

استخدام شبكة الارتباط الذاتي الغير خطي العصبية في تقدير التبخر اليومي: ومقارنتها مع شبكات عصبية بخوارزميات مختلفة

شهلة عبدالوهاب عبدالقادر ، نجم عبدالله عبدالقادر ، كفاء هادي ذنون

قسم أنظمة الحاسوب ، المعهد التقني ، الموصل ، العراق

الملخص:

تم في هذا البحث تطبيق نموذج الشبكة العصبية الديناميكية اللاخطية Dynamic NN المتمثلة بشبكة Nonlinear Autoregressive (NARX) لتقدير التبخر اليومي لمدينة الموصل بدلالة بعض المعلمات المناخية (درجة الحرارة العظمى والصغرى، الامطار، الرطوبة النسبية، سرعة الرياح والاشعاع الشمسي) لاي يوم من السنة ومقارنتها مع الشبكات العصبية الساكنة Static NN: مثل شبكة الانتشار العكسي للخطا Feed Forward Back Propagation (FFBPNN)، وشبكة التتابع (شبكة انتشار عكسي بروابط اضافية) Cascade-Forward Back Propagation (CFBPNN)، ولكل شبكة معماريتين: معمارية ب(4) طبقات و(5) خلايا في الطبقات الخفية، ومعمارية ب(5) طبقات و(5) خلايا في الطبقات الخفية.

تم استخدام خوارزميات مختلفة للتدريب كخوارزمية ليفن بيرك ماركوت (LM)، خوارزمية نيوتن (BFGS)، خوارزمية الميل الصرفية (CFG)، خوارزمية الانحدار التدريجي (GD) وخوارزمية الانحدار التدريجي المعجل (GDM). استخدمت البيانات المأخوذة من محطة الانواء الجوية في منطقة الرشيدية / محافظة نينوى للفترة من (1995-2008) حيث استخدمت بيانات (10) اعوام لتطوير النماذج وللفترة من (1995-2004) وبيانات اربعة اعوام لتقييم النماذج ومقارنة مخرجاتها مع البيانات المقاسة وللفترة من (2005-2008)، كما استخدمت مقاييس معامل التحديد (R^2 determination coefficient)، جذر متوسط مربع الاخطاء (Root Mean Square Error RMSE) في الحكم على مدى تطابق البيانات المقاسة ومخرجات الشبكات العصبية لاختيار افضل نموذج تنبؤي من بين النماذج المطبقة، بينت النتائج ان شبكة NARX وبخوارزمية (LM) ذات كفاءة في تحسين نموذج تنبؤي لتقدير التبخر اليومي حيث وصلت قيمة معامل التحديد الى (0.99) وهي افضل واسرع خوارزمية اذا ما توفرت درجات حرارة الامطار، الرطوبة النسبية، سرعة الرياح والاشعاع الشمسي لاي يوم من ايام السنة.

الكلمات المفتاحية: الشبكات الساكنة، الشبكات الديناميكية، الانحدار التدريجي، الانحدار التدريجي المعجل، الميل الصرفية، خوارزمية نيوتن، خوارزمية ليفن بيرك ماركوت.

1- المقدمة

تشكل عملية التبخر احدى المركبات المهمة للدورة الهيدرولوجية للمياه في الطبيعة وعلى هذا الاساس اهتم الكثير من الباحثين في دراسة هذه العملية وذلك بتخمينها بطرائق مختلفة منها استخدام الأحواض ومعاملات الأنواء الجوية وهي الأكثر شيوعاً، وأخرى تجريبية التي معظمها تم استنتاجها بالطرائق الإحصائية التقليدية، وحيث ان معاملاتها الثابتة لمناطق محددة تجعل تطبيقها على مناطق أخرى مصحوباً بالخطأ بالإضافة الى عدم إمكانية استخدامها لتقدير التبخر اليومي [9][8]، وكان التوجه إلى استخدام الشبكات العصبية الاصطناعية لتقدير التبخر اليومي لأي منطقة وبدقة عالية إذا ما توفرت المعلمات المناخية من درجات الحرارة، الرطوبة، السطوع الشمسي، الامطار وسرعة الرياح [10].

التجارب في السنوات الاخيرة اثبتت كفاءة الشبكات العصبية الاصطناعية في بناء النماذج لتقدير التبخر للحصول على قياسات متوقعة ذات دقة عالية مع القياسات الحقيقية، فقد استخدم الباحثون [9] درجات الحرارة، الرطوبة، سرعة الرياح والسطوع الشمسي لقياس التبخر باستخدام شبكة التغذية الامامية (FFBPNN) وبمعماريات مختلفة ووصل الى معامل تحديد مساو 0.72، ومربع خطأ مساو 0.07، اما الباحثون [11] [10] فقد استخدموا شبكة البيروسترون

في العقود القليلة الأخيرة طورت العديد من نماذج التنبؤ اللاخطية مثل أنظمة الذكاء الاصطناعي، الذي يمكن ان يلائم البيانات وبشكل أفضل من النماذج الخطية التي غالباً ما يكون التنبؤ بها ضعيف. أنظمة الذكاء الاصطناعي ذات تقنيات واسعة وتوفر طرق بديلة لمعالجة المشاكل المعقدة وهو قادر على معالجة البيانات غير المتكاملة والتي تحتوي على ضوضاء وكذلك حل المشاكل اللاخطية التنبؤية بسرعة عالية. استخدمت هذه التقنيات في تطبيقات متنوعة مثل الإنسان الآلي، تمييز الأنماط، التنبؤ، الطب، الأنظمة الكهربائية، التصنيع، الامتلية ومعالجة الإشارة [2][1].

أن أنظمة الذكاء الاصطناعي تشمل الأنظمة الخبيرة، الشبكات العصبية الاصطناعية [1]، الخوارزميات الجينية، المنطق المضبط، وأنظمة هجينة أخرى تجمع بين واحد أو أكثر من التقنيات أعلاه [2]. الشبكات العصبية الاصطناعية (ANN) تعتبر من أنظمة الذكاء الاصطناعي، التي تعالج بيانات هائلة بشكل متوازي لتدريبها، كسابها معرفة ملائمة لغرض التطبيق، والفائدة الأكبر قدرة الشبكات العصبية على تشكيل العلاقات اللاخطية المعقدة بدون فرضيات استنتاجية عن طبيعة العلاقة [7][6][5][4][3].

والمنطق المضطرب لتقدير التبخر اليومي وتوصلوا الى معامل تحديد لم يتجاوز 0.97 .

الهدف من هذا البحث هو بناء نماذج من الشبكات العصبية المختلفة بخوارزميات متعددة لتقدير التبخر اليومي لمدينة الموصل بدلالة بعض المعلومات المناخية سهلة القياس. ما تبقى من البحث كما يلي :القسم الثاني يتضمن تفاصيل الطريقة المستخدمة في البحث. القسم الثالث يشرح المواد ومنهجية البحث المقترحة. القسم الرابع النتائج ومناقشتها . القسم الخامس الاستنتاجات.

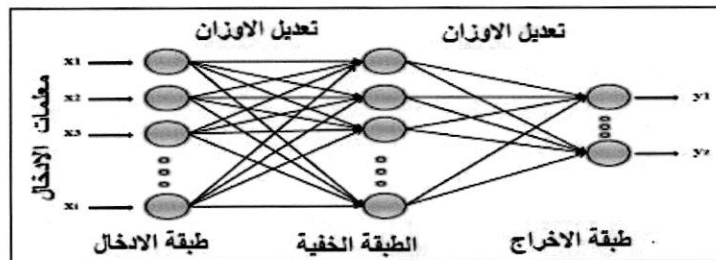
2- الشبكات العصبية

الشبكات العصبية تمتلك فائدة رئيسية هي قابليتها للاخطية المرنة لصياغة النماذج وليس هناك حاجة لتحديد نموذج قياسي ثابت، بالاحرى النموذج يتكيف حسب ميزات البيانات. قادت هذه النظرة الى صياغة العديد من البحوث التجريبية وبناء النماذج لمعالجة البيانات بانواع أكثر شيوعا من الشبكات العصبية مثل الشبكات الساكنة SNN ، والشبكات الديناميكية DNN [20].

2-1 الشبكات العصبية الساكنة

الشبكات العصبية الساكنة SNN ذات التغذية الامامية هي نوع من الشبكات التي لا تحتوي على روابط رجوع بصورة عكسية ولا تحتوي على عناصر تاخير Delay، الاخراج يحسب مباشرة من الادخال باتجاه تغذية امامية من خلال روابط الشبكة منها شبكة الانتشار العكسي او ما تسمى الشبكة ذات التغذية الامامية (FFBPNN) يوضحها الشكل رقم (1)، وشبكة التتابع (CFBPNN) يوضحها الشكل رقم (2). شبكة الانتشار العكسي استعملت على نحو واسع في الكثير من المجالات، ويمرور الوقت استعملت في السلاسل الزمنية التنبؤية لانها قادرة على حل انواع عديدة ومختلفة من المشاكل.

شبكة الانتشار العكسي بتغذية امامية (FFBPNN) مكونة عادة من عدة طبقات : طبقة الادخال، طبقة الاخراج واحدة او اكثر من الطبقات الخفية كما موضح في الشكل رقم (1). اما شبكة التتابع (CFBPNN) فانها شبكة ذات انتشار عكسي بروابط اضافية من المدخلات باتجاه الطبقات الخفية وطبقات الاخراج لتسريع عملية التدريب يوضحها الشكل رقم (2) [21].

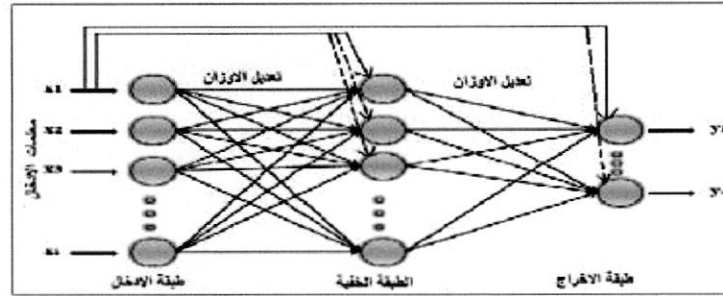


الشكل رقم (1): شبكة الانتشار العكسي المثالية

المتعددة الطبقات MultiLayers Perceptron (MLP) بخوارزمية التغذية العكسية (BP) لتقدير التبخر باستخدام درجات الحرارة العظمى والصغرى والاشعاع الشمسي وتوصلوا الى معامل تحديد لم يتجاوز 0.91 ، في حين قام الباحثين [12] باستخدام الشبكة العصبية ذات التغذية العكسية (BBNN) لتقدير التبخر اليومي من خلال بعض المعلومات المناخية من درجات حرارة العظمى والصغرى، وقام الباحثون [13] باستخدام شبكة البيروسترون (MLP) لتقدير التبخر اليومي باستخدام عدة معاملات من درجة حرارة الهواء، سرعة الرياح، كمية الامطار والاشعاع الشمسي وتوصلوا الى معامل التحديد 0.95. اما الباحثون [14] قاموا باستخدام الشبكات العصبية الاصطناعية لتقدير التبخر اليومي لمدينة لوس انجلوس وتوصلوا الى معامل تحديد 0.95. وتوصل الباحث [15] الى معامل تحديد مساو لـ 0.96 من خلال استخدام شبكة التغذية الامامية (FFBPNN) لتقدير التبخر الشهري لبعض المعلومات المناخية. اما الباحثين [16] فقد استخدموا شبكة البيروسترون المتعددة الطبقات (MLP) بخوارزمية الانتشار العكسي للخطأ (BP) وشبكة التتابع (CCBPNN) بخوارزمية كالمن لتقدير التبخر في اليونان باستخدام درجات الحرارة العظمى والصغرى وظهرت النتائج ان شبكة التتابع (CCBPNN) افضل من شبكة التغذية الامامية (FFBPNN) حيث توصلوا الى معامل تحديد مساو لـ 0.95. اما الباحثون

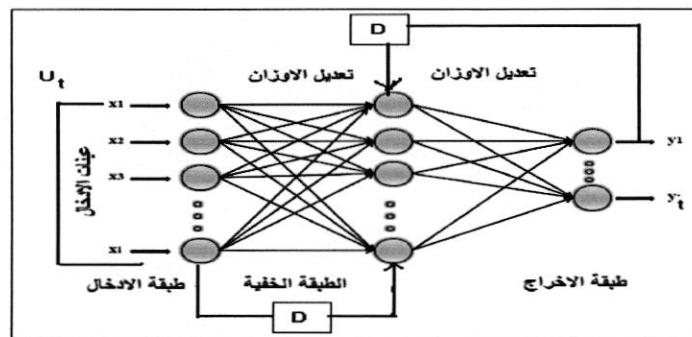
[1] قاموا باستخدام شبكة الانتشار العكسي بخوارزميات مختلفة لتقدير التبخر اليومي لمدينة الموصل/شمال العراق باعتماد معلومات مناخية يومية لمدة (21) سنة مؤثرة على التبخر من درجات حرارة، الاشعاع الشمسي، سرعة الرياح والرطوبة النسبية وتوصلوا الى معامل تحديد 0.95. في حين قام الباحث [2] بمقارنة الشبكات العصبية الاصطناعية مع نظام الاستدلال المضطرب (Fuzzy Logic) لتقدير التبخر اليومي لمدينة الموصل باعتماد معلومات مناخية لمدة (21) سنة من درجات حرارة، الاشعاع الشمسي، سرعة الرياح والرطوبة النسبية وتوصل الى معامل تحديد 0.91. وقام الباحثون [17][18][19][3] باستخدام انواع مختلفة من الشبكات العصبية لتقدير التبخر ولم يتجاوز معامل التحديد في تقديراتهم لـ 0.97 ، لكن الباحثين [8][4] قاموا باستخدام نظام هجين من الشبكات العصبية

وطبقة الاخراج بدالة خطية (Purline AF). شبكة الانتشار العكسي تمر بمرحلتين : مرحلة التغذية الامامية التي تدخل فيها انماط التدريب (البيانات) الى الشبكة وترسل خلال الطبقات المتعاقبة وصولاً الى الاخراج، ومرحلة التدريب او ماتسمى (التعليم) تعدل من خلالها الاوزان او البارامترات بشكل تكراري لكل انماط التدريب (ادخال - اخراج) على اساس تقليل الاخطاء بين مخرجات الشبكة والمخرجات الحقيقية بانحدار تدريجي للأخطاء Gradient Descent (GD) أو ماتسمى نشر الأخطاء العكسي بواسطة خوارزمية التدريب [21][5].



الشكل رقم (2): شبكة CFBPN

الشبكات الديناميكية تدرب بنفس الخوارزمية BP المستخدمة في الشبكات الساكنة، إلا أن أداءها مختلف تماماً من خلال الانحدار التدريجي للخطأ الذي يكون على نحو أكثر تعقيداً، والأوزان في هذا النوع من الشبكات لها تأثيران مختلفان على ناتج التدريب: الأول تأثير مباشر للأوزان والذي يسبب تغيير فوري في الناتج للخطوة الحالية (هذا الناتج يحسب باستخدام BP القياسية). التأثير الثاني غير مباشر وذلك من خلال بعض الإدخالات مثل $D(t-1)$ التي تؤثر في حساب الأوزان بصورة غير مباشرة باستعمال خوارزمية BP الديناميكية لحساب الانحدار التدريجي للخطأ وصولاً الى الحدود الدنيا المحلية للخطأ (Local Minim error)، لذلك سطوح الأخطاء الديناميكية تكون أكثر تعقيداً وتحتاج الى وقت أطول للتدريب للحصول على المطابقة للأخراج [24][23][21].



الشكل رقم (3): شبكة NARX

المدخلات في الشبكة لها وزن w ، تجمع المدخلات الموزونة مع الانحياز b ، بعدها تطبق دالة التنشيط Activation Function (AF). الخلايا العصبية يمكن ان تستعمل انواع من دوال التنشيط لتوليد الاخراج وبشكل عام دوال التنشيط اللاخطية تعتبر ملائمة لأكثر التطبيقات مثل الدالة اللوغارتمية (logsigmoid) والدالة الظلية (Tansigmoid)، والشبكات ذات التغذية الامامية التي تمتلك طبقة (طبقات) خفية في اغلب الاحيان تكون طبقاتها الخفية بدالة لاخطية (Sigmoid AF)

2-2 الشبكات العصبية الديناميكية:

الشبكات العصبية الديناميكية DNN هي نوع من الشبكات التي لا تعتمد فقط على الإدخال الحالي للشبكة بل تعتمد على الإدخالات أو الاخراجات السابقة أو اي حالات اخرى للشبكة. وبصورة عامة الشبكات الديناميكية أقوى من الشبكات الساكنة (على الرغم من صعوبتها في التدريب) وذلك لانها تمتلك ذاكرة تستطيع من خلالها أن تدرب الانماط التسلسلية أو أنماط الوقت، شبكة (NARX) هي نوع من الشبكات الديناميكية بارتباطات عكسية وممكن ان تحتوي على عدة طبقات خفية في هيكليتها وهي تعتمد بالاساس على نموذج الارتباط الذاتي الاحصائي (Auto Regressive ARX) الذي يستخدم في السلاسل الزمنية حيث أن القيمة التالية تستخرج من الاخراج الحقيقي بدل من الرجوع العكسي للأخراج المحسوب وكما موضح في الشكل رقم (3)، أي أن الاخراج $Y(t)$ يدخل في الحسابات الجديدة للقيمة التالية التي تعتبر ادخال للشبكة [21][7][4].

3-3 خوارزميات التدريب

$$\beta_k = \frac{g_k^T g_k}{g_k^T g_{k-1}} \dots\dots\dots(5)$$

• خوارزمية نيوتن (BFGS) Quasi-Newton :

طريقة نيوتن غالبا ماتكون اسرع من GD ولكن حساباتها مكلفة للغاية نظرا لحسابات واسعة لمصفوفة (Hessian) الاتية من المشتقات الثنائية (second derivatives) الاسلوب الاكثر نجاحا في الدراسات هو خوارزمية BFGS المحدثة والتي تتغير الاوزان بها حسب المعادلة (6) : [21][22]

$$W_{k+1} = W_k - H^{-1} g_k \dots\dots\dots(6)$$

حيث ان :

Hessian matrix (second derivatives) : H_k

• خوارزمية ليفن بيرك ماركوت Levenberg-Marquardt (LM) :

مشابهة لطريقة نيوتن ، لكنه صمم على نهج سرعة التدريب بدون الحاجة لحساب مصفوفة Hessian في ظل فرضية ان دالة الخطا هي مجموع المربعات، وان مصفوفة Hessian تكون بصورة تقريبية بالطريقة : $H = J^T J$ ، ويمكن حساب الميل على النحو الاتي $g = J^T e$ حيث ان J هي مصفوفة (Jacobian) التي تحتوي على الاخطاء التي تتعلق باوزان الشبكة ، مصفوفة Jacobian هي اقل حساب وكلفة من مصفوفة Hessian، e هي متجه الاخطاء للشبكة لذا الاوزان في الخوارزمية تتعدل حسب المعادلة (7) :

$$W_{k+1} = W_k - [JJ^T + \mu I]^{-1} W^T e \dots\dots\dots(7)$$

حيث ان : μ هو ثابت للسيطرة على سلوك الخوارزمية ، فعندما $\mu=0$ تتبع الخوارزمية طريقة نيوتن وتستعمل مصفوفة Hessian ، لكن عندما μ ذات مستوى عالي يصبح هبوط الميل عالي لحجم خطوات صغير [21][22].

3- مواد البحث وطرائقه

تم استخدام البيانات المأخوذة من محطة الانواء الجوية في منطقة الرشيدية / محافظة نينوى للفترة من 1995-2008 وحسب الجدول رقم (1) (درجة الحرارة العظمى والصغرى ، الامطار ، الرطوبة النسبية ، سرعة الرياح والاشعاع الشمسي) لاي يوم من السنة، تم بناء النماذج باستخدام خمس خوارزميات لتدريب الشبكات، قسمت البيانات المستخدمة الى مجموعتين، مجموعة التدريب وكانت بنسبة 70% من البيانات، ومجموعة الاختبار بنسبة 30% من البيانات، ولقياس أداء النماذج تم استخدام معامل التحديد R^2 ، جذر متوسط مربع الأخطاء RMSE [3].

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (O_i - T_i)^2}{n}} \dots\dots\dots(8)$$

واستخدمت هذه المعايير في مقارنة البيانات الحقيقية المقاسة والمحسوبة من نماذج الشبكة العصبية لاربع سنوات كاملة بين

هذا البحث مستند على الخوارزميات التكرارية المختلفة لتعليم شبكة الانتشار العكسي حيث استخدمت خوارزمية الانحدار التدريجي (GD)، الانحدار التدريجي المعجل (GDM)، خوارزمية الميل الصرفية (CFG)، خوارزمية نيوتن (BFGS)، خوارزمية ليفن بيرك ماركوت (LM) وكالاتي :-

• خوارزمية الانحدار التدريجي (GD) Gradient Descent :

خوارزمية الانحدار التدريجي تسمح لأوزان الشبكة ان تتحرك على الجانب العكسي من دالة التنشيط، ودور الانتشار العكسي يعود الى الطريقة التي يتم بها حساب الميل لطبقات الشبكة المتعددة اللاخطية ويتم خلاله ضبط اوزان الشبكة والذي يحقق سرعة للخوارزمية، وتمثل الخوارزمية لتكرار واحد لتعديل الاوزان بالمعادلة (1) :

$$W_{k+1} = W_k - \alpha_k g_k \dots\dots\dots(1)$$

حيث ان : α_k : معدل التعلم learning rate ، g_k :

الميل الحالي [21].

• خوارزمية الانحدار التدريجي المعجل Gradient Descent with Momentum (GDM) :

شبكة الانتشار العكسي تتطلب العديد من التكرار ، ولتخفيف هذه المشكلة يمكن ان تضاف معلمة تعجيل Momentum الى التدريب بواسطة جعل الوزن مساوي لمجموع جزء صغير من الوزن في الاوزان الاخيرة، ويكون تعديل الاوزان فيها حسب المعادلة (2):

$$W_{k+1} = W_k - \alpha_k g_k + \mu(W_k - W_{k-1}) \dots\dots\dots(2)$$

حيث ان : μ : التعجيل Momentum والذي تكون قيمته بين 1-0 [22].

• خوارزمية الميل الصرفية Conjugate gradient BP (CGF) :

خوارزمية BP القياسية تعدل الاوزان باتجاه هبوط حاد للميل الذي ليس بالضرورة ينتج عنه استقرارية للشبكة وهو حساس جدا لمعدل التعلم الذي قد يسبب عدم استقرارية للشبكة او تاخذ وقت طويل

للوصول الى الحل لذلك استخدم مؤخرًا عدة خوارزميات لحساب الميل

لتعليم الشبكات العصبية بطريقة ينتج عنها تسريع الخوارزميه للوصول الى الحل . ففي خوارزمية الميل الصرفية التي اقترحت من Fletcher and Reeves ، تبدأ بالبحث عن الهبوط الحاد في التكرار الاول [7]

$$P_0 = -g_0$$

واتجاه البحث في كل تكرار يحدد بواسطة تغيير الاوزان بالمعادلة (3) :-

$$W_{k+1} = W_k - \alpha_k g_k + \mu_k P_k \dots\dots\dots(3)$$

حيث ان :

$$P_k = -g_k + \beta_k P_{k-1} \dots\dots\dots(4)$$

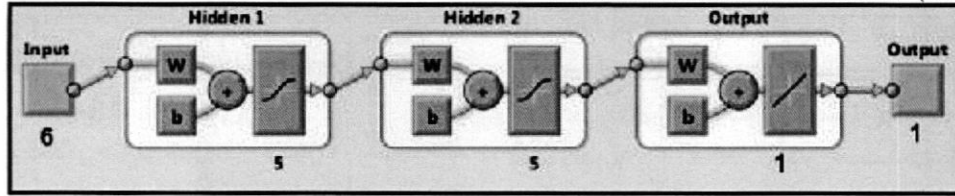
وان (Fletcher and Reeves) يغير الثابت β_k بالمعادلة التي هي نسبة تربيع الميل الحالي على تربيع الميل السابق وحسب المعادلة (5) :

(2005-2008) وذلك لضمان انطباق النماذج لجميع المراحل السنوية وظروفها ، علما بان هذه البيانات لم تدخل في عملية بناء نماذج الشبكات العصبية. تم تصميم النماذج باستخدام (Matlab 2012) لتحديد نموذج الشبكة العصبية الاصطناعية الذي يعطي تقدير جيد للتبخر اليومي، وكانت هذه النماذج للشبكات الساكنة: شبكة FFBPNN، وشبكة CFBPNN، والشبكات الديناميكية: شبكة NARX وبمعماريتين (6-5-5-1)، (6-5-5-1) والموضحة بالشكل رقم (4) لشبكة FFBPNN ، والشكل رقم (5) لشبكة CFBPNN، والشكل رقم (6) لشبكة NARX، وبالخوارزميات الآتية لكل معمارية:

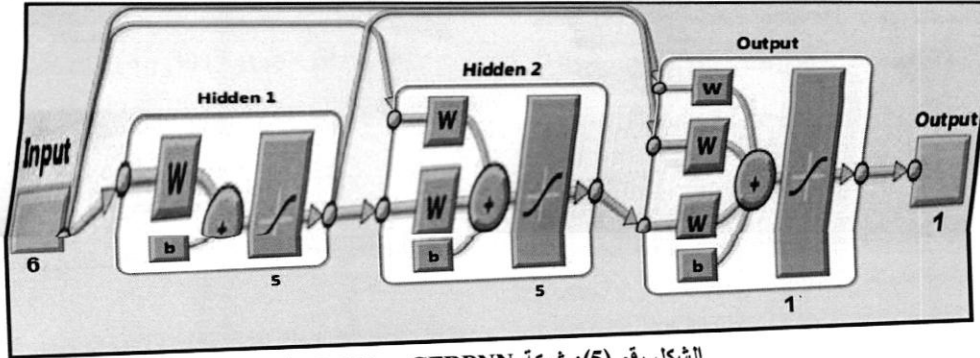
- 1- Levenberg-Marquardt algorithm (TRAINLM)
- 2- algorithm TRAINBFG
- 3- TRAINCGF algorithm
- 4- Gradient descent algorithm (TRAINGD)
- 5- Gradient descent with momentum (TRAINGDM)

جدول رقم (1):البيانات الحقيقية المستخدمة لتقدير التبخر

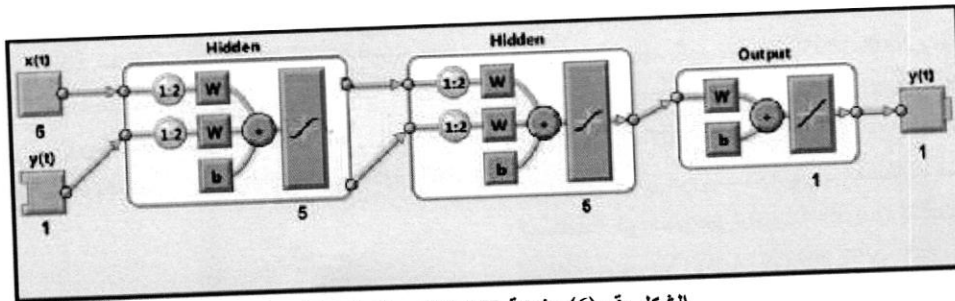
البيانات الحقيقية	
T_{Max}	درجة الحرارة العظمى
T_{Min}	درجة الحرارة الصغرى
W	Wind سرعة الرياح
S	Sunshine الاشعاع الشمسي
R	Rain الامطار
Rh	الرطوبة النسبية
Ev	Evapotranspiration التبخر



الشكل رقم (4): شبكة FFBPNN - (6-5-5-1)



الشكل رقم (5): شبكة CFBPNN - (6-5-5-1)



الشكل رقم (6): شبكة NARX - (6-5-5-1)

درجة الحرارة العظمى والصغرى، الامطار، الرطوبة النسبية والإشعاع الشمسي لوحده غير كافي لإعطاء تقديرات صحيحة وان نسبة الخطأ فيها كبير جدا، لكن عند بناء نماذج يؤثر بها كل العوامل أعطت

نتائج صحيحة كانت دقتها متقاربة من نموذج إلى آخر يوضحها الجدول رقم (2) الذي يبين فيه مقارنة ما بين البيانات المقاسة والمحسوبة من استخدام نماذج الاختبار لاختيار أفضل نموذج منها.

4- النتائج والمناقشة

تفسير النتائج التي تم الحصول عليها من تطبيق نماذج الشبكات العصبية على بيانات اربعة اعوام كاملة لم تدخل ضمن عملية بناء النماذج بإمكانية تقدير التبخر اليومي باستخدام بيانات الإدخال (درجة الحرارة العظمى والصغرى، الرطوبة النسبية، الامطار، سرعة الرياح والاشعاع الشمسي) والتبخر كإخراج، كما بينت النتائج ان استعمال

زيادة في دقة النتائج باستخدام خوارزمية (LM) لجميع نماذج الشبكات فقد تراوحت قيمة معامل التحديد بين (0.99-0.95) وانخفاض قيمة متوسط مربع الخطأ الذي تراوح بين (0.07-0.00005) .

يلاحظ من النتائج ان خوارزمية التدريب لها تأثير واضح على دقة النتائج فمع بقاء عدد الخلايا الخفية بعدد (5) لكل نموذج من نماذج التدريب بمعمارية (6-5-5-1) كانت قيم معامل التحديد متفاوتة من نموذج الى اخر مع اختلاف خوارزميات التدريب حيث كان افضلها عند استخدام خوارزمية (LM) لنموذج شبكة NARX الذي وصل الى 0.981، لكن عند تغيير المعمارية بزيادة عدد الطبقات الخفية من (2) الى (3) طبقات خفية من خلال معمارية (6-5-5-5-1) فمع زيادة الحسابات ويطئ التدريب لكل نموذج كان هناك دقة عالية في الحصول على النتائج أفضلها عند استخدام خوارزمية (LM) لنموذج شبكة (NARX) الذي وصل معامل التحديد فيها الى (0.99).

لوحظ من النتائج ان معامل التحديد للنماذج الذي استخدم شبكة (FFBPNN) تراوح بين (0.95-0.91) وجذر متوسط مربع الأخطاء تراوح بين (0.22-0.08) ولكل خوارزميات التدريب، اما عند تغيير النماذج باستخدام شبكة (CFBPNN) ازداد معامل التحديد بحيث وصل الى 0.96 وانخفض معه جذر متوسط مربع الأخطاء ووصل الى 0.01 عند استخدام خوارزمية (LM)، وبنفس الوقت حصلت زيادة في معامل التحديد لنماذج شبكة NARX الذي وصل الى 0.981 مع انخفاض ملحوظ لجذر متوسط مربع الخطأ وصل الى 0.0002 عند استخدام خوارزمية (LM).

ولغرض قياس كفاءة النماذج واختيار أفضل نموذج لتقدير التبخر تم تطبيق النماذج بنفس الإدخالات السابقة وإخراج لاي خوارزمية من الخوارزميات الخمسة وبمعمارية مختلفة من خلال زيادة عدد الطبقات الخفية (6-5-5-5-1) التي يوضحها الجدول رقم (3)، حيث لوحظ

جدول رقم (2): اداء النماذج بمعمارية (6-5-5-1)

النماذج				معامل التحديد R^2
NRAX	CFBP	FFBP	الخوارزميات	
0.981	0.96	0.95	LM	معامل التحديد R^2
0.96	0.923	0.91	BFGS	
0.975	0.925	0.92	CGF	
0.971	0.934	0.93	GD	
0.981	0.944	0.94	GDM	
0.0002	0.01	0.08	LM	جذر متوسط مربع الأخطاء RMSE
0.007	0.08	0.12	BFGS	
0.005	0.09	0.15	CGF	
0.002	0.07	0.22	GD	
0.001	0.05	0.19	GDM	

جدول رقم (3): اداء النماذج بمعمارية (6-5-5-5-1)

النماذج				معامل التحديد R^2
NARX	CFBP	FFBP	الخوارزميات	
0.99	0.967	0.955	LM	معامل التحديد R^2
0.97	0.938	0.92	BFGS	
0.978	0.937	0.93	CGF	
0.972	0.942	0.94	GD	
0.983	0.953	0.95	GDM	
0.00005	0.006	0.07	LM	جذر متوسط مربع الأخطاء RMSE
0.0005	0.05	0.11	BFGS	
0.0004	0.06	0.12	CGF	
0.0003	0.03	0.18	GD	
0.0002	0.01	0.16	GDM	

التدريبية بين (32-37) دورة تدريبية وذلك بسبب طبيعة الخوارزمية التي تبدأ بالبحث عن الهبوط الحاد للميل من التكرار الاول للخوارزمية والذي يسرع التدريب، في حين تراوحت عدد الدورات التدريبية لخوارزمية نيوتن BFGS بين (27-30) دورة تدريبية فعلى الرغم من الحسابات الواسعة لمصفوفة (Hessian) الا انه كان هناك سرعة في التدريب لاستخدامها خوارزمية (BFGS) المحدثة والتي تغير الاوزان

ولوحظ أيضاً ان الخوارزمية لها تأثير واضح في سرعة الحصول على النتائج فقد كانت خوارزمية GD أبطأ خوارزمية حيث تراوحت عدد الدورات التدريبية فيها بين (36-41) دورة يوضحها الجدول رقم (4) للنماذج في معمارية (6-5-5-5-1)، لكن عند استخدام خوارزمية GDM وبقيمة تعجيل فرضت (0.3) قلت سرعة التدريب وتراوح بين (35-39) دورة تدريبية، اما عند استخدام خوارزمية CGF تراوح عدد الدورات

التنفيذ عند زيادة عدد الخلايا الى (7) خلايا في الطبقات الخفية ولكل النماذج الذي تراوح بين (26-42) دورة تدريبية في حين ان معامل التحديد لم يقل عن (0.99) لجميع النماذج ، يلاحظ من هذه النتائج انه كلما ازداد عدد الطبقات وعدد الخلايا بين الطبقات ازدادت دقة النتائج مع تقليل الاخطاء الى اقل ما يمكن على حساب زيادة الوقت لتعليم النماذج، هذه الدقة سوف تؤدي الى تجاوز عندالوزن ان الداخلة في هيكليّة كل نموذج عن عدد نماذج بيانات التدريب والتي من المفروض أن تكون عدد هذه الاوزان اقل او مساوية لعدد بيانات التدريب في حين انه يمكن الحصول على هذه الدقة بعدد مناسب للخلايا وعدد طبقات خفية لا يتجاوز (3) طبقات ولا تتجاوز عدد الاوزان عن بيانات التدريب.

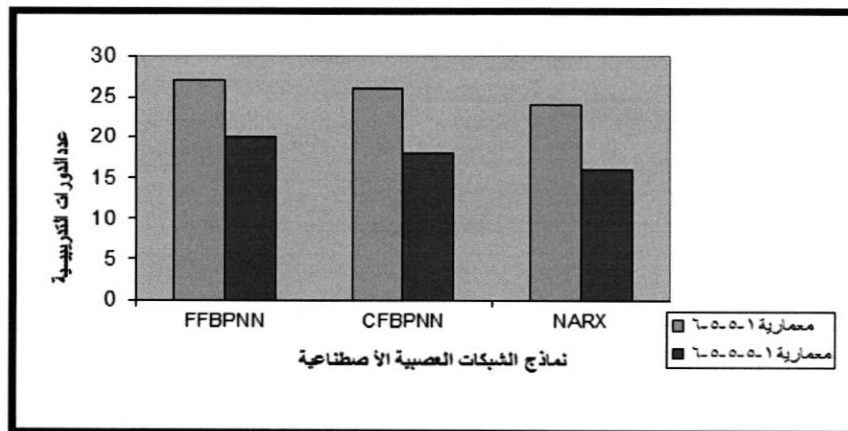
لذلك يمكن الاستنتاج من النتائج أن شبكة (NARX(6-5-5-1) بخوارزمية (LM) أفضل وأسرع خوارزمية بمعامل تحديد (0.99) وسرعة لم تتجاوز (16) دورة تدريبية لتقدير التبخّر اليومي اذا ما توفرت العوامل المناخية من درجة الحرارة العظمى والصغرى ، الامطار، الرطوبة النسبية، سرعة الرياح والاشعاع الشمسي والموضحة نتائجها في الشكل رقم (8) .

حسب المعادلة (6) لاسراع التدريب، أما خوارزمية LM كانت أسرع خوارزمية تدريبية فقد تراوحت عملية التدريب فيها بين (24-27) دورة تدريبية هذه السرعة جاءت من خلال دخول مصفوفة (Jacobian) في حساب الأوزان التي هي اقل كلفة من مصفوفة (Hessian) والتي تسرع التدريب.

لكن عند ازدياد عدد الطبقات الخفية من خلال معمارية (6-5-5-1) وزيادة الحسابات ويطى التنفيذ لاستقرارية كل نموذج لوحظ أن هناك تسريع لكل نماذج الاختبار وكان أسرعها باستخدام خوارزمية (LM) فقد تراوحت عملية التدريب فيها بين (16-20) دورة تدريبية لكل الشبكات يوضحها الشكل رقم (7). ولغرض اختيار افضل نموذج تنبؤي من بين نماذج الاختبار، تم تطبيق (8) نماذج بنفس الادخلات السابقة والاخراج لشبكة NARX وخوارزمية (LM) وبمعماريات مختلفة من خلال زيادة عدد الطبقات الخفية و عدد الخلايا في الطبقات الخفية يوضحها الجدول رقم (5)، فمع زيادة عدد الطبقات الخفية من (4-7) وبقاء عدد الخلايا ثابت بعدد (5) خلايا في كل طبقة خفية حصل بطى في التنفيذ من خلال زيادة عدد الدورات التدريبية الذي تراوح بين (20-26) دورة تدريبية ، وكذلك لوحظ ان هناك بطى في

جدول رقم (4): عدد الدورات التدريبية للنماذج

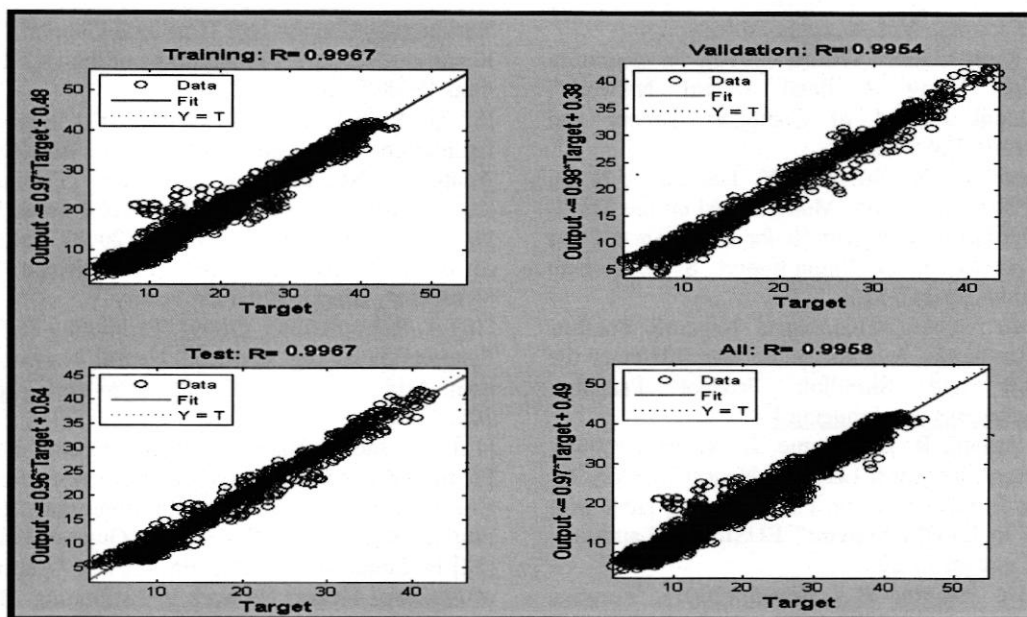
الشبكات العصبية			
NARX	CFBP	FFBP	الخوارزميات
24	26	27	LM
27	29	30	BFGS
32	34	37	CGF
36	39	41	GD
35	37	39	GDM
16	18	20	LM
25	26	28	BFGS
27	29	32	CGF
33	35	38	GD
31	32	35	GDM



الشكل رقم (7): تسريع خوارزمية (LM) للنماذج ANN المقترحة

جدول رقم (5): نماذج الاختبار لشبكة NARX بخوارزمية LM

النماذج	عدد الطبقات الخفية	عدد الخلايا في كل طبقة خفية	معامل التحديد R^2	عدد الدورات التدريبية
1	4	5	0.994	20
2	5	5	0.996	25
3	6	5	0.997	26
4	7	5	0.995	32
5	4	7	0.993	26
6	5	7	0.997	29
7	6	7	0.998	35
8	7	7	0.996	42



الشكل رقم (8): التدريب والاختبار لشبكة (1-5-5-6) NARX بسرعة (16) دورة تدريبية

في الوقت والجهد والكلفة المادية بقيم تزيد أو تقل عن قيمة واحدة لقيم التبخر اليومي المقاس، ولوحظ من خلال المقارنة بين نماذج الشبكات العصبية الاصطناعية الساكنة والديناميكية بخوارزمياتها ان جميعها يمكن استخدامها لاغراض تنبؤية تعتمد على طبيعة البيانات، وان الشبكة الديناميكية (ARX) بخوارزمية LM تعطي افضل واسرع نموذج يمكن اعتماده لتقدير التبخر اليومي اذا ما توفرت العوامل المناخية درجة الحرارة العظمى والصغرى، الامطار، الرطوبة النسبية، سرعة الرياح والاشعاع الشمسي لاي يوم من ايام السنة.

[2] الشلاوي سنوس، (2011)، "مقارنة تطبيقات الشبكات العصبية الاصطناعية والنظام المضرب لتقدير التبخر الاثنائي اليومي لمدينة الموصل" مجلة الكوفة للرياضيات والحاسبات، المجلد 1، العدد 3، ص 23-32 .

5- الاستنتاجات

الشبكات العصبية الاصطناعية اثبتت نجاحها في عملية بناء نموذج لتقدير التبخر اليومي باستخدام درجة الحرارة العظمى والصغرى، الامطار، الرطوبة النسبية، سرعة الرياح والاشعاع الشمسي لاي يوم من ايام السنة. أعطت النتائج تقديرات ذات دقة عالية للتبخر اليومي حيث وصلت قيمة معامل التحديد فيها الى 0.99، جاءت هذه الدقة من القراءات الحقيقية للبيانات والتي استخدمت في عملية التدريب والاختبار، وهذه الدقة تكون مقبولة كبديل عن القياسات الحقلية المكلفة

6- المصادر

[1] العاني. افتخار.، الحسن. كامل. (2011)، "تقدير التبخر -النوع المرجعي اليومي لمنطقة الموصل باستخدام الشبكات العصبية الاصطناعية".

- [15] C. Archana, (2010), "Reference Crop Evapotranspiration Estimation Using Artificial Neural Networks", International Journal of Engineering Science and Technology, Vol. 2(9), pp 4205-4212.
- [16] M.J. Diamantopoulou, P.E. Georgiou, (2011) "Performance Evaluation of Artificial Neural Networks in Estimating Reference Evapotranspiration with Minimal Meteorological Data", Global NEST Journal, Vol 13, No 1, PP 18-27.
- [17] S.A. Moasheri, P. Afrasiab, S. Sarani, (2010), "Estimating of Reference Evapotranspiration by Using Artificial Neural Networks", ICTECE.
- [18] V. Nourani, M.S. Fard, (2012), "Sensitivity Analysis of the Artificial Neural Network Outputs in Simulation of the Evaporation Process at Different Climatologic Regimes", Advances in Engineering Software Journal, Vol. 47, pp 127-146.
- [19] P. Kumar, A.K. Tiwari, (2010), "Evaporation Estimation Using Artificial Neural Network", International Journal of Computer Theory and Engineering, Vol.4 No. 1.
- [20] Haoffi Z., X. Guoping, Y. Fagting, Y.H an, (2007), "Neural Network Model Based on the Multi-Stage Optimization Approach for Short-Term Food Price Forecasting in China. Expert Systems with Applications, pp 347-356.
- [21] Math Works, (2010), Neural Network Toolbox 7.0, Math Works Announces Release 2010a of the MATLAB and Simulink Product Families, www.mathworks.com/trademarks.
- [22] R. Zayani, R. Bouallegue, D. Roviras, (2008), "Levenberg-Marquardt Learning Neural Network for Adaptive Predistortion for Time-Varying HPA with Memory In OFDM Systems" EUSIPCO, Lausanne, Switzerl and pp 25-29 .
- [23] O. De. Jesus and M.T. Hagan, (2001), "Forward Perturbation Algorithm for a General Class of Recurrent Network," in Proc. of the International Joint Conference on Neural Networks, Washington, DC, vol.4, pp.2626–2631.
- [24] O. De Jesus, J.M. Horn, and M.T. Hagan, (2001), "Analysis of Recurrent Network Training and Suggestions for Improvements," in Proc. of the International Joint Conference on Neural Networks, Washington, DC, vol.4, pp.2632–2637.
- [3] S.A. Kalogirou, (2003) "Artificial intelligence for the modeling and control of combustion processes: a review". Progress in Energy and Combustion Science, vol. 29, Issue 6, pp 515–566.
- [4] O. De Jesus and M.T. Hagan, (2007), "Backpropagation Algorithms for a Broad Class of Dynamic Networks," IEEE Transactions on Neural Networks, vol.18, no.1, pp.14 -27.
- [5] H. Beigy and M. R. Meybodi, (2009). "A learning Automata-Based Algorithm for Determination of the Number of Hidden Units for Three-layer Neural Networks," International Journal of Systems Science, vol. 40, no.1, pp.101–118.
- [6] L.R. Medsker, L.C. Jain, (2000). "Recurrent neural networks: design and applications". Boca Raton, FL: CRC Press.
- [7] De Jesús O., M.T. Hagan, (2001), "Backpropagation through Time for a General Class of Recurrent Network". Proceedings of the IJCNN, Vol. 4, pp 2638-2643.
- [8] D. Kumar, A.K. Tiwari, (2012), "Evaporation Estimation Using Artificial Neural Network and Adaptive Neuro-Fuzzy Interface Techniques", Pakistan Journal of Meteorology Vol.8, Issue 16.
- [9] A. Ariapour, M.N. Zavareh, (2008), "Estimation of Daily Evaporation Using of Artificial Neural Networks", Islamic Azad University.
- [10] A. Rahimikhoob, (2009), "Estimating Daily Pan Evaporation Using Artificial Neural Network in a Semi-Arid Environment", Springer-Verlag pp101-105.
- [11] V. Jadeja, (2011), "Artificial Neural Network Estimation of Reference Evapotranspiration from Pan Evaporation in a Semi-Arid Environment", B.V.M. Engineering College, V.V. Nagar, Gujarat, India.
- [12] H. Jahanbani, A. Hussein, (2010), "Application of Artificial Neural Network in Estimating Monthly Time Series Reference Evapotranspiration with Minimum and Maximum Temperatures", Vol. 9, pp 207-220.
- [13] Tabari, S. Marofi, A. Sabziparvar, (2009), "Estimation of Daily pan Evaporation Using Artificial Neural Network and Multivariate Non-Linear Regression". pp 399-406 .
- [14] C. Ozkan, O. Kisi, B. Akay, (2010), "Neural Network with Artificial Bee Colony Algorithm for Modeling Daily Reference Evapotranspiration" pp 431-441.

Using nonlinear autoregressive neural network for estimation daily evaporation: a comparison of neural networks with different algorithms

Shahla A. AbdAlKader_Najem A.AbdulKader_Kifaa H. Thanoon

Dept. of Computer System Technical Institute Mosul , Iraq

Abstract

In this research a model of Dynamic NN(NARX) was applied to estimate the daily Evaporation of Mosul city using certain climate parameters(the maximum and the minimum temperature ,rain ,relative humidity ,wind speed and the sun shine)for any day in the year , and comparison for Static NN like FFBPNN, CFBPNN . Each of these networks has two architecture: an architecture with four layers and five cells in hidden layers from one hand, and an architecture with five layers and five cells in the hidden layers from the other.

Different algorithm were used for the training like: Levenberg-Marquardt algorithm (LM), Quasi-Newton algorithm (BFGS), Conjugate Gradient algorithm (CFG), Gradient Descent algorithm (GD) and Gradient Descent with Momentum algorithm (GDM). Data was obtained from the forecast Directorate in AlRashedeyyah district in Nineveh Province for the period (1995-2008) and used in the research. Data of ten years for the period (1995-2004)was employed to develop the models and the data of four years was used to evaluate the models and to compare their outputs with the data measured for the period(2005-2008). Moreover; determination coefficient R_square (R^2) and the Root Mean Square Error (RMSE) methods were used to estimate the level of correspondence for the measured data and NN outputs to select the best prediction model from the models applied.

Results showed that the NARX with(LM) algorithm is efficient in improving a prediction model to estimate the daily Evaporation as the value of coefficient estimation was 0.99, and this is considered the best and the fastest algorithm if temperature, rain, relative humidity ,wind speed and sunshine data available for any day in the year.

Keywords: Dynamic NN, Static NN, Gradient Descent, Gradient Descent with Momentum, Conjugate Gradient, Quasi - Newton,Levenberg - Marquardt .