

Prediction of Common Pollution Parameters Values in Wastewater Treatment Plants by Using Artificial Neural Networks

Dr. Raed JAFAR*

(Received 21 / 12 / 2022. Accepted 1 / 2 / 2023)

□ ABSTRACT □

The measurement of the BOD₅ level in wastewater treatment plant (WWTPs) influent takes five days, and using a prediction model to estimate BOD₅ saves time and allows the use of an online control system. This study investigates the application of artificial neural networks (ANNs) in predicting the influent BOD₅ concentration and the (WWTPs) performance in terms of (BOD₅, COD, TSS) effluent concentrations. To determine the best performing ANN network structure and configuration, sensitivity analysis was performed. The results revealed that the ANN model developed to predict the COD concentration performed the best among the three parameters. The best performing ANN models yielded R² values of (0.92, 0.99, 0.94) for the prediction of the BOD₅, COD and TSS effluent concentrations, respectively. The optimal performing models were obtained (four inputs – one output), which indicated that the influent pH, TDS, TSS and BOD₅ greatly affect the (WWTPs) performance as inputs in all models. The developed prediction model for the influent BOD₅ concentration achieved a high accuracy (R²= 0.974) which indicates that the model is viable as an accurate tool for online control and management systems for (WWTPs). Generally, the ANN model provides a simple approach for the prediction of the (WWTPs) complex processes.

Keywords: wastewater treatment plants, modeling, artificial neural networks, forecasting, biochemical oxygen demand (BOD₅).

* Associate Professor, Department of Environmental Engineering, Faculty of Civil Engineering, Tishreen University, Lattakia, Syria. r-jafar@tishreen.edu.sy

التنبؤ بقيم بارامترات التلوث الشائعة في محطات معالجة مياه الصرف الصحي باستخدام الشبكات العصبونية الصناعية

د. رائد جعفر*

(تاريخ الإيداع 21 / 12 / 2022. قُبل للنشر في 1 / 2 / 2023)

□ ملخص □

يتطلب قياس مستوى BOD_5 لمياه الصرف الصحي خمسة أيام، بينما يوفر استخدام نموذج تنبؤ لتقدير BOD_5 الوقت ويسمح باعتماد نظام تحكم بشكل مباشر في الوقت الحقيقي. تبحث هذه الدراسة في تطبيق الشبكات العصبونية الصناعية (ANNs) للتنبؤ بتركيز BOD_5 الداخل بالإضافة إلى أداء محطات معالجة مياه الصرف الصحي من حيث تراكيز بعض بارامترات التلوث الشائعة مثل (BOD_5 , COD, TSS) الخارجة من المحطة. تم إجراء تحليل الحساسية لتحديد بنية شبكة ANN الأفضل أداءً وتكويناً. أظهرت النتائج أن نموذج ANN الذي تم تطويره للتنبؤ بتركيز COD الخارج كان الأفضل بين النواتج الثلاثة. أعطت نماذج ANN الأفضل أداءً من خلال قيم معامل التحديد R^2 التي بلغت (0.92، 0.99، 0.94) للتنبؤ بتركييزات (BOD_5 , COD, TSS) على التوالي. تم الحصول على نماذج الأداء الأمثل (أربع مدخلات - مخرج واحد)، مما يشير إلى أن الأس الهيدروجيني، المواد الصلبة المنحلة الكلية، المواد الصلبة المعقولة الكلية والطلب على الأوكسجين الحيوي الداخلة إلى محطة المعالجة تؤثر بشكل كبير على أداء المحطة كمدخلات في جميع النماذج. حقق نموذج التنبؤ المطور لتركيز BOD_{5_inf} الداخل إلى محطة المعالجة دقة عالية جداً ($R^2 = 0.974$)، مما يعني أن النموذج قابل للتطبيق كأداة دقيقة يمكن استخدامها في أنظمة التحكم والإدارة لمحطات معالجة مياه الصرف الصحي في الوقت الحقيقي. بشكل عام، يوفر نموذج ANN نهجاً بسيطاً للتنبؤ بالعمليات المعقدة لمحطات معالجة مياه الصرف الصحي.

الكلمات المفتاحية: محطات معالجة مياه الصرف الصحي، النمذجة، الشبكات العصبونية الصناعية، التنبؤ، الطلب على الأوكسجين الحيوي.

* أستاذ مساعد - قسم الهندسة البيئية - كلية الهندسة المدنية - جامعة تشرين - اللاذقية - سورية. r-jafar@tishreen.edu.sy

مقدمة:

يمكن أن تتأثر نوعية المياه سلباً بالظواهر الطبيعية، لكن يكمن السبب الرئيس في تدهورها التلوث الناتج عن الأنشطة البشرية المختلفة، حيث يؤدي التضخم السكاني والتطور العمراني السريع إلى زيادة معدلات تدفق المياه العادمة، بنوعيتها مياه الصرف الصحي والصناعي والتي تعد من أبرز منتجات الأنشطة البشرية الملوثة لمصادر المياه. الأمر الذي يتطلب وتيرة عالية من التقدم في طرق المعالجة والاهتمام ببرنامج مراقبة وتحسين عمل محطات معالجة مياه الصرف الصحي والصناعي بشكل مستمر من أجل الحفاظ على نوعية المياه وإدارتها بشكل صحيح. من المهم أن نقوم بعملية نمذجة محطات معالجة مياه الصرف الصحي ومحاكاتها من أجل تحديد الظروف المثلى لعمليات المعالجة وتجنب فشلها مستقبلاً، ومع ذلك، فإن الخصائص والمتغيرات التي تتطوي عليها محطات المعالجة عديدة وتظهر درجة عالية من التعقيد، مما يؤدي إلى صعوبات في النمذجة من خلال نماذج الانحدار الخطي [1]. قام Ferrer وآخرون [2] بتطبيق برنامج تصميم ومحاكاة أنظمة الحمأة المنشطة (DESASS) لتحسين أداء محطات معالجة مياه الصرف الصحي في الحالة الثابتة. تتضمن البرامج الأخرى نماذج الهضم اللاهوائي [3] ونماذج الحمأة المنشطة ASM [4]. إلا أنه، تحتاج هذه النماذج كمية كبيرة من بيانات الإدخال المتعلقة بعملية المعالجة في كل مرحلة من مراحل التشغيل. يتطلب تحقيق هذا الهدف جمع العينات بانتظام من الموقع وإجراء التحاليل المخبرية اللازمة لمؤشرات التلوث الهامة، الأمر الذي يساعد الباحثين في وضع نماذج للنتيؤ بهذه المؤشرات وتطويرها في محطات المعالجة.

تم تطوير العديد من النماذج الإحصائية التقليدية في المراحل الأولى، حيث تتطلب معظمها عدداً كبيراً من بيانات الإدخال المختلفة والتي لم تكن متاحة بسهولة، إلا أن تنفيذها على أرض الواقع يعدّ عملية مكلفة، يستغرق وقتاً طويلاً ويتطلب جهداً كبيراً. لقد ألهم هذا الأمر الباحثين لتطوير نماذج مبتكرة من أجل محاكاة نوعية المياه بشكل أكثر سرعة ودقة وموثوقية، وذلك باستخدام تقنيات الذكاء الصناعي الحديثة مثل الأنظمة الخبيرة (ES)، المنطق الضبابي (FL)، الخوارزميات الجينية (GA) والشبكات العصبونية الصناعية (ANN).

يمكن تطبيق شبكات ANN للنتيؤ بأداء محطات معالجة مياه الصرف الصحي، ونظراً لدقة نتبؤها العالية، تم تطوير شبكات ANN لنمذجة عملية معالجة المياه [5]، تعتمد هذه النمذجة بشكل كبير على جودة البيانات الإحصائية [6]. قد يؤدي ضعف جودة البيانات الإحصائية هذه إلى ضعف أداء نماذج ANN. ومع ذلك، تتطلب نمذجة ANN كمية صغيرة نسبياً من البيانات لتقديم نتائج نتبؤ مقبولة. يوفر استخدام نماذج الشبكة العصبونية لمحاكاة محطات الصرف الصحي إطاراً لرصد ومراقبة تشغيل المحطة [7]، حيث يؤدي نهج المراقبة إلى تقليل تكلفة التشغيل وتحديد جودة الاستقرار البيئي.

جمع Mjalli وآخرون [7] البيانات على مدى عام واحد (أجريت القياسات كل خمسة أيام). تم تحديد مخرجات مياه المعالجة الثانوية كمدخلات لنموذجهم. قاموا بتقسيم مجموعة البيانات الإجمالية بمعدل (1:2:4) إلى مجموعات بيانات التدريب (Training) والتحقق من الصحة (Validation) والاختبار (Testing) على التوالي. تجدر الإشارة إلى أن Mjalli وآخرون (2007) استخدموا شبكة Feedforward العصبونية (FFNN)، والتي تم تطويرها باستخدام برنامج MATLAB في جميع شبكات الإدخال الفردية، تتكون هيكلية الشبكة من ثلاث طبقات (طبقات الإدخال والمخفية والمخرجات). أما في الشبكات متعددة المدخلات والأكثر تعقيداً، تتكون بنية الشبكة من أربع طبقات (طبقتان مخفيتان). تم الحصول على بنية الشبكة المثلى لتكون (1-40-1) مع COD كمتغير مُدخل وحيد، حيث كان معامل التحديد ($R^2=0.987$) ومتوسط الخطأ التربيعي ($MSE=0.021$). استند تابع مخرجات التدريب إلى متوسط الخطأ

التربيعي (MSE) بين تنبؤ الشبكة والقيم المقاسة. طبق Nasr وآخرون (2012) [8] الشبكات العصبونية الصناعية ANN للتنبؤ بأداء محطات معالجة مياه الصرف الصحي في الإسكندرية. حيث تمت دراسة أداء المحطة من خلال إجمالي المواد الصلبة العالقة (TSS)، والطلب الحيوي على الأوكسجين (BOD_5)، والطلب الكيميائي للأوكسجين (COD) على مدار عام واحد. تم تصنيف البيانات التي تم جمعها إلى أربع مجموعات، حيث تمثل كل مجموعة ثلاثة أشهر من السنة. تم تطوير شبكة التغذية الأمامية وتدريبها باستخدام الانتشار العكسي، كانت القيم المتوقعة مرتبطة ارتباطاً وثيقاً بالقيم المقاسة حيث بلغ معامل الارتباط ($R=0.90317$).

طبق Guo وآخرون (2015) [9] التعلم الآلي من خلال التغذية المسبقة لشبكات ANN وآلة شعاع الدعم Support Vector Machine (SVM) للتنبؤ بالتركيز الكلي للنيتروجين في مياه الصرف الصحي في محطة المعالجة في مدينة Ulsan، كوريا. كانت مدخلات النماذج: النيتروجين الكلي، والفوسفور الكلي، والمواد الصلبة المنحلة للتدفق الداخل. تم التقييم بناءً على معامل التحديد (R^2) وكفاءة Nash-Sutcliffe (NSE) ومعايير الكفاءة النسبية (d_{rel})، أظهرت النتائج أن كلا النموذجين كانا فعالين في التنبؤ ($d_{rel}=0.8$ ، $NSE=0.56$ ، $R^2=0.55$) على الرغم من أن نموذج SVM حقق كفاءة تنبؤ أعلى، إلا أن نموذج ANN المكون من ثلاث طبقات (طبقات الإدخال والمخفية والمخارج) كان أكثر فاعلية في ربط قيم متغيرات الإدخال بتراكيز TN.

استخدم Nezhad وآخرون (2016) [10] مجموعة أدوات ANN في بيئة عمل MATLAB لتطبيق تقنيات التعلم الآلي للتنبؤ بمؤشر جودة مياه الصرف الصحي الخارجة من محطات المعالجة في طهران. تم تطوير شبكة عصبونية ذات تغذية أمامية وانتشار عكسي للخطأ تتكون من ثلاث طبقات. تضمنت مدخلات النموذج (BOD, COD, TDS, pH, PO_4 , NH_4 , TSS, FC). أشارت النتائج إلى أن هيكلية الشبكة المثلى كانت (1-7-8)، مما أدى إلى كفاءة عالية في التنبؤ بجودة مياه الصرف الصحي الخارجة ($R=0.96$, $MSE=0.1$).

تتبعاً Abba & Elkiran (2007) [11] بقيم الـ COD للتدفق الخارج من محطات معالجة مياه الصرف الصحي باستخدام شبكة التغذية الأمامية العصبونية (FFNN). يعتبر COD الخارج من المحطة من البارامترات الرئيسية لتقييم أداء محطات معالجة مياه الصرف الصحي. تم تطبيق (FFNN) للتنبؤ بـ COD، BOD، pH، T-N، T-P، TSS، EC لمياه الصرف الخارجة من محطات معالجة مياه الصرف الصحي. تم دراسة العديد من الهيكلية ومجموعات المدخلات؛ حيث أن نموذج (FFNN) بهيكلية شبكة (1-8-8)، باستخدام جميع بارامترات الإدخال، حقق أداء جيد وأعلى دقة في التنبؤ بـ COD الخارج حيث أعطى ($R^2=0.7$, $RMSE=0.0108$).

تم تطوير العديد من نماذج الانحدار والشبكات العصبونية الصناعية للتنبؤ بقيم الـ (BOD_5) [12, 13] تتبع أهمية نمذجة BOD_5 من تقليص الإجراءات المخبرية المكثفة التي يتم تطبيقها لقياس تركيز BOD_5 ، حيث تتطلب هذه الاختبارات خمسة أيام. تم تطوير نموذج انحدار يعتمد على العلاقة بين خصائص مياه الصرف الصحي و BOD_5 ، وثبت أنه يحقق دقة عالية ($R^2=0.7966$) [14]. كما طور Kasem وآخرون (2018) [12] جهاز استشعار برمجي (software sensor) لمراقبة تركيز BOD_5 في نهر Sefidrood في إيران باستخدام ANN تغذية أمامية كدالة لمستوى الأوكسجين المنحل. كان أداء ANN الذي تم تطويره مناسباً حيث حقق قيمة R^2 عالية (تصل إلى 0.89). في محطات معالجة مياه الصرف الصحي، تم أيضاً استخدام الشبكات العصبونية العميقة والخوارزميات الجينية لتصميم software sensor لـ BOD_5 [15]. تم اختبار البرنامج المطور في ظل ثلاثة ظروف جوية الجافة والممطرة والعاصفة على منصة محاكاة BSM1، وأظهرت النتائج أداءً جيداً في ظل الظروف الجوية القاسية. على الرغم من

وجود عدد قليل من الدراسات التي طورت برامج استشعار لرصد BOD_5 ، إلا أن هناك نقصاً في تحليل البارامترات وتأثيرها على التنبؤ بتركيز BOD_5 الداخل إلى محطة المعالجة.

أهمية البحث وأهدافه:

نظراً لأهمية الموارد المائية وحماية البيئة من التلوث الناتج عن صرف مياه محطات المعالجة التي لا تعمل بشكل صحيح، الأمر الذي يدفعنا إلى مراقبة، تقييم وتحسين عمل هذه المحطات بشكل مستمر لكافة منشآتها. يمكن إنجاز ذلك باستخدام تقانات الذكاء الصناعي الحديثة وبدقة عالية. تتكون محطات معالجة مياه الصرف الصحي عادةً من سلسلة من العمليات المعقدة التي لا يمكن نمذجتها باستخدام تقنيات الانحدار البسيطة نظراً للعدد الكبير من بارامترات الإدخال ونقاط البيانات المطلوبة لتحديد أداء المحطة. وبالتالي، فإن قدرة ANN تجعلها حلاً موثوقاً في نمذجة محطات معالجة مياه الصرف الصحي. الهدف من هذا البحث هو تطبيق الشبكات العصبونية الصناعية في نمذجة محطات معالجة مياه الصرف الصحي وتوفير خوارزمية بسيطة لتحديد التكوين الأمثل للشبكات العصبونية الصناعية التي تحدد السلوك المعقد لعمل محطات المعالجة. بالإضافة إلى ذلك، يمكن استخدام النموذج الناتج في تصميم أنظمة التحكم ومراقبة أداء المحطة ومعايير جودة المياه الخارجة. تهدف هذه الدراسة أيضاً إلى تطوير نموذج ANN قابل للتطبيق ليكون بمثابة أداة لتحديد قيم الـ BOD_5 الداخل إلى محطات معالجة مياه الصرف الصحي. تقلل هذه الأداة فترة قياس BOD_5 الداخل إلى المحطة من خمسة أيام إلى عدة ساعات، وكما تقلل أيضاً من تكلفة التشغيل، مما سيسمح باستخدام أنظمة التحكم في المحطة بشكل مباشر في الوقت الفعلي (online) وبالتالي ضمان جودة الاستقرار البيئي.

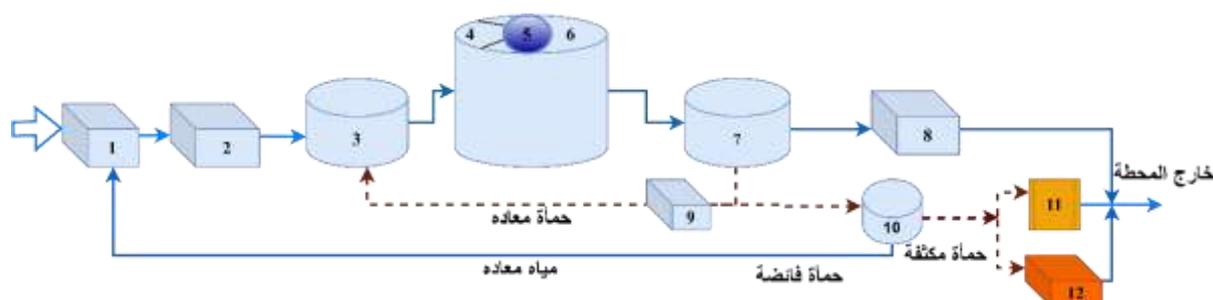
طرائق البحث ومواده:

1. منطقة الدراسة:

تقع محطة الرويمية في الريف الجنوبي لمحافظة اللاذقية باتجاه الجنوب الشرقي لمركز مدينة اللاذقية، تبعد حوالي 4km عن الطريق الدولي (اللاذقية-دمشق)، تخدم هذه المحطة تجمع قرى الرويمية والقطرية يقدر عدد سكانها وفق احصائيات عام 2020 بحوالي (7700) نسمة، الغزارة التصميمية للمحطة ($1100m^3/day$).

تعمل المحطة بنظام المعالجة البيولوجية (تهوية مديدة مع إزالة الفوسفور والنتروجين). تتم عملية المعالجة وفق المراحل التالية:

- معالجة ابتدائية تهدف للتخلص من المواد الصلبة العالقة والطافية.
 - معالجة بيولوجية تهدف للتخلص من المواد العضوية أو تخفيضها إلى الحد المقبول وفق المعايير المعتمدة.
 - التعقيم بالكlor: من التخلص من الكائنات الدقيقة والمرضة.
 - معالجة الحمأة الناتجة.
- يبين الشكل (1) أجزاء المحطة وآلية عملها.



الشكل (1) آلية عمل محطة الرومية

1- محطة ضخ المياه والحواجز القضبانية الخشنة، 2- حوض حجز الرمال والزيوت، 3- حوض لاهوائي، 4- حوض منقوص الأوكسجين، 5- حوض ترسيب أولي، 6- حوض الحمأة المنشطة، 7- حوض الترسيب النهائي، 8- حوض التعقيم، 9 محطة ضخ الحمأة، 10 حوض تكثيف الحمأة، 11 ساحات تجفيف الحمأة، 12 مكبس الحمأة.

من أهم المشاكل التشغيلية في المحطة والتي تمت ملاحظتها من خلال الجولات الميدانية مشكلة انقطاع التيار الكهربائي التي تعدّ من أبرز المشاكل التي تعاني منها المحطة حيث تخرج عن العمل وبشكل يومي لساعات عديدة. من مشاكل برنامج المراقبة (الاعتيان والقياس) أيضاً، أنه تم التوقف في العام 2013 بشكل جزئي عن إجراء التحاليل المخبرية وذلك للترشيد في استهلاك المواد المخبرية، حيث عدّل برنامج أخذ العينات إلى 4 عينات شهرية فقط، الأمر الذي أثر على مراقبة مؤشرات التلوث في المحطة.

مواصفات المياه المعالجة

تعد المعالجة البيولوجية الحل الأفضل والأكثر اقتصادياً وشيوعاً في معالجة مياه الصرف الصحي وكذلك الصرف الصناعي الناتجة عن الصناعات الغذائية، تضع معظم الكودات العالمية شروطاً خاصة لهذه المياه حتى تؤمن سير عملية المعالجة البيولوجية بشكل صحيح وأكثر جدوى. من هذه الاشتراطات:

- ألا تكون درجة حرارة المياه أقل من 6°C أو أكبر من 30°C .
- أن تكون قيمة الأس الهيدروجيني pH في المجال (6.5-8.5).
- ألا يتجاوز التركيز الكلي للأملاح المنحلة TDS عن 10 g/l .
- ألا تحتوي المياه على زيوت أو مواد نفطية أو مواد صعبة الأكسدة بيولوجياً أو مواد ضارة.
- ألا تتجاوز قيمة BOD_5 القيمة التصميمية لمحطة المعالجة ($100\text{--}500\text{ mg/l}$) للمياه الداخلة و ($20\text{--}40\text{ mg/l}$) للمياه الخارجة.
- ألا تكون قيمة COD أكبر من قيمة BOD_5 بأكثر من (1.5-2) مرة.
- نسبة التركيز الأدنى للمغذيات العضوية (P, N) إلى تركيز BOD_5 تكون على الشكل: $\text{BOD}_5:\text{N}:\text{P}=100:5:1$.

2. جمع البيانات وإعدادها

قمنا بإجراء سلسلة من التحاليل المخبرية لعينات المياه المأخوذة من مدخل ومخرج محطة المعالجة لدورة هيدرولوجية واحدة (2021)، ثم تم دمجها مع البيانات الأرشيفية التي تم تجميعها في هذه الدراسة من محطة معالجة مياه الصرف الصحي في الرومية، على مدى ثمان سنوات (2012 إلى 2020). تم اعتماد خمسة بارامترات لمياه الصرف الصحي كمدخلات نموذجية، بما في ذلك درجة الحموضة pH، المواد الصلبة المنحلة الكلية TDS، تراكيز المواد الصلبة المعلقة TSS، COD، BOD_5 . تصف بارامترات الإدخال والإخراج المختارة أداء محطات معالجة مياه الصرف الصحي بشكل عام.

يوضح الجدول (1) الوصف الاحصائي للبيانات المعتمدة من أجل تقييم المحطة ونمذجة أداءها.

الجدول (1) الوصف الاحصائي للبيانات المعتمدة

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
pH_inf	180	5.0	9.4	7.097	1.0266
TDS_inf (mg/l)	180	333	1368	839.81	243.303
TSS_inf (mg/l)	180	33	198	121.41	33.637
COD_inf (mg/l)	180	70	253	162.78	39.374
BOD ₅ _inf (mg/l)	180	11	205	111.02	37.556
pH_eff	180	5.1	9.6	7.244	0.9239
TDS_eff (mg/l)	180	327	1298	800.63	237.069
TSS_eff (mg/l)	180	6	26	14.57	4.524
COD_eff (mg/l)	180	20	48	32.29	6.214
BOD ₅ _eff (mg/l)	180	9	46	25.35	7.157

3. الشبكات العصبونية الصناعية

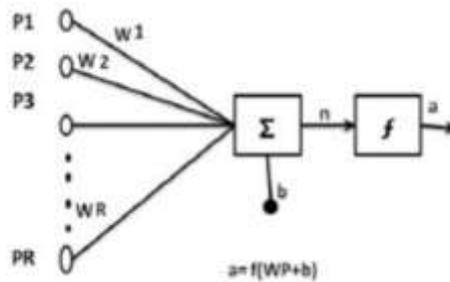
تعرف الشبكات العصبونية الصناعية على أنها نظام معالجة المعلومات بشكل يشابه ويحاكي طريقة عمل الجملة العصبية البيولوجية عند الانسان، لأنها تتصف بنفس خصائصها التي تعتمد على فكرة تخزين المعلومات وتعلمها والاستفادة منها لاحقاً. لقد زاد الاهتمام بالشبكات العصبونية الصناعية نظراً للخصائص المميزة لها وقدرتها على تجاوز محددات الحلول التقليدية. من خصائص هذه الشبكات التوازي (Parallelism) وهي أساس الشبكات العصبونية الصناعية، إذ تعمل كافة العصبونات في الشبكة بشكل متزامن، أي أن التوازي يمثل قدرة الشبكات على معالجة البيانات بسرعة عالية جداً مهما بلغت ضخامة تلك البيانات.

يمكن للشبكات أن تتيح لنفسها تعلم الحالة غير الخطية (Non-Linear Behavior) أو أن تتيح لنفسها تعلم السلوك الديناميكي (Dynamic Behavior) من خلال المعطيات المقاسة. وطالما أن معرفة الإجراءات الداخلية ليست ضرورية، لذا فإن النمذجة للنظام يمكن أن تتم بمعلومات قليلة ارشيفية سابقة، حيث تستبدل النماذج الرياضية المعقدة من خلال فترة التدريب بعمليات نظامية تقوم على معطيات متوفرة حالية.

الخلية العصبية الحية Biological Neural Cell: يتكون الجهاز العصبي البشري من كمّ هائل ومعقد من الخلايا العصبية (تقريباً 10 بليون خلية عصبية Neuron cell)، تكون بارتباطها شبكة كبيرة من الأعصاب، تتكون الخلية العصبية النموذجية من أربعة أجزاء رئيسية جسم الخلية (Soma)، الألياف العصبونية للمدخل (Dendrites)، الليف العصبي للمخرج أو المحور (Axon)، والعقد العصبية أو المشبك العصبي (Synapses) الذي يربط الليفة العصبية للمخرج مع الليفة العصبية للمدخل لنيرون آخر.

يعتبر العصبون البيولوجي عنصر الحساب الأساسي في الأنظمة البيولوجية الحية، حيث أن العصبون هو خلية صغيرة جداً تستقبل تنبيهاً إلكترو كيميائياً من منابع عديدة وتستجيب بتوليد نبضات كهربائية ترسل إلى عصبونات أخرى أو خلايا مؤثرة.

- الخلية العصبونية الصناعية Artificial Neural Cell:** صممت الشبكات العصبونية الصناعية لتحاكي عمل الشبكات العصبونية الطبيعية كما هو موضح في الشكل (2) وفق ما يلي:
- تحدث عمليات المعالجة ضمن عدة عناصر معالجة بسيطة (Process Element) تدعى العصبونات (Neurons).
 - تعبر الإشارات الناتجة عن تفعيل العصبون إلى العصبون الآخر من خلال وصلة بينهما (Connection).
 - لكل وصلة وزن مرتبط بها (Weight) يضرب بالإشارة المرسله عبر الوصلة.
 - تجمع وحدات المعالجة المدخلات الموزونة لتشكل مدخلاً صافياً للعصبون (Net input).
 - يطبق العصبون على مدخله الصافي دالة تفعيل (Activation Function) ليعطي قيمة وحيدة تمثل خرج العصبون ليعاد انتشارها إلى العصبونات التالية.



الشكل (2) الخلية العصبونية الصناعية

ينقسم عمل الوحدة العصبونية الصناعية إلى مرحلتين رئيسيتين:

المرحلة الأولى: يتم فيها معالجة المدخلات بضرب كل مدخل منفرد بوزن معين (W)، ثم تجرى عملية جمع المدخلات الموزونة، ويضاف إليها عامل انحياز (b)، يعبر عنها رياضياً بالمعادلة التالية

$$\text{Net} = \sum_{i=1}^{i=q} W_i \cdot P_i + b \quad \text{Eq.1}$$

حيث:

Net : ناتج المرحلة الأولى

P_i : المدخلات

W : أوزان المدخلات

b : عامل الانحياز (هو بمثابة الحد الثابت من المعادلات الرياضية في كثيرات الحدود الذي يساعد في حل تلك المعادلة بشكل أسرع وأسهل).

q : عدد المدخلات

المرحلة الثانية: تتولى معالجة مخرجات المرحلة السابقة باستخدام دوال رياضية تدعى دوال التفعيل (Activation Function) وفق المدى المحدد لكل نوع من أنواع هذه الدوال لينتج عنها مخرجات يعبر عنها بالمعادلة التالية:

$$\text{Output} = f(\text{Net}) = a \quad \text{Eq.2}$$

ترتب العصبونات ضمن عدد من الطبقات كما هو موضح في الشكل (3) وهي:

1- طبقة الدخل Input layer تتكون من مجموعة من الوحدات أو عناصر المعالجة Nodes التي تعمل على توزيع القيم المدخلة إليها من الوسط الخارجي إلى الطبقة الخفية التي تليها عبر الارتباطات الموجودة بينهما، فهي بذلك لا تقوم بأي عملية حسابية.

2- الطبقات الخفية Hidden layers هي مجموعة من الطبقات التي تقع بين طبقتي الدخل والخرج، وتتكون من طبقة واحدة أو أكثر تبعاً لحجم وطبيعة الظاهرة المدروسة، مع العلم أن (85%) من التطبيقات المطورة تم الاكتفاء بطبقة واحدة أو طبقتين. إن عناصر أو وحدات المعالجة الموجودة في هذه الطبقات هي التي تقوم بالعمليات الحسابية حيث تتلقى الإشارات من وحدات طبقة الدخل لتعالجها وتطبق تابع التنشيط على تلك الإشارات ثم تحويلها إلى الطبقة الخفية التالية أو إلى طبقة الخرج.

3- طبقة الخرج Output layer وهي الطبقة النهائية التي تقع عند مخرج الشبكة، حيث تقوم عناصر المعالجة فيها باستلام الإشارات من الطبقة الخفية السابقة لتعالجها وتقوم بإجراءات العمليات الحسابية بشكل مشابه تماماً للطبقات الخفية لتعطي بعد ذلك النتيجة النهائية.

يحدد عدد عصبونات الدخل والخرج تبعاً لأعداد بارامترات الدخل والخرج للشبكة، في حين يحدد عدد عصبونات الطبقة الخفية تجريبياً، ويدعى الإطار العام لمخطط الارتباط الداخلي بطوبولوجية الشبكة، وهو يحدد عدد الطبقات وعدد العصبونات في كل طبقة.

دالة تفعيل العصبون Neuron's activation function: تعالج مدخلات العصبون باستخدام دالة تفعيل رياضية تعبر عن تابع رياضي لتحويل إشارة الخرج وفقاً للمعادلة السابقة Eq.2، يوجد عدد من دوال التفعيل ذات الوظائف المختلفة، وكل منها يصلح لنوع معين من التطبيقات اعتماداً على المدى المحدد لها، نذكر منها:

- دالة سيغمويد اللوغاريتمية (Log-sigmoid Transfer Function): هي واحدة من الدوال الواسعة الاستخدام لأنها تقوم بمعالجة القيم بشكل انسيابي بمدى يتراوح بين (0) و (1) مما يوفر توزيعاً ملائماً لهذه القيم، حيث تساعد في تحقيق قيم المخرجات والمدخلات كافة وفق المعادلة التالية

$$f(\text{Net}) = \frac{1}{1+e^{(-\text{Net})}} \quad \text{Eq.3}$$

- دالة سيغمويد لظل الزاوية Tan-sigmoid Transfer Function:

تشابه دالة سيغمويد اللوغاريتمية، وتتمتع بالمزايا نفسها من حيث الانسيابية، إلا أن القيم الخارجة تكون بمدى يتراوح بين (1) و (-1)، ومعادلة الدالة على الشكل التالي:

$$f(\text{Net}) = \frac{2}{1+e^{(-\text{Net})}} - 1 \quad \text{Eq.4}$$

- الدالة الحدية Hard-Limit Transfer Function:

هي من الدوال التي تقوم بتقريب قيم المخرجات إما إلى (1) عندما تكون قيمة أي معامل موجبة أو إلى (0) عندما تكون قيمة أي معامل سالبة، كما هو مبين في المعادلة التالية:

$$f_x = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x \leq 0 \end{cases} \quad \text{Eq.5}$$

- الدالة الخطية Linear Transfer Function:

تقع قيم المخرجات بين (1) و (-1) ولذلك تدعى بدالة التحفيز الثنائي، معادلتها كما يلي:

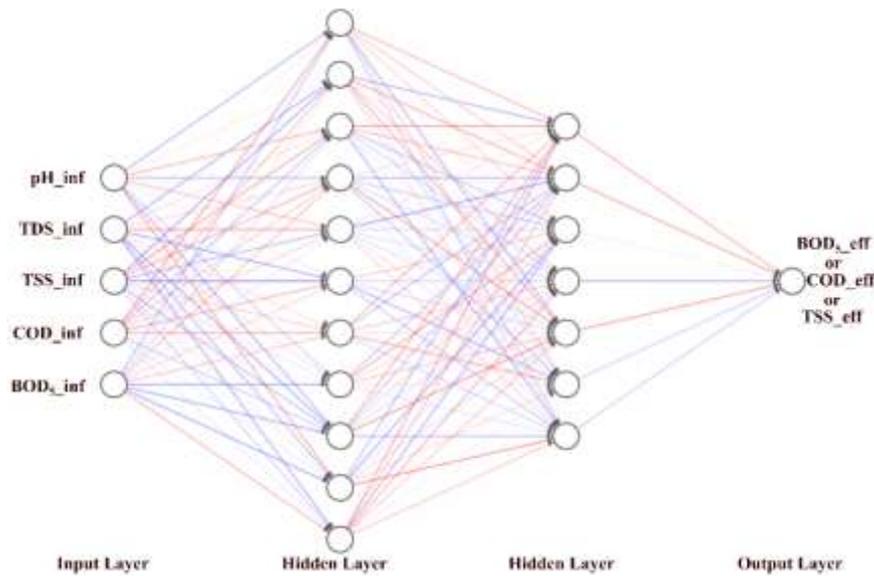
$$f_x = x \text{ for all } x \quad \text{Eq.6}$$

يتم تدريب الشبكة من خلال الإضافة المستمرة للبيانات. بعد ذلك، يتم تعميم الشبكة من خلال إدخال بيانات جديدة غير مرئية من قبل النموذج. تكمن المزايا الرئيسية لشبكات ANN في سرعة التعلم العالية ومعالجة البيانات وأيضاً قدرتها على تمثيل الأنظمة غير الخطية بشكل فعال للغاية.

يوضح الشكل (3) بنية (FFNN) مشتركة، تتكون من ثلاث طبقات شبكة رئيسية. تسمى الطبقة الأولى طبقة الإدخال، والتي تتلقى بيانات الإدخال وترسل هذه البيانات إلى الطبقات المخفية. الطبقات المخفية هي قلب الشبكة العصبونية وتتكون من خلايا عصبونية متعددة. يمكن في شبكة واحدة أن يكون هناك العديد من الطبقات المخفية مع العديد من الخلايا العصبونية. عادة، يتناسب عدد الطبقات المخفية مع مدى تعقيد النظام المراد نمذجته. مبدأ عمل ANN هو نفسه بالنسبة لأنواع الشبكات العديدة. يتلقى عنصر المعالجة الأساسي، العصبون، إشارات الإدخال ثم يعالجها من خلال وظيفة التنشيط ويوفر إشارة خرج. بالإضافة إلى ذلك، فإن وزن كل خلية عصبونية وتوابع التنشيط مسؤولة عن تمرير الإشارات من طبقة إلى الطبقة التالية. يوضح التعبير الرياضي مبدأ عمل الشبكة العصبونية في المعادلة:

$$Y_i = f(\sum_{j=1}^M W_{ij}X_j + b_i) \quad \text{Eq.7}$$

حيث Y_i هي قيمة الناتج المتوقع i ، f هي تابع التنشيط، W_{ij} هو الوزن المخصص لكل مُدخل j ، M هو العدد الإجمالي للمدخلات، b_i هو الانحياز لكل مخرج. أكثر أنواع توابع التنشيط شيوعاً هو تابع سيغمويد [16, 17] كما ورد سابقاً.



الشكل (3) بنية التغذية الأمامية النموذجية لشبكة عصبونية صناعية.

يوضح الشكل (3) هيكلية تغذية نموذجية لشبكة عصبونية صناعية. حيث X هو مدخل تابع التنشيط. في هذه الدراسة، تم تطبيق أداة الشبكة العصبونية في بيئة عمل MATLAB 2017b لتطوير وتدريب الشبكة العصبونية الأمامية باستخدام تابع سيغمويد لظل الزاوية tan-sigmoid المعادلة (4) لجميع الطبقات. يتم تدريب ANN باستخدام مصفوفة إدخال واحدة تحتوي على ما يصل إلى خمسة مدخلات بشكل متدرج لكل من المخرجات الموافقة للموديل المدروس وفق الهيكلية المعتمدة. يلخص الجدول (2) الهيكليات والتكوينات المستخدمة في هذه الدراسة. تتم معالجة بيانات الإدخال مسبقاً قبل التدريب لتحسين أداء تدريب الشبكة.

الجدول (2) تكوينات المدخلات المعتمدة في الدراسة

	Input	Parameter's combination
Two input parameters	input 01	pH_inf & TDS_inf
	input 02	pH_inf & TSS_inf
	input 03	pH_inf & COD_inf

	input 04	pH_inf & BOD ₅ _inf
	input 05	TDS_inf & TSS_inf
	input 06	TDS_inf & COD_inf
	input 07	TDS_inf & BOD ₅ _inf
	input 08	TSS_inf & COD_inf
	input 09	TSS_inf & BOD ₅ _inf
	input 10	COD_inf & BOD ₅ _inf
Three input parameters	input 11	TSS_inf & pH_inf & TDS_inf
	input 12	TSS_inf & COD_inf & BOD ₅ _inf
	input 13	TSS_inf & pH_inf & COD_inf
	input 14	TSS_inf & pH_inf & BOD ₅ _inf
	input 15	TSS_inf & TDS_inf & COD_inf
	input 16	TSS_inf & TDS_inf & BOD ₅ _inf
	input 17	pH_inf & TDS_inf & COD_inf
	input 18	pH_inf & TDS_inf & BOD ₅ _inf
	input 19	pH_inf & COD_inf & BOD ₅ _inf
	input 20	TDS_inf & COD_inf & BOD ₅ _inf
Four input parameters	input 21	pH_inf & TDS_inf & TSS_inf & COD_inf
	input 22	pH_inf & TDS_inf & TSS_inf & BOD ₅ _inf
	input 23	TDS_inf & TSS_inf & COD_inf & BOD ₅ _inf
	input 24	TSS_inf & pH_inf & COD_inf & BOD ₅ _inf
	input 25	pH_inf & TDS_inf & COD_inf & BOD ₅ _inf
Five input parameters	input 26	pH_inf & TDS_inf & TSS_inf & COD_inf & BOD ₅ _inf

لتحديد العدد الأمثل للطبقات والخلايا العصبونية المخفية في كل طبقة، يجب اختبار العديد من هيكليات الشبكة، ويجب تقييم أداء كل شبكة. على الرغم من أن عدداً قليلاً من الطبقات والخلايا العصبونية المخفية يمكن أن يؤدي إلى تعميم مناسب، إلا أنه قد يؤدي أيضاً إلى شبكة غير ملائمة. في المقابل، يمكن لعدد كبير من الطبقات المخفية أن تقلل من خطأ التدريب ولكن يمكن أن تزيد من فرط تعلم الشبكة [18]. تم اختبار التكوينات المختلفة لهيكلية الشبكة، وتم تحديد الهيكلية المثلى لكل ناتج خرج من خلال التجربة والخطأ. تضمنت بارامترات التكوين (تشكيل الموديل) عدد المدخلات والطبقات المخفية والخلايا العصبونية في الطبقة المخفية.

النتائج والمناقشة:

- مصفوفة الارتباط

تشير مصفوفة الارتباط الملخصة في الجدول (3)، إلى وجود علاقة خطية ضعيفة بين معظم البارامترات المدروسة، ومع ذلك، لا يشير هذا إلى عدم وجود أي أنواع أخرى من العلاقات، حيث يشير معامل الارتباط الصغير إلى أن طرق

الانحدار التقليدية غير مناسبة للتنبؤ بمثل هذا النظام المعقد. بالإضافة إلى ذلك، تقتصر بارامترات الإدخال إلى أي ارتباط فيما بينها، وهو أمر مطلوب لتدريب نماذج ANN موثوقة. حيث يؤدي وجود بارامترات الإدخال المرتبطة إلى تحيز ANN وانحرافها نحو تأثير هذه البارامترات.

الجدول (3) مصفوفة الارتباط للبارامترات المعتمدة

Correlations (Pearson Correlation)										
	pH_inf	TDS_inf	TSS_inf	COD_inf	BOD ₅ _inf	pH_eff	TDS_eff	TSS_eff	COD_eff	BOD ₅ _eff
pH_inf	1									
TDS_inf	0.284**	1								
TSS_inf	0.333**	0.706**	1							
COD_inf	0.075	0.186*	0.309**	1						
BOD ₅ _inf	0.240**	0.465**	0.513**	0.687**	1					
pH_eff	0.140	0.003	0.123	-0.091	0.037	1				
TDS_eff	0.088	0.857**	0.604**	0.029	0.340**	0.193**	1			
TSS_eff	0.159*	0.325**	0.374**	-0.026	0.167*	0.149*	0.263**	1		
COD_eff	-0.067	0.082	0.240**	0.100	0.168*	0.204**	0.138	0.334**	1	
BOD ₅ _eff	-0.048	0.207**	0.242**	0.172*	0.268**	0.106	0.207**	0.330**	0.578**	1

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

– النمذجة والتنبؤ بواسطة ANN

تم تطوير نموذج ANN باستخدام هيكليات مختلفة للحصول على نماذج ANN ذات الأداء الأمثل. أولاً، تم تطبيق كل من بارامترات التلوث الداخلة إلى المحطة كمدخلات للتنبؤ بكل من البارامترات الخارجة (مدخلين لمخرج واحد). بعد ذلك، تمت زيادة عدد المدخلات تدريجياً حتى تم تطبيق جميع المدخلات (لغاية خمسة مدخلات) للتنبؤ بكل ناتج. تم أيضاً فحص عدد الطبقات المخفية وعدد الخلايا العصبونية بحثاً عن بارامترات الشبكة المثلى. تم تقسيم مجموعة البيانات إلى مجموعات التدريب والتحقق والاختبار بنسبة (10:10:80) على التوالي. مؤشر الأداء الرئيسي في هذه الدراسة هو معامل التحديد (R^2) المعادلة (8)، ومتوسط الخطأ التربيعي المعادلة (9) بين القيم المتوقعة والقيم المقاسة لكل نماذج ANN التي تم تطويرها :

$$R^2 = \frac{\left[\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \right]^2}{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2} \quad \text{Eq.8}$$

$$\text{MSE} = \frac{1}{n} \cdot \sum_{i=1}^n (X_i - Y_i)^2 \quad \text{Eq.9}$$

حيث n هو العدد الإجمالي للنقاط، X_i هي القيمة المقاسة للمتغير، Y_i هي القيمة المتوقعة للمتغير، \bar{X} متوسط قيم المتغير X ، \bar{Y} متوسط قيم المتغير Y .

تم تلخيص نتائج تطوير نماذج مدخلين لمخرج واحد ANN وطبقة مخفية واحدة من أجل تقدير تراكيز (BOD_5 , COD, TSS) الخارجة من محطة معالجة مياه الصرف الصحي في الجدول (4). حيث تشير قيم R^2 المرتفعة نسبياً إلى قدرة الشبكات العصبونية على نمذجة النظام باستخدام بارامترتي إدخال فقط (pH_inf , TSS_inf) وفق المعاريات الموضحة في الجدول (4).

الجدول (4) أداء نماذج ANN لمُدخلين ومُخرج واحد من أجل التنبؤ بـ (BOD₅, COD, TSS).

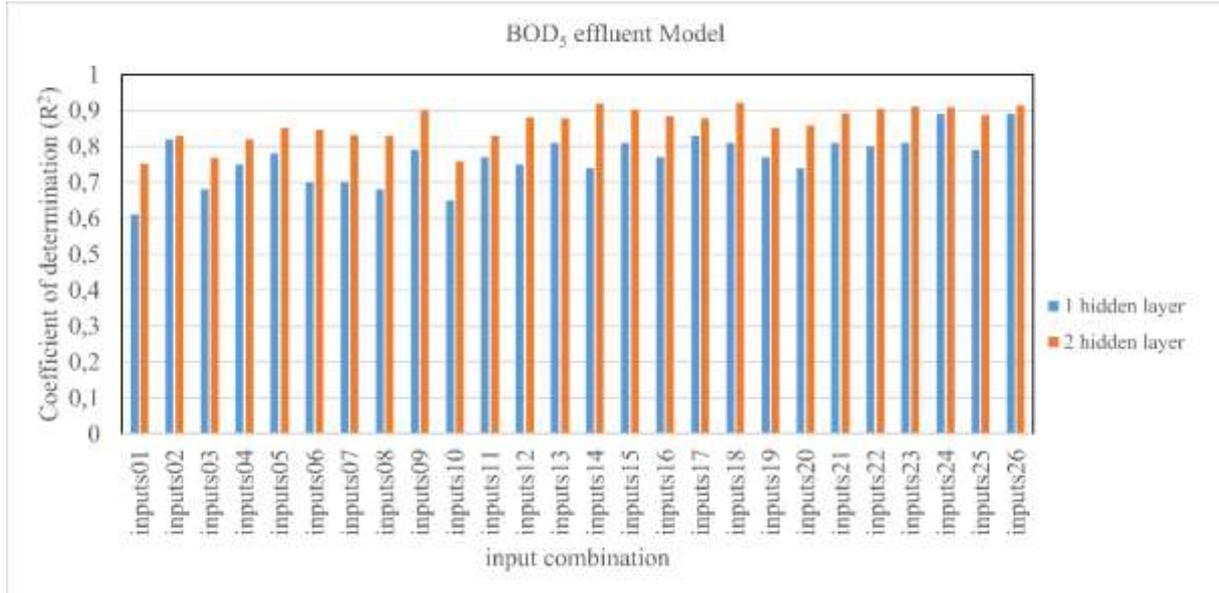
المدخلات	BOD ₅ effluent Model			COD effluent Model			TSS effluent Model		
	هيكلية الـ ANN	R ²	MSE	هيكلية الـ ANN	R ²	MSE	هيكلية الـ ANN	R ²	MSE
pH _{inf} & TDS _{inf}	2-58-1	0.61	33.359	2-99-1	0.76	18.012	2-61-1	0.8	7.4123
pH _{inf} & TSS _{inf}	2-97-1	0.82	17.333	2-60-1	0.77	15.626	2-79-1	0.86	5.3247
pH _{inf} & COD _{inf}	2-41-1	0.68	27.806	2-43-1	0.67	21.4	2-89-1	0.82	7.4341
pH _{inf} & BOD ₅ _{inf}	2-95-1	0.75	23.459	2-67-1	0.74	17.681	2-46-1	0.73	9.5753
TDS _{inf} & TSS _{inf}	2-73-1	0.78	21.269	2-96-1	0.72	20.783	2-95-1	0.82	6.8102
TDS _{inf} & COD _{inf}	2-43-1	0.7	26.723	2-58-1	0.71	19.414	2-56-1	0.74	9.3981
TDS _{inf} & BOD ₅ _{inf}	2-72-1	0.7	27.202	2-92-1	0.71	19.926	2-70-1	0.72	10.081
TSS _{inf} & COD _{inf}	2-31-1	0.68	28.533	2-51-1	0.73	18.603	2-99-1	0.85	6.461
TSS _{inf} & BOD ₅ _{inf}	2-48-1	0.79	22.88	2-97-1	0.77	15.664	2-82-1	0.76	8.6196
COD _{inf} & BOD ₅ _{inf}	2-60-1	0.65	31.678	2-58-1	0.71	20.556	2-45-1	0.65	11.886

تم الحصول على أفضل أداء للنماذج المذكورة أعلاه في التنبؤ بالطلب على الأوكسجين الحيوي (BOD₅) الخارج من المحطة، هذه المدخلات هي الأس الهيدروجيني pH الداخل إلى المحطة والمواد العالقة الكلية TSS الداخلة إلى المحطة من خلال المعمارية (1-97-2). بينما كانت أفضل ثلاثة عوامل للتنبؤ بـ COD هي الأس الهيدروجيني pH الداخل إلى المحطة والمواد العالقة الكلية TSS الداخلة إلى المحطة من خلال المعمارية (1-60-2)، وكذلك الطلب على الأوكسجين الحيوي الداخل إلى المحطة والمواد العالقة الكلية TSS الداخلة إلى المحطة من خلال المعمارية (1-97-2). أما بالنسبة للمواد الصلبة المعلقة الكلية الخارجة من المحطة فكان أفضل أداء للنماذج من خلال تركيبة المدخلين الأس الهيدروجيني pH الداخل إلى المحطة والمواد العالقة الكلية TSS الداخلة إلى المحطة وفق المعمارية (1-79-2).

كانت الخطوة التالية هي تدريب نماذج من ثلاثة وأربعة وخمسة مدخلات وفق التكوينات المحتملة ومخرج واحد (BOD₅ أو COD أو TSS) الخارجة من محطة المعالجة. وتجدر الإشارة إلى أنه تم النظر في كل مجموعة ممكنة من بارامترات الإدخال الخمسة باستخدام طبقة مخفية واحدة وكذلك طبقتين مخفيتين اثنتين.

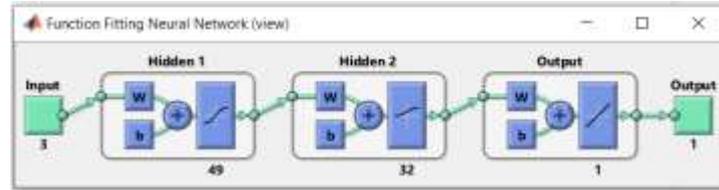
- النموذج الأول: التنبؤ بالطلب على الأوكسجين الحيوي الخارج (BOD₅ Effluent)

يوضح الشكل (4) قيم R² للتنبؤ بالطلب على الأوكسجين الحيوي الخارج من محطة المعالجة لكل مجموعة إدخال باستخدام (طبقة واحدة وكذلك طبقتين) من الطبقات المخفية. لقد وجدنا تفوق تكوين المدخلات رقم (18) في الجدول (2) وهي (pH_{inf}, TDS_{inf}, BOD₅_{inf}) على تكوينات المدخلات الأخرى في حالة وجود طبقتين مخفيتين. تظهر النتائج أيضاً أنه بالنسبة لكافة مجموعات الإدخال، تبلغ دقة التنبؤ ذروتها عند استخدام طبقتين مخفيتين.



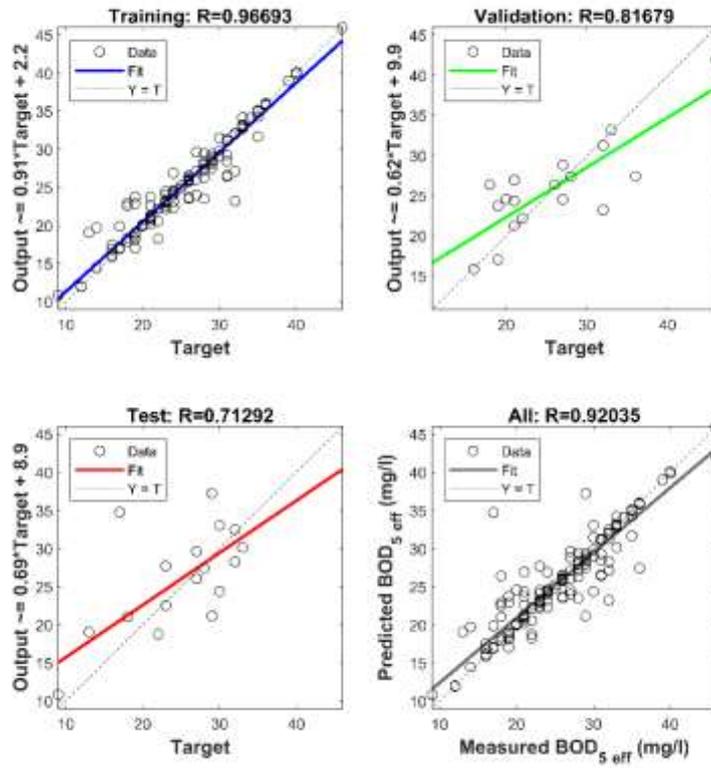
الشكل (4) دقة التنبؤ بنموذج BOD₅ الخارج لجميع مجموعات المدخلات باستخدام طبقة واحدة وطبقتين مخفيتين.

تم تدريب مجموعة الإدخال (18) على هيكليات مختلفة من الشبكات العصبونية، حققت هيكلية ANN المثلى معامل تحديد ($R^2=0.92$) ومتوسط الخطأ التربيعي ($MSE=7.812$) في حال وجود طبقتين مخفيتين مع 49 خلية عصبونية للطبقة المخفية الأولى و 32 عصبون للطبقة المخفية الثانية وفق المعمارية (3-49-32-1)، كما هو موضح في الشكل (5).

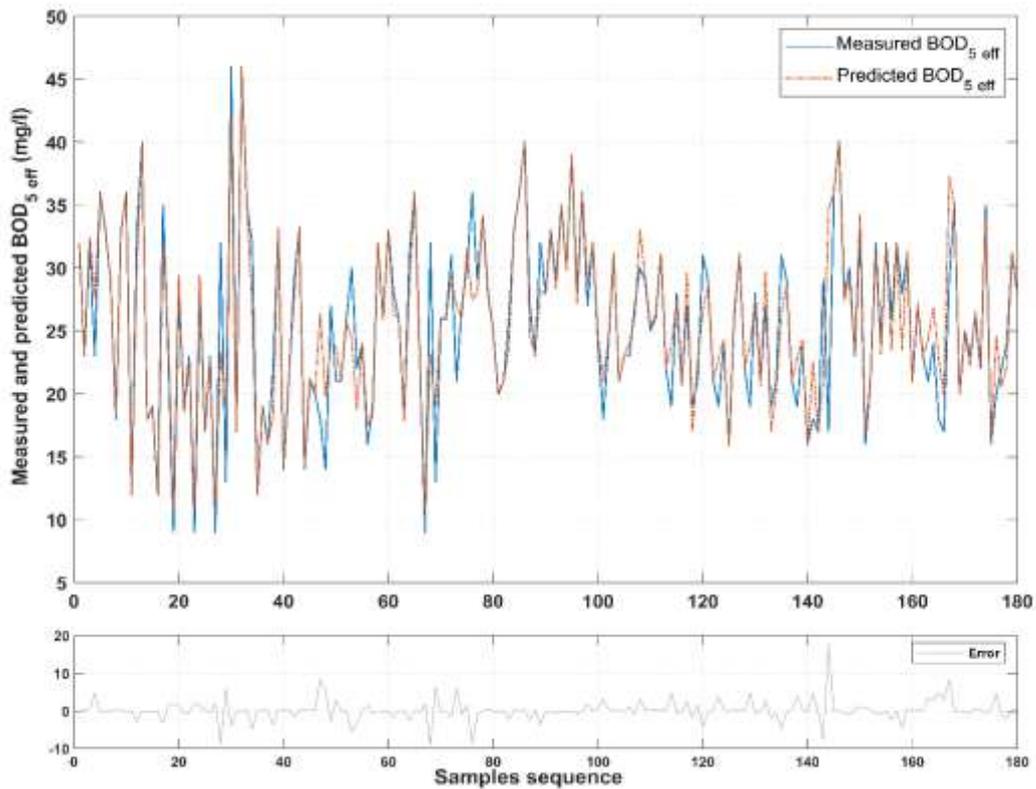


الشكل (5) هيكلية ANN المثلى للتنبؤ بالطلب على الأوكسجين الحيوي (BOD₅_eff).

يوضح الشكلان (6) و (7) أداء نموذج ANN ثلاثي المدخلات كخطط مقارنة بين القيم المقاسة والمتوقعة على مدى تسلسل البيانات، على التوالي. حيث تعتبر قيمة R^2 عالية، والقيم المحسوبة تتطابق مع القيم المقاسة بشكل وثيق باستثناء قيم معينة. قد تكون هذه القيم ناتجة عن مجموعة البيانات الشاذة نسبياً.



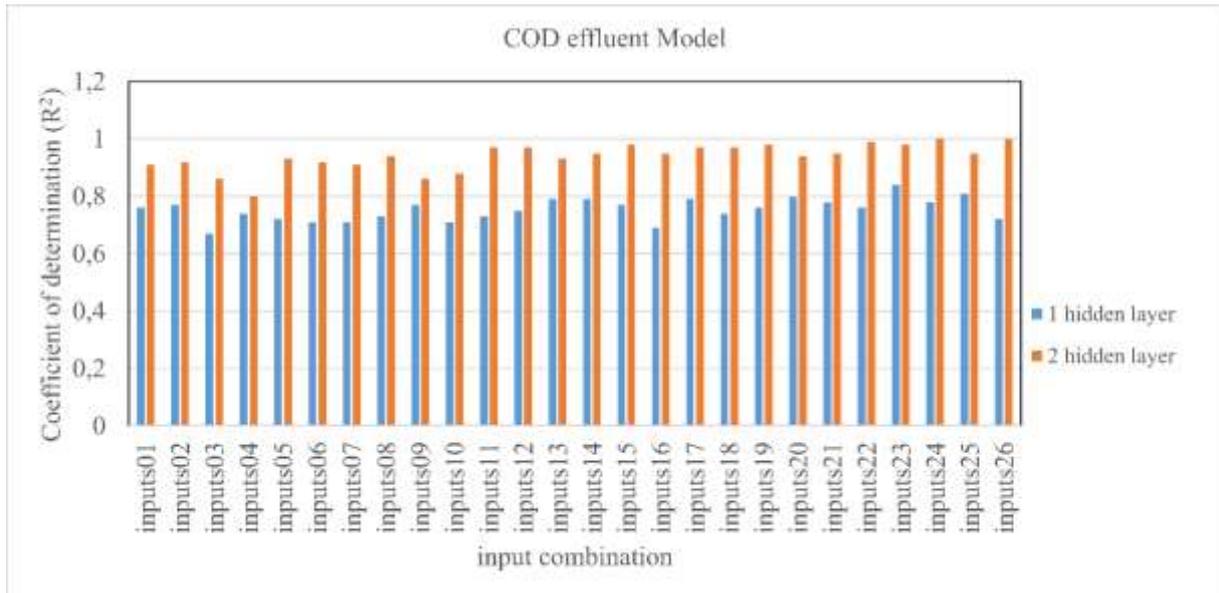
الشكل (6) مخطط تطابق القيم المقاسة مع القيم المحسوبة لنموذج التنبؤ بـ BOD_5



الشكل (7) مقارنة القيم المقاسة والقيم المحسوبة لنموذج التنبؤ بـ BOD_5

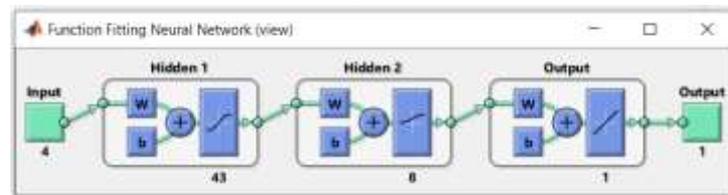
- النموذج الثاني: التنبؤ بالطلب الكيميائي للأوكسجين الخارج (COD Effluent)

تظهر نتائج الدقة لكل مجموعة إدخال في الشكل (8) للتنبؤ بالـ COD الخارج من المحطة باستخدام ANN بطبقة واحدة وكذلك طبقتين من الطبقات المخفية. نلاحظ تفوق تكوينة المدخلات (23) (TDS_inf & TSS_inf & COD_inf & BOD₅_inf) على التكوينات الأخرى في حالة طبقة مخفية واحدة، بينما تفوقت التكوينة (22) (pH_inf & TDS_inf & TSS_inf & BOD₅_inf) وكذلك التكوينة (24) و(26) على جميع تكوينات المدخلات الأخرى سواء في حالة طبقة مخفية واحدة وطبقتين مخفيتين أيضاً. ولوحظ أيضاً أن التكوينات التي تحتوي على (pH_inf, TSS_inf, TDS_inf, COD_inf) في تكوينات المدخلات (15, 19, 23, 24, 26) أدت إلى أداء تنبؤ أعلى من أداء التكوينات الأخرى.



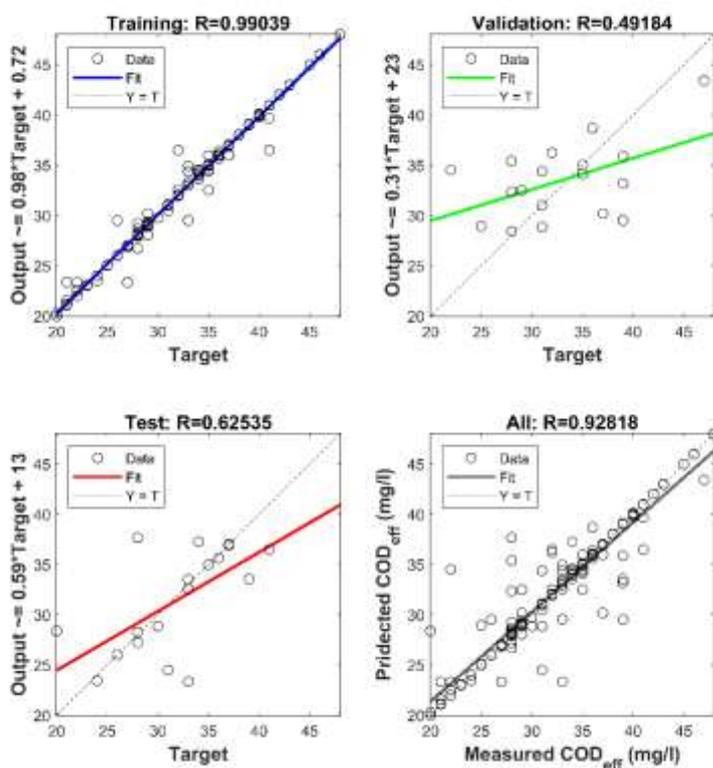
الشكل (8) دقة التنبؤ بنموذج COD لجميع مجموعات المدخلات باستخدام طبقة مخفية واحدة وكذلك طبقتين اثنتين.

تم تدريب التكوينة (22) والتي تتضمن الأس الهيدروجيني، المواد الصلبة المنحلة الكلية، المواد الصلبة المعلقة الكلية والطلب على الأوكسجين الحيوي على هيكليات الشبكات المختلفة، حققت هيكلية ANN المثلى ($R^2=0.99$) و ($MSE=5.372$) بوجود طبقتين مخفيتين تضمنت الطبقة الأولى على 43 خلية عصبونية والطبقة المخفية الثانية على 8 عصبونات وفق المعمارية (1-8-43-4)، كما هو موضح في الشكل (9).

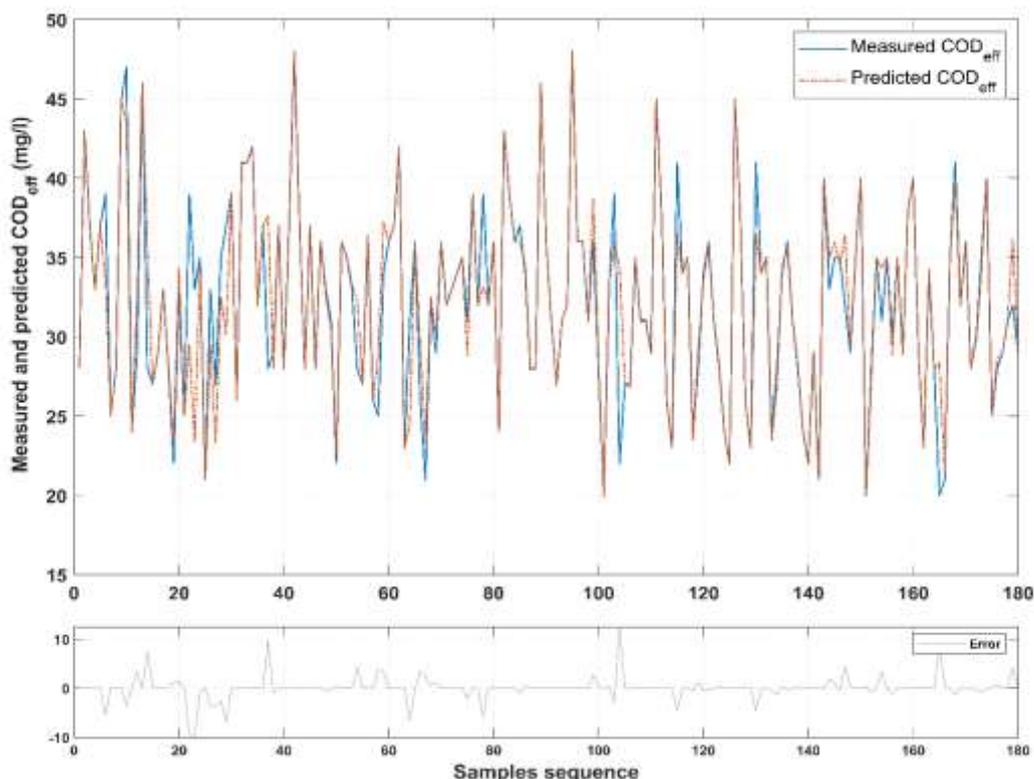


الشكل (9) هيكلية ANN المثلى للتنبؤ بـ COD

يوضح الشكلان (10) و (11) أداء نموذج ANN رباعي المدخلات كمخطط مقارنة بين القيم المقاسة والمتوقعة لكافة البيانات، على التوالي.



الشكل (10) مخطط تطابق القيم المقاسة مع القيم المحسوبة لنموذج التنبؤ بـ COD

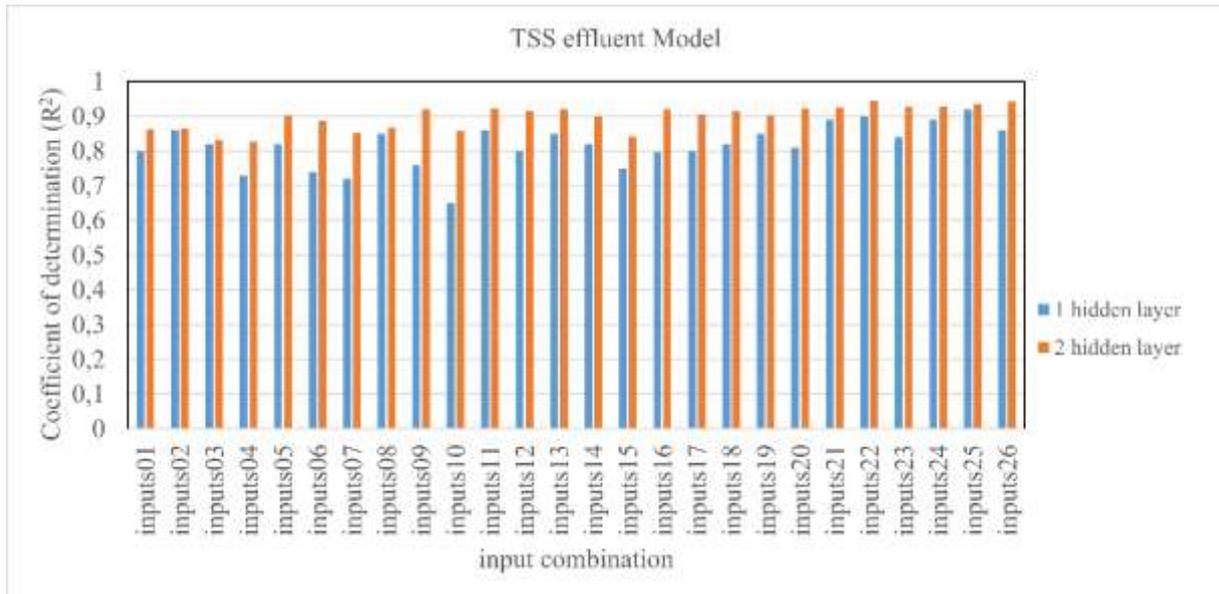


الشكل (11) مقارنة القيم المقاسة والقيم المحسوبة لنموذج التنبؤ بـ COD.

يوضح الشكل (11) أن دقة التنبؤ عالية جداً، حيث يقدم نموذج ANN أفضل أداء مع عدد قليل نسبياً من نقاط البيانات وبارامترات إدخال سهلة القياس مخبرياً ولا تحتاج إلى وقت طويل في إنجازها. بالإضافة إلى ذلك، فإن أداء ANN رباعي المدخلات الأمثل ($R^2=0.99$) يمكن مقارنته بأداء نموذج ANN المطور بثلاثة مدخلات حيث نجد أن نموذج التكوين رقم (15) مع المدخلات (TSS_inf, TDS_inf, COD_inf) قد حقق نتائج ممتازة بقيمة ($R^2=0.98$) و ($MSE=8.097$) بالإضافة إلى العدد القليل من المدخلات هي أيضاً سهلة وسريعة القياس.

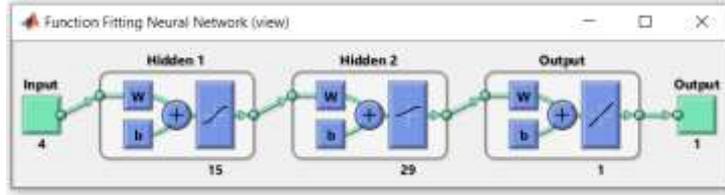
– النموذج الثالث: التنبؤ بالمواد الصلبة المعلقة الكلية الخارجة (TSS Effluent)

بالنسبة لنموذج الشبكات العصبونية الصناعية ANN بهدف التنبؤ بالمواد الصلبة المعلقة الكلية TSS الخارجة، أعطى تكوين المدخلات رقم (25) ($pH_inf, TDS_inf, COD_inf, BOD_5_inf$) في حالة طبقة مخفية واحدة نتائج أفضل من التكوينات الأخرى، كما هو موضح في الشكل (12). بالإضافة إلى ذلك، حققت تكوينة الإدخال التي تحتوي على المواد الصلبة المعلقة الكلية TSS_inf بدلاً من الـ COD_inf دقة تنبؤ عالية وقريبة من النموذج السابق. في حالة استخدام طبقتين مخفيتين للنموذج الثاني نجد أنه قد تحسن الأداء بشكل ملحوظ مقارنة بالنماذج ذات الطبقة المخفية الواحدة.

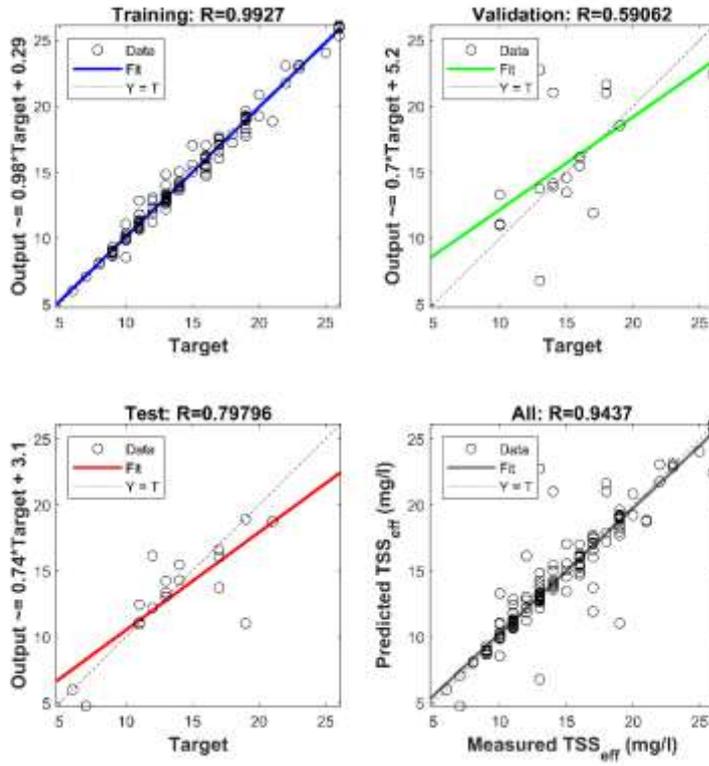


الشكل (12) دقة التنبؤ بنموذج TSS لجميع مجموعات المدخلات باستخدام طبقة مخفية واحدة وكذلك طبقتين اثنتين.

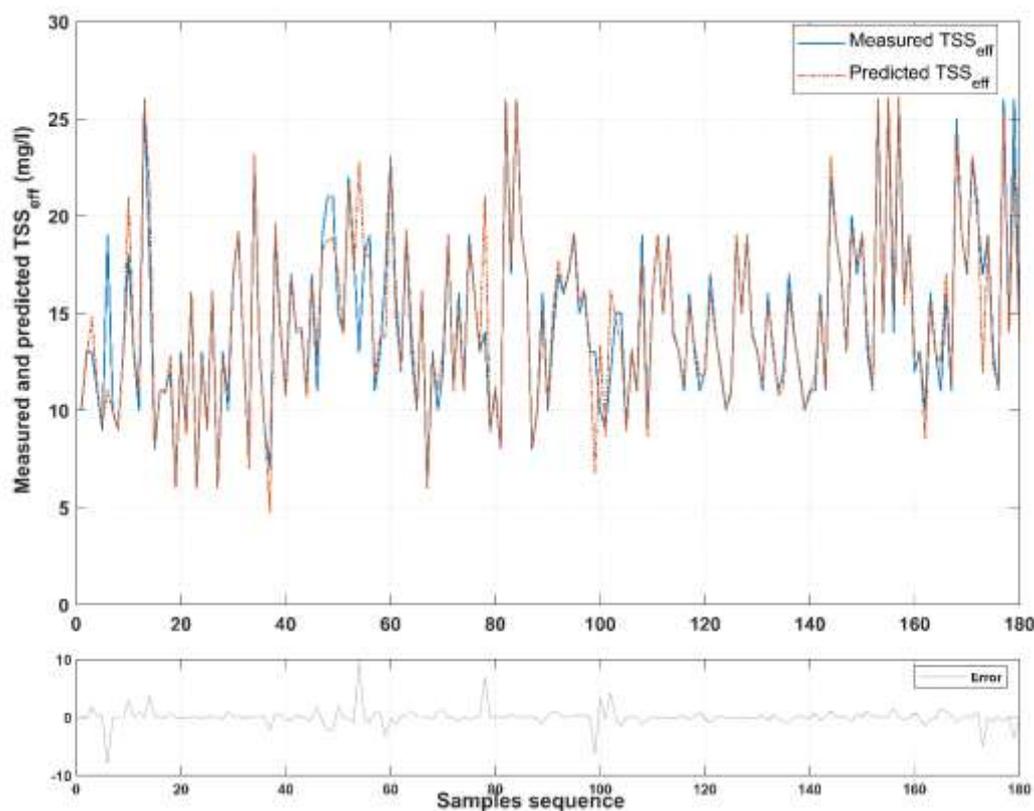
تم تدريب التكوين رقم (22) ($pH_inf, TDS_inf, TSS_inf, BOD_5_inf$) على هيكليات الشبكات المختلفة، وحققت بنية ANN المثلى ($R^2=0.944, MSE=2.303$) والمكونة من طبقتين مخفيتين حيث تحتوي الأولى على 15 خلية عصبونية والثانية على 29 خلية عصبونية (1-29-15-4)، كما هو موضح في الشكل (13). يوضح الشكلان (14) و (15) أداء نموذج ANN رباعي المدخلات كمخطط للمقارنة بين القيم المقاسة والمتوقعة لكافة البيانات، على التوالي. تمكنت شبكة ANN المطورة من التنبؤ بـ TSS بدقة كافية فيما يتعلق بالقيم المقاسة، كما هو الحال مع تنبؤات نموذجي الـ BOD₅ و COD.



الشكل (13) هيكلية ANN المثلى للتنبؤ بـ TSS .



الشكل (14) مخطط تطابق القيم المقاسة مع القيم المحسوبة لنموذج التنبؤ بـ TSS .

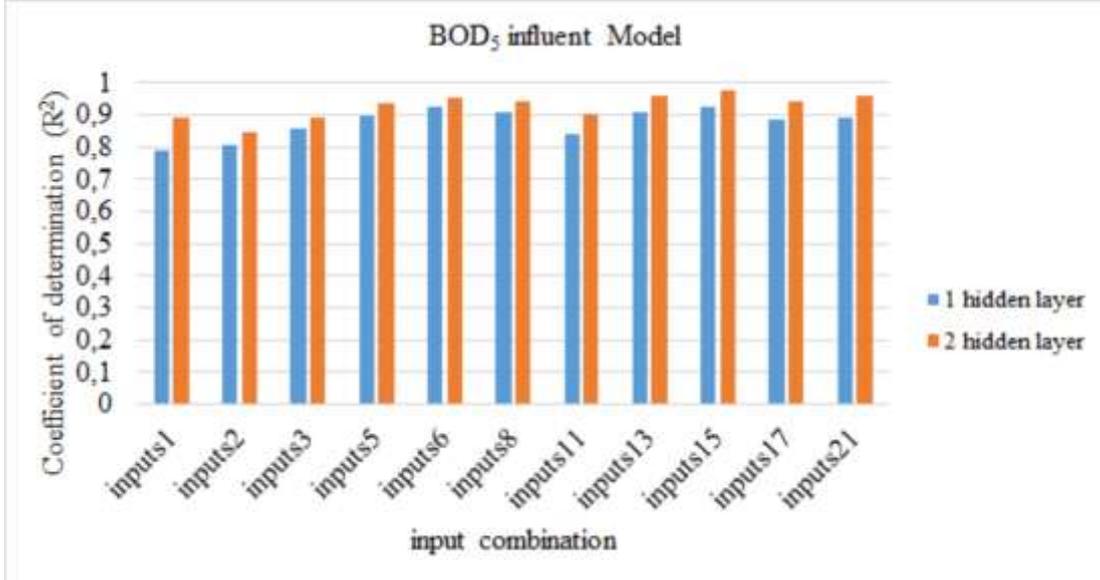


الشكل (15) مقارنة القيم المقاسة والقيم المحسوبة لنموذج التنبؤ بـ TSS.

بالإضافة إلى ذلك، تم تدريب خمسة مدخلات من الشبكات العصبونية الصناعية (pH_inf, TDS_inf, TSS_inf, COD_inf, BOD₅_inf) للتنبؤ بـ TSS_inf؛ ومع ذلك، لم يطرأ أي تحسن في الأداء (MSE = 0.942, R² = 2.307) وفق الهيكلية (1-23-39-5)، الواضح هنا أن زيادة عدد المدخلات ليست دائماً ذات فائدة [19, 20].

– النموذج الرابع: التنبؤ بالطلب الحيوي للأوكسجين الداخل (BOD₅ influent)

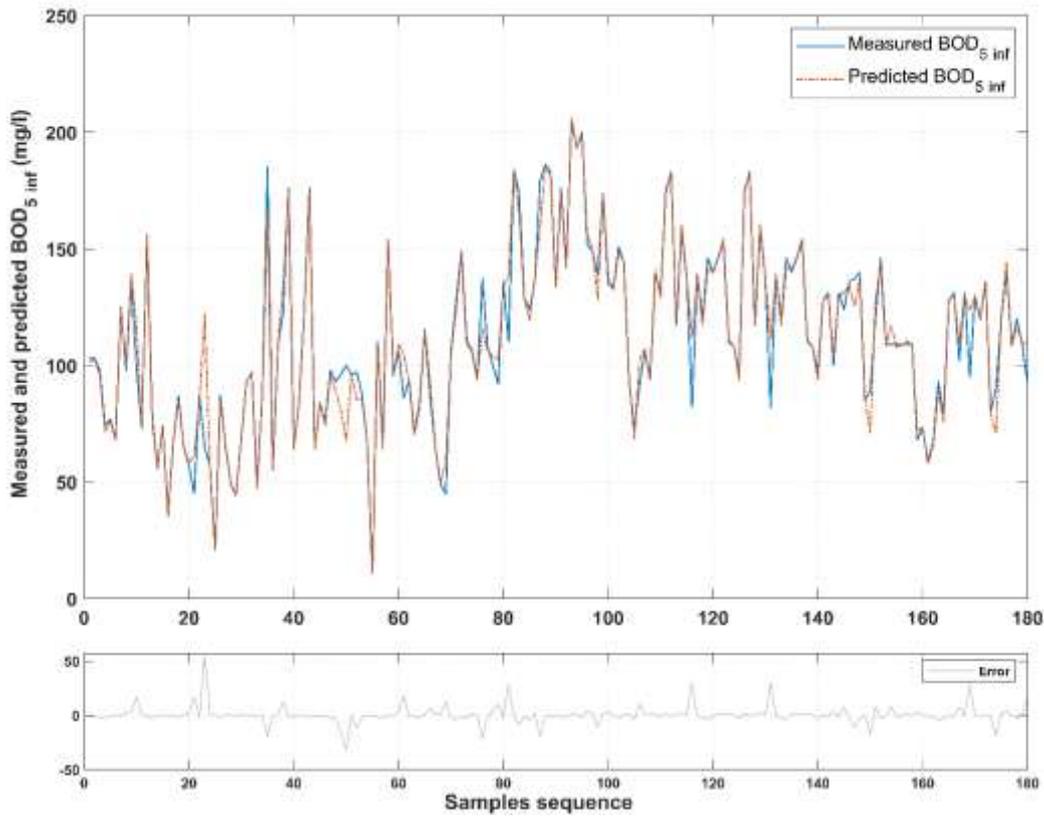
تم تطوير واختبار العديد من تكوينات ANN (مجموعات المدخلات المختلفة وعدد الطبقات المخفية وعدد الخلايا العصبونية في تلك الطبقات) للتنبؤ بتركيز BOD₅_inf الداخل إلى محطة معالجة مياه الصرف الصحي. يوضح الشكل (16) أفضل نتائج التنبؤ لمجموعة بيانات الاختبار لكل تكوينة ANN فيما يتعلق بمعامل التحديد (R²).



الشكل (16) دقة التنبؤ بنموذج BOD₅_inf لأفضل مجموعات المدخلات باستخدام طبقة مخفية واحدة وكذلك طبقتين اثنتين.

بالإضافة إلى ذلك، تم إجراء تحليل الحساسية لبارامترات الإدخال باستخدام دقة التنبؤ ANN في ظل مجموعات الإدخال المختلفة. يلخص الجدول (5) أفضل النتائج التي تم اختيارها (مرتبة بشكل تنازلي) والتي تم الحصول عليها وفقاً لكل عدد من المدخلات وتشكيل نماذج ANN. يشير الجدول إلى أن تركيز COD يفرض أكبر تأثير على دقة التنبؤ ANN، متبوعاً بالمواد الصلبة المنحلة الكلية، والمواد الصلبة المعلقة الكلية، والرقم الهيدروجيني في المرتبة الأخيرة. وبالتالي، لكي يتنبأ ANN بشكل كافٍ بـ BOD₅_inf الداخل، يلزم على الأقل معرفة قياسات تراكيز COD_inf والمواد الصلبة المنحلة الكلية TDS_inf وكذلك المواد الصلبة المعلقة الكلية TSS_inf كبارامترات إدخال. ومنه نجد، إن تقليل الوقت اللازم لتقدير BOD₅ من خمسة أيام إلى عدة ساعات هو الوقت المطلوب فقط لقياس تركيز COD لأي عينة من التكوينات المختلفة المدرجة في الجدول (2)، تم اعتماد النموذج ANN وفق المعمارية (3-35-27-1) للتنبؤ بـ BOD₅_inf، وتظهر النتائج في الشكل (17) حيث حقق النموذج أداءً عالياً ممثلاً بـ (R²=0.9744, MSE= 71.245). توضح النتائج الدقة العالية لشبكة ANN في التنبؤ بـ BOD₅_inf الداخل باستثناء عدد قليل جداً من نقاط البيانات. الجدول (5) تكوينات ANN بأعلى دقة تنبؤية لـ BOD₅_inf الداخل إلى المحطة للتكوينات المختلفة من بارامترات الإدخال.

النموذج	بارامترات الإدخال	هيكلية الشبكة	معامل التحديد R ²
M01	Input 15: TSS_inf, TDS_inf, COD_inf	(3-35-27-1)	0.97437
M02	Input 21: pH_inf, TDS_inf, TSS_inf, COD_inf	(4-12-41-1)	0.95957
M03	Input 13: TSS_inf, pH_inf, COD_inf	(3-36-47-1)	0.9591
M04	Input 06: TDS_inf, COD_inf	(2-49-34-1)	0.95283
M05	Input 08: TSS_inf, COD_inf	(2-35-33-1)	0.94144
M06	Input 17: pH_inf & TDS_inf & COD_inf	(3-49-31-1)	0.94111
M07	Input 05: TDS_inf, TSS_inf	(2-44-27-1)	0.93979
M08	Input 06: TDS_inf, COD_inf	(2-43-1)	0.92792
M09	Input 15: TSS_inf, TDS_inf, COD_inf	(3-23-1)	0.92405
M10	Input 13: TSS_inf, pH_inf, COD_inf	(3-41-1)	0.90893
M11	Input 08: TSS_inf, COD_inf	(2-73-1)	0.90757



الشكل (17) مقارنة القيم المقاسة والقيم المحسوبة لنموذج التنبؤ بالطلب على الأوكسجين الحيوي الداخل إلى المحطة BOD_5_{inf} (M01).

الاستنتاجات والتوصيات:

تنبأت نماذج ANN المطورة بأداء محطة معالجة مياه الصرف الصحي لقرية الرويمية والتي تم تحديدها بناءً على تراكيز المواد الصلبة العالقة الكلية (TSS)، والطلب على الأوكسجين الحيوي (BOD_5)، والطلب الكيميائي للأوكسجين (COD) الداخلة إلى محطة المعالجة بدرجة عالية من الموثوقية. أظهرت النتائج أن التحليل الأولي للبيانات وتحضيرها ضروريان لتدريب ANN، حيث يوصى بزيادة المتغيرات تدريجياً (عدد المدخلات، الطبقات المخفية، والخلايا العصبونية) لتكوين الشبكة المثلى حيث لا يتم فيها الوصول إلى أي تحسن إضافي. في هذه الدراسة، كان تكوين ثلاثة مدخلات ومخرج واحد كافياً، كما أشارت النتائج إلى أن زيادة عدد المدخلات لم تكن مفيدة دائماً. نستنتج من النماذج المطورة أهمية بارامترات المدخلات في عملية معالجة مياه الصرف الصحي، حيث يمكن تفسير أهمية البارامتر من خلال تكوينه للإدخال للنموذج والتي تؤدي إلى أعلى دقة للتنبؤ. في هذه الدراسة، يعمل نموذج التنبؤ ANN المطور لتركيز BOD_5_{inf} الداخل إلى المحطة بشكل جيد ويحقق دقة أداء عالية ($R^2 = 0.9744$). توضح هذه النتيجة جدوى وأهمية استخدام هذا النموذج كأداة سهلة وهامة لأنظمة التحكم والإدارة في محطات معالجة مياه الصرف الصحي. هناك توصية أخرى تتمثل في تنفيذ نظام تحكم أكثر دقة يستلزم قياسات متكررة عديدة لمزيد من ضبط النموذج ببيانات تمثيلية بدرجة عالية لظروف التشغيل في الوقت الفعلي.

References:

1. Hamed, M.M., M.G. Khalafallah, and E.A. Hassanien, *Prediction of wastewater treatment plant performance using artificial neural networks*. Environmental Modelling & Software, 2004. **19**(10): p. 919-928.
2. Ferrer, J., et al., *DESASS: a software tool for designing, simulating and optimising WWTPs*. Environmental Modelling & Software, 2008. **23**(1): p. 19-26.
3. Batstone, D.J., et al., *The IWA anaerobic digestion model no 1 (ADM1)*. Water Science and technology, 2002. **45**(10): p. 65-73.
4. Henze, M., et al., *Activated sludge model no. 2d, ASM2d*. Water science and technology, 1999. **39**(1): p. 165-182.
5. Alver, A. and Z. Kazan, *Prediction of full-scale filtration plant performance using artificial neural networks based on principal component analysis*. Separation and Purification Technology, 2020. **230**: p. 115868.
6. Newhart, K.B., et al., *Data-driven performance analyses of wastewater treatment plants: A review*. Water research, 2019. **157**: p. 498-513.
7. Mjalli, F.S., S. Al-Asheh, and H. Alfadala, *Use of artificial neural network black-box modeling for the prediction of wastewater treatment plants performance*. Journal of Environmental Management, 2007. **83**(3): p. 329-338.
8. Nasr, M.S., et al., *Application of Artificial Neural Network (ANN) for the prediction of EL-AGAMY wastewater treatment plant performance-EGYPT*. Alexandria engineering journal, 2012. **51**(1): p. 37-43.
9. Guo, H., et al., *Prediction of effluent concentration in a wastewater treatment plant using machine learning models*. Journal of Environmental Sciences, 2015. **32**: p. 90-101.
10. Nezhad, M.F., et al., *Artificial neural network modeling of the effluent quality index for municipal wastewater treatment plants using quality variables: south of Tehran wastewater treatment plant*. Journal of Water Supply: Research and Technology-Aqua, 2016. **65**(1): p. 18.
11. Abba, S.I. and G. Elkiran, *Effluent prediction of chemical oxygen demand from the astewater treatment plant using artificial neural network application*. Procedia Computer Science, 2017. **120**: p. 156-163.
12. Kasem, R., et al., *A software sensor for in-situ monitoring of the 5-day biochemical oxygen demand*. Rudarsko-geološko-naftni zbornik (The Mining-Geological-Petroleum Engineering Bulletin), 2018. **33**(1): p. 15-23.
13. Najafzadeh, M. and A. Ghaemi, *Prediction of the five-day biochemical oxygen demand and chemical oxygen demand in natural streams using machine learning methods*. Environmental monitoring and assessment, 2019. **191**(6): p. 1-21.
14. Baki, O., et al., *Biochemical oxygen demand prediction in wastewater treatment plant by using different regression analysis models*. Desalination and Water Treatment, 2019. **157**.
15. Qiu, Y., Y. Liu, and D. Huang, *Date-driven soft-sensor design for biological wastewater treatment using deep neural networks and genetic algorithms*. Journal of chemical engineering of Japan, 2016. **49**(10): p. 925-936.
16. Gangi Setti, S. and R. Rao, *Artificial neural network approach for prediction of stress-strain curve of near β titanium alloy*. Rare Metals, 2014. **33**(3): p. 249-257.
17. Haykin, S., *Neural networks and learning machines, 3/E*. 2009: Pearson Education India.
18. Geman, S., E. Bienenstock, and R. Doursat, *Neural networks and the bias/variance dilemma*. Neural computation, 1992. **4**(1): p. 1-58.

19. Hamada, M., H. Adel Zaqoot, and A. Abu Jreiban, *Application of artificial neural networks for the prediction of Gaza wastewater treatment plant performance-Gaza strip*. Journal of Applied Research in Water and Wastewater, 2018. **5**(1): p. 399-406.
20. Zare, A., V. Bayat, and A. Daneshkare, *Forecasting nitrate concentration in groundwater using artificial neural network and linear regression models*. International agrophysics, 2011. **25**(2).