive indique une relation fonctionnelle croissante entre deux variables et inversement lorsque cette valeur est négative. Ce dernier concerne la relation entre les variables considérées deux à deux, il va permettre de déceler toutes les variables explicatives potentiellement intéressantes pour la simulation de la STEP. Les résultats obtenus sont présentés dans le Tableau 6.

PARTIE 03 : RESULTATS ET DISCUSSIONS

| | | | | | | | | _ | 1 | - | , v | 1 | - | 1 | | | | | | | 0 | 0 | | DEC |
|-------------------|-----------|------------|-----------|----------|-----------|-------------------|--------|-------------------|-------------------|--------|--------|-----------|--------|----------|--------|----------|-------------------|--------|-------------------|-------------------|--------|--------|-----------|-------------------|
| | DBO5 e | DCO e | NH4e | O2e | MES e | PO ₄ e | Te | NO ₂ e | NO ₃ e | PHe | CEe | DBO5 s | NH4s | DCO s | O2s | MES s | PO ₄ s | Ts | NO ₂ s | NO ₃ s | PHs | CEs | MES BA | REC YCL AGE |
| DBO5e | 1 | 0,641 | 0,593 | -0,549 | 0,468 | 0,647 | 0,614 | 0,354 | 0,341 | 0,205 | 0,233 | 0,467 | 0,204 | 0,579 | -0,328 | 0,686 | 0,176 | 0,527 | -0,053 | -0,102 | 0,333 | 0,134 | 0,226 | 0,249 |
| DCOe | 0,641 | 1 | 0,477 | -0,474 | 0,524 | 0,536 | 0,419 | 0,321 | 0,357 | 0,198 | 0,317 | 0,405 | 0,086 | 0,629 | -0,171 | 0,535 | 0,199 | 0,319 | -0,017 | 0,079 | 0,193 | 0,209 | 0,164 | 0,165 |
| NH4e | 0,593 | 0,477 | 1 | -0,548 | 0,596 | 0,646 | 0,379 | 0,231 | 0,341 | 0,067 | 0,182 | 0,232 | 0,094 | 0,378 | -0,191 | 0,429 | 0,235 | 0,332 | 0,057 | 0,013 | 0,134 | 0,126 | 0,178 | 0,168 |
| O2e | -0,549 | -0,474 | -0,548 | 1 | -0,393 | -0,578 | -0,371 | -0,184 | -0,227 | -0,027 | -0,323 | -0,339 | -0,091 | -0,327 | 0,190 | -0,449 | -0,240 | -0,347 | 0,004 | -0,109 | -0,158 | -0,257 | -0,143 | -0,098 |
| MESe | 0,468 | 0,524 | 0,596 | -0,393 | 1 | 0,607 | 0,263 | ,331 | 0,351 | 0,135 | 0,340 | 0,262 | 0,153 | 0,396 | -0,218 | 0,464 | 0,119 | 0,200 | -0,030 | -0,034 | 0,029 | 0,058 | 0,078 | 0,084 |
| PO ₄ e | 0,647 | 0,536 | 0,646 | -0,578 | 0,607 | 1 | 0,306 | 0,195 | 0,196 | 0,018 | 0,341 | 0,366 | 0,150 | 0,449 | -0,230 | 0,546 | 0,267 | 0,272 | -0,037 | 0,060 | 0,037 | 0,152 | 0,109 | 0,097 |
| Te | 0,614 | 0,419 | 0,379 | -0,371 | 0,263 | 0,306 | 1 | 0,325 | 0,236 | 0,054 | 0,019 | -0,043 | -0,152 | 0,088 | 0,112 | 0,331 | 0,326 | 0,873 | -0,187 | 0,032 | 0,451 | 0,344 | 0,608 | 0,606 |
| NO ₂ e | 0,354 | 0,321 | 0,231 | -0,184 | 0,331 | 0,195 | 0,325 | 1 | 0,382 | 0,133 | 0,083 | 0,093 | -0,029 | 0,206 | -0,021 | 0,283 | 0,039 | 0,258 | -0,063 | 0,032 | 0,237 | 0,106 | 0,173 | 0,229 |
| NO ₃ e | 0,341 | 0,357 | 0,341 | -0,227 | 0,351 | 0,196 | 0,236 | 0,382 | 1 | 0,112 | 0,104 | 0,026 | -0,041 | 0,280 | -0,075 | 0,177 | -0,011 | 0,114 | -0,028 | 0,129 | 0,282 | 0,107 | 0,137 | 0,201 |
| РНе | 0,205 | 0,198 | 0,067 | -0,027 | 0,135 | 0,018 | 0,054 | 0,133 | 0,112 | 1 | 0,140 | ,108 | 0,015 | 0,258 | -0,028 | 0,138 | -0,127 | -0,077 | 0,216 | -0,054 | 0,347 | -0,088 | -0,043 | 0,023 |
| CEe | 0,233 | 0,317 | 0,182 | -0,323 | 0,340 | 0,341 | 0,019 | 0,083 | 0,104 | 0,140 | 1 | 0,238 | 0,180 | 0,138 | -0,019 | 0,315 | 0,243 | -0,045 | 0,005 | 0,152 | -0,008 | 0,203 | -0,020 | -0,003 |
| DBOs | 0,467 | 0,405 | 0,232 | -0,339 | 0,262 | 0,366 | -0,043 | 0,093 | 0,026 | 0,108 | 0,238 | 1 | 0,620 | 0,531 | -0,676 | 0,701 | 0,047 | 0,014 | 0,188 | -0,141 | 0,045 | -0,056 | -0,535 | -0,582 |
| NH ₄ s | 0,204 | 0,086 | 0,094 | -0,091 | 0,153 | 0,150 | -0,152 | -0,029 | -0,041 | 0,015 | 0,180 | 0,620 | 1 | 0,355 | -0,576 | 0,463 | -0,011 | -0,130 | 0,218 | -0,148 | 0,068 | -0,108 | -0,536 | -0,508 |
| DCOs | 0,579 | 0,629 | 0,378 | -0,327 | 0,396 | 0,449 | 0,088 | 0,206 | 0,280 | 0,258 | 0,138 | 0,531 | 0,355 | 1 | -0,493 | 0,539 | -0,165 | 0,012 | 0,078 | -0,162 | 0,174 | -0,142 | -0,165 | -0,140 |
| O2s | -0,328 | -0,171 | -0,191 | 0,190 | -0,218 | -0,230 | 0,112 | -0,021 | -0,075 | -0,028 | -0,019 | -0,676 | -0,576 | -0,493 | 1 | -0,486 | 0,348 | 0,092 | -0,090 | 0,422 | -0,010 | 0,343 | 0,525 | 0,475 |
| MESs | 0,686 | 0,535 | 0,429 | -0,449 | 0,464 | 0,546 | 0,331 | 0,283 | 0,177 | 0,138 | 0,315 | 0,701 | 0,463 | 0,539 | -0,486 | 1 | 0,194 | 0,361 | 0,045 | -0,111 | 0,176 | 0,091 | -0,113 | -0,134 |
| PO ₄ s | 0,176 | 0,199 | 0,235 | -0,240 | 0,119 | 0,267 | 0,326 | 0,039 | -0,011 | -0,127 | 0,243 | 0,047 | -0,011 | -0,165 | 0,348 | 0,194 | 1 | 0,336 | 0,366 | 0,558 | 0,031 | 0,583 | 0,198 | 0,151 |
| Ts | 0,527 | 0,319 | 0,332 | -0,347 | 0,200 | 0,272 | 0,873 | 0,258 | 0,114 | -,077 | -0,045 | 0,014 | -0,130 | 0,012 | 0,092 | 0,361 | 0,336 | 1 | -0,200 | -0,028 | 0,310 | 0,300 | 0,627 | 0,531 |
| NO ₂ s | -0,053 | -0,017 | 0,057 | 0,004 | -0,030 | -0,037 | -0,187 | -0,063 | -0,028 | 0,216 | 0,005 | 0,188 | 0,218 | 0,078 | -0,090 | 0,045 | 0,366 | -0,200 | 1 | 0,223 | 0,035 | 0,040 | -0,258 | -0,207 |
| NO ₃ s | -0,102 | 0,079 | 0,013 | -0,109 | -0,034 | 0,060 | 0,032 | 0,032 | 0,129 | -0,054 | 0,152 | -0,141 | -0,148 | -0,162 | 0,422 | -0,111 | 0,558 | -0,028 | 0,223 | 1 | 0,128 | 0,708 | 0,152 | 0,156 |
| PHs | 0,333 | 0,193 | 0,134 | -0,158 | 0,029 | 0,037 | 0,451 | 0,237 | 0,282 | 0,347 | -0,008 | 0,045 | 0,068 | 0,174 | -0,010 | 0,176 | 0,031 | 0,310 | 0,035 | 0,128 | 1 | 0,202 | 0,127 | 0,224 |
| CEs | 0,134 | 0,209 | 0,126 | -0,257 | 0,058 | 0,152 | 0,344 | 0,106 | 0,107 | -0,088 | 0,203 | -0,056 | -0,108 | -0,142 | 0,343 | 0,091 | 0,134 | 0,126 | 0,209 | -0,257 | 0,058 | 1 | 0,344 | 0,106 |
| MESBA | 0,226 | 0,164 | 0,178 | -0,143 | 0,078 | 0,109 | 0,608 | 0,173 | 0,137 | -0,043 | -0,020 | -0,535 | -0,536 | -0,165 | 0,525 | -0,113 | 0,198 | 0,627 | -0,258 | 0,152 | 0,127 | 0,279 | 1 | 0,904 |
| RECYC LAGE | 0,249 | 0,165 | 0,168 | -0,098 | 0,084 | 0,097 | 0,606 | 0,229 | 0,201 | 0,023 | -0,003 | -0,582 | -0,508 | -0,140 | 0,475 | -0,134 | 0,151 | 0,531 | -0,207 | 0,156 | 0,224 | 0,232 | 0,904 | 1 |
| Note : les | paramèt | res de l'o | entrée so | nt menti | ionnés pa | ar e | 1 | | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | I. | | |
| les | paramèt | res de la | sortie so | ont ment | ionnés p | ar s | | | | | | | | | | | | | | | | | | |

Tableau 6 : Matrice de corrélation des paramètres physico-chimiques à l'entrée et à la sortie de la STEP de Touggourt

CHAPITRE V

D'après la matrice de corrélation, on a remarqué que le coefficient R varie entre 0.904 et 0.003. On trouve que la corrélation entre plusieurs variables de système est inférieure à 0,5. Cela peut être attribué aux conditions de fonctionnement non uniformes caractérisées par la variabilité en R, aux relations non linéaires entre les paramètres de système, ainsi qu'aux erreurs dans les mesures. L'analyse entre les paramètres physico-chimiques étudiés montre qu'il existe :

- Une bonne corrélation positive entre la DBO₅e et (DCOe, NH₄e, MESe, PO₄e,T, DBO₅s, DCOs, MESs et Ts) avec un coefficient de corrélation (R= 0.641 ,0.593, 0.468, 0.647, 0.614, 0.467, 0.579, 0.686 et 0.527) respectivement ;
- Une bonne corrélation positive entre la DCOe et (le NH₄e, MESe, PO₄e, T, DBO₅ s, DCOs et MESs) avec un coefficient de corrélation (R= 0.477 ,0.524, 0.536, 0.419, 0.405, 0.629 et 0.535) respectivement ;
- Le NH₄ est bien positivement corrélé avec MESe, PO₄e, et MESe avec un coefficient de corrélation (R= 0.596, 0.646 et 0.429) respectivement ;
- Le OD est négativement corrélé avec tous les paramètres où on remarque un coefficient de corrélation élevé avec les paramètres de pollution organique notamment la DBO₅e, DCOe, MESe et PO₄e ;
- La MESe est bien corrélé avec le PO₄e par un coefficient (R=0.607) et avec la MESs par un coefficient (R=0.464) ;
- La Te et Ts sont très corrélés (R= 0,873). On note aussi une bonne corrélation positive entre la Te et le RECY d'une part et la Te et la matière en suspension dans le bassin d'aération d'autre part avec un coefficient de corrélation (R= 0.606 et 0.608) respectivement ;
- La DBO₅s, DCOs, NH₄s et MESs sont plus liés entre elles. Ceci suggère l'existence des relations communes.
- La plus forte corrélation signalée entre les paramètres physico-chimique de cette station est celle entre le RECY et la MESBA avec un coefficient de corrélation (R=0.904) ;
- L'absence de corrélation significative entre les différentes formes azotées ce qui pourrait suggérer que les formes azotées pourraient avoir plusieurs origines, une origine exogène (rejets industriels et domestiques) et une origine interne par le processus de nitrification et de dénitrification ;

- Une légère corrélation positive entre la CEe et la MESe et également entre la CEe et PO₄e avec un coefficient de corrélation (R=0.340 et 0.341) respectivement ;
- Il est à noter également que la MESBA a une corrélation significative avec la DBO_5s et NH_4s par un coefficient négatif (R= -0.535 et -0.536) respectivement ;
- Il existe une corrélation positive entre le PHs et Te avec un coefficient (R=0.451).

On constate, que les paramètres de pollution DBO₅, DCO, MES, NH₄ ont une bonne attirance entre eux.

Ces résultats montrent d'une part, qu'il existe une certaine redondance d'information, et d'autre part les différents paramètres ne présentent pas des hautes significations entre elles. Une analyse multidimensionnelle aura pour objectif d'étudier simultanément les variables. L'ACP permet de réduire le nombre initial des variables, par un plus petit nombre possible de variables qui correspondent à une combinaison linéaire des variables de base, en vue de les intégrer par la suite dans un modèle. (Lebart *et al.*, 2006).

L'approche de l'ACP utilise toutes les variables d'origine pour obtenir un ensemble plus petit de nouvelles variables, appelées composantes principales (CP). Le nombre de CP requis pour expliquer les données dépend du degré de corrélation entre l'ensemble de données - plus le degré de corrélation entre les variables d'origine est grand, plus le nombre de nouvelles variables requises est petit (Al-Ghazzawi et Lennox, 2008; Refaat, 2007).

V.3 Analyse des facteurs

Le tableau 7 permet de dégager une première approche typologique des différentes variables selon leurs affinités et leur regroupement sur les composantes principales à partir de leur contribution. Analytiquement on prend uniquement les axes ayant des valeurs les plus élevées en matière de valeurs propres et leur taux de variances. On remarque que les six premières composantes principales, expliquent 72,389 % de la variance totale. Le portrait des pondérations des variables originales sert à interpréter chaque composante principale, alors que la variance associée indique quel pourcentage de la variance totale de l'ensemble des variables originelles chaque composante principale représente.

| Tableau | 7 : Matrice de | es composa | ntes | | | | | |
|--------------------|------------------|---------------|---------------------|------------------------|-------------|---------------|---------|--|
| | | | | Comp | osante | | | |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| DBO5 | | <u>0,884</u> | -0,014 | -0,169 | -0,004 | 0,134 | 0,074 | |
| DCO | | <u>0,780</u> | -0,013 | 0,008 | 0,086 | -0,169 | 0,015 | |
| MESs | | <u>0,767</u> | -0,342 | 0,043 | -0,096 | 0,222 | -0,017 | |
| PO ₄ | | <u>0,753</u> | -0,069 | 0,102 | -0,291 | -0,250 | 0,126 | |
| NH ₄ | | <u>0,721</u> | 0,028 | 0,003 | -0,137 | -0,221 | 0,130 | |
| O2 | | <u>-0,690</u> | -0,022 | -0,140 | 0,190 | 0,071 | -0,031 | |
| MES | | <u>0,673</u> | -0,090 | -0,029 | -0,086 | -0,429 | 0,023 | |
| Т | | <u>0,625</u> | 0,560 | -0,211 | -0,032 | 0,379 | 0,016 | |
| DCOs | | <u>0,610</u> | -0,449 | -0,180 | 0,192 | -0,123 | 0,043 | |
| Ts | | 0,552 | 0,526 | -0,206 | -0,221 | 0, 471 | 0,022 | |
| NO ₂ | | 0,443 | 0,146 | -0,176 | 0,298 | -0,108 | -0,399 | |
| CE | | 0,372 | -0,085 | 0,354 | -0,049 | -0,357 | 0,109 | |
| MESBA | | 0,217 | 0,859 | -0,242 | -0,101 | -0,040 | 0,151 | |
| RECYCL | LAGE | 0,217 | 0,835 | -0,275 | 0,040 | -0,075 | 0,142 | |
| O2s | | -0,308 | 0,763 | 0,763 0,222 0,0 | | -0,174 | 0,106 | |
| DBO ₅ s | | 0,501 | -0,711 0,173 | | -0,059 | 0,244 | -0,066 | |
| NH ₄ s | | 0,237 | -0,682 | -0,682 0,182 | | 0,304 | -0,100 | |
| NO ₃ s | | 0,046 | 0,401 | 0,746 | 0,216 | -0,093 | -0,217 | |
| PO ₄ s | | 0,301 | 0,366 | 0,725 | -0,124 | 0,178 | 0,127 | |
| CEs | | 0,268 | 0,483 | 0,621 | 0,040 | 0,130 | -0,263 | |
| NO ₂ s | | 0,001 | -0,220 | 0,511 | 0,364 | 0,126 | 0,401 | |
| PH | | 0,192 | -0,109 | -0,111 | 0,699 | -0,071 | 0,485 | |
| PHs | | 0,346 | 0,185 | -0,106 | 0,608 | 0,397 | -0,097 | |
| NO ₃ | | 0,434 | 0,111 | -0,125 | 0,394 | -0,332 | -0,448 | |
| X 7 1 | Total | 6,403 | 4,573 | 2,353 | 1,573 | 1,451 | 1,021 | |
| Valeurs | % de la | 76 670 | 10.054 | 0.002 | 6 550 | 6 047 | 1 252 | |
| initiales | variance | 20,078 | 19,050 | 9,003 | 0,352 | 0,047 | 4,233 | |
| minuales | % cumulés | 26,678 | 45,734 | 55,537 | 62,089 | 68,136 | 72,389 | |
| Méthode | d'extraction : A | Analyse en d | composante | s principale | es à 6 comp | osantes ext | raites. | |

La qualité d'information récupérée dans I 'axe est présentée par les valeurs propres provenant de la matrice de corrélation. La variance et le pourcentage cumulé permettent de déterminer le poids relatif de chaque facteur. Pour cette étude on peut dire que la première valeur propre représente la variance expliquée par la première composante principale (6,403), elle correspond à 26,678 % de la variance totale, donc il représente l'axe prédominant , la deuxième valeur propre est égale à (4,573) explique 19,056 % de l'inertie total représente la deuxième composante principale , la troisième composante principale a une valeur propre (2,353) qui explique 9,803% de l'inertie totale et ainsi de suite jusqu'à la vingt-quatrième valeur .

Alors, on peut constater que seules les trois premières composantes principales qui expliquent près de 55.537 % seront prises en considération dans cette étude.

V.4 Analyse des variables et des individus

La projection des variables et des individus a été effectuée sur 6 axes, qui représentent 72% de la variance totale.

La projection sur le plan I-II : montre l'existence de quatre groupes (figure 26) :

Le premier et le deuxième se développe sur l'axe I (26,678 % de la variance). Le premier groupe présente la température et les paramètres de la pollution organique à l'entrée de la STEP notamment la DBO₅e, DCOe, MESe, NH₄e en plus deux paramètres de la sortie qui sont la DCOs et MESs, tout est corrélé positivement avec cet axe, le deuxième est corrélé négativement avec l'axe, représente l'OD à l'entrée de la STEP. Le troisième et le quatrième ensemble se positionnent sur l'axe II (19,056% de la variance) les valeurs négatives représentent la pollution ammoniacale NH4s et la DBO5s à la sortie de la STEP. Par contre les valeurs positives représentent les paramètres de contrôle de la STEP : le RECY et la MESBA en plus de l'OD à la sortie. La projection des individus sur le plan montre que les mois de la période estival ont des teneurs élevés en pollution organique et contrairement pour la période hivernale, cela est expliqué par la position de la T et les paramètres de la pollution organique par rapport au plan I. Généralement la charge polluante implique la diminution de l'OD, on peut le déduire par la position de ces paramètres dans l'axe I, où on remarque la diminution de l'oxygène au cours des mois les plus chauds puis il s'augmente progressivement si on va vers la période hivernale où la charge polluante est faible par rapport les mois chauds, cela est inversement proportionnel avec les paramètres de pollution organique.



Figure 26. Projection des variables et des individus sur le plan I-II (ACP)

On remarque que les mois froids ont des valeurs élevées de la DBO₅s et NH₄s à la sortie de la STEP par rapport aux autres mois, cela est attribué à la baisse de la Température qui influe négativement sur l'activité bactérienne et également sur leur rendement épuratoire.

La projection des individus montre que les mois février, Mars, avril nécessite un rapport important de RECY qui conduit vers la diminution de la charge polluante.

La Projection sur le plan III- IV : ce plan représente 16.355% de la variance, il est caractérisé par la présence de deux groupes (figure 27) :



Figure 27. Projection des variables et des individus sur le plan III- IV (ACP)

-Le premier groupe sur l'axe III (9.803% de la variance) indique l'évolution de la salinité à la sortie de la STEP exprimé par la CEs, les PO₄s et les NO₃s, ces dernières représentent la forme minérale azotée à la sortie de la STEP, tous ces paramètres sont corrélés positivement avec l'axe III. D'après la projection des individus sur ce plan, on note que les mois de printemps se caractérisent par des concentrations élevées en CEs, NO₃s à la sortie de la STEP, cela est dû à la non optimisation des paramètres d'oxygénation au niveau de la filière intensive de la STEP qui conduisent vers l'inhibition de l'étape de dénitrification, ce qui conduit à l'augmentation des teneurs de nitrates.

-Le deuxième est construit autour du pH à l'entrée et à la sortie de la STEP qui sont positivement corrélés avec l'axe IV (6.552% de la variance). Cela indique qu'il y a une évolution du pH entre eaux brutes et traités. La projection des individus sur cet axe ne donne aucune interprétation présentable en terme du pH, cela est dû aux valeurs de ces paramètres qui sont variées dans une marge très courte (entre 6.94 et 7.96).

-La Projection sur le plan V-VI : ce plan représente 10.3% de la variance (figure 28) : D'après ce plan on remarque que tous les paramètres n'ont aucune corrélation significative avec les deux axes de ce plan, ce qui a conduit à la disparition des sous-groupes caractérisant les axes. Ce résultat est dû à la faible variance qu'il contient.



Figure 28. Projection des variables et des individus sur le plan 5-6 (ACP)

V.5 Conclusion

L'application de l'ACP a été basée sur les indicateurs de pollution des eaux brutes et des eaux épurées, elle présente l'existence des degrés différents des corrélations entre les paramètres :

Il existe une corrélation significativement positive entre les paramètres d'entrées notamment la DBO₅e, la DCOe, NH₄e, PO₄e MESe, et la Te. Pour les corrélations qui existent entre la MES, la DCO et la DBO₅, on peut dire que la MES est de nature organique en grande partie puisque cette étude a reflété que cette eau usée est biodégradable, cela veut dire que la matière organique incluse dans la MES sera consommée par les microorganismes et demande plus l'oxygène (DCO et DBO₅), tandis qu'on note qu'il existe une corrélation significativement négative entre les paramètres précédemment cités et l'oxygène dissout, cela est dû à la présence de l'activité bactérienne qui absorbe l'oxygène de l'eau lors de la dégradation de la matière polluante . On observe qu'il existe presque la même corrélation entre les paramètres à la sortie de la STEP sauf les PO₄ qui n'ont pas une corrélation avec les autres paramètres de sortie. On note qu'il n'existe pas une corrélation significative entre les paramètres de sortie. No note qu'il n'existe pas une corrélation significative entre les paramètres de sortie. No note qu'il n'existe pas une corrélation significative entre les paramètres de sortie. No note qu'il n'existe pas une corrélation significative entre les paramètres de sortie. No note qu'il n'existe pas une corrélation significative entre les paramètres de sortie. No note qu'il n'existe pas une corrélation significative entre les paramètres de sortie. No note qu'il n'existe pas une corrélation significative entre les paramètres notamment le NH₄, NO₃ et NO₂ à l'entrée et à la sortie de la STEP.

L'application de l'ACP sur les eaux usées brutes et traitées informe qu'il existe une relation étroite entre la DBO₅e, DCOe, MESs, PO₄e, NH₄e, Ode, MESe, Te et DCOs sur l'axe F1. Cet axe peut être assimilé à un axe reflétant le degré de la pollution organique à l'entrée de la STEP et les conditions climatiques représentées par la température. Les paramètres MESBA, RECY, ODs, DBO₅s, NH₄s sont bien corrélés à l'axe F2, ce dernier représente les paramètres de contrôle et leur résultant en matière de la dégradation de la pollution organique notamment la pollution ammoniacale. Tandis qu'on n'observe aucune corrélation significative pour les autres axes. Finalement on peut conclure que l'analyse de l'ACP a réduit le nombre des variables d'origine qui sont affectés la STEP de Touggourt.

Chapitre VI

VI Simulation de la STEP de Touggourt par le RNA

VI.1 Introduction

La simulation de la station d'épuration de boue activée de Touggourt joue un rôle important dans l'amélioration de ses performances. Dans ce chapitre, le RNA-PMC est développé pour simuler la STEP de Touggourt. Les RNA-PMC consistent un ensemble de neurones artificiels appelés nœuds, et ils ont des connexions entre eux, appelés poids. Les valeurs optimales pour ces poids sont obtenues en formant du réseau. L'utilisation d'un modèle neuronale (RNA-PMC) passe à la fois par le choix des variables d'entrée et aussi par l'optimisation de l'architecture du réseau de neurones lui-même. Ceci signifie qu'il faut trouver à la fois le type de fonction d'activation qui convient au problème et le nombre de neurones qui constitue la couche cachée. L'algorithme d'apprentissage utilisé pour tous les modèles est le Levenberg - Marquardt. Sur la base des études précédentes, il est constaté que cet algorithme fournit une vitesse d'apprentissage rapide et des performances élevées par rapport à d'autres algorithmes d'optimisation. Les détails de cet algorithme sont rapportés par Hagan et al. (1996). En général, une PMC, formées à l'aide de l'algorithme de Levenberg-Marquardt, peut approximer n'importe quelle fonction avec une précision suffisante (Hagan et al., 1996; Daliakopoulos et al., 2005). La fonction d'activation choisie est la Tangente hyperbolique pour les couches cachées et la couche de sortie.

VI.2 Le choix des variables d'entrées

Le choix des variables d'entrée est basé sur l'examen de la matrice de corrélation et les résultats de l'ACP. Une analyse de corrélation, basée sur les caractéristiques des données, est réalisée entre les variables (chapitre précédent). Les variables les plus corrélées avec la DBO₅s, la DCOs et les MESs des eaux traités sont :

La DBO₅e, DCOe, NH₄e, ODe, MESe, PO₄e, MESBA, Recy, NH₄s, DCOs, ODs et MESs qui présentent des coefficients de corrélation de 0,467, 0,405, 0,232, -0,339, 0,262, 0,366, -0,535, -0,582, 0,620, 0,531, -0,676 et 0,701 respectivement avec la DBO₅s des eaux traitées, aussi bien que le facteur de la température T qui a été montré par l'analyse de l'ACP.

D'autre part la DBO₅e, DCOe, NH₄e, ODe, MESe, PO₄e, DBO₅s, ODs et MESs ont des coefficients de corrélation de 0,579, 0,629, 0,378, -0,327, 0,396, 0,449, 0,531, -0,493 et 0,539 avec la DCOs de l'effluent respectivement plus la température Te. La MES de l'effluent

est corrélée avec la DBO₅e, DCOe, NH₄e, ODe, MESe, PO₄, Te, DBO₅s, DCOs, ODs et NH₄s par des coefficients de corrélation de 0,686, 0,535, 0,429, -0,449, 0,464, 0,546, 0,331, 0,701, 0,539, -0,486, 0,463 respectivement. Celles-ci ont été choisies comme des variables d'entrée possibles dans les modèles neuronaux à réaliser.

Le but principal de la construction de ces modèles est de simuler l'effet de la variation des quantités et de qualité des influents à l'entrée de la station d'épuration de Touggourt sur la qualité des effluents, prenant en considération tous les paramètres qui influent sur les performances de cet procédé d'épuration.

VI.3 Prédiction de la DBO5s par le modèle PMC-RNA

VI.3.1 Prédiction de la DBO5s par l'utilisation des paramètres d'entrée de la STEP (model simple RNA-PMCSDBO5s)

Comme indiqué précédemment, les paramètres de l'influent utilisée pour le modèle simple RNA-PMCSDBO₅s sont comme suit : DBO₅e, DCOe, NH₄e, ODe, MESe, PO₄e, Te, MESBA et Recy, qui sont présentés dans le tableau ci-dessous.

Tableau 8 : Données d'entrée-sortie du modèle simple RNA-PMCSDBO5 pour la prédiction de la DBO5 à la sortie de la STEP

| Modèle | | | | Param | ètres de l | 'entrée | | | | Paramètre de la sortie |
|----------------------------------|---------------------------|-------------|-------------|------------|-------------|--------------------------|--------------|-------------|-----------|---------------------------|
| RNA- PMCSDBO5 (150 donnés) | DBO ₅ e (1) | DCOe (2) | NH4e (3) | ODe (4) | MESe (5) | PO ₄ e (6) | MESBA (7) | Recy (8) | Te (9) | DBO ₅ s |

Il est observé que la mise en œuvre du modèle a commencé par la détermination des combinaisons de variables présentant des meilleurs résultats dans la sortie du modèle RNA (Mjalli *et al.*, 2007). À cette fin, dans cette partie, la DBO₅s est modélisée séparément en considérant diverses variables uniques comme entrées de RNA-PMCSDBO₅s afin d'examiner l'effet de chaque variable sur la variation des valeurs de DBO₅s. De même, des modèles différents ont été réalisés afin de montrer l'effet des variables d'entrée multiple sur la variation des valeurs de la DBO₅s. Ces entrées sont utilisées pour former les RNA-PMCSDBO₅s en groupes de deux, trois, quatre, cinq, six, sept, huit et neuf variables. Le tableau 9 présente les détails de 45 modèles formés avec les données brutes. Comme indiqué précédemment, la technique d'arrêt précoce est utilisée dans laquelle le processus

d'apprentissage de modèle est arrêté lorsque l'erreur de validation commence à augmenter. Cela garantit que l'apprentissage excessif ne se produit pas.

| | | N | | | | | Racine de l'erreur | | | | |
|------|--------------------|----------|----------|----------------|--------------|---------|---------------------|-------|------|--|--|
| Scé | | IN. N | (| Coefficient de | e corrélatio | n | quadratique moyenne | | | | |
| na- | Paramètre d'entrés | N. | | | | | (REQM) | | | | |
| rios | | C. | Apprenti | Validatio | Test | Total | Appren | Valid | Test | | |
| | | C | ssage | n | 1031 | Total | tissage | ation | Test | | |
| 1 | (1) | 35 | 0,53389 | 0,8675 | 0,67877 | 0,595 | 2,97 | 1,51 | 2,88 | | |
| 2 | (2) | 40 | 0,68729 | 0,82563 | 0,7038 | 0,69023 | 2,78 | 1,48 | 2,76 | | |
| 3 | (3) | 60 | 0,4416 | 0,49303 | 0,62235 | 0,49865 | 2,98 | 1,75 | 3,23 | | |
| 4 | (4) | 40 | 0,55094 | 0,72219 | 0,69881 | 0,58825 | 2,89 | 2,01 | 2,92 | | |
| 5 | (5) | 25 | 0,45621 | 0,69096 | 0,50513 | 0,48101 | 3,15 | 1,85 | 3,29 | | |
| 6 | (6) | 40 | 0,57876 | 0,66221 | 0,47045 | 0,57108 | 2,88 | 1,96 | 3,11 | | |
| 7 | (7) | 50 | 0,69352 | 0,84814 | 0,74402 | 0,71636 | 2,40 | 1,41 | 2,57 | | |
| 8 | (8) | 45 | 0,69404 | 0,84891 | 0,73075 | 0,71626 | 2,51 | 1,64 | 2,99 | | |
| 9 | (9) | 40 | 0,65928 | 0,84266 | 0,61792 | 0,67677 | 2,78 | 1,59 | 2,25 | | |
| 10 | (7)(1) | 30 | 0,93275 | 0,98921 | 0,98072 | 0,94724 | 1,22 | 0,51 | 0,81 | | |
| 11 | (7)(2) | 35 | 0,83019 | 0,9665 | 0,8884 | 0,86613 | 1,84 | 1,16 | 1,14 | | |
| 12 | (7)(3) | 30 | 0,73921 | 0,88724 | 0,91275 | 0,78167 | 2,12 | 1,23 | 2,09 | | |
| 13 | (7) (4) | 20 | 0,83873 | 0,8146 | 0,89517 | 0,84078 | 1,92 | 1,30 | 1,59 | | |
| 14 | (7) (5) | 25 | 0,85346 | 0,85263 | 0,84064 | 0,84887 | 1,78 | 1,45 | 1,77 | | |
| 15 | (7)(6) | 30 | 0,75853 | 0,91437 | 0,75522 | 0,76774 | 2,35 | 1,30 | 1,88 | | |
| 16 | (7) (8) | 50 | 0,7614 | 0,90073 | 0,76819 | 0,77892 | 2,22 | 1,58 | 2,09 | | |
| 17 | (7) (9) | 20 | 0,80279 | 0,8708 | 0,78758 | 0,80998 | 2,23 | 1,26 | 1,31 | | |
| 18 | (7) (1) (2) | 30 | 0,91541 | 0,97887 | 0,95059 | 0,92881 | 1,44 | 0,58 | 1,55 | | |
| 19 | (7) (1) (3) | 20 | 0,93587 | 0,97914 | 0,95512 | 0,94066 | 1,20 | 0,57 | 1,15 | | |
| 20 | (7) (1) (4) | 20 | 0,92742 | 0,98718 | 0,95245 | 0,94203 | 1,24 | 0,60 | 1,20 | | |
| 21 | (7) (1) (5) | 25 | 0,79809 | 0,98078 | 0,97552 | 0,83147 | 2,30 | 0,55 | 1,56 | | |
| 22 | (7) (1) (6) | 25 | 0,90996 | 0,97682 | 0,9207 | 0,91685 | 1,45 | 0,61 | 1,28 | | |
| 23 | (7) (1) (8) | 30 | 0,92668 | 0,97775 | 0,98607 | 0,94258 | 1,28 | 0,58 | 1,70 | | |
| 24 | (7) (1) (9) | 20 | 0,93792 | 0,97029 | 0,9013 | 0,93645 | 1,26 | 0,61 | 1,42 | | |
| 25 | (7) (1) (3) (2) | 30 | 0,9126 | 0,97201 | 0,95936 | 0,92562 | 1,38 | 0,48 | 0,94 | | |
| 26 | (7) (1) (3) (4) | 25 | 0,94642 | 0,99083 | 0,94976 | 0,95248 | 1,12 | 0,48 | 1,82 | | |
| 27 | (7) (1) (3) (5) | 35 | 0,95162 | 0,97975 | 0,91218 | 0,94693 | 1,03 | 0,60 | 1,56 | | |
| 28 | (7) (1) (3) (6) | 15 | 0,92154 | 0,98869 | 0,95347 | 0,93569 | 1,29 | 0,52 | 1,01 | | |
| 29 | (7) (1) (3) (8) | 15 | 0,94542 | 0,98217 | 0,98421 | 0,95521 | 1,11 | 0,43 | 0,74 | | |
| 30 | (7) (1) (3) (9) | 30 | 0,93517 | 0,98432 | 0,94133 | 0,94217 | 1,09 | 0,59 | 1,48 | | |

Tableau 9 : Performances de la prédiction de la DBO5s par le modèle simple RNA-PMCSDBO5

| 31 | (7) (1) (3) (8) (2) | 40 | 0,96094 | 0,98486 | 0,97611 | 0,96559 | 0,92 | 0,46 | 0,87 |
|------|-----------------------------|---------|-------------|---------|---------|---------|------|------|-------|
| 32 | (7) (1) (3) (8) (4) | 45 | 0,9503 | 0,99004 | 0,97705 | 0,95961 | 1,03 | 0,45 | 0,75 |
| 33 | (7) (1) (3) (8) (5) | 30 | 0,94654 | 0,98148 | 0,98272 | 0,95749 | 1,05 | 0,55 | 0,79 |
| 34 | (7) (1) (3) (8) (6) | 25 | 0,94285 | 0,98759 | 0,98044 | 0,95501 | 1,11 | 0,44 | 0,76 |
| 35 | (7) (1) (3) (8) (9) | 45 | 0,96838 | 0,98931 | 0,97009 | 0,97112 | 0,88 | 0,45 | 0,70 |
| 36 | (7) (1) (3) (8) (9) (2) | 45 | 0,96306 | 0,99209 | 0,94012 | 0,96043 | 0,90 | 0,43 | 1,33 |
| 37 | (7) (1) (3) (8) (9) (4) | 30 | 0,9669 | 0,99274 | 0,97235 | 0,96998 | 0,89 | 0,39 | 0,70 |
| 38 | (7) (1) (3) (8) (9) (5) | 45 | 0,9615 | 0,9882 | 0,98711 | 0,9684 | 0,85 | 0,42 | 0,56 |
| 39 | (7) (1) (3) (8) (9) (6) | 55 | 0,96132 | 0,9894 | 0,9753 | 0,9672 | 0,94 | 0,41 | 0,81 |
| 40 | (7) (1) (3) (8) (9) (5) (2) | 40 | 0,946 | 0,982 | 0,984 | 0,959 | 1,73 | 0,60 | 0,79 |
| 41 | (7) (1) (3) (8) (9) (5) (4) | 35 | 0,959 | 0,983 | 0,961 | 0,963 | 1,94 | 0,58 | 0,81 |
| 42 | (7) (1) (3) (8) (9) (5) (6) | 45 | 0,966 | 0,986 | 0,9783 | 0,9698 | 1,82 | 0,51 | 0,59 |
| 43 | (7) (1) (3) (8) (9) (5) (6) | 40 | 0.967 | 0.9826 | 0.9805 | 0.9718 | 0.87 | 0.50 | 0.58 |
| 15 | (2) | 10 | 0,707 | 0,7020 | 0,7005 | 0,7710 | 0,07 | 0,50 | 0,50 |
| 44 | (7) (1) (3) (8) (9) (5) (6) | 45 | 0.949 | 0.989 | 0.948 | 0.955 | 1.01 | 0.51 | 0.94 |
| | (4) | | 3,5 .5 | 0,203 | 3,2.0 | 0,500 | 1,01 | 0,01 | 0,5 . |
| 45 | (7) (1) (3) (8) (9) (5) (6) | 45 | 0 96959 | 0 98094 | 0 98792 | 0 97232 | 0.84 | 0.49 | 0.45 |
| 1.5 | (2)(4) | 10 | 0,70757 | 0,20024 | 0,70772 | 0,77252 | 0,01 | 0,12 | 0,13 |
| N.N. | C.C : Nombre de neurones da | ns la c | couche cach | ée | | | | | |

Les critères d'évaluation (c'est-à-dire REQM et R) sont calculés pour chaque scénario, et le nombre de neurones dans la couche cachée -correspondant à la meilleure performance- a été sélectionné d'après plusieurs tentatives. Étant donné que l'apprentissage a été arrêté en raison de l'erreur de validation, le nombre d'époques s'est varié pour chaque architecture. Le tableau 9 présente les performances de l'architecture optimale pour chaque scénario et le nombre de neurones dans la couche cachée correspondant au modèle sélectionné.

La DBO₅s est modélisée séparément en prenant en compte différentes variables uniques comme entrées des RNA afin d'examiner les effets de chaque variable sur les concentrations de la DBO₅s. L'effet de chaque variable sur les modèles simples RNA-PMCSDBO₅s par rapport aux autres variables est déterminé par son importance dans l'ordre des critères de performance. Afin de trouver le meilleur modèle, les critères d'évaluation obtenus lors de la phase d'apprentissage, validation et de test pour le meilleur modèle dans chaque combinaison ont été comparés. Les valeurs mises en surbrillance sont les performances les plus élevées atteintes, c'est-à-dire l'erreur REQM la plus faible et le coefficient de corrélation le plus élevé. Les résultats du meilleur modèle de différentes combinaisons sont représentés sous forme graphique dans la figure 29.


Figure 29 : Comparaison entre les meilleurs modèles simples RNA-PMCSDBO5s des différentes combinaisons

De même, différents modèles ont été établis afin de montrer les effets des variables d'entrée multiples sur la concentration de la DBO₅s. Les résultats montrent que le MESBA parmi les variables d'entrée simple, la MESBA et le DBO₅e parmi les groupes de deux variables, la MESBA, le DBO₅e et le NH₄e parmi les groupes de trois variables ,la MESBA, le DBO₅e et le NH₄e parmi les groupes de trois variables ,la MESBA, le DBO₅e et le NH₄e parmi les groupes de trois variables ,la MESBA, le DBO₅e le NH₄e et Recy parmi les groupes de quatre variables les plus considérables avaient un effet sur le modèle simple de DBO₅s. En outre, la MESBA, le DBO₅e le NH₄e Recy la Te et la MESE parmi les groupes de six variables, la MESBA, le DBO₅e le NH₄e Recy la Te, MESe et le PO₄e parmi les groupes de sept variables, la MESBA, le DBO₅e le NH₄e Recy la Te, MESe , PO₄e, et le DCOe parmi les groupes de huit variables qui ont les effets les plus importants sur le modèle simple de DBO₅e.

Finalement le groupe de neuf variables, y compris toutes les paramètres de l'influent considérées comme des d'entrée de modèle neuronale simple, est le plus efficace en matière de prédire la DBO₅s. Ce dernier donne des valeurs minimales de REQM et des valeurs maximales de R. Le RNA-PMCSDBO₅ donne des résultats assez bien pour le modèle RNA-PMCDBO₅s avec des paramètres d'entrée simples et des paramètres multiples, ils varient selon le degré de corrélation entre ces paramètres et la DBO₅s d'une part et entre ces paramètres eux-mêmes d'autre part. La figure 30 présente l'architecture finale du meilleur modèle simple RNA-PMCSDBO₅s qui est 9-45-1.



Figure 30 : L'architecture du meilleur modèle simple RNA-PMCSDBO5s

Comme le scénario 45 représente le modèle simple RNA-PMCDBO₅s avec 45 neurones dans la couche cachée comprenant tous les paramètres d'entrée et il offre les meilleures performances, une analyse plus approfondie est effectuée uniquement avec ce modèle. Cependant, le meilleur moyen pour évaluer la prédiction du modèle consiste à examiner des courbes prédictives et analyser le taux de congruence. Dans la figure 31, une comparaison graphique entre les valeurs de la DBO₅s prédites par le modèle simple RNA-PMCSDBO₅s et les valeurs de la DBO₅s mesuré.





Les résultats du modèle simple RNA-PMCSDBO₅s montrent que lors de la première phase (phase d'apprentissage), le REQM est de 0.84 mg/l et le coefficient de corrélation R=0,9695. Ces paramètres indiquent que le modèle simple RNA-PMCSDBO₅s est en mesure de bien répondre aux données d'apprentissage et capable de les rapprocher. Le modèle simple RNA-PMCSDBO₅s est donc en mesure de résoudre le problème particulier de cartographie les données entrées-sorties.

Lors de la deuxième phase (phase de validation), la REQM de la validation est de 0.49 mg/l et le coefficient de corrélation R=0,9809. Ces paramètres indiquent que le modèle réagit bien dans cette phase où il donne des bons résultats. Lors de la troisième phase (phase de test), les résultats du modèle simple RNA-PMCSDBO₅s montrent que la REQM de la phase de test est de 0.45 mg/l, le coefficient de corrélation R= 0,9879. Ces résultats indiquent que les performances de généralisation du modèle simple RNA-PMCSDBO₅ sont bonnes. Le modèle simple RNA-PMCSDBO₅ est en mesure à faire des prédictions précises. D'après ces figures on peut dire que le modèle est capable de prédire la DBO₅s pendant les conditions de fonctionnement car les deux lignes (prédites et observées) sont identiques dans la majorité des points de courbe.



Figure 32 : Diagramme de régression entre les valeurs prédites et mesurées pour le meilleur modèle simple RNA-PMCSDBO5s

La Figures 32 montre le diagramme de régression des résultats de la DBO₅s mesurée et prédit lors de la phase d'apprentissage, validation et de test. On remarque qu'il y a une très bonne concordance entre valeurs mesurées et valeurs prédites.

VI.3.2 Prédiction de la DBO₅s par l'utilisation des paramètres d'entrée et de sortie de la STEP (model extensif RNA-PMCEDBO₅s)

Bien qu'il ne soit pas courant d'utiliser un paramètre de la sortie de la STEP comme entrée dans la modélisation, cette approche peut être utile pour un système expert où des modifications peuvent être apportées à l'opération non seulement en fonction des paramètres d'influent, mais également en fonction des paramètres de l'effluent. En outre, dans les cas où les conditions de fonctionnement hétérogènes ont une incidence sur les paramètres du système, les paramètres d'effluent peuvent contribuer à l'amélioration des prévisions des modèles.

Aussi bien que l'utilisation des paramètres d'entrée donnant le meilleur modèle simple, dans cette partie on utilise les paramètres de sortie de la STEP qui ont une bonne corrélation avec la DBO₅s. Ces paramètres sont choisis d'après l'analyse de l'ACP et la matrice de corrélation. Les paramètres globaux utilisés dans ces modèles sont comme suit : DBO₅e, DCOe, NH₄e, ODe, MESe, PO₄e, Te, MESBA, Recy, DCOs, NH₄s, ODs et MESs qui sont présentés dans le tableau ci-dessous.

Tableau 10: paramètres d'entrée-sortie du modèle extensif RNA-PMCEDBO₅s pour la prédiction de la DBO₅s a la sortie de la STEP

| Modèle | | Paramètres d'entrée | | | | | | | | | | |
|--------------|--------------------|---------------------|-------------------|------|------|-------------------|-------|------|-----|-------|--|--|
| RNA- | DBO ₅ e | DCOe | NH ₄ e | ODe | MESe | PO ₄ e | MESBA | Recy | Te | | | |
| PMCEDBO | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | | | |
| 58 | DCOs | NH ₄ s | ODs | MESs | | | - | | | DB058 | | |
| (150 donnés) | (10) | (11) | (12) | (13) | | | | | | | | |

Les performances de la prédiction de la concentration de la DBO₅s par le modèle extensif RNA-PMCEDBO₅s sont présentées dans le tableau ci-dessous.

| Scé | | N.N.C | C | Coefficient | de corrélatio | on | Raci quadra | ne de l'en atique mo (REQM) | rreur yenne | |
|-------------|--|-------|-------------------|----------------|---------------|--------|-----------------------|-----------------------------------|----------------|--|
| na- rios | Paramètres d'entrée | .C | Appren tissage | Validati on | Test | Total | Appre ntissa ge | Valid ation | Test | |
| 1 | (7) (1) (3) (8) (9) (5) (6) (2) (4) (10) | 30/10 | 0,9481 | 0,9954 | 0,9937 | 0,9609 | 1,071 | 0,297 | 0,435 | |
| 2 | (7) (1) (3) (8) (9) (5) (6) (2) (4) (11) | 25/10 | 0,9714 | 0,9924 | 0,9959 | 0,9780 | 0,813 | 0,366 | 0,332 | |
| 3 | (7) (1) (3) (8) (9) (5) (6) (2) (4) (12) | 30/10 | 0,9689 | 0,9935 | 0,99167 | 0,9737 | 0,866 | 0,315 | 0,325 | |
| 4 | (7) (1) (3) (8) (9) (5) (6) (2) (4) (13) | 35/10 | 0,9542 | 0,9927 | 0,9939 | 0,9676 | 0,840 | 0,404 | 0,420 | |
| 5 | (7) (1) (3) (8) (9) (5) (6) (2) (4) (12) (10) | 30/10 | 0,9529 | 0,9945 | 0,9899 | 0,9668 | 0,939 | 0,589 | 0,409 | |
| 6 | (7) (1) (3) (8) (9) (5) (6) (2) (4) (12) (11) | 20/10 | 0,9854 | 0,9943 | 0,9951 | 0,9878 | 0,556 | 0,411 | 0,274 | |
| 7 | (7) (1) (3) (8) (9) (5) (6) (2) (4) (12) (13) | 30/10 | 0,9493 | 0,9917 | 0,9919 | 0,9656 | 0,939 | 0,606 | 0,470 | |
| 8 | (7) (1) (3) (8) (9) (5) (6) (2) (4) (12) (11) (10) | 30/10 | 0,9908 | 0,9968 | 0,997 | 0,9936 | 0,393 | 0,258 | 0,364 | |
| 9 | (7) (1) (3) (8) (9) (5) (6) (2) (4) (12) (11) (13) | 30/10 | 0,9782 | 0,9898 | 0,9905 | 0,9802 | 0,727 | 0,463 | 0,514 | |
| 10 | (7) (1) (3) (8) (9) (5) (6) (2) (4) (12) (11) (10) (13) | 30/10 | 0,9934 | 0,9979 | 0,9961 | 0,9946 | 0,351 | 0,218 | 0,282 | |
| N.N. | N.N.C.C : Nombre de neurones dans les couches cachées | | | | | | | | | |

Tableau 11: Performances de la prédiction de la DBO5s par modèle extensif RNA-PMCEDBO5

Suite à la section précédente, on a essayé d'améliorer les modèles simples RNA-PMCSDBO₅s par l'intégration des paramètres de la sortie de la STEP de Touggourt suivant les mêmes démarches précédentes. Différents modèles sont réalisés afin de montrer l'effet des paramètres de la sortie simple et multiple sur la variation et l'amélioration du modèle extensif RNA-PMCEDBO₅s. Les paramètres d'entrée de ces modèles sont les neuf paramètres d'entrée de la STEP qui donnaient le meilleur modèle simple, en plus des paramètres de la sortie de la STEP. Le tableau11 présente les détails de 10 scénarios des modèles formés avec les paramètres d'eau brute et traitée utilisée dans laquelle le processus de formation est arrêté lorsque l'erreur de la phase de validation commence à augmenter. Les critères d'évaluation ont été calculés pour chaque architecture et le nombre de neurones dans les couches cachés est sélectionné d'après plusieurs tentatives. Due au nombre important des paramètres d'entrée de ces modèles, on utilise dans ce cas deux couches cachées à afin d'améliorer le fonctionnement et la précision de modèle. Tandis que les résultats pour une couche cachée dans le cas des modèles extensifs ont donné des résultats médiocres.

La DBO₅s est simulée par l'intégration des paramètres de la sortie de la STEP avec des différentes combinaisons. Afin de trouver le meilleur modèle, les critères d'évaluation obtenus lors de la phase d'apprentissage, validation et de test pour le meilleur modèle dans chaque groupe sont comparés. La figure 33 montre les performances des différents modèles extensifs RNA-PMCEDBO₅s développés.



Figure 33 : Comparaison entre les meilleurs modèles extensifs RNA-PMCEDBO5s des différentes combinaisons

Les résultats montrent que le ODs donne la meilleure performance par rapports les autres paramètres notamment la DCOs, NH₄s et MESs dans le cas où on utilise un seul paramètre de la sortie de la STEP. Dans le cas où on a utilisé deux paramètres de la sortie de la STEP on observe que la meilleure combinaison est donnée par l'utilisation d'ODs et NH₄s dans les paramètres d'entrée du modèle. L'utilisation de l'ODs, NH₄s et DCOs de la sortie de la STEP dans les paramètres d'entrée du modèle extensif RNA-PMCEDBO₅s donne les meilleures performances dans le cas où on a utilisé trois paramètres de la STEP. Les performances du meilleur modèle de chaque combinaison sont montrées dans la figure 33

Finalement l'utilisation de tous les quatre paramètres de la sortie de la STEP dans les paramètres d'entrée du modèle neuronal extensif RNA-PMCEDBO₅s donne le meilleur résultat en matière de prédire de la DBO₅s. La figure 34 montre l'architecture 13-30-10-1 qui

représente le modèle extensif RNA-PMCEDBO₅s comprenant tous les paramètres d'entrée et de sortie considérés, offre les meilleures performances. Une analyse approfondie est effectuée uniquement avec cette architecture. Dans la figure 35, une comparaison graphique entre les valeurs de la DBO₅s prédites par le modèle extensif RNA-PMCEDBO₅s et les valeurs de la DBO₅s mesuré.



Figure 34 : L'architecture du meilleur modèle extensif RNA-PMCEDBO5s



Figure 35 : Comparaison entre les valeurs prédites et les valeurs mesurées pour le meilleur modèle extensif RNA-PMCEDBO₅s

Les résultats du modèle extensif RNA-PMCEDBO5s montrent que lors de la première phase (phase d'apprentissage), le REQM de la phase d'apprentissage est de 0,351 mg/l et le coefficient de corrélation R=0.9934. Ces paramètres indiquent que le modèle extensif RNA-PMCEDBO₅s est en mesure de bien répondre de façon excellente aux données d'apprentissage et capable de les rapprocher, cela est dû aux valeurs rapprochées de la DBO₅s où on observe un écart-type plus faible de ce paramètre. Lors de la deuxième étape (phase de validation), le REQM est de 0.218 mg/l et le coefficient de corrélation R=0,9979. Ces paramètres indiquent que le modèle réagit bien dans cette phase où il donne des bons résultats. Lors de la troisième phase (phase de test), les résultats du modèle extensif RNA-PMCEDBO5s montrent que REQM de la phase de test est de 0.282 mg/l, le coefficient de corrélation R=0.9961. Les résultats de ces deux dernières phases indiquent que les performances de généralisation du modèle extensif RNA-PMCEDBO₅s sont très bonnes. Le modèle extensif RNA-PMCEDBO₅s est en mesure de faire des prédictions précises. On conclut que l'utilisation des paramètres de la sortie de la STEP notamment le ODs, DCOs, NH4s et MESs donne une forte amélioration de modèle neuronale simple malgré l'existence d'une forte non linéarité entre les paramètres d'entrée et le paramètre prédit d'une part et le nombre important des paramètres d'entrée de modèle d'autre part. D'après ces figures, on peut dire que le modèle est capable de prédire la DBO₅s pendant les conditions de fonctionnement car les deux lignes (prédites et observées) sont presque identiques.

Une analyse de régression est également réalisée et les résultats de ces analyses pour les phases d'apprentissage, validation et Test sont présentés dans la figure 36. L'accord parfait est atteint lorsque R est égal à 1. Les coefficients de corrélation montrent une bonne correspondance entre les valeurs prédites par le modèle extensif RNA-PMCEDBO₅s et les valeurs mesurées. Les fortes valeurs de coefficients de corrélation dans la phase de validation et la phase de Test impliquent que le modèle donne une très bonne prédiction avec des nouvelles valeurs (cela n'a pas été reconnu aux phases d'apprentissage). Dans ce cas, la réponse du réseau est largement satisfaisante et la simulation peut être utilisée pour des nouvelles entrées.



Figure 36 : Diagramme de régression entre les valeurs prédites et mesurées pour le meilleur modèle extensif RNA-PMCEDBO5s

VI.4 Prédiction de la DCOs par le modèle PMC-RNA

VI.4.1 Prédiction de la DCOs par l'utilisation des paramètres d'entrée de la STEP (model simple RNA-PMCSDCOs)

Afin de prédire la DCOs par RNA-PMCSDCOs, la DBO₅e, DCOe, NH₄e, ODe, MESe, PO₄e, et Te sont utilisés comme entrées de modèle RNA (tableau 12). Les données opérationnelles sur 300 jours (150 points de données au total) sont utilisées dans les modèles RNA-PMCSDCO.

 Tableau 12: paramètres d'entrée-sortie de modèle simple RNA-PMCSDCO pour la prédiction de la DCO a la sortie de la STEP

| Modèle | |] | | Paramètre de la sortie | | | | |
|---------------------------------|--------------|-------------|-------------|------------------------|-------------|-------------|-----------|------|
| RNA- PMCSDCO (150 donnés) | DBO5e (1) | DCOe (2) | NH4e (3) | ODe (4) | MESe (5) | PO4e (6) | Te (7) | DCOs |

Afin de simuler la concentration de la DCO à la sortie de la STEP de Touggourt par le RNA, la DBO₅e, DCOe, NH₄e, ODe, MESe, PO₄e, et Te sont considéré comme des entrées de réseau. Le RNA-PMCSDCOs est formé par l'algorithme d'apprentissage Levenberg – Marquardt qui a donné les modèles les plus précis pour les trois phases notamment : apprentissage, validation et test.

Le tableau 13 présente les détails de 28 scénarios formés avec les données brutes.

| | | | | | Racine de l'erreur | | | | |
|------|---------------------|-----|---------|---------------|--------------------|--------|----------|----------|-------|
| Scé | | NN | Co | pefficient of | le régressi | on | quadrati | ique moy | venne |
| nar | Paramètres d'entrée | | | | , | | (1 | REQM) | 1 |
| i-os | | C.C | Appren | Validati | Test | Total | Appren | Valid | Test |
| | | | tissage | on | 1050 | Total | tissage | ation | 1050 |
| 1 | (1) | 40 | 0,6063 | 0,3797 | 0,6867 | 0,5986 | 5,05 | 5,50 | 4,99 |
| 2 | (2) | 20 | 0,7082 | 0,8797 | 0,7303 | 0,7414 | 4,60 | 2,79 | 3,96 |
| 3 | (3) | 30 | 0,6221 | 0,4127 | 0,6246 | 0,5953 | 5,00 | 4,99 | 5,05 |
| 4 | (4) | 35 | 0,6401 | 0,5552 | 0,6102 | 0,6167 | 4,60 | 5,09 | 6,07 |
| 5 | (5) | 20 | 0,5836 | 0,5940 | 0,5677 | 0,5754 | 5,24 | 4,53 | 5,39 |
| 6 | (6) | 40 | 0,6245 | 0,6899 | 0,6002 | 0,6344 | 5,16 | 5,10 | 3,70 |
| 7 | (7) | 30 | 0,6571 | 0,6840 | 0,5060 | 0,6512 | 5,01 | 4,85 | 5,01 |
| 8 | (2)(1) | 30 | 0,7359 | 0,8535 | 0,8001 | 0,7703 | 4,20 | 3,86 | 3,54 |
| 9 | (2)(3) | 25 | 0,7152 | 0,7445 | 0,6942 | 0,7210 | 4,56 | 3,93 | 5,53 |
| 10 | (2) (4) | 25 | 0,7180 | 0,7974 | 0,7205 | 0,7288 | 4,51 | 3,96 | 3,96 |
| 11 | (2) (5) | 30 | 0,7427 | 0,7465 | 0,6470 | 0,7367 | 4,25 | 4,14 | 4,01 |
| 12 | (2) (6) | 25 | 0,7385 | 0,7955 | 0,6845 | 0,7417 | 4,38 | 3,88 | 4,25 |
| 13 | (2)(7) | 30 | 0,7698 | 0,9249 | 0,8842 | 0,8035 | 4,15 | 2,20 | 3,19 |
| 14 | (2)(7)(1) | 20 | 0,8359 | 0,9513 | 0,8494 | 0,8532 | 3,43 | 2,04 | 3,54 |
| 15 | (2) (7) (3) | 25 | 0,8377 | 0,9257 | 0,8811 | 0,8477 | 3,43 | 2,38 | 3,70 |
| 16 | (2) (7) (4) | 30 | 0,8222 | 0,8465 | 0,8079 | 0,8199 | 3,91 | 2,03 | 3,81 |
| 17 | (2) (7) (5) | 40 | 0,8357 | 0,8627 | 0,9293 | 0,8545 | 3,65 | 2,59 | 2,50 |
| 18 | (2) (7) (6) | 35 | 0,8188 | 0,8836 | 0,9191 | 0,8380 | 3,81 | 2,80 | 2,48 |
| 19 | (2) (7) (5) (1) | 20 | 0,8635 | 0,8774 | 0,8807 | 0,8669 | 3,43 | 2,49 | 2,50 |
| 20 | (2) (7) (5) (3) | 30 | 0,8288 | 0,8954 | 0,8913 | 0,8436 | 3,65 | 2,61 | 3,07 |
| 21 | (2) (7) (5) (4) | 35 | 0,8378 | 0,9088 | 0,9213 | 0,8595 | 3,65 | 2,34 | 2,47 |
| 22 | (2) (7) (5) (6) | 25 | 0,8430 | 0,8568 | 0,9043 | 0,8566 | 3,49 | 2,33 | 3,29 |
| 23 | (2) (7) (5) (1) (3) | 25 | 0,8830 | 0,9722 | 0,9126 | 0,9025 | 3,07 | 1,62 | 2,34 |
| 24 | (2) (7) (5) (1) (4) | 25 | 0,8811 | 0,9129 | 0,8085 | 0,8700 | 3,43 | 2,08 | 3,19 |
| 25 | (2) (7) (5) (1) (6) | 30 | 0,8635 | 0,9201 | 0,8196 | 0,8641 | 3,43 | 2,49 | 3,31 |

Tableau 13: Performances de la prédiction de la DCOs par le modèle simple RNA-PMCSDCOs

| 26 | (2) (7) (5) (1) (3) (4) | 40 | 0,8618 | 0,9481 | 0,9026 | 0,8761 | 3,25 | 2,41 | 2,80 |
|--|-----------------------------|----|--------|--------|--------|--------|------|------|------|
| 27 | (2) (7) (5) (1) (3) (6) | 35 | 0,8705 | 0,9729 | 0,8676 | 0,8866 | 3,19 | 1,75 | 2,80 |
| 28 | (2) (7) (5) (1) (3) (6) (4) | 30 | 0,9041 | 0,9596 | 0,9097 | 0,9127 | 2,76 | 1,71 | 2,73 |
| N.N.C.C : Nombre de neurones dans la couche cachée | | | | | | | | | |

Dans le modèle simple RNA-PMCSDCOs, la précision prédictive des réseaux dépend du nombre de neurones dans la couche cachée, des fonctions d'apprentissage et du taux d'apprentissage (Ghoreishi et Heidari, 2013). Donc Ces variables sont choisies pour optimiser la structure RNA. Sur la base de cette partie, les meilleurs modèles sont obtenus avec un nombre de neurones dans la couche cachée variant entre 20 et 45 neurones.

La DCO est simulée séparément en prenant en compte les différents variables uniques cités précédemment comme des variables d'entrée de RNA-PMCSDCOs afin d'examiner les effets de chaque variable sur les concentrations de la DCOs.

De même, d'autres modèles sont introduits afin de montrer les effets des variables d'entrée multiple sur la concentration de la DCOs. Les différents paramètres d'entrée utilisés dans cette partie sont groupés en deux, trois, quatre, cinq, six et sept variables. Les résultats montrent que la DCOe parmi les variables d'entrée simples, la DCOe et la Te parmi les groupes de deux variables, la DCOe ,Te et la MESe parmi les groupes de trois variables ,la DCOe ,Te, MESe et la DBO₅e parmi les groupes de quatre variables, la DCOe, Te, MESe, DBO₅e et le NH₄e parmi les groupes de cinq variables, la DCOe, Te, MESe, DBO₅e le NH₄e et le PO₄ parmi les groupes de six variables, qui ont les effets les plus importants sur les modèles simples RNA-PMCDCOs. Finalement on note que le modèle qui comprend tous les paramètres d'entrés est le plus performant dans ce contexte avec une architecture 7-30-1 (figure 37).





L'influence de chaque variable sur les modèles simples RNA-PMCSDCOs par rapport aux autres variables est citée dans le tableau 13. Le modèle le plus efficace correspond à une meilleure concordance entre les valeurs mesurées et les valeurs prédites, et également moins de RMSE et plus de valeurs de R. La figure 38 représente la comparaison entre les performances des meilleurs modèles (REQM et R) développés dans les trois phases de modélisation : apprentissage, validation et la phase de test de la DCOs.



Figure 38 : Comparaison entre les meilleurs modèles simples RNA-PMCSDCOs des différentes combinaisons

Le RNA-PMCSDCOs montre des modèles plus fiables selon les données expérimentales pour la DCOs. Les résultats de RNA-PMCSDCOs n'ont pas fluctué considérablement pour tous les modèles avec des variables d'entrée simples et des variables multiples. On a conclu qu'avec l'augmentation du nombre des paramètres d'entrée destinés à former les réseaux de neurones, le RNA-PMCSDCOs donne des modèles plus fiables.

Dans la figure 39, une comparaison graphique entre les concentrations de la DCOs prédit par le modèle RNA-PMCDCOs et les concentrations mesurées durant la période d'études.

Lors de la première étape (phase d'apprentissage) les résultats du modèle RNA-PMCSDCOs montrent que la REQM de l'apprentissage est de 2,76 mg/l. Le coefficient de régression R=0,90. Ces paramètres indiquent que le modèle RNA-PMCDCOs a bien répondu aux données d'apprentissage et qu'il est capable de les rapprocher. Le modèle simple RNA-PMCSDCOs est en mesure de cartographier les données entrées-sorties.



Figure 39 : Comparaison entre les valeurs prédites et les valeurs mesurées pour le meilleur modèle simple RNA-PMCSDCOs

Les résultats du modèle simple RNA-PMCSDCOs lors de la deuxième étape (phase de validation) montrent une légère baisse de la REQM de la validation estimée à 1,71mg/l. Ainsi une légère augmentation dans Le coefficient de régression R= 0,96. Cette amélioration de la

performance du modèle est due en fait que les données de validation sont peu étendues (l'écart types des données de validation sont moins importants que ceux des données d'apprentissage). Lors de la troisième étape (phase de test), la REQM de l'apprentissage est de 2.73 mg/l et le coefficient de corrélation R= 0,91. Ces paramètres indiquent que le modèle réagit bien dans cette phase où il donne des bons résultats. D'après ces résultats les performances de généralisation du modèle sont bonnes, et le modèle simple RNA-PMCSDCOs est en mesure de faire des prédictions moins précises que le DBO₅s. Une analyse de régression de résultat du modèle simple RNA-PMCSDCOs entre la sortie de réseau et les valeurs mesurées est effectuée. La régression linéaire entre les sorties du réseau et les concentrations mesurées correspondantes montre que les sorties du réseau de neurones (données prévisionnelles) pendant toutes les trois phases de modèle apprentissage, validation et test étaient bien concordées avec les valeurs expérimentales.



Figure 40 : Diagramme de régression entre les valeurs prédites et mesurées pour le meilleur modèle simple RNA-PMCSDCOs

La figure 40 présente la régression entre les valeurs mesurées et les valeurs prédites de meilleur modèle simple RNA-PMCSDCOs.

D'après la comparaison entre les résultats obtenus par le modèle de DBO₅s et ces résultats, on a conclu que la précision des modèles RNA-PMC est déterminée non seulement par le nombre de données d'entrée servant à former les réseaux, mais également par la corrélation entre ces données. La faible corrélation des données d'entrée affecte le modèle simple RNA-PMCSDCOs.

Les modèles RNA-PMC constituent un outil de prédiction robuste en sens que l'erreur de prédiction varie légèrement sur la plage de tailles de données utilisées. Cependant, si on collecte plus de données, si les données sont moins bruitées et si d'autres paramètres supplémentaires sont mesurés, cela va entraîner une amélioration de la capacité de prévision du réseau.

VI.4.2 Prédiction de la DCOs par l'utilisation des paramètres d'entrée et de sortie de la STEP (model extensif RNA-PMCEDCOs)

L'utilisation des paramètres d'entrée de la STEP dans le modèle simple pour prédire la concentration de la DCOs à la sortie de la STEP donne des résultats assez bien, mais il nécessite une amélioration. Pour cela dans cette partie, on a essayé d'intégrer les paramètres de la sortie de la STEP qui ont une bonne corrélation avec la DCOs d'après l'analyse en composantes principales aussi que la matrice de corrélation qui donnent les paramètres suivants: DBO₅e, DCOe, NH₄e, ODe, MESe, PO₄e, Te DBO₅s, ODs et MESs présentés dans le tableau ci-dessous.

Tableau 14: paramètres d'entrée-sortie de modèle extensif RNA-PMCEDCOs pour laprédiction de la DCOs à la sortie de la STEP

| Modèle | | | | Paramètre de sortie | | | | |
|--------------|--------------------|------|-------------------|---------------------|------|-------------------|-----|------|
| DNA | DBO5e | DCOe | NH ₄ e | ODe | MESe | PO ₄ e | Te | |
| MA- | (1) | (2) | (3) | (4) | (5) | (6) | (7) | DCOs |
| | DBO ₅ s | ODs | MESs | | | | | DCOS |
| (150 donnés) | (8) | (9) | (10) | | | | | |

Les performances de la prédiction de la concentration de la DCOs par le modèle extensif RNA-PMCEDCOs sont présentées dans le tableau ci-dessous.

| Scé | Doromàtres d'entrés | N. N. | С | Racine de l'erreur quadratique moyenne (REQM) | | | | | |
|-------|--|-----------|-------------------|---|---------|---------|-----------------------|----------------|-------|
| OS | | C. C | Apprenti ssage | Validati on | Test | Total | Appre ntissa ge | Valid ation | Test |
| 1 | (2) (7) (5) (1) (3) (6) (4) (8) | 40/ 10 | 0,9321 | 0,9705 | 0,9885 | 0,9452 | 2,309 | 1,366 | 1,029 |
| 2 | (2) (7) (5) (1) (3) (6) (4) (9) | 40/ 10 | 0,9259 | 0,9732 | 0,9488 | 0,9388 | 1,726 | 1,762 | 2,326 |
| 3 | (2) (7) (5) (1) (3) (6) (4) (10) | 35/ 40 | 0,8948 | 0,9374 | 0,9363 | 0,9066 | 2,169 | 2,167 | 2,800 |
| 4 | (2) (7) (5) (1) (3) (6) (4) (8) (9) | 30/ 10 | 0,949 | 0,9904 | 0,9960 | 0,9665 | 1,960 | 0,830 | 0,613 |
| 5 | (2) (7) (5) (1) (3) (6) (4) (8) (10) | 35/ 10 | 0,92102 | 0,9692 | 0,96011 | 0,93184 | 2,504 | 1,872 | 1,815 |
| 6 | (2) (7) (5) (1) (3) (6) (4) (8) (9) (10) | 35/ 10 | 0,95 | 0,9950 | 0,9961 | 0,9641 | 1,940 | 0,639 | 0,539 |
| N.N.C | N.N.C.C : Nombre de neurones dans les couches cachées | | | | | | | | |

Tableau 15: Performances de la prédiction de la DCOs par modèle extensif RNA-PMCEDCOs

Vu aux résultats de la simulation de la DCOs par le modèle simple RNA-PMCSDCOs qui donne une performance peu adéquate en matière de REQM et R, il est nécessaire d'améliorer cette prédiction par l'utilisation des autres paramètres de la sortie de la STEP affectants la DCOs notamment la DBO₅s, ODs et MESs. Des tentatives sont effectuées afin de développer des modèles extensifs RNA-PMCEDCOs. On utilise les paramètres de la sortie de la STEP par des groupes combinés avec les sept paramètres d'entrés qui ont donné le meilleur modèle simple RNA-PMCSDCOs. Les résultats obtenus de six architectures effectuées sont présentés dans le tableau 15. Contrairement au modèle simple, on utilise dans ce modèle extensif deux couches cachées avec 10 neurones dans la deuxième couche afin d'améliorer le fonctionnement et la précision du modèle. La DCOs est prédite par l'intégration des paramètres de la sortie de la STEP par différentes combinaisons.



Figure 41 : Comparaison entre les meilleurs modèles extensifs RNA-PMCEDCOs des différentes combinaisons

La figure 41 montre les performances des différents modèles RNA-PMCEDCOs développés.

Les résultats montrent que le modèle qui contient la DBO₅s donne le meilleur résultat par rapports aux autres paramètres de sortie intégrés dans ce modèle dans le cas où on a utilisé un seul paramètre de la sortie de la STEP. Dans le cas où on a utilisé deux paramètres de la sortie de la STEP, on observe que le meilleur modèle est donné par l'utilisation de la DBO₅s et ODs parmi les paramètres d'entrée du modèle. Finalement l'utilisation de tous les trois paramètres de la sortie de la STEP dans les paramètres d'entrée du modèle neuronal extensif RNA-PMCEDCOs donne le meilleur résultat. On note que les contributions de chaque paramètre de la sortie de la STEP dans la prédiction de la DCOs sont distinguées, où on observe que la DBO₅s et l'OD donnent une forte amélioration du modèle RNA-PMCEDCOs contrairement à la MES, cela est attribué à la forte corrélation entre la DBO₅s et l'OD avec la DCOs.

Comme indique la figure 42, l'architecture 9-35-10-1 représente le meilleur modèle extensif RNA-PMCEDCOs comprenant tous les paramètres d'entrée et de sortie considérée.



Figure 42 : L'architecture du meilleur modèle extensif RNA-PMCEDCOs



Figure 43 : Comparaison entre les valeurs prédites et les valeurs mesurées pour le meilleur modèle extensif RNA-PMCEDCOs

Dans la figure 43, une comparaison graphique entre les valeurs de la prédites DCOs par le modèle extensif RNA-PMCEDCOs et les valeurs mesuré de la DCOs. Les résultats du modèle extensif RNA-PMCEDCOs montrent que lors de la première phase (phase d'apprentissage), le REQM de la phase d'apprentissage est de 1,94 mg/l et le coefficient de corrélation R=0,95. Ces paramètres indiquent que le modèle extensif RNA-PMCEDCOs est en mesure de répondre aux données d'apprentissage d'une façon excellente. Lors de la deuxième étape (phase de validation), le REQM est de 0,639 mg/l et le coefficient de corrélation R=0,9950. Ces paramètres indiquent que le modèle réagit bien dans cette phase où il donne des bons résultats. Lors de la troisième phase (phase de test), les résultats du modèle extensif RNA-PMCEDCOs montrent que le REQM est de 0,539 mg/l et le coefficient de corrélation R= 0,9961. Les résultats de ces deux dernières phases indiquent que les performances de généralisation du modèle extensif RNA-PMCEDCO5 sont très bonnes. L'utilisation des paramètres de la sortie de la STEP notamment la DBO₅s, ODs, et MESs donne une forte amélioration sur le modèle neuronale simple. Les résultats montrent que le modèle neuronal se base beaucoup plus sur le degré de corrélation entre les paramètres. On conclut que contrairement au modèle extensif, la prédiction de la DCOs par l'utilisation des paramètres d'entrée de la STEP reste peu adéquate à cause de la nature de l'eau et le type de procédés d'épuration (boue activée) qui se base sur la biodégradabilité. D'après ces résultats on peut dire que le modèle extensif est capable de prédire la DCOs pendant les conditions de fonctionnement.

La figure 44 représente l'analyse de régression entre les valeurs prédites et les valeurs mesurées de la DCOs par le modèle extensif RNA-PMCEDCO5 dans les trois phases apprentissage, validation et Test. Les coefficients de corrélation montrent une bonne correspondance entre les valeurs prédites par le modèle extensif RNA-PMCEDBO₅s et les valeurs mesurées. Les fortes valeurs de coefficients de corrélation dans la phase de validation et la phase de test impliquent que le modèle donne une très bonne prédiction avec des nouvelles valeurs.



Figure 44 : Diagramme de régression entre les valeurs prédites et mesurées pour le meilleur modèle extensif RNA-PMCEDCOs

VI.5 Prédiction de la MESs par le modèle PMC-RNA

VI.5.1 Prédiction de la MESs par l'utilisation des paramètres d'entrée de la STEP (model simple RNA-PMCSMESs)

Comme pour les paramètres précédents, on suit les mêmes démarches. Les paramètres d'entrée utilisées pour le modèle simple RNA-PMCSMESs sont comme suit : DBO5e, DCOe, NH4e, ODe, MESe, PO4e, et Te présentés dans le tableau ci-dessous.

Tableau 16 : paramètres d'entrée-sortie du modèle simple RNA-PMCSMES pour laprédiction de la MES à la sortie de la STEP

| Modèle | |] | | Paramètre de sortie | | | | |
|---|---------------------------|-------------|-------------|---------------------|-------------|--------------------------|-----------|------|
| RNA- PMC _{MESs} (150 donnés) | DBO ₅ e (1) | DCOe (2) | NH4e (3) | ODe (4) | MESe (5) | PO ₄ e (6) | Te (7) | MESs |

Afin de prédire les MESs à la sortie de la STEP de Touggourt par le RNA, la DBO₅e, DCOe, NH₄e, ODe, MESe, PO₄e, et Te sont considérés comme des paramètres d'entrée du réseau. Les performances de prédiction de 28 scénarios effectués pour les trois phases de simulation : apprentissage, validation et test sont présentés dans le tableau 17. Les résultats montrent que la DBO₅e parmi les variables d'entrée simples, la DBO₅e et la Te parmi les groupes de deux variables, la DBO₅e, Te et la MESe parmi les groupes de trois variables, la DBO₅e, Te, MESe et la NH₄e parmi les groupes de quatre variables, la DBO₅e, Te, MESe, NH₄e et le PO₄e parmi les groupes de cinq variables, la DBO₅e, Te, MESe, NH₄e le PO₄e et le OD parmi les groupes de six variables qui ont des effets considérables sur les modèles simple RNA-PMCSMESs. En outre, il est déterminé que le modèle qui comprend tous les paramètres d'entrée avait l'effet le plus important sur le modèle simple RNA-PMCMESs en matière de leur coefficient de corrélation plus élevé et de REQM plus faible. D'après les résultats des analyses, les modèles formés ont montré des résultats convergents pour la plus part des modèles.

| G | | | С | oefficient | de régressi | ion | Raci quadra | ne de l'er atique mo | reur yenne |
|-------|-------------------------|-----|-------------------|----------------|-------------|--------|-----------------------|--------------------------|---------------|
| arios | ios Paramètres d'entrée | C.C | Appren tissage | Validati on | Test | Total | Appre ntissa ge | (REQM) Valida tion | Test |
| 1 | (1) | 20 | 0,7229 | 0,8438 | 0,8719 | 0,7509 | 5,97 | 3,89 | 3,67 |
| 2 | (2) | 25 | 0,6287 | 0,7220 | 0,6410 | 0,6386 | 6,58 | 4,10 | 6,51 |
| 3 | (3) | 30 | 0,5439 | 0,9062 | 0,4949 | 0,6101 | 6,93 | 3,70 | 5,88 |
| 4 | (4) | 30 | 0,7048 | 0,8419 | 0,5723 | 0,7087 | 6,00 | 3,87 | 12,79 |
| 5 | (5) | 40 | 0,6186 | 0,8360 | 0,6437 | 0,6458 | 6,79 | 3,20 | 5,19 |
| 6 | (6) | 20 | 0,5603 | 0,8463 | 0,6962 | 0,6201 | 6,87 | 4,11 | 5,55 |
| 7 | (7) | 35 | 0,5758 | 0,8285 | 0,6142 | 0,6170 | 6,79 | 4,23 | 6,58 |
| 8 | (1)(2) | 20 | 0,7715 | 0,7888 | 0,8398 | 0,7795 | 5,28 | 3,80 | 5,46 |
| 9 | (1) (3) | 20 | 0,8025 | 0,8284 | 0,8862 | 0,8150 | 5,00 | 3,62 | 4,34 |
| 10 | (1)(4) | 30 | 0,7194 | 0,7288 | 0,7236 | 0,7300 | 5,88 | 4,35 | 5,19 |
| 11 | (1) (5) | 25 | 0,8037 | 0,8611 | 0,7076 | 0,8037 | 5,19 | 4,08 | 3,92 |
| 12 | (1)(6) | 20 | 0,7393 | 0,8828 | 0,8886 | 0,7882 | 5,37 | 4,31 | 3,54 |
| 13 | (1)(7) | 20 | 0,8237 | 0,8138 | 0,8586 | 0,8254 | 4,90 | 3,76 | 3,28 |
| 14 | (1)(7)(2) | 30 | 0,8192 | 0,8343 | 0,8303 | 0,8196 | 4,80 | 3,80 | 4,78 |
| 15 | (1) (7) (3) | 30 | 0,8430 | 0,7943 | 0,8682 | 0,8379 | 4,60 | 3,88 | 4,39 |

 Tableau 17: Performances de la prédiction de la MESs par le modèle simple RNA-PMCSMES

| 16 | (1) (7) (4) | 35 | 0,8095 | 0,9319 | 0,8242 | 0,8282 | 4,70 | 3,12 | 5,37 |
|--|-----------------------------|----|--------|--------|--------|--------|------|------|------|
| 17 | (1) (7) (5) | 25 | 0,8300 | 0,8700 | 0,8800 | 0,8400 | 4,60 | 3,52 | 3,40 |
| 18 | (1)(7)(6) | 25 | 0,8085 | 0,8866 | 0,8378 | 0,8214 | 5,00 | 3,35 | 4,80 |
| 19 | (1) (7) (5) (2) | 35 | 0,8572 | 0,8733 | 0,8448 | 0,8506 | 3,92 | 3,51 | 4,16 |
| 20 | (1) (7) (5) (3) | 35 | 0,8374 | 0,8930 | 0,9154 | 0,8529 | 4,16 | 3,28 | 3,37 |
| 21 | (1) (7) (5) (4) | 30 | 0,8112 | 0,8905 | 0,8480 | 0,8225 | 5,00 | 3,34 | 3,40 |
| 22 | (1) (7) (5) (6) | 35 | 0,8480 | 0,8266 | 0,8884 | 0,8479 | 4,39 | 3,92 | 3,67 |
| 23 | (1) (7) (5) (3) (2) | 25 | 0,8271 | 0,9421 | 0,9322 | 0,8717 | 4,11 | 3,27 | 3,10 |
| 24 | (1) (7) (5) (3) (4) | 35 | 0,8656 | 0,8331 | 0,8724 | 0,8628 | 4,16 | 3,90 | 3,67 |
| 25 | (1) (7) (5) (3) (6) | 40 | 0,8282 | 0,8391 | 0,8998 | 0,8412 | 4,39 | 3,62 | 3,59 |
| 26 | (1) (7) (5) (3) (2) (4) | 30 | 0,8590 | 0,9439 | 0,9061 | 0,8828 | 4,09 | 2,99 | 3,01 |
| 27 | (1) (7) (5) (3) (2) (6) | 35 | 0,8732 | 0,9553 | 0,9295 | 0,8972 | 3,80 | 2,82 | 2,88 |
| 28 | (1) (7) (5) (3) (2) (6) (4) | 35 | 0,8989 | 0,9592 | 0,9515 | 0,9228 | 3,22 | 2,58 | 2,77 |
| N.N.C.C : Nombre de neurones dans la couche cachée | | | | | | | | | |

La figure 45 montre la comparaison entre les critères de performance des meilleurs modèles (REQM et R) développés dans les trois phases de simulation : Apprentissage, validation et la phase de test de la MESs.



Figure 45 : Comparaison entre les meilleurs modèles simples RNA-PMCSMESs des différentes combinaisons



Figure 46 : L'architecture du meilleur modèle simple RNA-PMCSMESs

Le modèle RNA-PMCMESs le plus efficace dans cette étude prend l'architecture suivante : 7-30-1 comme l'indique la figure 46.

Lors de son apprentissage le modèle simple RNA-PMCSMESs affiche une valeur de REQM de l'ordre de 3,22 mg/l, Une bonne valeur du coefficient de corrélation R égale à 0,8989. Les paramètres de performances du modèle simple RNA-PMCMESs indiquent que le modèle RNA-PMCMESs est capable de cartographier les données entrées-sorties des données d'apprentissage. Les résultats du modèle simple RNA-PMCSMESs dans la phase de validation et de Test sont les suivants : REMQ= 2.58 mg/l et 2.77 mg/l et R= 0,9592 et 0,9515 pour les phases de validation et de test respectivement. Ces résultats montrent que les performances de généralisation du modèle simple sont très bonnes, et que ce modèle est en mesure de fournir des prédictions assez précises.

La figure 47 montre graphiquement une comparaison entre les concentrations de MESs prédits et mesurées. Ces graphiques indiquent que les valeurs de MESs prédits suivent les valeurs observées. En outre, certaines parties du graphique ont montré des différences entre les valeurs mesurées et prédits. Ces différences pourraient être attribuées à l'influence des autres facteurs et paramètres sur la MESs. Il est bien connu que le nombre des facteurs affectant MESs n'est pas seulement 7 paramètres d'entrées utilisées dans ce travail, mais

d'autres facteurs pourraient être impliqués. Par conséquent, l'utilisation de plus de paramètres, peut réduire la différence entre les valeurs mesurées et les valeurs prédites (Abyaneh, 2014).



Figure 47 : Comparaison entre les valeurs prédites et les valeurs mesurées pour le meilleur modèle simple RNA-PMCSMESs

Les résultats du test de régression linéaire effectué sur les sorties du réseau et les valeurs mesurées correspondantes pour le modèle optimal montrent une compatibilité élevée via le coefficient de corrélation. La Figure 48 présente la corrélation entre les sorties RNA et les données mesurées lors des phases d'apprentissage, de validation et de test. La divergence entre les valeurs de MESs observées et prédites pourrait être due aux mauvaises caractéristiques de sédimentation dans le décanteur secondaire. Ainsi, une amélioration de la capacité de prévision de MESs peut être obtenue en utilisant des informations plus détaillées sur le décanteur secondaire tel que l'indice de boue.



Figure 48 : Diagramme de régression entre les valeurs prédites et mesurées pour le meilleur modèle simple RNA-PMCSMESs

VI.5.2 Prédiction de la MESs par l'utilisation des paramètres d'entrée et de sortie de la STEP (model extensif RNA-PMCEMESs)

Dans cette partie on va suivre les mêmes aspects utilisés pour la DBO₅s et la DCOs afin de prédire les MESs par l'intégration des paramètres de la sortie de la STEP. Le tableau 18 présente les paramètres utilisés dans ces modèles qui sont comme suit : DBO₅e, DCOe, NH₄e, ODe, MESe, PO₄e, Te, DBO₅s, DCOs ODs et NH₄.

| Tableau 18 : paramètres d'entrées-sortie du modèle | extensif RNA-PMCEMESs pour la |
|--|-------------------------------|
| prédiction du MESs à la sortie de la STEP | |

| Modèle | | | | Paramètre de sortie | | | | |
|--------------|--------------------|------|-------------------|---------------------|------|-------------------|-----|---------|
| DNA | DBO ₅ e | DCOe | NH ₄ e | ODe | MESe | PO ₄ e | Te | |
| DMCEMES | (1) | (2) | (3) | (4) | (5) | (6) | (7) | MES |
| FIVICEIVIESS | DBO ₅ s | DCOs | ODs | NH ₄ s | | | | IVIL'58 |
| (150 donnés) | (8) | (9) | (10) | (11) | | | | |

Les performances de prédiction de la concentration de la MESs par le modèle extensif RNA-PMCEMESs sont présentées dans le tableau ci-dessous.

| Scé na- rios | Paramètres d'entrée | N.N.C .C | Coefficient de corrélation | | | | Racine de l'erreur quadratique moyenne (REQM) | | |
|---|--|-------------|----------------------------|----------------|--------|--------|---|----------------|-------|
| | | | Appre ntissag e | Valida tion | Test | Total | Appre ntissa ge | Valid ation | Test |
| 1 | (1) (7) (5) (3) (2) (6) (4) (8) | 25/10 | 0,9198 | 0,9775 | 0,9278 | 0,9332 | 3,101 | 2,317 | 2,402 |
| 2 | (1) (7) (5) (3) (2) (6) (4) (9) | 30/10 | 0,8752 | 0,9035 | 0,8995 | 0,8814 | 4,044 | 3,166 | 3,311 |
| 3 | (1) (7) (5) (3) (2) (6) (4) (10) | 30/10 | 0,9123 | 0,917 | 0,9336 | 0,9125 | 3,397 | 2,812 | 2,876 |
| 4 | (1) (7) (5) (3) (2) (6) (4) (11) | 25/10 | 0,9133 | 0,9098 | 0,9303 | 0,9140 | 3,340 | 2,795 | 3,132 |
| 5 | (1) (7) (5) (3) (2) (6) (4) (8) (9) | 25/10 | 0,9499 | 0,9813 | 0,9841 | 0,9597 | 2,402 | 1,606 | 1,754 |
| 6 | (1) (7) (5) (3) (2) (6) (4) (8) (10) | 35/10 | 0,9616 | 0,9917 | 0,9818 | 0,97 | 2,057 | 1,241 | 1,581 |
| 7 | (1) (7) (5) (3) (2) (6) (4) (8) (11) | 30/10 | 0,9649 | 0,9777 | 0,9812 | 0,9701 | 2,057 | 1,876 | 1,699 |
| 8 | (1) (7) (5) (3) (2) (6) (4) (8) (10) (9) | 40/10 | 0,9789 | 0,9945 | 0,9960 | 0,9842 | 1,581 | 0,905 | 0,760 |
| 9 | (1) (7) (5) (3) (2) (6) (4) (8) (10) (11) | 30/10 | 0,9802 | 0,9922 | 0,9947 | 0,9840 | 1,641 | 0,760 | 0,877 |
| 10 | (1) (7) (5) (3) (2) (6) (4) (8) (10) (9) (11) | 35/10 | 0,9891 | 0,9964 | 0,9980 | 0,9918 | 1,118 | 0,582 | 0,707 |
| N.N.C.C : Nombre de neurones dans les couches cachées | | | | | | | | | |

Tableau 19: Performances de la prédiction du MESs par modèle extensif RNA-PMCEDCO5

Différentes architectures sont testées afin de prédire la MESs de la STEP de Touggourt par un modèle extensif RNA-PMCEMESs.

Les performances de prédiction de 10 scénarios des meilleurs modèles pour chaque combinaison dans les trois phases : apprentissage, validation et test sont présentés dans le tableau 19. On a utilisé dans cette partie deux couches cachées avec nombre fixé de neurones dans la deuxième couche (10 neurones). Comme les cas précédents, on a utilisé dans cette partie différentes combinaisons entre les paramètres de l'influent de la STEP qui a donné le meilleur modèle simple et les quatre paramètres de l'effluent de la STEP qui avait une influence sur la MESs. La figure 49 montre la comparaison entre les critères de performances des meilleurs modèles développés (REQM et R) dans les trois phases : apprentissage, validation et test. Le meilleur résultat est obtenu avec l'architecture 11-35-10-1(figure 50).



Figure 49 : Comparaison entre les meilleurs modèles extensifs RNA-PMCEMESs des différentes combinaisons



Figure 50 : L'architecture du meilleur modèle extensif RNA-PMCEMESs

Les résultats des différentes combinaisons montrent que la DBO₅s parmi les paramètres de la sortie de la STEP qui donne la meilleure performance, quand on a intégré un seul paramètre de l'effluent de la STEP. Parmi les groupes de deux variables, la DBO₅s et le NH₄s est la combinaison à deux paramètres de l'effluent de la STEP qui a donné la meilleure performance. Parmi les deux groupes de trois paramètres de l'effluent, la DBO₅s, le NH₄s et

ODs ont l'effet le plus considérable sur les modèles extensifs de RNA-PMCEMESs. En outre, il est déterminé que le scénario qui comprend tous les paramètres de l'influent et de l'effluent considérés de la STEP donne la meilleure performance prédictive en matière de prédire la MESs par le modèle extensif RNA-PMCEMESs.

On note que les différents paramètres de la sortie de la STEP donnent différentes contributions sur l'optimisation des performances de prédiction de la MESs. D'après les résultats des analyses, les modèles extensifs formés montrent des meilleures performances de modèle extensif RNA-PMCEMESs par rapport au modèle simple RNA-PMCSMESs. Lors de son apprentissage, le modèle extensif RNA-PMCEMESs affiche une valeur de REQM de l'ordre de 1,118 mg/l et une très bonne valeur du coefficient de corrélation R égale à 0,9891. Les résultats du modèle extensif RNA-PMCEMESs en phases de validation et de Test sont comme suit : REMQ= 0,582 mg/l et 0,707 mg/l et R= 0,9964 et 0,9980 pour les phases de validation et de test respectivement. Ces résultats montrent que les performances de généralisation du modèle extensif sont très bonnes et que ce modèle est en mesure de fournir des prédictions précises.



Figure 51 : Comparaison entre les valeurs prédites et les valeurs mesurées pour le meilleur modèle extensif RNA-PMCEMESs

La figure 51 représente graphiquement une comparaison entre les valeurs de MESs mesurées et prédits par le modèle extensif RNA-PMCEMESs. Ce graphique montre que les

valeurs de MESs prédits suivent les valeurs observées tout au long de la courbe : ce qu'indique l'efficacité de ce modèle. Aussi l'importance de l'intégration des paramètres de la sortie de la STEP dans la prédiction peut donner des valeurs ajoutées surtout qu'ils représentent la pollution organique. En outre, certaines parties du graphique montrent des différences entre les valeurs mesurées et les valeurs prédites, ces différences pourraient être attribuées à l'influence d'autres facteurs et paramètres sur le MESs.



Figure 52 : Diagramme de régression entre les valeurs prédites et mesurées pour le meilleur modèle extensif RNA-PMCSEMESs

La Figure 52 représente la régression entre les valeurs mesurées et les valeurs prédites durant les trois phases. Les résultats du test de régression linéaire effectué entre les valeurs mesurées et les valeurs prédites par le modèle extensif RNA-PMCEMESs correspondantes au modèle optimal montrent une compatibilité très élevée entre les sorties du modèle neuronal et les valeurs mesurées via un coefficient de corrélation très élevé dans les trois phases qui a dépassé 0,99.

VI.6 Conclusion

Les résultats montrent que les RNA-PMC sont capables de modéliser la station d'épuration des eaux usées municipales de la ville de Touggourt. Les résultats des modèles dans les trois phases : Apprentissage, validation et test montrent une correspondance presque parfaite entre les valeurs prédites et mesurées pour les trois paramètres simulés. En conséquence, pour tous les modèles RNA-PMC, le coefficient de corrélation des meilleurs modèles dans les trois phases de simulation variait de 0,9 à 0,999 et le REQM était minimal. Les résultats montrent que les meilleurs modèles pour les trois paramètres prédits sont les modèles qui comprennent tous les paramètres utilisés comme entrées du modèle. Les architectures sélectionnées assurent un apprentissage à une vitesse raisonnable et à un temps de simulation court pour des performances spécifiques de réseau. Il est observé que pour la prédiction de la DBO₅s, les critères de performance de modèle sont très fiables que les deux autres paramètres prédits (DCOs et MESs), cela est dû à la forte corrélation entre les variables d'entrées utilisée dans ce modèle. Bien qu'il soit rare de prédire un paramètre d'effluent en utilisant un autre comme entrée du modèle, les paramètres d'effluent peuvent être utilisés pour contrôler les performances du système, en particulier lorsque les données mesurées de l'influent ou les données du système présentent des biais et des erreurs potentielles. De plus, cette étude a indiqué qu'un paramètre d'effluent facilement mesurable peut être utilisé dans un modèle RNA pour prédire en permanence un paramètre d'effluent.

D'après ces résultats, on a conclu que la précision des modèles RNA-PMC est déterminée non seulement par le nombre des données d'entrée servant à former les réseaux, mais également par les corrélations entre ces données. Le RNA utilisé dans cette étude est un outil précieux pour le contrôle de l'exploitation d'une station d'épuration, car ils peuvent être adaptés aux conditions d'exploitation locales.

CONCLUSION GENERALE ET PERSPECTIVES

Conclusion générale

Cette étude a permis de faire un diagnostic de l'état de la pollution des eaux usées dans la ville de Touggourt. Cette évaluation comprend une évaluation de la qualité physicochimique des effluents des eaux usées à l'entrée et à la sortie de la station d'épuration. Souvent Les STEP connaissent des problèmes de dysfonctionnements et des difficultés de traitement par la biomasse au niveau du bassin d'aération. Pour les données d'entrée de la station, on constate que les valeurs de l'écart-type des paramètres de pollutions sont assez variées, dû principalement à la variabilité moyenne de la charge polluante d'entrée.

L'analyse de l'écart-type annuel de la DBO₅ à la sortie montre que la station de Touggourt est statistiquement stable pour la période d'étude. (Une station est statistiquement stable si l'écart-type pour la DBO₅ des effluents est inférieur à 10 mg/l, et si l'écart-type pour les MES des effluents est inférieur à 70 mg/l) (Niku *et al.*, 1982).

Les résultats indiquent des rendements épuratoires satisfaisants pour l'élimination des matières polluantes en accord avec les normes de rejet des effluents de boues activées. Quant à la concentration DBO₅, DCO elle montre qu'il y a une bonne biodégradation. Aucun changement n'est obtenu sur la CE par la purification.

Il semble utile de signaler que malgré un traitement maximal, le traitement de PO_4 et des NO_3 n'est pas efficace et reste inchangé ou même les valeurs sont plus élevées à la sortie. L'environnement est menacé par les concentrations élevées de ces deux derniers, donc la station nécessite un traitement complémentaire en matière de NO_3 et PO_4 . Les résultats montrent que la STEP respecte les normes de rejet pour la plupart des paramètres. Par ailleurs, les résultats montrent que les concentrations des paramètres de pollution augmentent dans la période estivale et diminuent dans la période hivernale.

L'application de l'ACP sur les eaux usées brutes et traitées a informé de l'existence d'une relation étroite entre la DBO₅e, DCOe, MESs, PO₄e, NH₄e, ODe MESe, Te et DCOs d'une part et entres la MESBA RECY, ODs, DBO₅s et NH₄s d'autre part. Finalement on peut conclure que l'analyse de l'ACP a réduit le nombre des variables originales qui peut affecter la STEP de Touggourt. La nécessité d'un contrôle efficace d'une station de traitement des eaux usées utilisant le procédé de boues activées est primordiale pour assurer son efficacité dans la protection de l'environnement.

Le réseau de neurone artificiel constitue un modèle fiable pour toute STEP des eaux usées afin de fournir des outils puissants permettant de prédire ses performances et de constituer une base pour contrôler le fonctionnement de processus. Il peut traiter des problèmes de non-linéarité élevés dans le cas de la modélisation de processus d'une STEP. Cela minimise les coûts d'exploitation et évalue la stabilité du bilan environnemental. L'objectif principal de cette étude était de développer des modèles RNA-PMC pour prédire avec précision les trois paramètres clés de la pollution organique à la sortie de la STEP de TOUGGOURT notamment la DBO₅s, DCOs et MESs. Les paramètres de fonctionnement, y compris la DBO₅e, DCOe, NH₄e, ODe, MESe, PO₄e, MESBA, RECY, Te, DCOs, NH₄s, ODs et MESs, sont utilisés pour prédire la DBO₅s. La DBO₅e, DCOe, NH₄e, ODe, MESe, PO₄e, Te, DBO₅s, ODs et MESs sont utilisés pour prédire la MESs.

L'architecture du réseau neuronal qui a donné les modèles optimaux était la RNA-PMC à rétro propagation avec une fonction d'activation de tangente hyperbolique, une seule couche cachée pour le modèle simple et deux couches cachées pour le modèle extensif et un algorithme d'apprentissage LM. Le coefficient de corrélation des meilleurs modèles pour les trois phases variait de 0,9 à 0,999 et le REQM était minimal. Il est observé que pour la prédiction de la DBO₅s, les paramètres de performance de modèle sont très fiables que les deux autres (DCOs et MESs), cela est dû à la forte corrélation entre les variables d'entrées utilisée dans ce modèle.

La simulation par le RNA a montré une correspondance presque parfaite entre les valeurs mesurées et les valeurs prédites pour tous les trois paramètres simulés. Les résultats indiquent que dans la phase d'apprentissage, et avec faibles valeurs expérimentaux des données d'entrée, le RNA-PMC donne des résultats précis avec des valeurs de R plus élevées et valeurs de REQM plus faibles. Dans cette étude on a remarqué que l'intégration des paramètres de l'effluent de la STEP donne une amélioration très importante en matière de prédiction des trois paramètres étudiés. La comparaison de ces résultats avec autres études a montré que, même les paramètres utilisés dans cette étude sont minimaux, les résultats étaient

supérieurs aux études précédentes. Ces résultats suggèrent que l'utilisation de plus de paramètres d'entrée ne conduit pas nécessairement à une amélioration des résultats prévus, mais le type des paramètres d'entrée est plus important que son nombre. Bien que le RNA puisse gérer les données bruyantes et erronées, les paramètres de traitement biologique des eaux usées sont nombreux et une erreur relative peut se propager lors de la modélisation neuronale. Par conséquent, de telles erreurs dans les données d'observation, ainsi que la nature complexe des réactions biologiques et leur reflexe dans différentes conditions de fonctionnement, peuvent avoir une incidence sur les résultats de cette étude. La technique de modélisation RNA présente de nombreuses caractéristiques favorables telles que l'efficacité, la généralisation et la simplicité, ce qui en fait un choix attrayant pour la simulation de systèmes complexes, tels que les processus de traitement des eaux usées. Pour la STEP considérée dans cette étude, les paramètres d'entrée de modèle ayant des coefficients de corrélation élevée avec les paramètres prédits, offrent des influences positives sur la simulation par le réseau de neurones artificiel. Par conséquent, une analyse de corrélation semble être un outil à utiliser pour la sélection des paramètres d'entrée pour la modélisation RNA du système.

Finalement on peut conclure que les résultats des modèles réalisés confirment l'importance et l'utilité de la modélisation intelligente en tant qu'outil rapide, facile à utiliser, non invasif et peu coûteux. Ainsi, le modèle RNA représente un outil d'analyse et de diagnostic efficace pour comprendre et simuler le comportement non linéaire de la STEP. Cela peut constituer un outil d'évaluation des performances utile pour les exploitants des STEP et les décideurs. Une exploitation et un contrôle plus sûrs de la STEP peuvent être atteints par le développement d'un modèle RNA pour prédire la performance de la STEP sur la base des observations passées de certains paramètres clés de la qualité des eaux.

Perspectives

Malgré le succès enregistré dans ce travail, certains aspects ont été identifiés et mériteront d'être approfondis. Par conséquent, les points suivants sont suggérés comme domaines de travaux futurs :

- Des travaux similaires devront être exécutés pour simuler les autres paramètres qui n'ont pas été pris en considération dans ce travail.
- Des nouvelles variables d'entrée susceptibles d'affecter les performances de la STEP constitueront une extension logique de ces travaux.
- Le présent travail peut être étendu en collectant des données à long terme, pour toutes les étapes du traitement biologique des eaux usées, et en développant des modèles prédictifs intelligents pour prédire les performances de la STEP à chaque étape.
- Les recherches réalisées dans ce domaine ont indiqué qu'un réseau hybride peut être une meilleure solution. Cela peut inclure d'autres options de calcul comme un réseau de neurones artificiels à logique floue ou un réseau de neurones artificiels combiné avec un algorithme génétique, qui feront l'objet de futurs travaux de recherche, et une continuité du présent travail pour optimiser les modèles construits.
- Les recherches devront se concentrer sur le transfert des connaissances issues de la recherche à la station d'épuration. Alors un opérateur de la STEP peut lire les informations et prendre des décisions pour gérer la station de manière optimale.

Références bibliographique

- Abyaneh, HZ. (2014), "Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters", Iranian J Environ Health Sci Eng 12:40.
- Adeloye, A.J., De Munari A. (2006)."Artificial neural network based generalized storageyield-reliability models using the Levenberg-Marquardt algorithm". Journal of Hydrology, 326(1-4): 215-230.
- Al-Ghazzawi, A., et Lennox, B. (2008). "Monitoring a complex refining process using multivariate statistics".Control Eng Pract, 16: 294-307.
- Baughman, D.R., et Liu, Y.A. (1995). "Neural networks in Bioprocessing and chemical Engineering", Academic Press 435p.
- Bekkari, N., et Zeddouri, A. (2019). "Using artificial neural network for predicting and controlling the Effluent chemical oxygen demand in wastewater treatment plant effluent chemical oxygen demand in wastewater treatment plant". Management of Environmental Quality, 30(3): 593-608.
- Belanche, L.B., Valdés, J.J., Comas, J., Ignasi, R., Rodab, I.R., et Poch M. (1999). "Towards a model of input–output behaviour of wastewater treatment plants using soft computing techniques", Environmental Modelling & Software Volume 14, Issue 5, pp.: 409-419.
- Bhat, N.V., et Minderman P.A., Mcavoy, T. et Wang, N.S. (1990)."Modeling chemical process systems via neural computation", IEEE Control Systems Magazine, Vol. 10.
- Bishop, C. M. (1995). "Neural Networks for Pattern Recognition"; Clarendon Press, Oxford
- Bonjoch, X., Balleste´, E., et Blanch, A. R. (2004). "Multiplex PCR with 16S RRNA genetargeted Primers of Bi.dobacterium spp. to identify sources of fecal pollution". Appl. Environ. Microbiol. 70: p.3171–3175.Botton *et al.*; 1990
- Boukerroucha. (2011). "Modélisation des stations d'épuration a boues activée cas de la station de Baraki (Alger) ". Thèse de Magister, ENA, Algerie, 173 pages.Burton, 2006
- Chapman, D., et Kimstach, V. (1996). "Selection of water quality variables. Water quality assessments: a guide to the use of biota, sediments and water in environment monitoring", Chapman edition, 2nd ed. E & FN Spon, London, pp. 59-126
- Chen J., et Beck M.B. (1993). "Modelling, control and offline estimation of activated sludge bulking". Water science and technology, 28 (11, 12), pp. 249-256.
- Choi, D. J., et Park, H. (2001). "A hybrid artificial neural network as a software sensor for optimal control of a wastewater treatment process", Water Research Volume 35, Issue 16, pp.: 3959-3967.
- Cors, M. (2007). "Techniques extensives d'épuration des eaux usées domestiques". Le meilleur choix environnemental en zone rurale Dossier IEW Inter-Environnement Wallonie.
- Cote, M., Grandijean, B.P.A., Lessard, P., et Yhibault, J.(1995). "Dynamic modeling of the activated sludge process: improving prediction using neural networks", Water Res. 29 , pp. 995–1004.
- Daliakopoulos, I.N., Coulibaly, P., et Tsanis, I.K.(2005). "Groundwater level forecasting using artificial neural networks". Journal of Hydrology, 309, (1-4): 229-240.
- Demuth, H., Beale, M., et Hagan, M. (2013). "Neural Network Toolbox For Use with MATLAB, Users Guid, R2013, 430 pages.
- Djeddou, M. (2014) "Prévision du taux d'échec avec les réseaux neurones artificiels dans une station de traitement des eaux résiduaires" Thèse de Doctorat, Université Mohamed Khider Biskra, 175 pages.
- Dogan, E., Ates, A., Yilmaz EC., et Eren B (2008) ."Application of artificial neural networks to estimate wastewater treatment plant inlet biochemical oxygen demand .Environ", Prog 27: 439-446. DOI: 10.1002/ep.10295
- Dufournet, R. (1974). "Traitement des eaux usées", Techniques de l'ingénieur, C670, Editions T. I., Paris, France, 38 pages.

- El Kebir, M. (2009). "Modélisation et commande par modèle inverse neuronale d''un four ventile", Thèse de Magister, Université d''Oran.
- Fox, K. K., Cassani, G., Facchi, A., Schröder, F.R., Poelloth, C., et Holt, M. S. (2002). "Chemosphere", 47, 499-505.
- Gaïd, A. (1993). "Traitement des eaux usées urbaines", Techniques de l'ingénieur, C5220, Editions T. I., Paris, France, 28 pages.
- Garvey, E.B. (1997). "On-line Quality Control of Injection Molding Using Neural Networks.Minor Thesis Department of Computer Science, Royal Melbourne Institute of Technology, Melbourne, Australia, 82 pages
- Gernaey, K. V., van Loosdrecht, M. C., Henze, M., Lind, M., et Jørgensen, S. B. (2004). "Activated sludge wastewater treatment plant modelling and simulation: State of the art". Environmental Modelling and Software: Environmental Sciences and Artificial Intelligence, 19(9), 763-783.
- Ghoreishi,S.M., et Heidari, E. (2013). "Extraction of epigallocatechin-3-gallate from green tea via supercritical fluid technology: neural network modeling and response surface optimization". J. Supercrit. Fluids, 74 (2013), pp. 128-136.
- Giustolisi, O. (2006). "Using a multi-objective genetic algorithm for SVM construction", J Hydroinf 8:125-139. DOI: 10.2166/hydro.2006.016
- Gleick, P. H. (1996). "Water resources, In Encyclopedia of Climate and Weather, ed. by S. H. Schneider", Oxford University Press, New York, vol. 2, pp.817-823.
- Govindaraju, R.S. (2000)."Artificial neural network in hydrology. II: hydrologic application, ASCE task committee application of artificial neural networks in hydrology", Journal of Hydrologic Engineering 5, pp.: 124-137.
- Graupe, D. (2007). "Principal of Artificial Neural Networks, Advanced Seris In Circuits And Systems", Vol. 6, World Scientific Publishing Co. Pte. Ltd., 2nd Edition, Singapore, Thailande, 320 pages.

- Graves, A., Liwicki, M., Fernandez, S., et al. (2009), "A Novel Connectionist System for Improved Unconstrained Handwriting Recognition." IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 31, N° 5, pp.: 855-868.
- Gray, N.F. (2004). "Biology of wastewater treatment", 2nd Ed. Imperial College, UK.
- Grosclaude, .G. (1999). "L'eau : usage et polluant", Tome II. 4eme Edition: INRA, Paris. 11pp.
- Güçlü, D., et Dursun, S. (2008). "Amelioration of carbon removal prediction for an activated sludge process using an artificial neural network (ANN) ", CLEAN – Soil Air Water, 36: 781-787.
- Güçlü, D., et Dursun, S. (2010). "Artificial neural network modelling of a large-scale wastewater treatment plant operation", Bioprocess Biosyst Eng 33:1051–1058. DOI: 10.1007/s00449-010-0430-x
- Guerree, H., et Gomella, C. (1982). " les eaux usées dans les agglomérations urbaines ou rurales",vol,1,la collecte,Eyrolles,Paris,250p.
- Häck, M., et Köhne, M. (1996). "Estimation of wastewater process parameters using neural networks", Water Science and Technology Volume 33, Issue 1, pp.: 101-115
- Hagan, M., Demuth, H., et Beale, M. (1996). "Neural Network Design, PWS, Boston.
- Hamed, M., Khalafallah, M.G., et Hassanein, E.A. (2004). "Prediction of wastewater treatment plant performance using artificial neural network", Environmental Modeling and Software, 19, pp.: 919-928.
- Hamoda, M.F., Al—Ghusain, I.A., et Hassan, A.H. (1999). "Integrated wastewater treatment plant performance evaluation using ANN, Water Sci. Tech., Vol. 40, No. 7, pp. 55-65.
- Harremoës, P., Capodaglio, A. G., Hellstrom, B. G., Henze, M., Jensen, K. N., Lynggaaard-Jensen, A., Otterpohl, R., et Soeborg, H. (1993). "Wastewater treatment plants under transient loading-performance, modeling and control. Wat. Sci. Tech., 27 (12): 71-79.
- Haykin, S. (1999). "Neural Networks A Comprehensive Foundation", Second Edition, Rearson Education, Inc., India, 823 pages.

- Heaton, J. (2008), "Introduction to Neural Networks for C#", 2nd Edition, Heaton Research, Inc., Saint Louis, USA, 430 pages.
- Hebert, S., et Légre, S. (2000). "Direction du suivi de l'état de l'environnement". Ministère de l'Environnement Gouvernement du Québec, vol. 5.
- Henze, M., Grady, CPL., Gujer, W., Marais, GvR., et Matsuo, T. (1987). "Activated sludge model no 1", International Association on Water Pollution Research and Control. Scientific and Technical Reports No 1, London.
- Henze, M., Harremoës, P., La Cour Jansen J., et Arvin, E. (1997). "Wastewater treatment.Biological and chemical processes". 2nd Edition, Germany, Springer, 383p.
- Hotelling, H. (1933). "Analysis of a complex statistical variable into principal component", J.Edu. Psy., Vol 24, pp 417-441 and pp 498-520.
- Hu, J. (1997). "Research on Hybrid Black-Box Modeling for Nonlinear Systems and Its Applications", thèse de Doctorat, Waseda university, Japan, 139 pages.
- Jackson, J.E. (1991). "A user's guide to principal components". New York Wiley, 351p
- JORA. (1996). Journal Officiel De La République Algérienne N° 26 23 Avril 2006
- Karima, T. (2006). "Etude des performances épuratoires d'une station d'épuration à boues activées ", mémoire de PFE DEUA hydraulique, université de Tlemcen,
- Khan, R., et Ondr öek, . (2000). "Peak electric load forecasting using an artificial neural network", submitted for publishing in Engineering Mechanics International Journal of Theoretical and Applied Mechanics, Technical University of Brno, Czech Republic.
- Krenker, A., Bešter, J., et Kos, A. (2011). "Artificial Neural Networks-Methodological Advances and Biomedical Applications", Publisher InTech, Shanghai, China, 362 pages.
- Kriger, C. (2007). "Prediction of the influent wastewater variables Using Neural Network theory", thèse de Doctorat, Cape Peninsula University of Technology, RSA, 331 pages.

- Laucelli, D., Savic, DA., et Giustolisi, O. (2009). "Data-based modelling: a comparison between artificial neural networks and evolutionary polynomial regression", in: (IAHR) IAfH-EEaR, editor. 8th International Conference on Hydroinformatics. Concepcion, Chile Curran Associates, Inc., Red Hook, NY; p. 929
- Lebart, L., Piron, M., et Morineau. (2006). "Statistique exploratoire multidimensionnelle Visualisation et inférence en fouilles de données". 4e édition Paris: Dunod.
- Lee, D.S., et Park, J.M. (1999). "Neural network modeling for on-line estimation of nutrient dynamics in a sequentially-operated batch reactor", Journal of Biotechnology 75, pp.: 229-239.
- Leynaud, G. (1968). "Les pollutions thermiques, influence de la température sur la vie aquatique". BTI Ministère de l'agriculture, 224-881.
- Lynch, P.A., B.J. Gilpin, L.W. Sinton, et M.G. Savill. (2002). "The detection of Bi.dobacterium adolescentis by colony hybridization as an indicator of human faecal pollution. J. Appl. Microbiol. 92: p.526–533.
- Maier, H.R., et Dandy, G.C. (2000). "Neural networks for prediction and forecasting of water resources variables: a review of modeling issues and applications", Water Resources Research 15, pp.:101-124.
- Mathieu, C., Pieltain, F., et Jeanroy, E. (2003). Analyse chimique des sols: Méthodes choisies. Tec & doc.
- Meinck, F., Stooff, H., et Kohlschutter, H. (1977). "Les eaux résiduaires industrielles", 2ème Ed. Masson, Paris, 863p.
- Metcalf., et Eddy, (2003). "Wastewater Engineering : Treatment and reuse, 4th Ed, revisé par Tchobanoglous, F.L. Burton et H.D. Stensel, McGraw-Hill Inc, New-York, p, 616, 928, 933, 969.

Metcalf., et Eddy. (1978). "Wastewater Engineering: Treatment, Disposal and Reuse", 2nd Ed

Metcalf., et Eddy. (1991). "Wastewater engineering: treatment, disposal and reuse, McGraw-Hill, Inc.

- Mjalli, F., Alasheh S., et Alfadala, H.E. (2007). "Use of artificial neural network black- box modelling for the prediction of wastewater treatment plants performance" Journal Of Environment Management 83 (2007) 329-338.
- Moshiri, S., et Cameron, N. (2000). Neural network versus econometric models in forecasting inflation. Journal of forecasting, 19(3), 201-217.
- Nasr, M. S., Medhat, A.E. Moustafa, M. A. E., Seif, H. A. E., et El Kobrosy G. (2012). "Application of Artificial Neural Network (ANN) for the prediction of EL-AGAMY wastewater treatment plant performance-EGYPT, Alexandria Engineering Journal, 51, pp: 37-43
- Neelakantan, T.R., Brion, T.R., et Lingireddy S. (2001). "Neural network modeling of cryptoposporidium and giardia concentrations in Delaware River, USA", Water Science and Technology 43, pp.:125-132.
- Ng, G. W. (1997)."Application of Neural Networks to Adaptive Control of Nonlinear Systems", Research Studies Press Ltd, John Wiley and Sons Inc, .Singapore, 198 pages.
- Niku, S., Schroeder, E. D., et Haugh. R. S. (1982). "Reliability and Stability of Trickling Filter Processes", Journal Water Pollution Control Federation, Vol. 54, No. 2, pp.: 457-470.
- Olden, J. D., et Jackson D. A. (2002). "Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks", Ecological Modelling, 154, pp.: 135-150.
- Oliveira-Esquerre, K.P., Mori M., et Bruns, R.E. (2002). "Simulation of an industrial wastewater treatment plant using artificial neural networks and principal components analysis. Brazilian Journal of Chemical Engineering, 19, pp.: 365-370.
- Olsson, G., et Nowell, B. (1999). Wastewater Treatment Systems. Model. Diag. and Contr., London: IWA Publishing.

- Orhon, D., et Artan, N. (1994). "Modelling of activated sludge systems", Technomic Publishing Co., Inc., pp. 498- 507.
- Oukhellou, L. (1997). " Para métrisation et classification de signaux en contrôle non destructif. Application à la Reconnaissance des Défauts de Rails par Courant de Foucault". Thèse de doctorat université Paris sud, pp211.
- Pai, T.Y. (2008). "Grey and neural network prediction of effluent from the wastewater treatment plant of industrial park using influent quality", Environ. Eng. Sci 25: 757-766.
- Palm, R. (1998). "L'analyse en composantes principales principes et applications. Notes de statistique et d'informatique. Faculté universitaire des Sciences agronomiques Centre de Recherches agronomiques GEMBLOUX (Belgique). 98/233p.
- Parizeau, M. (2006). "Réseaux de neurones GIF-21140 et GIF-64326", Université de Laval, Canada, 116 pages.
- Park, M. H. (1996). "Neural Network Control of a Chlorine Basin", University of California, USA, pp.:20-52.
- Payment, P., Godfree, A., et Sartory D. (2002). "Encyclopedia of Environmental Microbiology", pp. 861–871, Wiley-Interscience, N.Y.
- Pearson, K. (1901). "On lines and planes of closest fit to systems of points in space", Phil.Mag.,2, 11, pp 559-572.
- Pelmont, J. (1993). "Bactéries et environnement, adaptations physiologiques", [eds] Presses Universitaires de Grenoble. Presses universitaires de Grenoble, Grenoble. 898 pp
- Pendashteh, A.R., Fakhru'l-Razi, A., Chaibakhsh, N., Abdullah, L.C., Madaeni, S.S., et Abidin, Z.Z. (2011). "Modeling of membrane bioreactor treating hypersaline oily wastewater by artificial neural network", J. Hazard. Mater. 192: 568–575. DOI: 10.1016/j.jhazmat.2011.05.052
- Pescod, M.B. (1985). "Design, operation and maintenance of wastewater stabilization ponds in treatment and use of sewage effluent for irrigation". Ed Pescodand Arar, 93-114.

- Petry, J., Soulsby, C., Malcolm, I. A., et Youngson, A. F. (2002). "The Science of the Total Environment", 294, 95110
- Plazl, I., Pipus, G., Drolka, M., et Koloini T. (1999). "Parametric sensitivity and evaluation of a dynamic model for single-stage wastewater treatment plant. Acta Chimica Slovenica 42, pp.:289-300.
- Prevost, L. (2012). "Modèles Connexionnistes Apprentissage Fusion d'informations". 30 pages.
- Refaat, H. A.R., *et al.* (2007). Labor productivity: Benchmarking and variability in Egyptian projects. International Journal of project management, 25(2), 189-197.
- Rejsek, F. (2002). Analyse des eaux: aspects réglementaires et techniques. Centre régional de documentation pédagogique d'Aquitaine.
- Rodier, J. (1984). L'analyse de l'eau; Eaux naturelles, eaux résiduaires, eau de mer, 7ème Edition, Ed. Dunod, Paris.
- Rodier, J. (1996). "L'analyse de l'eau naturelle, eaux résiduaires, eau de mer", 8ème Edition, Dénod, Paris, 1383 p.
- Rodier, J. (2005). L'analyse de l'eau naturelle, eaux résiduaires, eaux de mer. 8ème Edition DUNOD technique. Paris, 1008-1043.
- Rui, Y., et El-Keib, A. A. (1995). "A Review of ANN-Based Short-Term Load Forecasting Models", Proceeding of 27th IEEE Southeastern Symposium on System Theory, MS, USA, pp.: 78-82.
- Rustum, R. (2009). "Modelling Activated Sludge Wastewater Treatment Plants Using Artificial Intelligence Techniques (Fuzzy Logic and Neural Networks)", Thèse de Doctorat, Heriot-Watt University, UK, 271 pages.
- Rustum, R., Adeloye, A.J., et Scholz M. (2008a). "Applying Kohonen Self- organizing Map as a Software Sensor to Predict the Biochemical Oxygen Demand". Water Environment Research, 80 (1): 32-40.

- S.T.E.P Touggourt. (2008). Station d'épuration des beaux usés de Touggourt.
- Saporta, G. (1990). "Probabilités, Analyse des Données et Statistique. Éditions Technip, 1990.
- Singh, KP., Basant, A., Malik, A., et Jain G. (2009). "Artificial neural network modeling of the river water quality a case study", Ecol Model 220:888–895.
- Spearman, C. (1904). "General intelligence objectively determined and measured".
- Spellman, F. R. (2003). "Hand book of water and wastewater treatment plant operation". Lewise publishers, USA.
- Steel, E. W., et McGhee, T. J. (1979). "Water Supply and Sewerage", 5th ed., McGraw-Hill, New York, pp.: 456–467.
- Suh, CW., Lee, JW., Hong, YST., et Shin, HS. (2009). "Sequential modeling of fecal coliform removals in a full-scale activated-sludge wastewater treatment plant using an evolutionary process model induction system", Water Res 43:137–47.
- Tarassenko, L. (1998). "A Guide to Neural Computing Applications", John Wiley and Sons, New York, USA, 151 pages.
- Thurstone, L. (1947). "Multiple factor analysis". Chicago: University of Chicago Press.
- Touzet, C. (1992). "Les réseaux de neurones artificiels: Introduction au connexionnisme", 129 pages.
- Urios, L. (2005). "Technique D'épuration des eaux usées". Technique et documentation. Paris.
- Vasel, J. L. (2007). "Evolution de l'assainissement individuel : Perspectives et questions en suspens", Tribune de l'eau, Vol. 60, N°641, pp.: 3-16.
- Vilagines, R. (2003). "Eau, environnement et santé publique. Introduction à l'hydrologie", 2è édition, Editions Tec&Doc, 198 pages.
- Villebrun, J. F. (1989). La déphosphatation biologique appliquée à la station d'épuration de Craon, Rapport de la DDAF de la Mayenne.

- Vyas, M., Modhera, B., Vyas, V., et Sharma, A. K. (2011). "Performance Forecasting of Common Effluet Treatment Plant Parameters by Artificial Neural Network", ARPN Journal of Engineering and Applied Sciences, Vol. 6, N°1, pp.: 38-42.
- Zeng, GM., Lin, YP., Qin, XS., Huang, GH., Li, JB., et Jiang, R. (2004). "Optimum municipal wastewater treatment plant design with consideration of uncertainty", J. Environ. Sci. 16:126–131.
- Zhu, J., Zurcher, J., Rao, M., *et al.* (1998). "An on-line wastewater quality prediction system based on a time-delay neural network", Engineering Application of Artificial Intelligence 11, pp.: 747-758.
- Zupan, J., et Gasteiger, J. (1993). "Neural Networks For Chemists: An Introduction". VCH, New York.



Annexe 01

|--|

| N | PARAMETRES | UNITE | VALEURS LIMITES | TOLERANCES AUX VALEURS LIMITES ANCIENNES INSTALLATIONS | |
|----|--|-------|--------------------|--|--|
| 1 | Température | °C | 30 | 30 | |
| 2 | PH | - | 6,5 - 8,5 | 6,5 - 8,5 | |
| 3 | MES | mg/l | 35 | 40 | |
| 4 | Azote Kjeldahl | " | 30 | 40 | |
| 5 | Phosphore total | " | 10 | 15 | |
| 6 | DCO | " | 120 | 130 | |
| 7 | DBO5 | | 35 | 40 | |
| 8 | Aluminium | | 3 | 5 | |
| 9 | Substances toxiques bioaccumulables | " | 0,005 | 0,01 | |
| 10 | Cyanures | " | 0,1 | 0,15 | |
| 11 | Fluor et composés | " | 15 | 20 | |
| 12 | Indice de phénols | | 0,3 | 0,5 | |
| 13 | Hydrocarbures totaux | | 10 | 15 | |
| 14 | Huiles et graisses | " | 20 | 30 | |
| 15 | Cadmium | " | 0,2 | 0,25 | |
| 16 | Cuivre total | | 0,5 | 1 | |
| 17 | Mercure total | " | 0,01 | 0,05 | |
| 18 | Plomb total | .11 | 0,5 | 0,75 | |
| 19 | Chrome Total | " | 0,5 | 0,75 | |
| 20 | Etain total | " | 2 | 2,5 | |
| 21 | Manganèse | | 1 | 1,5 | |
| 22 | Nickel total | | 0,5 | 0,75 | |
| 23 | Zinc total | " | 3 | 5 | |
| 24 | Fer | | 3 | 5 | |
| 25 | Composés organiques chlorés | ." | 5 | 7 | |

| Paramètres | Bonne ou trés bonne qualité | Qualité acceptable | Qualité médiocre | Mauvaise ou trés mauvaise |
|-----------------------------|-----------------------------------|--------------------|------------------|------------------------------|
| O ₂ dissous mg/l | >5 | ≥3 | ≥1 | <1 |
| O2 dissous % | ≥70 | ≥50 | ≥10 | <10 |
| DBO ₅ mg / l | ≤5 | ≤10 | 25 | >25 |
| DCO mg / I | ≤25 | ≤40 | 80 | >80 |
| NO3 mg / 1 | ≤25 | ≤50 | 80 | >80 |
| NH4 ⁺ mg / 1 | ≤0.5 | ≤2 | 8 | >8 |
| $NO_2 mg/1$ | ≤0.3 | ≤1 | >1 | F |
| NTK mg/1 | ≤2 | ≤3 | 10 | >10 |
| PO3-4 mg / 1 | ≤0.5 | ≤1 | 2 | >2 |
| MES mg / 1 | ≤70 | × | >70 | ×. |
| Phosphore total mg / l | ≤0.3 | ≤0.6 | 1 | >1 |
| Conductivité | ≤2 | | 2000 | - |
| Ph | $\geq 6.5 \text{ et} \leq 8.5$ | | <6.5 ou >8.5 | - |

Tableau 02 : Les normes des eaux usées rejetées selon l'OMS (1971)

Annexe 02

Production scientifique:

• Publication

Bekkari, N. and Zeddouri, A. (2019), "Using artificial neural network for predicting and controlling the effluent chemical oxygen demand in wastewater treatment plant", Management of Environmental Quality, Vol. 30 No. 3, pp. 593-608. https://doi.org/10.1108/MEQ-04-2018-0084

• Communications Dans Des Manifestations Scientifiques

Bekkari Naceur eddine, Amiri khaled, Débbakh Abd Errezak, Zeddouri Azeiz (2019) Utilisation des réseaux de neurone artificiel pour la prédiction de la DBO₅ d'une station d'épuration à boue activée- Touggourt- in 2eme séminaire international sur l'hydrologie et l'environnement (SIHE), université de Ouargla, Algérie. Du15 au 17 octobre 2019.

Using artificial neural network for Effluent chemical predicting and controlling the effluent chemical oxygen demand in wastewater treatment plant

Naceureddine Bekkari

Laboratory of Underground Reservoirs: Oil. Gas and Aquifers. University of Ouargla, Ouargla, Algeria and Department of Management and Development of Water Resources in Arid Regions, Scientific and Technical Research Centre for Arid Areas (CRSTRA). Biophysical Station, Touggourt, Algeria, and

Aziez Zeddouri

Laboratory of Underground Reservoirs: Oil, Gas and Aquifers, University of Ouargla, Ouargla, Algeria

Abstract

Purpose – Modeling Wastewater Treatment Plant (WWTP) constitutes an important tool for controlling the operation of the process and for predicting its performance with substantial influent fluctuations. The purpose of this paper is to apply an artificial neural network (ANN) approach with a feed-forward back-propagation in order to predict the ten-month performance of Touggourt WWTP in terms of effluent Chemical Oxygen Demand (CODeff).

Design/methodology/approach – The influent variables such as (pHinf), temperature (TEinf), suspended solid (SSinf), Kjeldahl Nitrogen (KNinf), biochemical oxygen demand (BODinf) and chemical oxygen demand (CODinf) were used as input variables of neural networks. To determine the appropriate architecture of the neural network models, several steps of training were conducted, namely the validation and testing of the models by varying the number of neurons and activation functions in the hidden layer, the activation function in output layer as well as the learning algorithms.

Findings – The better results were achieved with an architecture network [6-50-1], hyperbolic tangent sigmoid activation functions at hidden layer, linear activation functions at output layer and a Levenberg – Marquardt method as learning algorithm. The results showed that the ANN model could predict the experimental results with high correlation coefficient 0.89, 0.96 and 0.87 during learning, validation and testing phases, respectively. The overall results indicated that the ANN modeling approach can provide an effective tool for simulating, controlling and predicting the performance of WWTP.

Originality/value - This work is the first of its kind in this region due to the remarkable development in terms of population and agricultural activity in the region, which drove to the increase of water pollutants, so it is necessary to use the modern technologies to modeling and controlling of WWTP.

Keywords Wastewater treatment plant, Activated sludge process, COD, Artifificial neural networks, Learning algorithm, Activation function

Paper type Research paper

1. Introduction

The quality of water is a very important issue because it is an essential factor for different aspects of life, both biotope and biocenosis. To develop any operational assessment for water resources management and planning, it is necessary to develop the evaluation system of water quality parameters.

This research was supported by the Scientific and Technical Research Centre for Arid Areas, CRSTRA. The authors thank the staff of Touggourt wastewater treatment plant for their technical and logistical assistance during this work. Many thanks are also going to Youcef Hallis, Moussa Hadjoudi, Med Hocine Benaissa, Med Lamine Ben haddya, Khaled Amiri and Abd Elaziz Boukhalkhal for helpful comments and advices.

oxygen demand in wastewater treatment plant 593

> Received 2 May 2018 Revised 20 June 2018 19 July 2018 Accepted 26 July 2018



Management of Environmental Quality: An International Journal Vol. 30 No. 3, 2019 © Emerald Publishing Limited 1477-7835 DOI 10.1108/MEQ-04-2018-0084