

دراسة مقارنة للتنبؤ بمعدلات الرطوبة النسبية RH% محافظة نينوى باستخدام نماذج السلاسل الزمنية الموسمية والتنبؤات العصبية الاصطناعية

م.د. سهاد علي شهيد *

المستخلص :

تعد الشبكات العصبية من أهم مجالات الذكاء الاصطناعي الذي يعكس تطوراً هاماً ملموساً في طريقة التفكير الإنساني، وتدور فكرة الشبكات العصبية حول محاكاة العقل البشري باستخدام الحاسب الآلي. وقد يعود التطور المنظور في هذا المجال إلى العديد من الدراسات التي تمت في مجال المعالجة العصبية المحاكاة عن طريق حل المشاكل التي تواجهه ، وذلك من خلال إتباع عمليات التعلم الذاتي والتي تعتمد على الخبرات المختزنة في الشبكة التي تحقق أفضل النتائج. ويقصد بالنماذج السلاسل الزمنية الموسمية هي مجموعة من القيم المشاهدة المرتبطة مع بعضها تولدت بشكل متعاقب مع استمرار الزمن وتشير إلى النمط المتماثل لحركة السلسلة الزمنية في الأشهر المتعاقبة خلال السنوات المتتالية. وقد توصل البحث إلى أن أفضل نموذج باستخدام الشبكات العصبية هو $NN(1)$ ، 10 ؛ 12 الذي يضم (12) متغيراً مزاح و(10) عقد مخفية ، وبالنسبة لنماذج السلاسل الزمنية الموسمية فقد كان النموذج $SARMA(1,0,1)$ ، $(0,1,1)_{12}$ وقد تم استخدام بعض المقاييس الاحصائية لاختبار افضلية الاسلوبين في التنبؤ والذي كانت نتيجته هو اسلوب السلاسل الزمنية الموسمية إذ حصل على أقل قيم بالنسبة للاختبارات $(MAD \cdot MAPE \cdot MAE \cdot RMSE)$ على التوالي ، بذلك كان أفضل طريقة للتنبؤ بالرطوبة النسبية (RH%) لمحطة الموصل هو باستخدام النموذج $SARMA(1,0,1)$ ، $(0,1,1)_{12}$.

Abstract:

The neural network is one of the most important fields of artificial intelligence, which reflects an important development in a significant way of human thinking, and spin around the idea of neural networks simulate the human mind using the computer. Foreseeable development in this area may be due to several studies, which has in the field of neural processing simulation by solving the problems that face, and by following the self-learning processes that rely on the expertise stored in the network, which yield better results. It is intended models of seasonal time series is a set of viewing values associated with each generated sequentially with the continuation of time and refer to replicate the pattern of movement of the time series in the opposite months during successive years. Research has found that the best model is the use of neural networks $NN(1; 10)$ which includes (12) variable backward in time and (10) hidden contract. For the time-series models seasonal has been the model $SARMA(1,0,1)$ ، $(0,1,1)_{12}$. we have been using some statistical measurements to test the advantage two modes in the prediction which was its result is the style of the time series as seasonal got the lowest values for tests $(RMSE \cdot MAE \cdot MAPE \cdot MAD)$ respectively. It was the best way to predict the relative humidity (RH%) of the province of Mosul is using the model $SARMA(1,0,1)$ ، $(0,1,1)_{12}$.

الفصل الاول المقدمة وهدف البحث

1-1 المقدمة :

تعد نماذج الانحدار الذاتي والوساط المتحركة التكاملية (Autoregressive And Moving Average (ARIMA) ركناً أساسياً من اركان علم الاحصاء واسلوباً مهماً من اساليب تحليل السلاسل الزمنية في دراسة مختلف الظواهر الاقتصادية والاجتماعية عبر فترات زمنية محددة. ومن اهم اهداف تحليل السلاسل الزمنية الحصول على وصف دقيق لسلوك السلسلة الزمنية من خلال نماذج يفسر ذلك السلوك، واستخدام النموذج المشخص لغرض التنبؤ بسلوك السلسلة في المستقبل. ولتحقيق ذلك ينبغي دراسة تحليلية لنماذج السلاسل الزمنية بالاعتماد على الاساليب الاحصائية والحاسوبية. ومن الاساليب الحديثة التي اخذت اهتماماً ملحوظاً، ولاسيما في علوم الحاسبات اسلوب الشبكات العصبية الاصطناعية (Artificial Neural Networks) (ANNS)، اذ تعد محاكاة متواضعة بفعلها وشكلها ومضمونها للشبكة العصبية الحيوية الموجودة في الدماغ البشري.

1-2 مشكلة الدراسة:

مدى تمكين التكامل بين اسلوبي (ARIMA) و (ANNS) من اجل بناء أنموذجين يمكن خلالهما التنبؤ بالرطوبة النسبية للسنوات القادمة في محطة الموصل.

1-3 هدف الدراسة:

1. استخدام انموذج الانحدار الذاتي والوساط المتحركة التكاملية (ARIMA) في التنبؤ بدرجات الرطوبة النسبية % (RH) الشهرية في محافظة نينوى للفترة (1980-2013).
2. استخدام اسلوب الشبكات العصبية الاصطناعية في التنبؤ بدرجات الرطوبة النسبية % (RH) الشهرية في محافظة نينوى للفترة (1980-2013).
3. المقارنة بين نتائج تطبيق كل انموذج الشبكات العصبية الاصطناعية وانموذج (ARIMA) للتنبؤ بكميات الرطوبة النسبية.

1-4 فروض البحث:

1. التنبؤ باستخدام انموذج (ARIMA) اكثر دقة من التنبؤ باستخدام الشبكات العصبية الاصطناعية (ANNS).
2. اسلوب الشبكات العصبية افضل من اسلوب (ARIMA) في بناء نماذج غير خطية لتمثيل بيانات السلسلة الزمنية قيد الدراسة.

الفصل الثاني الجانب النظري

1-2 الشبكات العصبية الاصطناعية (Artificial Neural Cell) (2)·(3)·(4)·(8) :

تعد الشبكات العصبية الاصطناعية احد حقول الذكاء الاصطناعي الذي شكل تطورات ملحوظة في طريق مكينة التفكير الانساني. والفكرة الاساس للشبكات العصبية هي محاكاة العقل البشري باستخدام الحاسب الالى. والفضل يعود الى الدراسات التي تمت في حقل المعالجة العصبية والذي يدرس النشاط العصبي في مخ الانسان. وهذه الدراسات تعمل على محاكاة العقل البشري في طريقة حله للمسائل التي تواجهه وذلك من خلال عملية التعليم الذاتية التي تحدث في العقل ويتم فيها الاستفادة من الخبرات التي مرت عليه في السابق لغرض الوصول الى نتائج افضل بالمستقبل .

وتقسم الشبكات العصبية الاصطناعية الى :

1. طريقة التدريب بمعلم.
2. طريقة التدريب من دون معلم.

تعتمد الطريقة الاولى على جمع بيانات ومن ثم تقسيمها الى قيم داخلية وقيم خارجية هي القيم المستهدفة التعرف عليها من خلال الشبكة وبعدها تقوم الشبكة بمقارنة النتائج عن طريق القيم التي تقدرها المتغيرات الخارجية لكل عينة داخلية بالقيم الفعلية (المستهدفة) حيث تعدل الاوزان وفق خوارزمية تعليم تسمى التدريب باشراف، وتكون مخرجات هذه الشبكة عبارة عن عناصر معالجة تاخذ الرقم (1) اذا كانت المدخلات تابعة لنفس الصنف او العينة وتاخذ الرقم (1-) في الحالة العكس وتسمى ثنائية القطب.

اما الطريقة الثانية وهي متشابهة لطريقة التدريب الاشرافي، الا انها تختلف عن الطريقة الاولى بانه لا توجد قيم المتغيرات الخارجية حيث تقسم البيانات الداخلة الى مجموعات تقوم الشبكة باكتشاف المميزات غير الظاهرة فيها، وبعد ذلك يتم تقسيم المدخلات الى مجموعات مختلفة فيما بينها ومتقاربة لكل مجموعة.

اما تعليم الشبكات العصبية فيعتمد اسلوب الشبكات العصبية على تعديل الاوزان النسبية لعقد الاتصال، حيث لا تحتاج الشبكات الى تزويدها بالمعلومات بصورة واضحة ومباشرة، بل لهذا النوع من الشبكات القدرة على انشاء قواعده التعليمية.

2-1-1 هيكل الشبكة العصبية الاصطناعية (2)(4)(7): Artificial Neural Network Structure

تتكون الشبكة العصبية (ANN) من ثلاث مستويات:

1. مستوى المدخلات (Input Level): ويحتوي على عدد من العقد التي تمثل عدد المتغيرات المزاحة زمنياً .
2. مستوى المخفي (Hidden Level): قد يكون في الشبكة العصبية أكثر من مستوى مخفي واحد ويرتبط مع مستوى المدخلات والمخرجات، والشبكة العصبية التي تحتوي على مستوى مخفي واحد تسمى (Single-Layer Network) اما الشبكة التي تحتوي على أكثر من مستوى مخفي واحد تسمى (Multi-Layer Network).

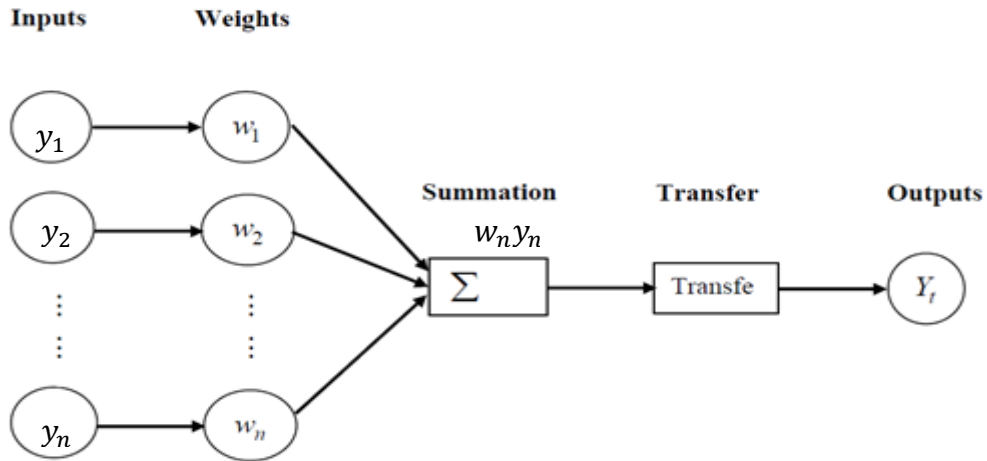
3. مستوى المخرجات (Output Level): هو عبارة عن مخرجات الشبكة العصبية الاصطناعية ويتضمن كل مستوى من المستويات الثلاثة الأنفة الذكر :

- a- العقد (Nodes): هي نقاط الارتباط العصبي بين مستويات الشبكة العصبية.
- b- المستوى (Level): هي مجموعة من العقد التي تستلم المدخلات ولها مخرجات.
- c- الاوزان (Weights): توضح الاوزان مدى قوة الارتباط العصبي بين مستويات الشبكة العصبية. إذ ان كل عقدة لها وزن يربطها مع المستوى السابق ووزن يربطها مع المستوى اللاحق، وتحتوي الشبكة العصبية ثلاث طبقات من الاوزان هي :

- طبقة اوزان مستوى المدخلات والمستوى المخفي (Input to Hidden Weight) .

- طبقة اوزان بين المستويات المخفية. (في حالة وجود أكثر من مستوى مخفي واحد) .

- طبقة اوزان المستوى المخفي ومستوى المخرجات (Hidden to Output Weight) .



الشكل (1)

يوضح أنموذج من الشبكات العصبية

إذ ان :

$y_1, y_2, \dots, y_n$ تمثل مدخلات الشبكة .

$w_1, w_2, \dots, w_n$ هي القيم الحقيقية للمعلمات وتسمى الاوزان .

$$net = w_1 y_1 + w_2 y_2 + \dots + w_n y_n$$

ويطلق عليه مجموع الاوزان .

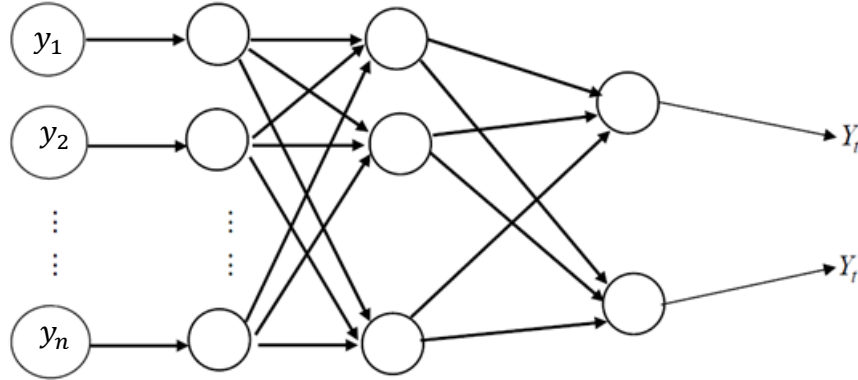
اما $Transfe$: فيطلق عليها دالة التفعيل (activation fun.) .

وان : $output = y = f(net)$ وهي مخرجات الشبكة العصبية .

2-2 انواع الشبكات العصبية :

تتألف الشبكة العصبية عادة من عدد من العصبونات المرتبطة داخلياً فيما بينها ولكن نوعية الارتباط بالنسبة للعصبونات الداخلية فضلاً عن طبيعتهما من جدران معمارية ونوع الشبكة. ومن انواع الشبكات العصبية هي:

1- الشبكات ذات التغذية الامامية (Feed forward networks):
وفيها يكون انتشار الدخلة الى الشبكة الى الامام دائما، وبذلك فان الإشارة الخارجة من اي خلية تعتمد على الإشارة الداخلة فقط .

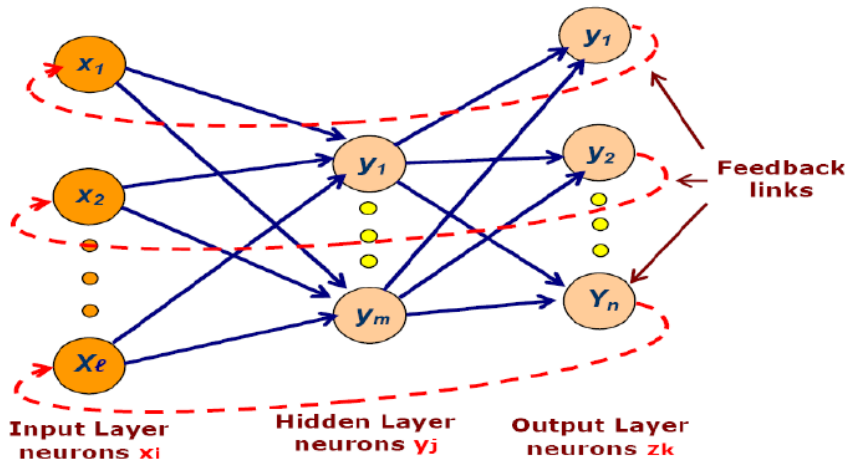


شكل (2) يوضح شبكات التغذية الامامية

• مميزات الشبكة الامامية

- المعلومات فقط تتبع طريق واحد .
- البيانات تقدم لطبقة الدخول .
- تمرر إلى الطبقة المخفية .
- تمرر لطبقة الخروج .
- المعلومات موزعة .
- معالجة المعلومات يكون بشكل متوازي .
- التمثيل الداخلي هو تفسير للبيانات .

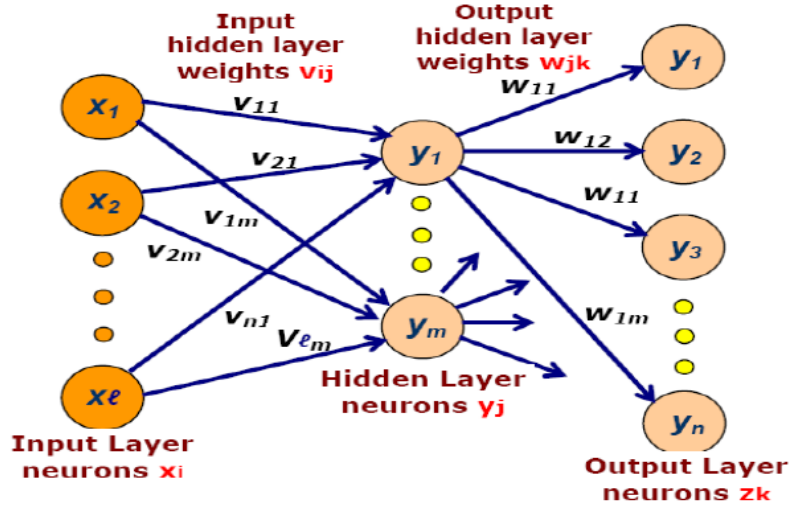
2- الشبكات العصبية ذات التغذية الخلفية (Feed Backword network):
هذا النوع يحوي على الاقل حلقة تغذية خلفية واحدة ويمكن ان يتالف من طبقة واحدة من العصبونات وكل عصبون يعود اخلاله الى ادخال كل العصبونات المتبقية وقد تكون هناك تغذية خلفية ذاتية اي ان اخراج العصبون يعود الى ادخاله منها بان جميع شبكات التغذية العكسية القليلة الاستخدام خاطئه .



الشكل (3)
وضح الشبكة العصبية ذات التغذية الخلفية

3- الشبكات المتكررة (Recurrent networks) :

في هذا النوع من الشبكات المخرجات لبعض العصبونات يتعلق بنتائج العصبونات ذاتها في لحظة سابقة او باخراج عصبونات اخرى، وبالتالي فان الإشارة في هذا النمط من الشبكات العصبية تتجه اماماً اي نحو الاخراج وايضاً بالاتجاه العكسي لتكون ادخالا في الوقت ذاته ويطلق عليها الشبكات الديناميكية لان الاخراج في لحظة ما يتعلق بالادخال باللحظة ذاتها وباللحظات السابقة اي تمتلك ذاكرة.



الشكل (4)
يوضح الشبكات المتكررة

مميزاتها :

- العقد تتصل ببعضها أو بعقد أخرى.
- تدفق المعلومات متعدد الاتجاهات.

2-2-1 التصميم الهيكلي لبناء الشبكة العصبية الاصطناعية (3) (4):

الخطوة الأولى:

يتم تجميع البيانات التي تستخدم في تدريب أو إختيار الشبكة.

الخطوة الثانية:

يتم تعريف البيانات الخاصة بالتدريب للشبكة، ووضع خطة التدريب والتعلم.

الخطوة الثالثة:

بناء هيكل الشبكة وتحديد ترأيب الشبكة في عدد المدخلات والطبقات ونوع الشبكة.

الخطوة الرابعة:

إختيار طريقة التعلم حسب ما هو متاح من أدوات أو تطوير الشبكة أو القرارات.

الخطوة الخامسة:

وضع قيم للأوزان ، ثم بعد ذلك يتم تعديل القيم في الأوزان عن طريق التغذية المرتدة والعكسية.

الخطوة السادسة:

تحويل البيانات إلى النوع المناسب للشبكة ويتم ذلك عن طريق كتابة معادلة لتجهيز البيانات والبرامج الجاهزة.

الخطوة السابعة والثامنة:

وفيهما تتم عمليتي التدريب والإختبار من خلال تكرار عرض المدخلات والمخرجات المرغوبة إلى الشبكة، ومنها مقارنة القيم الفعلية مع القيم المحسوبة، ثم حساب الفرق (الخطأ) ، ثم تعديل الأوزان لتقليل الفرق حتى يصبح فرقاً مقبولاً.

الخطوة الأخيرة:

وهنا تستطيع الشبكة الوصول إلى النتائج المرغوبة من خلال إستخدام مدخلات التدريب، وبهذا يمكن الإعتماد على الشبكة في الاستخدام لنظام مستقل قائم بذاته أو أجزاء من النظام.

4-2 استخدام الشبكات العصبية في التنبؤ (2)(3)(4)(8):

الهدف الاساس من استخدام أنموذج الشبكات العصبية هو حساب القيم المستقبلية للسلسلة الزمنية المدروسة. بفرض وجود سلسلة زمنية تحتوي على المشاهدات $(y_1, y_2, \dots, y_{n-1}, x_n)$ وللتنبؤ بالقيم المستقبلية بمعنى ايجاد $(y_{n+1}, y_{n+2}, \dots)$ وكانت السلسلة محددة بمعنى وجود عدداً حقيقياً (P) والذي يدعى البعد المظمو (Embedding Dimension) والدالة (f) بحيث يكون $(t > P)$.

$$Y_{(t)} = f(Y_{t-1}, Y_{t-2}, \dots, Y_{t-p})$$

وبالتالي يتكون التنبؤ من (N) قيمة لسلسلة زمنية معطاة ولايجاد قيم (P, f) وبفرض ان (P) معروفة ونرغب في ايجاد (f) ومن المعروف ان الشبكات العصبية الاصطناعية مكونة من عدد كبير من عناصر بسيطة خطية مترابطة فان هناك ثلاثة انواع من الوحدات:

1- وحدات ادخال : والتي تهين للقيم السابقة في السلسلة الزمنية

$$Y_{t-1}, Y_{t-2}, \dots, Y_{t-n}$$

حيث P تمثل البعد المظمو

2- وحدات مخفية : والتي تستخدم لحفظ التمثيل الداخلي للسلسلة الزمنية .

3- وحدات اخراج : والتي تعطى مخرجات الشبكة العصبية في حالة وحدة اخراج فقط لارجاع $X(t)$.

ففي حالة التنبؤ لخطوة واحدة (One-Step-ahead) يكون باستخدام المشاهدات الفعلية لكل المتغيرات المزاحة كمدخلات .

اما في حالة التنبؤ لعدة خطوات ولتكن (K) يكون بطريقتين :

الاولى: يستخرج التنبؤ لخطوة واحدة ثم ندخل هذه القيمة للتنبؤ لخطوة وهكذا الى ان نحصل على (L) من التنبؤات وهي طريقة تكرارية مثل ما يحصل في تنبؤات بوكس جينكيز وكما موضح :

$$\hat{Y}_{t+1} = f(Y_1, Y_{t-1}, \dots, Y_{t-n})$$

$$\hat{Y}_{t+2} = f(Y_{t+1}, Y_t, \dots, Y_{t-n+1})$$

$$\hat{Y}_{t+3} = f(Y_{t+2}, Y_{t-1}, \dots, Y_{t-n+2})$$

الثانية: باستخدام شبكة عصبية ذات (K) من عقد الاخراج تستخدم للتنبؤ الى تنبؤات متعددة مباشرة باستخدام كافة المشاهدات السابقة وعليه فان التنبؤات لـ (K) من الخطوات :

$$\hat{Y}_{t+1} = f_1(Y_t, Y_{t-1}, \dots, Y_{t-n})$$

$$\hat{Y}_{t+2} = f_2(Y_t, Y_{t-1}, \dots, Y_{t-n})$$

.

.

$$\hat{Y}_{t+k} = f_k(Y_t, Y_{t-1}, \dots, Y_{t-n})$$

اذ ان :

$$f_1, f_2, \dots, f_k \text{ دوال غير خطية في الشبكة العصبية}$$

5-2 استخدام نماذج السلاسل الزمنية الموسمية للتنبؤ (1)(2)(8):

يقصد بالسلسلة الزمنية الموسمية (Seasonal Time Series) : مجموعة من القيم المشاهدة المرتبطة مع بعضها والتي تولدت بشكل متعاقب مع استمرار الزمن وتحتوي على ظاهرة موسمية تستند الى النمط المتماثل لحركة السلسلة الزمنية في الاشهر المتقابلة خلال السنوات المتتالية. اي ان السلسلة تعيد نفسها بعد فترات زمنية ثابتة.

وتدعى هذه الفترة بالفترة الموسمية ونرمز لها بالرمز (S) وقد تكون (S) سنة او فصلاً او شهراً.

اي ان :

$$f(t + s) = f(t)$$

وبعد ان نحصل على بيانات مستقرة يتم تحديد الموسمية عن طريق فحص الارتباطات الذاتية للفتترات الزمنية، فاذا وجد ان تلك الارتباطات لها فروق معنوية عند فترات زمنية ثابتة (تمثل طول الموسم) فان السلسلة المستقرة تكون موسمية. اما اهم المعايير الاحصائية التي تستخدم في وصف سلوكية السلسلة الزمنية فهي :

1- الارتباط الذاتي : Autocorrelation

وهو مقياس لدرجة العلاقة بين قيم المتغير نفسه عند فترات ازاحة موسمية مختلفة. ويقدر معامل الارتباط الذاتي في حالات السلاسل الزمنية الموسمية عند الازاحة (S) حسب الصيغة الآتية :

$$\hat{\rho}_s = \frac{Cov(Y_t, Y_{t+s})}{\sqrt{Var(Y_t)Var(Y_{t+s})}} = \frac{\sum_{t=1}^{n-s} (Y_t - \bar{Y})(Y_{t+s} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2} \dots \dots \dots (1)$$

إذ أن :

Y_t : قيم مشاهدات السلسلة الزمنية

ويمكن ملاحظة ان قيم معاملات الارتباط الذاتي

$$\rho_1, \rho_2, \rho_3, \dots, \rho_{s-1}, \rho_s, \rho_{s+1}, \dots$$

اما بالنسبة لدالة الارتباط الذاتي (ACF) فتستخدم السلاسل الزمنية الموسمية لتعطينا معلومات حول سلوك السلسلة وعن مكوناتها الاساس ، وتساعد على تحديد استقرارية السلسلة وهل هي موسمية ام لا ؟

2- الارتباط الذاتي الجزئي : Partial Autocorrelation Function

وهو يقيس درجة العلاقة بين المشاهدين (Y_{t+s}, Y_t) بثبوت بقية المشاهدات الاخرى $(Y_{t+1}, \dots, Y_{t+s+1})$ ، وتعد (PACF) اداة مهمة تستعمل في تشخيص النموذج وتحديد درجته وفيما يخص ملائمة الأنموذج من خلال اختيار عشوائية اخطاء التنبؤ (البواقي) .

2-6 أنموذج المختلط (الاتحدار الذاتي- الاوساط المتحرك) الموسمي (5)(2) :

Seasonal Mixed (Autoregressive-Moving Average) Model (SARMA)

الصيغة العامة للأنموذج المختلط الموسمي من الدرجة (Q,P) سناخذ الشكل التالي :

$$Y_t = \phi_s Y_{t-s} + \phi_{2s} Y_{t-2s} + \dots + \phi_{ps} Y_{t-ps} + a_t - \theta_s a_{t-s} - \theta_{2s} a_{t-2s} \dots - \theta_{qs} a_{t-s} \dots \dots \dots (2)$$

حيث ان :

Y_{t-is} : قيم مشاهدات السلسلة الزمنية الموسمية $i=0,1,2,\dots,p$

S : طول الفترة الموسمية

ϕ_{is} : معاملات الاتحدار الذاتي الموسمي $i=1,2,\dots,p$

P : درجة النموذج الموسمي

a_t : الخطأ العشوائي ، حيث ان : $a_t \sim \text{NID}(0, \sigma_a^2)$

ولكي تحقق درجة الاستقرارية يشترط ان تكون جذور المعادلة :

$$\phi_s(B^S) = 1 - \phi_s B^S - \phi_{2s} B^{2s} \dots - \phi_{ps} B^{ps} = 0 \dots (3)$$

اي خارج دائرة الوحدة (Unit Circle)، (دائرة نصف قطرها يساوي واحد)

إذ أن (B) عامل الارتداد الخلفي (Back Shift Operator) ويعرف بالشكل :

$$B^S X_t = X_{t-s} \quad \forall \quad S = 1, 2, \dots \dots \dots (4)$$

θ_{is} : معاملات أنموذج الاوساط المتحركة الموسمي

Q : درجة النموذج الموسمي

ويرمز للأنموذج المختلط (الاتحدار الذاتي – الاوساط المتحركة)، الموسمي بالرمز $\text{ARMA}(P, Q)_s$.

ان النماذج الموسمية اعلاه تنطبق على السلاسل الزمنية المستقرة، اما اذا كانت السلسلة (Y_t) غير مستقرة فإنه يمكن ايجاد الأنموذج بعد ايجاد الفروق الموسمية المطلوبة لانتاج سلسلة مستقرة، إذ ان عامل الفرق الموسمي من الدرجة D هو :

$$\nabla_S^D (1 - B^S)^D$$

ويتكون لدينا أنموذج مختلط موسمي غير مستقر :

(Seasonal Autoregressive Integrated Moving Average Model)

ويكتب بالصيغة التالية:

$$\phi_P(B^S) \nabla_S^D Y_t = \theta_Q(B^S) a_t \dots (5)$$

إذ أن : (Q, D, P) تحدد درجة النموذج والذي يكتب اختصاراً $D \cdot ARIMA(P, Q)_S$. ويمكن استخدام النموذج بالتنبؤ بالقيم المستقبلية الى $(L=1, 2, \dots)$ فترة قادمة وذلك باخذ التوقع الشرطي عند الزمن (t) لنحصل على التنبؤات $\hat{Y}_t(L) = Y_{t+1}$ بمتوسط مربعات خطأ التنبؤ أقل ما يمكن. وباستخدام صيغة معادلة الفروق (Differences Equation Form) التي تحتوي على قيم حالية وسابقة لـ (Y_t) وقيم حالية وسابقة للخطأ العشوائي (a_t) يمكن حساب التنبؤات للنموذج المختلط الموسمي باخذ التوقع الشرطي ،

$$\hat{Y}_{t+L} = E \left((\hat{Y}_{t+L} | Y_1, Y_2, \dots, Y_n) \right)$$

فنحصل على النموذج وكالاتي :

$$Y_{t+L} = \hat{Y}_t(L) = \hat{\theta}_S Y_{t+L-S} + \hat{\theta}_{2S} Y_{t+L-2S} + \dots + \hat{\theta}_{PS} Y_{t+L-PS} + a_{t+L} - \hat{\theta}_S a_{t+L-S} - \hat{\theta}_{2S} a_{t+L-2S} - \dots - \hat{\theta}_{QS} a_{t+L-QS} \quad \dots(6)$$

إذ أن :

$$a_{t+L} = 0 \quad ; \quad Y_{t+L} = E(Y_{t+L})$$

الفصل الثالث الجانب التطبيقي

1-3 المقدمة :

يعد أسلوب التنبؤ نشاط متكامل من أنشطة صنع القرار في إدارة أي منظومة بسبب حاجة المنظومة عند التخطيط للمستقبل الى صيغة مناسبة للتنبؤ بسلوكية أو حجم العوامل المؤثرة في تحقيق الاهداف ومتطلبات النجاح.

تتطلب بيانات السلاسل الزمنية تحديد الفترة الزمنية للمشاهدات حيث يمثل المتغير المعتمد بيانات السلسلة الزمنية نفسها أما المتغيرات المستقلة فتتمثل مشاهدات سابقة لنفس السلسلة.

ان أول اجراء يعتمد في التحليل لمثل هذا النوع من البيانات هو رسم السلسلة مع الزمن لتمييز انماط السلسلة الزمنية، فقد تكون البيانات متذبذبة حول متوسط ثابت اي ليس هناك تزايد او تناقص متواصل مع الزمن ، او قد تكون قيم السلسلة في حالة تصاعد أو تناقص مسنمر مع الزمن ، او قد تكون قيم السلسلة متذبذبة وفق موسم محدد لفترة زمنية شهر او سنة او فصل .

وهو ما يمثل عامل الموسمية، او قد يكون هناك ارتفاعاً أو انخفاضاً في قيم السلسلة الزمنية ولكن ليس خلال فترة ثابتة وهو ما يمثل الدورية .

إذا لم يكن التشخيص واضحاً من خلال الرسم فيتم اختيار استقرارية هذه السلسلة من ناحية المتوسط وبالتالي تحديد عاملي الموسمية والاتجاه وذلك من خلال حساب دالتي الارتباط الذاتي والجزئي للسلسلة وتحقيق الاستقرارية بالاعتماد على أسلوب الفروقات ، ومن خلال فحص دالتي الارتباط الذاتي والجزئي يمكن تحديد الازاحات التي تجعل السلسلة أكثر استقراراً . وبذلك يعتمد عدد المتغيرات المزاحة زمنياً التي تمثل مدخلات أنموذج الشبكية على تشخيص مدى استقرارية السلسلة اي عدد الازاحات المطلوبة .

2-3 حدود تطبيق البحث :

تضمن البحث بيانات حقيقية تمثل الرطوبة النسبية ((Relative Humidity (RH%)) الشهرية في محافظة نينوى للسنوات (1980- 2013) .

3-3 النتائج والاختبارات :

1-3-3 اختبارات الاستقرارية للسلسلة الزمنية(5) :

Testing Stationarity of Time Series

1-اختبار جذر الوحدة : Unit Root Test:

تتمثل المرحلة الاولى للتحليل الاحصائي في هذه الدراسة في اختبار استقرار السلاسل الزمنية، اي الوقوف على ما اذا كانت السلسلة الزمنية للمتغيرات ساكنة في مستوياتها او عند حساب الفروق الاولى لها، وهناك العديد من اختبارات الاستقرار أهمها اختبار (ديكي- فولر الموسمي)، (ADF)، (Augmented Dickey Fuller Test) والذي يحسب بالصيغة التالية :

$$\Delta Y_t = \phi Y_{t-1} + \sum_{i=1}^p a_i (Y_{t-i} - Y_{t-i-1}) + \varepsilon_t \quad \dots(7)$$

إذ أن :

Δ : يمثل الفروق الاولى .

Y : يمثل المتغير المراد اختباره .

ϵ : عنصر الخطأ .
 p : عدد مرات الإبطاء .
 ϕ, ai : ترمز للمعلومات المراد تقديرها .
 إذ أن : $\phi = \rho - 1$
 بالتالي فإن فرضيات الاختبار :

$$H_0: \phi \geq 1$$

$$H_1: \phi < 1$$

وينطوي اختبار (ديكي- فوللر الموسع) على فريضة العدم ($\phi < 0$) مقابل الفريضة ($\phi \geq 0$) حيث يتم رفض فريضة العدم المتمثلة في وجود جذر الوحدة وإن السلسلة مستقرة في مستواها إذا كانت قيمة (t_c) المحسوبة أكبر من قيمة (t_t) الجدولية، ويتم قبول فريضة العدم إذا كانت قيمة (t_c) المحسوبة أقل من قيمة (t_t) الجدولية، ويعاد تقدير المعادلة أعلاه استخدام الفروق الأولى للمتغير لتحديد ما إذا كانت السلسلة الزمنية للمتغير ساكنة عند حساب الفروق الأولى .

ويوضح جدول (1) أدناه نتائج اختبار جذر الوحدة للسلسلة محل الدراسة والمتمثلة بالرطوبة النسبية % (RH) الشهرية لمحافظة نينوى للفترة (2013-1980) باستخدام اختبار (ديكي- فوللر الموسع (ADF)) لأختبار السلاسل الزمنية وكانت النتائج كما يلي :

جدول (1)

نتائج اختبار جذر الوحدة لـ (ADF)

ADF		t- Statistics	Pro.*
		-4.424838	0.0003
Test Critical Values	1% Level	-3.446734	
	5% Level	-2.868657	
	10 % Level	-2.570627	

المصدر: من اعداد الباحث، نتائج في التحليل الاحصائي باستخدام E-Views. Ver.7

ويلاحظ ان قيمة (ADF) بلغت (-4.4248) وعند مقارنتها مع قيم (ADF) الجدولية عند جميع المستويات (1% ، 5% ، 10%) فإن (ADF) المحسوبة اكبر من القيم احصاءة الاختبار بذلك ترفض فرضية العدم ($H_0: \phi \geq 1$) وتقبل الفرضية البديلة ، اي ان لها جذر وحدة بمعنى انها ذات اتجاه عام مستمر.

4-3 نتائج التنبؤ باستخدام الشبكات العصبية :

البيانات المدخلة للشبكة عبارة عن سلسلة زمنية للفترة (2013-1983) والتي تمثل الرطوبة النسبية الشهرية لمحافظة الموصل . والذي يمثل المتغير المستهدف (Target Variable) والذي يراد التنبؤ به أي المتغير الذي يجب انتاجه بوساطة الشبكة (Output Variable) .
 وتم تقسيم السلسلة الدخلة للشبكة الى ثلاثة اجزاء كما يلي :
 * 268 مشاهدة بنسبة 67.7 % من البيانات استخدمت لتدريب الشبكة (Training Set).
 * 128 مشاهدة بنسبة 32.3 % من البيانات استخدمت لاختبار الشبكة بعد الانتهاء من التدريب (Test Set).
 ومن اجل ملائمة البيانات المدخلة لمتطلبات الشبكة تم تحويل البيانات في المدى ما بين (1-1) للمدخلات، وبين (0.045 ، 2.828) للمخرجات .

4-3-1 تحديد عقد الادخال : Determination of input Nodes

تم الاعتماد على معاملات الارتباط الذاتي المعنوي لسلسلة الرطوبة النسبية المستقرة وبعد أخذ الفرق الموسمي لها وجد ان معاملات الارتباط الذاتي للرطوبة النسبية معنوية فقط عند الازاحة الاولى والثانية عشر كما في الشكل (7) ، وبذلك فإن أنموذج الخلية العصبية سيكون:

$$Y_t = f(Y_{t-1}, Y_{t-12}) \quad \dots (8)$$

اي ان عدد عقد الادخال هي اثنان والمتغيرات المزاحة هي (Y_{t-1}, Y_{t-12})

المصدر: تم استخدام برنامج Eviews.7 في ايجاد النتائج من قبل الباحث
 ان بيانات السلسلة الزمنية تمثل الرطوبة النسبية الشهرية والمتغيرات المزاحة

$$Y_{t-k}, \quad k = 1, 2, \dots, 16$$

اي ان أنموذج الخلية العصبية يحتوي على 16 عقدة ادخال وتمثل

$$Y_t = f(Y_{t-1}, Y_{t-2}, \dots, Y_{t-15}, Y_{t-16}) \quad \dots (9)$$

2-4-3 تحديد عقد الادخال المخفية : Determination of hidden Nodes

ان عدد العقد المخفية تعتمد على عدد عقد الادخال وبهذا فإن انموذج الخلية العصبية سيكون حسب عدد عقد الادخال كما في الجدول (2) :

جدول (2)

يوضح المعايير التي تم اعتمادها لاحتساب عدد عقد الادخال الامثل

النموذج	عدد المعلمات (N _p)	AIC	BIC	AIC _c
NN(1.12;4)	25	12.4	48.508	17.558
NN(1.12;5)	31	-19.3	11.1390	-11.519
NN(1.12;6)	37	-26.985	-23.428	3.626
NN(1.12;7)	43	-66.057	-4.439	-49.387
NN(1.12;8)	49	-113.89	-43.996	-94.0519
NN(1.12;9)	55	-95.62	-14.816	-68.953
NN(1.12;10)	61	-135.073	-47.858	-98.531
NN(1.12;11)	67	-115.129	-22.443	-69.89
NN(1.12;12)	73	-129.27	-24.517	-74.704

علماً ان :

$$AIC = N_t K \ln \hat{\sigma}^2 + 2N_p \quad \dots (10)$$

$$BIC = N_t \ln \hat{\sigma}^2 + N_p \ln(n) \quad \dots (11)$$

$$AIC_c = AIC + \frac{2N_p (N_p + 1)}{N_t - N_p - 1} \quad \dots (12)$$

$$N_p = (n_i + 2)n_u + 1 \quad \dots (13)$$

إذ أن :

N_p : عدد المعلمات في الخلية

N_t : حجم العينة لمجموعة التدريب (278)

K : عدد عقدة وحدة الاخراج

n_i : عدد المتغيرات الداخلة.

n_u : عدد العقد المخفية.

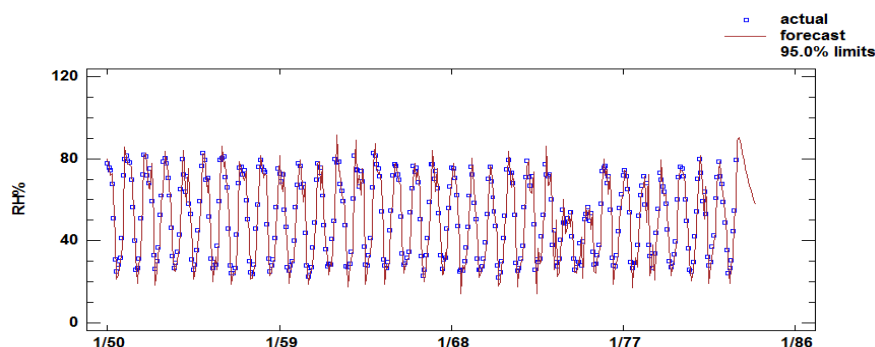
تم حساب المقاييس الاحصائية وكما موضح في جدول (2) وبملاحظة ان الانموذج NN(1.12;10) الذي يضم (12) متغيراً مزاح وهو (y_{t-1}، y_{t-2}، ...، y_{t-12}) و (10) عقد مخفية قد حصل على اقل

قيم حسب المعايير المحسوبة وهي (AIC، BIC، AIC_c) على التوالي، لذلك سنستخدم الانموذج لحساب التقديرات لسنة (2014-2015) .

جدول (3)

تقديرات الرطوبة النسبية % (RH) لسنة (2014-2015)

Month \ Yaer	2014	2015
Jan.	80.00	78.67
Feb.	73.42	71.63
Mar.	58.99	68.01
Apr.	63.96	58.92
May.	53.03	48.64
Jun.	31.89	35.62
Jul.	27.98	24.32
Aug.	29.56	26.50
Sep.	36.88	30.33
Oct.	42.47	44.67
Nov.	71.08	79.75
Dec.	77.2	79.33

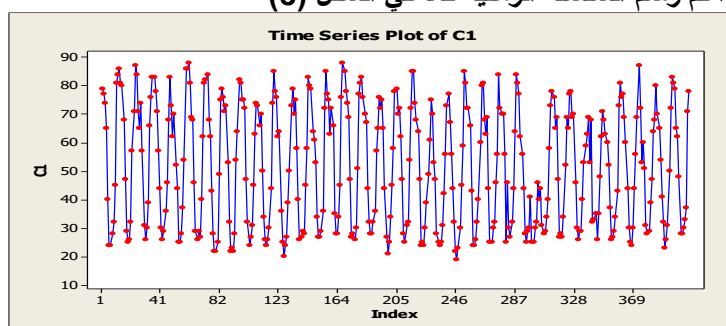


الشكل (5)

يوضح القيم الحقيقية والقيم المتنبأ بها باستخدام الشبكات العصبية حتى عام 2015

5-3 نتائج التنبؤ باستخدام $(ARIMA)(Q,D,P)_s$:

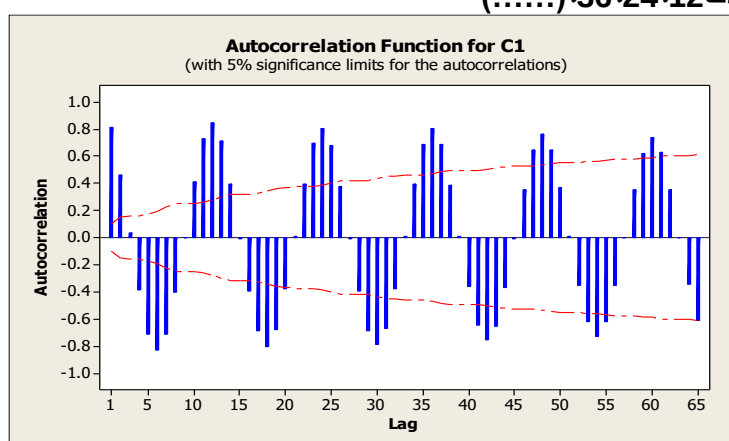
لغرض بناء نموذج للتنبؤ بالرطوبة النسبية في مدينة الموصل والتعرف على رتبة النماذج (AR)، (MA) ينبغي رسم السلسلة الزمنية لمعرفة فيما إذا كانت غير مستقرة بالمتوسط أو بالتباين وتحتوي على تأثيرات موسمية لذا تم رسم السلسلة الزمنية كما في الشكل (6)



الشكل (6)

يمثل رسم السلسلة الزمنية للمعدلات الشهرية للرطوبة النسبية في الموصل من الفترة 1980-2013

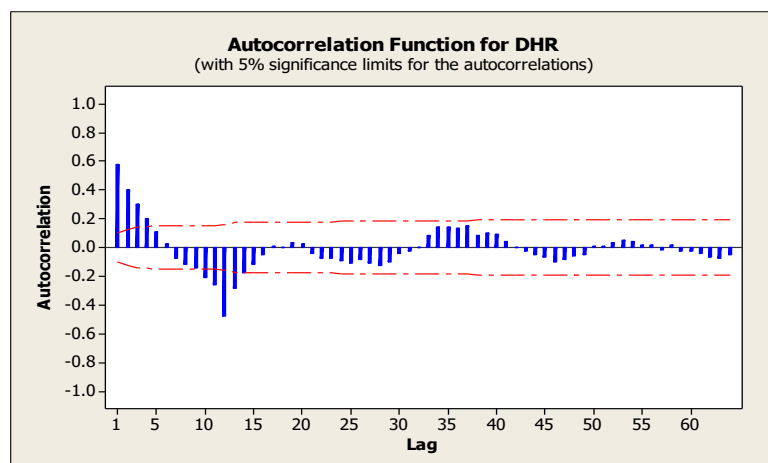
من خلال الرسم نلاحظ ان السلسلة مستقرة بالمتوسط اذ تتذبذب حول متوسط واحد وايضا ذات قيم متجانسة فهي مستقرة بالتباين لكنها تحتوي على تأثيرات موسمية اذ تعيد سلوكيتها كل 12 شهرا وهذا ماتؤكد دالة الارتباط الذاتي للسلسلة كما في الشكل (7) ادناه اذ نجد ان قيم معاملات الارتباط الذاتي معنوية بشكل كبير في الفترات 12، 24، 36،)



الشكل (7)

دالة الارتباط الذاتي للسلسلة الزمنية

لذا تم اخذ الفرق الموسمي للسلسلة الزمنية بموجب المعادلة (6)، وتم حساب معاملات الارتباط الذاتي وكانت دالة الارتباط الذاتي كما في الشكل الاتي:



الشكل (8)

دالة الارتباط الذاتي للسلسلة الزمنية بعد ازالة التأثيرات الموسمية

من خلال هذه الدالة يمكن ان نقترح النموذج الاتي $SARMA(1,0,1)(0,1,1)_{12}$ وقد تم تقديره كما في الجدول الاتي :

جدول (4)

تقدير النموذج المقترح للسلسلة الزمنية

Type	Coef.	SE	T	P-value
AR (1)	0.6975	0.0588	11.86	0.000
MA (1)	0.1531	0.0815	1.88	0.061
SMA (12)	0.9452	0.0224	42.14	0.000

Differencing: 0 regular, 1 seasonal of order 12

Number of observations: Original series 408, after differencing 396

Residuals: SS = 15241.4 (backforecasts excluded)

MS = 38.8 , DF = 393

كما تم اختباره كما في الجدول الاتي:-

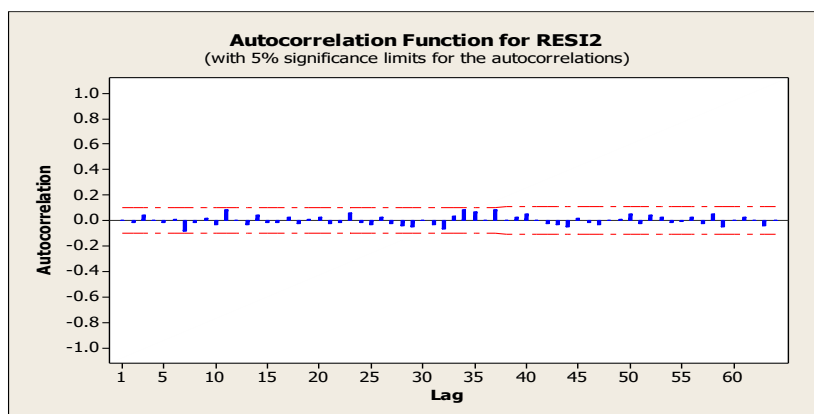
جدول (5)

يمثل اختبار النموذج المقترح

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	7.1	11.1	22.3	29.5
DF	9	21	33	45
P-Value	0.625	0.961	0.921	0.964

اذ يلاحظ من الجدول بان قيمة مربع كاي غير معنوية عند الازحاحات (12,24,36,48) وبالتالي تقبل الفرضية اي ان البواقي ماهي الا متغيرات عشوائية غير مترابطة وعليه فالنموذج جيد في تمثيل البيانات وهذا ما تؤكد دالة الارتباط الذاتي المستخرجة لهذه البواقي كما في الشكل الاتي:-



الشكل (9)

دالة الارتباط الذاتي لبواقي السلسلة الزمنية

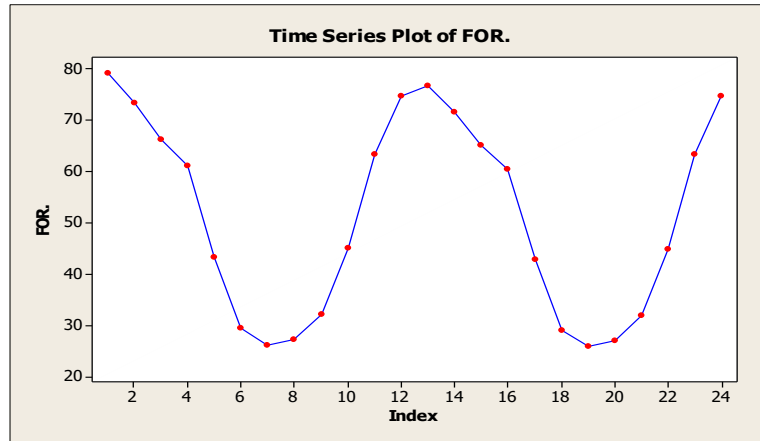
وقد تم حساب التنبؤات المستقبلية للمعدلات الشهرية للرطوبة النسبية في مدينة الموصل للفترة 2015-2014 مع حدود الثقة الدنيا والعليا وباحتمال ثقة 95%.

جدول (6)

القيم المتنبأ بها باستخدام النموذج المقترح

Period	Forecast	Lower	Upper
409	79.0949	66.8864	91.3033
410	73.2657	59.3656	87.1658
411	66.1412	51.4885	80.7938
412	61.1466	46.1414	76.1517
413	43.3230	28.1494	58.4967
414	29.3672	14.1123	44.6222
415	26.1073	10.8129	41.4017
416	27.2401	11.9266	42.5536
417	32.0264	16.7036	47.3492
418	44.9990	29.6716	60.3263
419	63.3558	48.0263	78.6853
420	74.7725	59.4420	90.1031
421	76.8000	61.4488	92.1512
422	71.6650	56.3072	87.0229
423	65.0247	49.6636	80.3858
424	60.3679	45.0052	75.7305
425	42.7799	27.4165	58.1433
426	28.9884	13.6246	44.3522
427	25.8430	10.4791	41.2070
428	27.0558	11.6917	42.4199
429	31.8978	16.5337	47.2619
430	44.9093	29.5452	60.2734
431	63.2933	47.9291	78.6574
432	74.7289	59.3648	90.0931

كما تم رسمها كما في الشكل ادناه ويتضح ان السلسلة لازالت موسمية مما يعني ان النموذج جيد وكفوء في تمثيل البيانات.



الشكل (9)

رسم التنبؤات لمعدلات الرطوبة النسبية للفترة 2014-2015

3-5-3 اختبار جودة الطريقتين :

للقوف على أي من الطريقتين هي الأفضل في التنبؤ بمعدلات الرطوبة النسبية (RH%) الشهرية في محافظة الموصل تم استخراج المقاييس الآتية وكما موضحة في جدول (7) ، والذي يوضح على أن استخدام السلاسل الزمنية الموسمية أفضل للتنبؤ بمعدلات الرطوبة النسبية .

جدول (7)

المقاييس الاحصائية المستخدمة في اختبار جودة الطريقتين

	RMSE	MAE	MAPE	MAD
أنموذج $SARMA(12,0,1)$ (1,1,1)	27.23863	5.560065	23.1669	0.23166
أنموذج الشبكات العصبية (12,10,1)	262.6302	9.079167	37.829	0.37829

إذ أن :

$$MAE = \sum_{t=1}^n |Y_t - \hat{Y}_t| / n \quad \dots (14)$$

$$MAD = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{n} \right| \quad \dots (15)$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| \times 100 \quad \dots (16)$$

$$RMSE = \sqrt{MSE} = \sqrt{\frac{\sum (Y_t - \hat{Y}_t)^2}{n}} \quad \dots (17)$$

الفصل الرابع

الاستنتاجات والتوصيات

1-4 الاستنتاجات :

1. استناداً إلى النتائج المستخرجة في استخدام السلاسل الزمنية الموسمية فقد كان أفضل نموذج موسمي من الدرجة $SARMA(12,0,1)$ وذلك بالاعتماد على معيار (AKIAKE) كما في جدول (2) .
2. المعادلة المقدرة بالاستناد إلى النموذج في (1) كانت معلماتها جميعها معنوية أي لها تأثير معنوي .
3. أما بالنسبة لنتائج أنموذج الشبكات العصبية الاصطناعية فقد كان أفضل أنموذج $NN(1,12,10)$ وكما موضح في الجدول (4)، إذ حسب معيار (AIC، BIC، AIC_c) حصل على أقل قيم بين النماذج الأخرى .
4. تعد نماذج السلاسل الزمنية الموسمية أفضل في التنبؤ من نماذج الشبكات العصبية وذلك يتضح من خلال الجدول (7) إذ حصلت نماذج السلاسل الزمنية الموسمية على أقل قيم الاختبار حسب معايير (MAE، MAD، MAPE، RMSE) على التوالي .

2-4 التوصيات :

على ضوء الاستنتاجات يمكن التوصل الى التوصيات الاتية :

1. نقترح استخدام أنموذج السلاسل الزمنية الموسمية $SARMA(1,0,1)(0,1,1)_{12}$ للتنبؤ بمعدلات الرطوبة النسبية (RH%) الشهرية لمدينة الموصل من قبل الدوائر ذات العلاقة .
2. توسيع الدراسة لتشمل جميع محافظات العراق للتنبؤ بمعدلات الرطوبة النسبية (RH%) الشهرية .
3. التنبؤ بعناصر المناخ الأخرى (الامطار، الضغط الجوي) باستخدام اسلوب نماذج السلاسل الزمنية الموسمية .

المصادر

1. المحمدي، أ.د. ناظم عبد الله، طعمة، سعدية عبد الكريم، "استخدام نماذج السلاسل الزمنية الموسمية للتنبؤ باستهلاك الطاقة الكهربائية في مدينة الفلوجة"، جامعة الانبار، كلية الادارة والاقتصاد/مجلة جامعة الانبار للعلوم الاقتصادية والادارية .
2. الساعدي، علي احمد حسن: "استخدام نماذج السلاسل الزمنية الهجينة وغير الهجينة للتنبؤ باعداد المسافرين لمطار بغداد الدولي"، رسالة ماجستير في علوم الاحصاء، الجامعة المستنصرية/كلية الادارة والاقتصاد، قسم الاحصاء، 2016.
3. عبد العزيز، أ.م.د. بثينة عبد الجادر؛ الوكيل، أ.م.د. خولة حسين: "التنبؤ بالمعدلات الشهرية لدرجات الحرارة العظمى والصغرى في مدينة بغداد باستخدام الشبكة العصبية المتعددة الطبقات"، قسم الاحصاء، كلية الادارة والاقتصاد/الجامعة المستنصرية، بحث منشور.
4. مقدمة في الشبكات العصبية الاصطناعية وتطبيقاتها في العلوم الاجتماعية باستخدام SPSS ، معهد الدراسات والبحوث الاحصائية، قسم الاحصاء الحيوي والسكاني ، جامعة القاهرة .
5. ناظم، م.م. ايفان علاء، " استخدام نماذج الشبكات العصبية الاصطناعية للتنبؤ والمقارنة "، وزارة التعليم العالي والبحث العلمي، دائرة البحث والتطوير .
6. الهيئة العامة للأنواء الجوية العراقية والرصد الزلزالي، وزارة النقل والمواصلات، بيانات معدلات الرطوبة النسبية (RH%) الشهرية للمدة (1980-2013) ، قسم المناخ ، بيانات غير منشورة .
7. Julian Faraway، chri Chatfield، 2004. "Time Series Forecasting with neural network: a comparative study using the airline data and ANN models"، African Journal of Science and Technology (AJST)، Science and Engineering series Vol.5، No.2، PP41-49.
8. Makridakis، S.، Wheelwright، S.C. and Hyndman، R.J.، (1998) ، "Forecasting Methods and Application"، 3rd . ed.، Johnwiley -Sons .
9. Rzepolunk، E.J.، (1998)، "Neural Network Data Analysis Using Stimulant "، Springer.