أستخدام نماذج السلاسل الزمنية الهجينة وغير الهجينة للتنبؤ بأعداد المسافرين لمطار بغداد الدولي

أ.م.د. بثينة عبد الجادر عبد العزيز* علي احمد حسن**

المستخلص

يعد التنبؤ في السلاسل الزمنية من الموضوعات المهمة في الأساليب الأحصائية ولكافة المجالات الحيوية إذ يساعد الإدارات في التخطيط الأقتصادي والستراتيجي وإتخاذ القرارات المستقبلية ، لذا تناول هذا البحث دراسة لأستخدام نماذج السلاسل الزمنية الهجينة المتولدة من دمج انموذج ARIMA مع أنموذج الشبكات العصبية المتعددة الطبقات او بالعكس على فرض أن السلسلة تضم المركبتين الخطية وغير الخطية اذ يستعمل أنموذج ARIMA لتمثيل المركبة الخطية للسلسلة الزمنية وأنموذج الشبكات العصبية المتعددة الطبقات لتمثيل المركبة غيرالخطية للسلسلة اذ أن الأنموذج الهجين الأول (ARIMA-ANN) والثاني (ANN-ARIMA) وقد تم استخدام هذين النموذجين مع النماذج المفردة للتنبؤ بأعداد المسافرين في مطار بغداد الدولي لغرض المقارنة بين هذه النماذج من خلال جملة من المقاييس الأحصائية وقد اظهرت النتائج المستخرجة تفوق الأنموذج الهجين (ANN-ARIMA) على النماذج الأخرى لأمتلاكه أقل القيم لهذه المقاييس وقد تم أستخدامه لحساب التنبؤات المستقبلية لأعداد المسافرين للفترة الزمنية من أيلول / 2015 ولغاية كانون الأول / 2016 وبشكل شهري ، وقد تم تحليل البيانات وأستخراج النتائج بالأعتماد على البرامج الأحصائية الجاهزة (Minitab 16 , SPSS 19) .

Abstract

Prediction in time series is one of the most important topics in statistical procedures and all vital areas it can help managements in the economic and strategic planning and decision-making . This research use time series hybrid models generated from the integration of Box- Jenkins (ARIMA) model and multi-layered neural networks as The first hybrid model (ARIMA-ANN) and second (ANN-ARIMA) model and assuming series includes two components linear and non-linear with the single models such as Box- Jenkins model in which time series is a linear combination , and multi-layered neural networks in which time series has a nonlinear combination to predict the number of passengers in International Baghdad Airport , when comparing these models through a number of statistical criteria ,they show that the hybrid model (ANN-ARIMA) is the superior model because it has lower values of these criteria , and it has been used to calculate the prediction of passenger numbers for the time period from September 2015 until December 2016 , where the data were analyzed and the results were obtained based on the statistical programs package (Minitab 16 , SPSS 19) .

* الجامعة المستنصرية / كلية الادارة والاقتصاد .

** باحث .

مقبول للنشر بتأريخ 2015/12/28

مستل من رسالة ماجستير

المقدمة Introduction

نظراً للتطور والنمو السكاني الحاصل في العراق وإزدياد حركة السياحة الدينية والترفيهية في الأونة الأخيرة فقد تم التطرق إلى موضوع التنبؤ لأعداد المسافرين لمطار بغداد الدولي لإكتسابه أهمية كبيرة من حيث حركة المسافرين فيه ، حيث أن المنشأة العامة للطيران المدني شأنها شأن كافة الدوائر والمؤسسات الحكومية تحتاج الى تنبؤات دقيقة لحركة المسافرين والتي تساعدها في عملية ادارة المطار والتخطيط للمستقبل للخدمات المقدمة والحاجة لأنشاء مطارات اخرى وحسب الحاجة التي تبينها حركة المسافرين .

توجد اساليب احصائية كثيرة تستخدم لحساب التنبؤات المستقبلية لظاهرة معينة حيث أهتم الكثير من الباحثين الأحصائين بدراسة تحليل السلسلة الزمنية إذ تم وضع أنموذج خطي لها واخر غير خطي لتمثيل السلسلة كأنموذج بوكس– جنكينز وأنموذج الشبكات العصبية على التوالي ، والتنبؤ بقيمها المستقبلية ، وقد شاع استخدام هذه النماذج في مجالات حيوية مختلفة ، إذ ساعدت البرامج الأحصائية الجاهزة مثل (MINITAB,SPSS) وغيرها من استعمال هذه النماذج .

وفي الوقت نفسه ظهرت بحوث ودراسات اخرى تهتم بتنبؤات السلسلة الزمنية تؤيد فكرة دمج الأنموذج الخطي (Linear model) مع الأنموذج غير الخطي (Non-Linear model) لأنتاج أنموذج جديد هو الأنموذج الهجين واستخدامه لأيجاد تنبؤات دقيقة للسلسلة ، أن فكرة الدمج بين الأنموذج الخطي وغير الخطي متأتية من الأفتراض بأن الأنموذج المفرد (single model) لايمكن أن يوظف البيانات الحقيقية لهذه السلسلة إذ أن الأنموذج المفرد ربما يكون غير كافٍ لتشخيص كل الخواص والمميزات (characteristic) للسلسلة الزمنية ، فالسلسلة الزمنية ربما تتضمن كلا المركبتين الخطية وغير الخطية خلال مدة من الزمن وفي هذه الحالة لايمكن للأنموذج الخطي وغير الخطي أن يعالج هاتين المركبتين معاً ، لذا تم اقتراح تقنية جديدة الأ وهي النماذج الهجينة (Hybrid models) .

درس عدد من الباحثين مسألة الدمج بين النماذج الخطية (Linear models) المستخدمة لتمثيل السلاسل الزمنية الخطية كنماذج بوكس- جنكينز مع النماذج غير الخطية (Non-Linear models) المستخدمة لتمثيل السلاسل الزمنية غير الخطية كنماذج الشبكات العصبية الأصطناعية (ANN) لأنتاج النماذج الهجينة (Hybrid models) ، وذلك للحصول على افضل وادق التنبؤات المستقبلية للسلسلة الزمنية ، واهم وابرز البحوث والدراسات الحديثة التي تناولت هذه النماذج من قبل الباحث(Zhang) [16] والباحثين [14](Šterba Ján and Hilovsda Katarína) والباحثين(N.Merh) واخرين [13] و الباحثين (Zheng and Zhong)[6] و غيرهم من الباحثين .

هدف البحث Objective of Research

يهدف البحث الى دراسة النماذج الهجينة وغير الهجينة للسلسلة الزمنية وتحديد الأنموذج الأفضل واستعماله للتنبؤ باعداد المسافرين لمطار بغداد الدولي بالأعتماد على جملة من المقاييس الأحصائية (MAD, RMSE , SEP , MAPE , MAE).

الجانب النظري

1.الأنموذج الهجين[16],[13] Hybrid Model

أن الأنموذج الهجين الذي سيتم دراسته هو الأنموذج الناتج والمتولد من دمج أنموذج خطي مع اخر غير خطي للسلسلة وبالعكس لغرض تحسين التنبؤات المستقبلية للسلسلة على افتراض أن السلسلة الزمنية Z_t مؤلفة من مركبتين الأولى خطية Linear (L) والثانية غير خطية Non-Linear(N) ، وكما موضح بالمعادلةالأتية:-

$$Z_t = L + N \quad , t = 1, 2, 3, 4, \dots \quad , \dots (1)$$

إذ أن Z_t : تمثل السلسلة الزمنية .

L : تمثل المركبة الخطية في السلسلة الزمينة ، N : تمثل المركبة غير الخطية في السلسلة الزمينة .

إذ سيتم دراسة نوعين من النماذج الهجينة المتولدة بهذه الطريقة وهي كالأتي :-

1.1. الأنموذج الهجين (ARIMA-ANN) Hybrid Model (ARIMA-ANN)

بافتراض السلسلة الزمينة مؤلفة من بنية الأرتباط الذاتي الخطي والعناصر غير الخطية ، ولتحليل هذه السلسلة بأستخدام الأنموذج الهجين يكون بمرحلتين :-

المرحلة الأولى :

يتم استخدام احد نماذج (ARIMA) لتمثيل الجزء الخطي للسلسلة الزمنية ومن ثم حساب التنبؤات ،وأن الأخطاء العشوائية المقدرة (البواقي) المستخرجة من أنموذج ARIMA اذ ان هذه البواقي تكون ضرورية جداً لتدقيق تشخيص الأنموذج الخطي .

أن الباحث (Zhang) هو أول من طرح فكرة دمج هذين الأنموذجين وسار عليه الباحثون الأخرون لذا فاعن هذه البواقي بأعتبارها الجزء غير الخطي للسلسلة إذ تمثل بأنموذج الشبكات العصبية المتعددة الطبقات لحساب التنبؤات لها.

اذ ان هذه البواقي تحسب من المعادلة الأتية :

$$a_t = Z_t - \widehat{Z}_t, \dots (2)$$

إذ أن $\widehat{Z}_t$: تمثل تقديرات السلسلة الزمنية بأستخدام انموذجARIMA . a_t : تمثل البواقي .

أن هذه البواقي (الأخطاء العشوائية) تمثل المركبة غيرالخطية (N) للسلسلة التي يتم تمثيلها بأنموذج الشبكات العصبية المتعددة الطبقات (ANN) للحصول على القيم المتوقعة لها وكما في المعادلة الأتية:

$$\mathrm{N} = f\left(a_{t-1}, a_{t-2}, \dots \dots \dots, a_{t-p}\right) + \varepsilon_t, \dots (3)$$

إذ أنε_t : الخطأ العشوائي ، f : دالة غير خطية تحدد من خلال الشبكة العصبية .

المرحلة الثانية :

في هذه المرحلة يتم جمع القيم المتنبأ بها للسلسلة بواسطة أنموذج (ARIMA) مع توقعات البواقي التي تم حسابها من قبل أنموذج (ANN) ، والتي بدورها تنتج القيمة المتوقعة النهائية للسلسلة الزمنية كما في المعادلة الاتية:

$$\widehat{Z}_t = \widehat{L} + \widehat{N} \qquad \dots (4)$$

2.1.الأنموذج الهجين Hybrid Model (ANN-ARIMA) (ANN-ARIMA)

هذا النوع من النماذج الهجينة يتكون من مرحلتين ولكن عكس الأنموذج الهجين السابق اذ يكون في:

المرحلة الأولى :

يتم استخدام أنموذج الشبكات العصبية الصناعية (ANN) لتمثيل الجزء غير الخطي للسلسلة الزمنية ومن ثم التنبؤ بالقيم المستقبلية ثم تستخرج البواقي من انموذج الشبكة العصبية وتمثيلها بأنموذج (ARIMA) وحساب القيم التنبؤية للأخطاء كما مبين ادناه :

المركبة غيرالخطية (N) تمثل بأنموذج الشبكات العصبية الصناعية (ANN) وحسب المعادلة الأتية .

$$\mathrm{N} = f\left(Z_{t-1}, Z_{t-2}, \dots \dots \dots, Z_{t-p}\right) + a_t, \dots (5)$$

إذ أن$\widehat{N}$: تمثل تقديرات أنموذج الشبكات العصبية الصناعية (ANN) .

أن البواقي المستخرجة (a_t) تمثل المركبة الخطية (L) للسلسلة الزمنية التي يتم تمثيلها بأنموذج ARIMA وحساب القيم المتنبأ بها .

المرحلة الثانية :

يتم جمع القيم التي تم التنبؤ بها بواسطة (ANN) مع القيم المتنبأ بها للاخطاء التي تم حسابها من قبل أنموذج (ARIMA)، والتي بدورها تنتج القيمة المتوقعة النهائية للسلسلة الزمنية كما في المعادلة الاتية(4).

2. تمثيل المُركبة الخطية للسلسلة الزمنية بأنموذج ARIMA

تعد نماذج بوكس- جنكينز من الأساليب الأحصائية المهمة إذ تستخدم لتمثيل بيانات سلسلة زمنية لظاهرة معينة والتنبؤ بقيمها المستقبلية، وتكون على نوعين وفق استقرارية السلسلة الزمينة

1.2. نماذج السلاسل الزمنية المستقرة

أنموذج الأنحدار الذاتي والأوساط المتحركة

{11} Autoregressive– Moving average model

أن هذا الأنموذج يجمع بين أنموذجي الأنحدار الذاتي(AR) والاوساط المتحركة(MA) إذ يرمز له بالرمز ARMA(p,q)، والصيغة العامة لهذا الأنموذج كالأتي :

$$Z_t = \emptyset_1 Z_{t-1} + \dots + \emptyset_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad , \dots (6)$$

إذ أن $\emptyset_1,\dots,\emptyset_p$: تمثل معلمات أنموذج الأنحدار الذاتي،(p) : رتبة الأنحدار الذاتي

θ1,...,θq : تمثل معلمات أنموذج الأوساط المتحركة، (q) : رتبة المتوسطات المتحركة

a_t: يمثل الخطاء العشوائي الذي يطلق عليها بالتشويش الأبيض (White noise)و يتوزع بمتوسط صفر وتباين $\sigma_a{}^2$.

وعندما تكون (q=0) نحصل على أنموذج الأنحدار الذاتي AR(p) وعندما (p=0) نحصل على أنموذج الأوساط المتحركة MA(q).

2.2.نماذج السلاسل الزمنية غير المستقرة[15]

أنموذج الأنحدار الذاتي والأوساط المتحركة التكاملي

Autoregressive integrated Moving average model [11]

عندما تكون السلسلة الزمنية غير مستقرة يمكن تحويلها الى سلسلة مستقرة بأخذ الفروق من الدرجة d عندئذ يرمز له بالرمز ARIMA(p,d,q) وتكتب معادلته :

$$\emptyset(B)(1-B)^d Z_t = \theta(B)a_t \ , \dots (7)$$

إذ أن : (d) : درجة الفرق المأخوذ للسلسلة الزمنية ، (B) : عامل الأرتداد الخلفي

3. مراحل بناء أنموذج السلسلة الزمنية [11]

أن مراحل بناء الأنموذج هي كما يلي :-

1.3 المرحلة الأولى: التشخيص Identification[15]

تتم مرحلة تشخيص أنموذج السلسلة الزمنية المستقرة من خلال :

1. دالةالأرتباط الذاتي والأرتباط الذاتي الجزئي
2. دالة الأرتباط الذاتي الموسعة للعينة The Extended Sample Autocorrelation Function

وعند فشل الاساليب اعلاه في تشخيص الانموذج نلجا الى الاعتماد على المقاييس الاتية:

معيار اكيكي Akaike Information (AIC) [15]

وهو معيار يستخدم لتشخيص درجة الأنموذج ARMA(p,q) وحسب الصيغة الأتية .

$$AIC = n\ln(\hat{\sigma}_a^2) + 2m \ , \dots (8)$$

إذ أنm : عدد المعلمات في الأنموذج، n : عدد المشاهدات ، $\hat{\sigma}_a^2$: تباين البواقي

معيار شوارز Schwarz Criterion (SC) [15]

ويعرف أيضاً بمعيار Beyesion Information Criterion (BIC) وحسب الصيغة الأتية:-

$$BIC(m) = n\ln(\hat{\sigma}_a^2) + m\ln(n) \ , \dots (9)$$

ولقد اقترح كل من (Brock Well and Davis . 1993) تصحيح حالة التحيز في معيار (AIC) فأصبح :

$$AIC_c(m) = n\ln(\hat{\sigma}_a^2) + \frac{2mn}{(n-m-1)} \ , \dots (10)$$

ووفق هذه المعايير فاءن الأنموذج الأفضل هو الأنموذج الذي يعطي اقل قيمة منAIC(m) و BIC(m) و $AIC_C(m)$.

2.3 المرحلة الثانية : التقدير Estimation [4] [15]

بعد أن تم التعرف على الأنموذج المقترح الذي يمثل بيانات السلسلة الزمنية قيد الدراسة وتحديد الرتبة المناسبة له ، يتم تقدير معلمات الأنموذج المختار، وهناك عدة طرائق للتقدير منها طريقة الأمكان الأعظم (Maximum likelihood).

3,3.المرحلة الثالثة : دقة التشخيص Diagnostic checking[11]

لغرض تدقيق صحة الأنموذج المقترح لتمثيل بيانات السلسلة الزمنية تستخرج البواقي والتي ينبغي أن تكون متغيرات عشوائية غير مترابطة عندئذ تكون فرضية العدم .

$$H0: \rho = 0 \ , \dots (11)$$

ولأختبار هذه الفرضية ينبغي استخراج مايلي :

1.3.3.احصاءة Box-Ljung(Q) [11]

ولقد قام كل من (Ljung ,G.M&Box,G.E.P.) بتعديل صيغة اختبار Q الأصلية بالشكل الأتي

$$Q = n(n+2)\sum_{k-1}^{m}(n-k)^{-1}r_k^2(\hat{a}) \sim x_{(m)}^2 \ , \dots (12) \qquad ,k=1,\dots,m$$

اذ ان: n : يمثل عدد المشاهدات للأنموذج المشخص ، و: n= N - d

N : عدد المشاهدات الأصلية للسلسلة الزمنية ،m : تمثل $\sqrt{n}$ و d معرفة سابقا.

2.3.3.إذا كانت معاملات الأرتباط الذاتي للبواقي واقعة ضمن حدود الثقة بأحتمال 95% أي أن:

$$-1.96\frac{1}{\sqrt{n}} \le \hat{r}_k(\hat{a}) \le 1.96\frac{1}{\sqrt{n}}, \dots (13)$$

4.3. المرحلة الرابعة : التنبؤ Forecasting {4}

أن احد الأهداف الأساسية لتحليل السلسلة الزمنية هو التنبؤ ،فمثلاً إذا اردنا التنبؤ بقيمة السلسلة الزمنية في الفترة (t+L) والتي هي $\widehat{Z}_t(L)$ اذ تستخرج هذه القيمة باخذ التوقع الشرطي الى Z_t عن الزمن(L+t)

$$\widehat{Z}_t^L = E_t(Z_{t+L} \backslash Z_1, Z_2, Z_3, \dots\dots, Z_n), \dots (14)$$

الذي يجب أن يكون له اقل متوسط مربعات خطاء التنبؤ (M.M.S.E.F.) والذي يساوي

$$\text{M. M. S. E. F} = \sigma_a^2 \left\{ 1 + \sum_{j=1}^{L-1} \varphi_j^2 \right\}, \dots (15)$$

إذ أن φ_j :- تمثل معلمات الأنموذج عند كتابته بصيغة الأخطاء العشوائية .

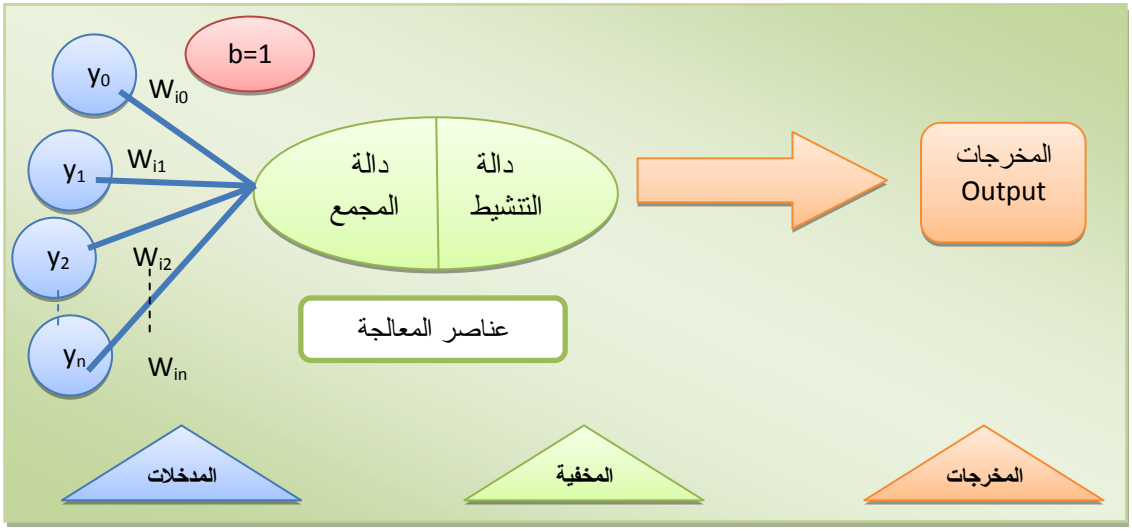
4. تمثيل المركبةغيرالخطية للسلسة الزمنية بأنموذج الشبكات العصبية المتعددة الطبقات

1.4. الشبكة العصبية الأصطناعية Artificial Neural Network {1},{12}

تعرف الشبكة العصبية الأصطناعية (ANN) على أنها نظام حسابي مكون من عدد من وحدات (عصبونات) المعالجة (Processing Elements) المترابطة مع بعضها وتتصف بطبيعتها الديناميكية والمتوازية في معالجة البيانات الداخلة اليها ، أن فكرة عمل الشبكات العصبية الأصطناعية هو محاكاة البيانات للوصول الى أنموذج لهذه البيانات لغرض التحليل أو التصنيف أو التنبؤ أو أي معالجة اخرى دون الحاجة الى أنموذج مقترح لهذه البيانات ، أن الشبكات العصبية الأصطناعية هي تقنيات حسابية مصممة لمحاكاة الطريقة التي يؤدي بها الدماغ البشري مهمة معينة وذلك عن طريق معالجة ضخمة موزعة على التوازي ومكونة من وحدات معالجة بسيطة ، هذه الوحدات ما هي الأ عناصر حسابية تسمى عصبونات أو عقد (Nodes , Neurons) والتي لها خاصية عصبية وكما في الشكل الأتي .

الشكل (1)

يوضح مفهوم بسيط للشبكة العصبية الأصطناعية



ويمكن الملاحظة من خلال الشكل (1) أن الشبكة العصبية الأصطناعية تتكون من عناصر المعالجة التي تحتوي على جزئين هما :

الأول - دالة المجمع Sum function

وتحدد طريقة وصيغة ادخال المعلومات الى الشبكة العصبية والتي تعرف بالمدخلات وغالباً ماتكون عبارة عن تركيبة خطية (Linear combination) بدلالة الأوزان وتوصف :-

$$n = b + \sum_{j=1}^{R} w_j y_j \quad , \dots (16)$$

إذ أن n :- نتائج مدخلات التركيبة الخطية ، b :- الحد المطلق (bias)

w_j :- الأوزان المرتبطة بالمدخلات وهي تقابل المعلمات في أنموذج الأنحدار، y_j تمثل متغيرات الأدخال

الثاني - دالة التنشيط Activation Functions{10}

تعرف بدالة التحويل (Transfer Function) لأنها تقوم بتحويل المدخلات عبر تفاعلها مع الأوزان من صيغة الى اخرى وتكون منها الخطية وغير الخطية ومنها الدوال الشائعة الأستخدام هي :-

الدالة اللوجستية Log-Sigmoid Function

تعد هذه الدالة الأكثر استخداماً عند عقد الطبقة المخفية للشبكة التي مدخلاتها ذات قيم حقيقية بين $(-\infty, \infty)$ أما مخرجاتها لكل عقدة تكون بين (0,1) والصيغة الرياضية لها هي :

$$f(x) = \frac{1}{1 + exp(-x)} \quad , \dots (17)$$

ومشتقتها

$$f'(x) = f(x)\left[1 - f(x)\right] \quad , \dots (18)$$

الدالة الخطية The Linear Function

تستخدم هذه الدالة عند طبقة الأخراج في حالة التنبؤ وصيغتها :-

$$Output = R(Net) \quad , \dots (19)$$

$$Net = X_1W_1 + X_2W_2 + \dots + X_nW_n \quad , \dots (20)$$

أذ أن (R):- دالة تنشيط خطية .

2.4. معايرة البيانات Data Normalization {7}

هي عملية تهيئة البيانات قبل معالجتها (Pre – Processing) لاستخدامها في عملية تدريب الشبكة العصبية، وتوجد عدة طرق للتهيئة منها :-

الصيغة المعيارية Standardized Formula

تستخدم هذه الصيغة عندما تكون دالة التنشيط في عقدة الأخراج هي الدالة الخطية وحسب الصيغةالأتية.

$$X_{new} = \frac{X - \overline{X}}{S}, \dots (21)$$

إذ أن $\overline{X}$:- الوسط الحسابي ، S :- الأنحراف المعياري

3.4.الشبكات العصبية المتعددة الطبقات Multi-layered neural networks{1},{7}

تتكون هذه الشبكات من طبقة واحدة أو اكثر من الطبقات المخفية (Hidden Layer) التي بدورها تحتوي على عدد من العقد المخفية (Hidden Nodes) ، ويمكن أن يبين أن الشبكة العصبية الأصطناعية تتكون من ثلاث مستويات وهي :-

مستوى المدخلات(Inputs Level) :

يمثل المستوى الأول في الشبكة ويحتوي على عدد من العقد التي تمثل عدد المتغيرات المفسرة (المدخلات) ويتم فيه تلقي المعلومات المطلوب معالجتها .

المستوى المخفي (Hidden Level) :

يمثل المستوى الثاني للشبكة العصبية والذي يلي المستوى الأول(مستوى المدخلات) وقد يحتوي على عدد من العقد المخفية (Hidden Nodes) ولكل عقدة وزن يربطها مع المستوى السابق (مستوى المدخلات) ووزن يربطها مع المستوى اللاحق (مستوى المخرجات).

مستوى المخرجات (Output Level) :

يمثل المستوى الأخير في الشبكة العصبية الأصطناعية الذي هو عبارة عن مخرجات الشبكة العصبية ويمكن أن يكون للناتج عقدة واحدة أو اكثر .

4.4. عمل الشبكة العصبية متعددة الطبقات Work of Multi-layered N.N.{9}

أن القيمة المراد التنبؤ بها والتي يطلق عليها بالنتائج (output) يمكن أن تستخرج بشكل عام فاعن أي أنموذج للشبكات العصبية بطبقة مخفية واحدة وبعدد عقد مخفية (r) وعدد عقد الأدخال (p) يكتب بالشكل الأتي :

$$\hat{y}_t = \varphi_0\left\{w_o + \sum_{j=1}^{r} w_j * f\left(w_{1j} + \sum_{i=1}^{p} w_{ij}y_{t-i}\right)\right\}, \dots (22)$$

إذ أن w_{1j} :- تمثل الأوزان التي تربط الثابت (Baise) وبين العقد المخفية .

w_o:الوزن الذي يربط الثابت مع الناتج output . w_j :الأوزان التي تربط بين العقد المخفية والناتج .
w_{ij} :- الأوزان التي تربط بين عقد المدخلات والعقد المخفية .
φ_0 :- دالة التنشيط (التفعيل في مستوى المخرجات) وتكون خطية .
f :- دالة التنشيط (التفعيل في مستوى الخفي) وعادة ماتكون لوجستية .
y_{t-i} :- تمثل متغيرات الأدخال ، r :- عدد العقد المخفية ، p :- عدد عقد الأدخال
ويستخدم الرمز ($j_1,\dots,j_p$; h) NN للتعبير عن أنموذج الشبكة العصبية إذ أن ($j_1,\dots,j_p$) تمثل المدخلات المزاحة زمنياً عند الأزاحة ($J_p,\dots,J_1$) و (h) عدد العقد المخفية ،أما عدد الأوزان (المعلمات) لأنموذج الشبكة العصبية فتستخرج بواسطة المعادلة الأتية :$P = (ni + 2)h + 1$, ... (23)
إذ أن ni :- عددعقد الأدخال ، h :- عدد العقد المخفية

5.4. خوارزمية الأنتشار العكسي للخطأ ومنهجيتها Error Back Propagation Algorithm and Approach[12]

أن منهجية الأنتشار العكسي (EBP) للخطأ طورت في عام 1970 من قبل عدة علماء إذ أن الخطوات الأساسية فيها هو حساب خطأ مستوى المخرجات لتحديث أوزان طبقة المستوى المخفي-المخرجات ، ثم حساب خطأ مستوى المخفي لتحديث أوزان طبقة مستوى المدخلات- المخفي ، وبعد ذلك نحسب مخرجات الشبكة بالأوزان الجديدة لتستمر العملية في حساب الخطأ وتحديث الأوزان للوصول الى اقل خطأ في الشبكة العصبية ، أن خوارزمية الأنتشار العكسي تستخدم في تدريب الشبكات العصبية كاملة الأرتباط ومتعددة الطبقات وغير الخطية ، ويتم تنفيذ هذه الخوارزمية من خلال ثلاث مراحل وكالأتي :-
1- مرحلة الأنتشار الأمامي للخطأ 2- مرحلة الأنتشار الخلفي للخطأ 3- مرحلة توليف أوزان الشبكة .
تبدأ العملية بحساب قيمة الناتج (output) بموجب أنموذج الشبكات العصبية ذات الطبقة المخفية الواحدة بالأعتماد على القيم الأبتدائية للأوزان وعلى المدخلات من دون أضافة الخطأ ، ثم يتم أيجاد الخطأ بين المخرج المطلوب والمخرج الفعلي وترجع بهذا الخطأ عكسيًا من الطبقة الأخيرة الى الطبقة الخلفية ثم اخيرًا الى طبقة المدخلات ، وفي اثناء الرجوع (الأرتداد) يتم تغير الأوزان في الأتجاه الذي يرجع الخطأ الى النقصان واحيانًا الى الصفر وهذه مرحلة الأنتشار الخلفي للخطأ ، وتستعمل هذه الطريقة التدريبية مع الشبكات ذات التغذية الأمامية إذ أن التغذية الأمامية يرجع الى بنية الشبكة أما أنتشار الخطأ العكسي فيرجع الى طريقة التدريب التي تستعمل مع هذه الشبكة .

6.4. التنبـــــــؤ Forecasting [7]

أن الهدف الأساس من استخدام أنموذج الشبكة العصبية هو حساب القيم المستقبلية للسلسلة الزمنية المدروسة ، ففي حالة التنبؤ لخطوة واحدة (One-Step-ahead) يكون بأستخدام المشاهدات الفعلية لكل المتغيرات المزاحة كمدخلات .
أما في حالة التنبؤ لعدة خطوات ولتكن (k-Step) فيكون بأستخرج التنبؤ لخطوة واحدة ثم تدخل هذه القيمة للتنبؤ لخطوة ثانية وهكذا الى أن نحصل على (k) من التنبؤات وهي طريقة تكرارية .

7.4. مراحل التنبؤ بواسطة الشبكات العصبية الأصطناعية [5],[8]

أولاً : تحديد المتغيرات Variables Selection:
أن النجاح في تصميم الشبكة العصبية يعتمد على فهم واضح للمشكلة،وأن السبب في الأعتماد على الشبكة العصبية هوالقدرة القوية للكشف عن العلاقات غير الخطية المعقدة بين عدد من المتغيرات.

ثانياً : جمع البيانات Data Collection:[8]
على الباحث أن يأخذ بنظر الأعتبار الوقت والكلفة والدقة في جمع البيانات للمتغيرات المفسرة عندما يكون أنموذج الشبكة العصبية أنموذج أنحدار متعدد غير خطي ، ولمتغير واحد (سلسلة زمنية) عندما يكون انموذج الشبكة العصبية انحدار ذاتي غير خطي وعليه أن يقوم بفحص هذه البيانات إذ أن بعضها قد يكون مفقود أو غير موجود.

ثالثاً : تحويل البيانات Data processing:[8]
نادرًا مايتم تغذية الشبكة العصبية بالمدخلات التي تم جمع بياناتها بشكل مواد خام وغالباً مايؤخذ لها اللوغاريتم الطبيعي إذا كانت كبيرة جداً ، أو يأخذ الفرق الأعتيادي الأول إذا كانت تعاني من اتجاه عام .

رابعاً: تجزئة البيانات Data partition :[5]
في الخلية العصبية الأصطناعية تجزأ البيانات الى ثلاث مجاميع لإستعمالها في التحليل والتنبؤ وكما يلي :
1- مجموعة التدريب أو التعلم Training Set : وهي مجموعة من بيانات السلسلة الزمنية وتستخدم لتقدير الأوزان وتدريب الشبكة على التنبؤ .

2- مجموعة الأختبار **Test Set** : وهي جزء من بيانات السلسلة الزمنية ولعينة لم تستعمل في التدريب تقوم بإختبار قابلية الشبكة على التنبؤ بعد عملية التعليم ولا يوجد أساس علمي لتحديد تجزئة البيانات إلى مجموعتين تعلم وإختبار، فقد تؤخذ 90% من بيانات السلسلة للتعلم و 10% للإختبار أو تكون 50% للتعلم و 50% للإختبار وبعض البرامجيات تحدد 70% للتعلم و30% للإختبار .

3- مجموعة التقويم **Validation Set** : تقوم بمراقبة نهائية لتقويم إداء تدريب الشبكة العصبية بصورة عامة وأن إختيار عدد بيانات هذه المجموعة يجب أن يكون مناسب لتقويم كلاً من التدريب والأختبار وذلك للتوصل إلى أفضل تنبؤ .

خامساً : تحديد معمارية الشبكة العصبية [9] Determine the Architecture of N.N.

لتحديد هيكلية الشبكة العصبية نتبع الخطوات الأتية :-

1- تحديد عدد عقد الأدخال Determination of Inputs Nodes [10]

أن المتغيرات في أنموذج السلسلة الزمنية تكون بدلالة المتغيرات المزاحة أي أن

$$Y_t = f(Y_{t-1}, Y_{t-2}, ..., Y_{t-p})$$

إذ أن :- (p) تمثل درجة الأنحدار الذاتي أي أنحدار Y_t على القيم السابقة له .

أن تحديد هذه المتغيرات تمثل مشكلة كبيرة تواجه مصمم الشبكة العصبية وقد اقترحت عدة طرق منها:-

الطريقة الأولى :

اقترح Kihoro واخرون عام 2004 طريقة تحديد الصلات اليا (Automatic Relevance Determation Method)- لتحديد عدد المتغيرات المزاحة زمنياً التي لها تأثير على المتغير المعتمد Y_t وذلك بالأعتماد على حساب المقدار الأتي

$$Inf\ (L) = \sum_{h=1}^{N_h} W_{ih}\, W_{ho} \quad ,...(24)$$

إذ أن $Inf\ (L)$:-مختصر لـ(Influence) يمثل التأثير الكلي على الناتج (Output) من قبل المتغير Y_{t-L}

W_{ih}:تمثل الأوزان بين عقد الأدخال والطبقة المخفية ،W_{h0}: تمثل الأوزان بين الطبقة المخفية والناتج .

وعلى فرض أن $r(L) = |\inf\ (L)| \quad ,...(25)$

فإذا كانت قيمة $r(L)$ صغيرة فاعنه يشير الى Y_{t-L} سوف لا يدخل في النموذج ويمكن تلخيص هذه القاعدة كما يلي :

تحديد عدد العقد المخفية ، نختار قيمة P التي نشك بأن يكون لها تأثير على القيمة المتنبأ بها ، ندرب الشبكة بأستخدام كل الأزاحات المختارة ، نحسب المقدار Inf، نختار قيمة k . فإذا كانت

$$P(Y(t-L) \in Y) = \begin{cases} 1 & if\ |r(L)| > k \\ & L = 1,2,.....P \quad ,...(26) \\ 0 & if\ |r(L)| \le k \end{cases}$$

مع مراعاة الأخذ بنظر الأعتبار مجموع مربعات الأخطاء لمجموعة التدريب SSE_t

الطريقة الثانية :

الأعتماد على معاملات الأرتباط الذاتي المعنوية للسلسلة الزمنية المستقرة.

الطريقة الثالثة :

الأعتماد على طريقة تجميع البيانات فإذا كانت شهرية فاعن عدد عقد الأدخال هي $P = 12$ وإذا كانت فصلية فاعن $P = 4$.

2- تحديد عدد الطبقات المخفية Determination of Hidden Layers

في اغلب تطبيقات التنبؤ بأستخدام الشبكات العصبيةالمتعددة الطبقات يفضل استخدام طبقة مخفية واحدة.

3- تحديد عدد العقد في الطبقة المخفية {9} Determination of Hidden Nodes

تواجه مصمم الشبكة العصبية مشكلة اخرى الأ وهي تحديد عدد العقد في الطبقة المخفية ، أن عدد العقد المخفية للشبكة العصبية يحدد من قبل المستخدم وللبدء بتدريب الشبكة العصبية لا بد من اعطاء تقدير أولي لعدد العقد المخفية .

الطريقة الأولى :

أن بعض الباحثين يختارها بعدد المتغيرات المزاحة زمنياً أي بعدد عقد الأدخال في حين البعض الأخر يعتمد على مقياس AIC,BIC,AIC_C .

الطريقة الثانية :

أيضاً أقترح الباحثKihoroواخرون طريقة لتحديد عدد العقد المخفية بموجب الصيغة الأتية

$$N_h \le \frac{(N_t - 1) * \varepsilon}{(N_I + 2)} \quad ,...(27)$$

إذ أن N_h : عدد العقد المخفية ، N_t : حجم العينة لمجموعة التدريب ، N_I : عدد عقد المدخلات

ε : كمية موجبة فيها $\varepsilon \le 0.125$

أن هذه الصيغة تعتمد على عدد عقد الأدخال كلما قلت عددها زاد عدد العقد المخفية إذ أن العلاقة بينهما عكسية ، كما يمكن تدريب الشبكة على عدد من العقد قد تكون (h,…,3,2,1) ويختار العدد الذي يكون فيه مجموع مربعات الأخطاء اقل ما يمكن .

سادساً : تحديد عقد الأخراج Determination of output nodes

اتفق الباحثون على أن وجود عقدة اخراج واحدة تكفي لعملية التنبؤ بخطوة واحدة نحو الأمام أما في حالة التنبؤ لعدة خطوات قادمة فاءن عدد العقد يكون مساوياً الى عدد خطوات التنبؤ.

سابعاً : إختيار دوال التنشيط Selection of Activation functions

وهي المعادلات الرياضية التي تحدد الناتج لمعالجة الشبكات العصبية ، وتسمى أيضًا دوال التحويل وغالبية الشبكات العصبية الحالية تستخدم الدالة اللوجستية في الطبقة المخفية والدالة الخطية عند طبقة المخرجات عندما يكون اكثر من عقدة للناتج .

ثامناً : تدريب الشبكة العصبية {8} Neural Network Training

أن عملية تدريب الشبكات العصبية على التنبؤ تتم من خلال تعديل أوزان الشبكة للحصول على أقل خطأ بالأعتماد على خوارزمية الأنتِشار الخلفي (BP) ، ومن أهم خطوات تدريب الشبكة هي :

1- تحديد قيم الأوزان الأولية Select the weight initialization

وتُعد نقطة البداية لعملية التدريب وغالباً ما يتم إختيار قيم عشوائية صغيرة للأوزان والحد المطلق تقع ما بين (1-,1).

2- عدد تكرارات التدريب {8} Number of training iteration

الباحث يجب أن يوقف التدريب فقط حتى يكون هناك أي تحسن في وظيفة الخطأ استنادا إلى عدد معقول من الأوزان التي تم اختيارها عشوائيا وتسمى النقطة التي لاتحسن الشبكة بالتقارب .

3- معدل التعلم (η)وعامل الزخم (α){1},{8} Learning Rate and Momentum

أن حجم الشبكة يؤثر على عملية التعلم حيث ان الشبكة الكبيرة تحتاج الى وقت اكبر للتعلم والوصول الى اقل خطا ، ويتاثر بمعدل التعلم (η) ومعامل الزخم (α)فكلما كانت عملية التعلم بطيئة نعطي قيمة اكبر الى (η)وعندما تكون عملية التخفيض في الخطا سريعة فنختار قيمة اصغر الى (η) لضمان الوصول الى اصغر قيمة للخطا ولكنها ستحتاج الى وقت طويل لتدريب الشبكة.

أما بالنسبة الى (α)فالقيمة الكبيرة لها ستؤدي الى زيادة سرعة التعلم ونوصي بها عندما تكون عملية الوصول الىاقل خطا بطيئة .

تاسعاً : التـنـفيذ{5} Implementation :

وهي من أهم المراحل حيث تختبر الشبكة من حيث قدرة التكيف مع حالة التغير في دورة وأمكانية أعادة التدريب والوصول الى أقل مربع خطأ عند تغير البيانات .

5. مقاييس الكفاءة على دقة النماذج {15}

لقياس كفاءة اداء النماذج واختيار الأفضل من خلال استخدام جملة من المقاييس الأحصائية التي تعتبر من المؤشرات التي توضح مدى ملائمة الأنموذج والتي هي :-

1- متوسط الانحراف المطلق Mean absolute deviation (MAD){13}

صيغته كماهو ات : $$MAD = \sum_{t=1}^{n} \left|\frac{Y_t - \hat{Y}_t}{n}\right| , \dots (28)$$

n : يمثل عدد المشاهدات ، Y_t : يمثل المشاهدات الحقيقية ، $\hat{Y}_t$: يمثل المشاهدات المتوقعة .

2- متوسط مطلق الخطأ النسبي Mean Absolute Percentage Error (MAPE){13}

وصيغة هي $$MAPE = \frac{1}{n}\sum_{t=1}^{n} \left|\frac{Y_t - \hat{Y}_t}{Y_t}\right| \ 100 , \dots (29)$$

3- الجذر التربيعي لمتوسط مربعات الخطأ Root mean square error (RMSE){13}

وصيغته هي $$RMSE = \sqrt{MSE} = \sqrt{\frac{\sum (Y_t - \hat{Y}_t)^2}{n}} , \dots (30)$$

4- الخطأ المعياري للتنبؤ Standard error of Prediction (SEP){13}

وصيغته هي $$SEP = \frac{100}{\bar{Y}} \ RMSE , \quad \dots (31)$$

حيث ان :- $\bar{Y}$:- يمثل الوسط الحسابي للمشاهدات .

الجانب التطبيقي

في هذا الجانب سيتم التطبيق لبيانات واقعية تمثل الاعداد الشهرية للمسافرين لمطار بغداد الدولي من الفترة (كانون الثاني 2006 ولغاية اب 2015) و بغية تحقيق هدف البحث فقد قسم التحليل الأحصائي لبيانات السلسلة أعداد المسافرين من الفترة (كانون الثاني 2006 ولغاية كانون الاول 2014) الى ثلاثة اجزاء رئيسية :-

1- بناء أنموذج بوكس- جنكينز لسلسلة أعداد المسافرين، إذ يتم تشخيص الأنموذج الملائم وتقدير معلماته .
2- بناء نماذج الشبكات العصبية الأصطناعية بالأعتماد على الفرض بأن السلسلة الزمنية تحوي فقط المركبة غير الخطية للسلسلة الزمنية للمسافرين وبإستعمال خوارزمية الأنتِشار الخلفي بعد تحديد معمارية مختلفة للشبكة.
3- بناء النماذج الهجينة التي تنص على الفرض بأن السلسلة الزمنية تحوي المركبتين الخطية وغير الخطية.
4- ويختم بأجراء مقارنة بين هذه النماذج (بوكس- جنكينز ، الشبكات العصبية ، الهجينة) من الفترة كانون الثاني ولغاية اب من العام 2015 بالأعتماد على جملة من المقاييس الأحصائية لأختيار الأنموذج الأكفأ الذي يستخدم لحساب التنبؤات المستقبلية لأعداد المسافرين للفترة الزمنية من (ايلول 2015 ولغاية كانون الأول 2016) .

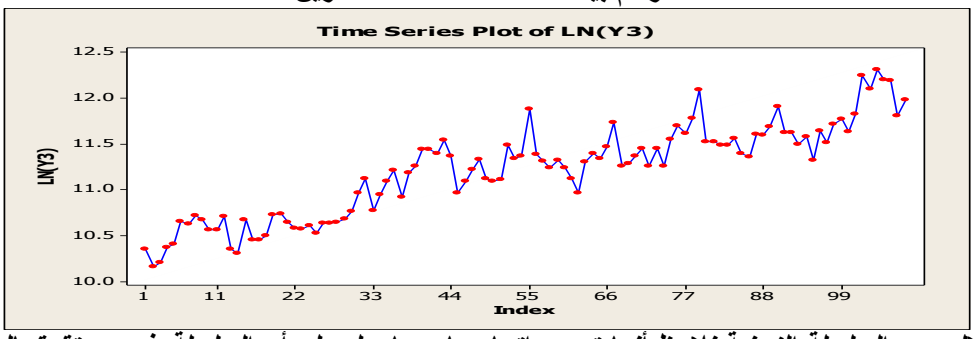
انموذج بوكس-جنكينز لتحليل السلسلة الزمنية لأعداد المسافرين لمطار بغداد الدولي

لغرض تحليل السلسلة الزمنية يتطلب بناء أفضل أنموذج للتنبؤ وتحديد معلماته وتقديرها والتأكد من ملائمة الأنموذج للبيانات موضوعة البحث وبالأعتماد على البرنامج الجاهز (Mintab 16)، اذ يعتمد تحليل السلسلة الزمنية على عدة خطوات وكما يلي :-

1.استقرارية السلسلة

ولغرض أنجاز الأستقرارية في التباين فقد أخذ اللوغاريتم الطبيعيلسلسلة البيانات بأعداد المسافرين للفترة الزمنية من كانون الثاني / 2006 ولغاية كانون الاول / 2014 وتم رسمها كما مبين ادناه .

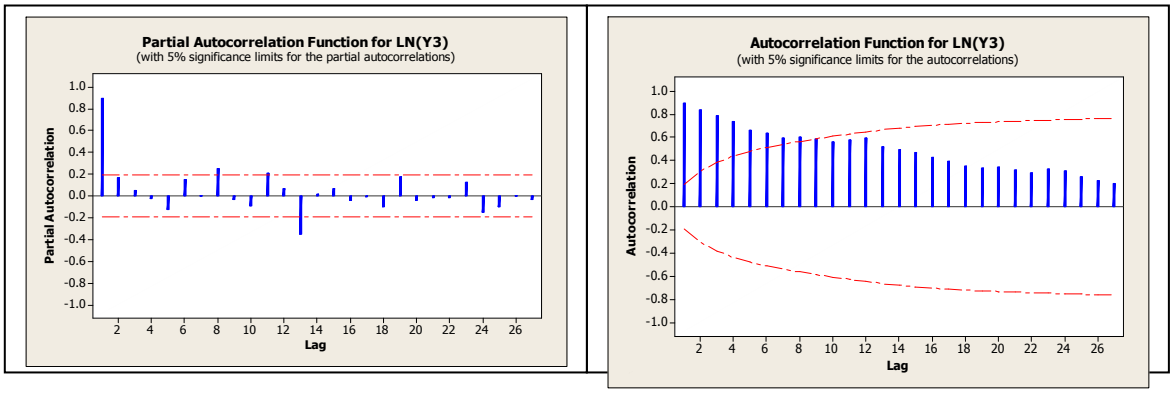
الشكل (2)
رسم بيانات سلسلة أعداد المسافرين



من خلال رسم السلسلة الزمنية نلاحظ أنها تحوي اتجاه عام مما يدل على أن السلسلة غير مستقرة بالمتوسط ولمزيد من الدقة تم استخراج معاملات الأرتباط الذاتي (ACF) والأرتباط الذاتي الجزئي (PACF) على التوالي وقد تم رسمهما كما في الأشكال الأتية.

الشكل (3)
يمثل(PACF) لسلسلة أعداد المسافرين

الشكل (4)
يمثل(ACF) لسلسلة أعداد المسافرين

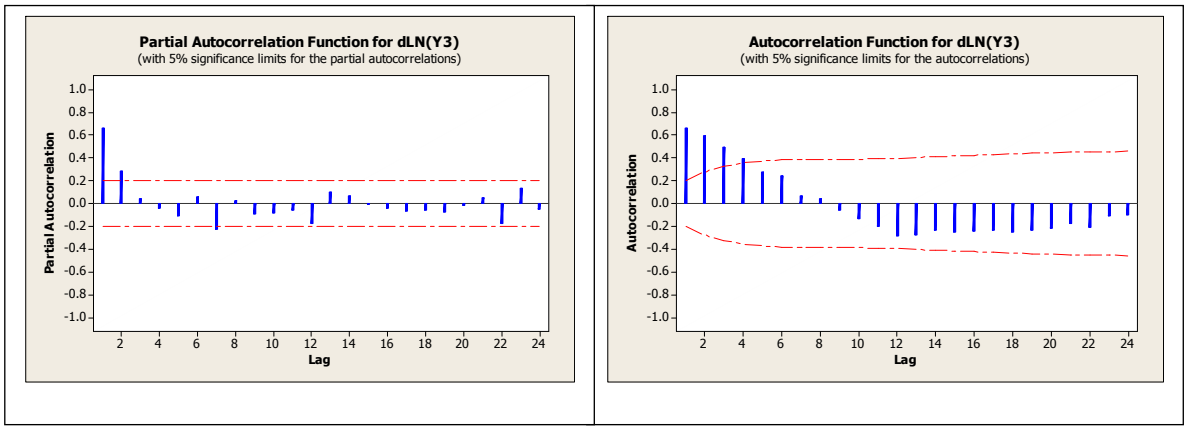


ومن خلال ملاحظة قيم معاملات الأرتباط الذاتي (ACF)نجد انها معنوية لعدد كبير من الأزاحات وتتنازل ببطئ مما يشير الى أن السلسلة الزمنية غير مستقرة وهذا مايؤيده الشكل (3)، اما الشكل (4)يشير الى أن الأزاحة الأكثر معنوية هي الأزاحة الأولى وهناك ازاحات اخرى غير معنوية ، لذلك كان لا بد من أخذ الفروقات اللازمة للسلسلة الأصلية لتحقيق الأستقرارية في المتوسط عندئذٍ تصبح السلسلة الزمنية لبيانات المسافرين مستقرة .

لذا تم أخذ الفرق الأول كما وتم حساب معاملات الأرتباط الذاتي والجزئي للسلسلة المستقرة وقد تم رسم كل من دالتي الأرتباط الذاتي والذاتي الجزئي (PACF, ACF) كما يظهر في الشكلين أدناه حيث تبين أن السلسلة اصبحت مستقرة .

الشكل (5)
يمثل(ACF) لسلسلة أعداد المسافرين

الشكل (6)
يمثل(PACF) لسلسلة أعداد المسافرين



وكذلك تم استخراج دالة الأرتباط الذاتي الموسعة للعينة (ESACF) لغرض تشخيص الأنموذج كما مبين في الجدول (1) .

جدول (1)

يبين دالة الأرتباط الذاتي الموسعة للعينة (ESACF) لسلسلة اعداد المسافرين

$$-0.2 \leq \hat{r}_k(\hat{a}) \leq 0.2$$

	MA								MA							
AR	0	1	2	3	4	5	•••	AR	0	1	2	3	4	5	•••
0	0.6602	0.5967	0.4932	0.394	0.2722	0.2446	•••	0	x	x	x	x	x	x	•••
1	-0.331	0.1127	0.0168	0.0631	-0.11	0.2451	•••	1	x	x	0	0	0	x	•••
2	-0.056	-0.089	0.0324	-0.01	-0.047	0.1689	•••	2	0	0	0	0	0	0	•••
3	-0.024	-0.025	-0.024	0.0096	-0.053	0.1657	•••	3	0	0	0	0	0	0	•••
4	-0.022	-0.021	-0.011	0.0015	-0.052	0.1654	•••	4	0	0	0	0	0	0	•••
5	-0.019	-0.041	-0.034	-0.045	-0.019	0.1648	•••	5	0	0	0	0	0	0	•••
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋱

حيث ان ($\hat{r}_k(\hat{a})$) :- تشير الى معاملات الارتباط الذاتي للبواقي

من الملاحظ في الجدول (1) أن رأس المثلث عند النقطة (2,0).

2. تشخيص وتقدير الأنموذج واختيار الأنموذج الأفضل

من خلال ملاحظة دالة الأرتباط الذاتي في الشكل(5)يبدو أن السلسلة الزمنية لبيانات المسافرين تتناقص بشكل اسي متناوب وأن دالة الأرتباط الذاتي الجزئي كما في الشكل(6) معنوية عند الأزاحة الثانية وبالأعتماد على النتائج المستخرجة من دالة الأرتباط الذاتي الموسعة للعينة كما مبين في الجدول (1) بالأمكان أن نقترح أنموذج الأنحدار الذاتي من الدرجة الثانية لبيانات المسافرين، ويكون الا نموذج هو (2) ARالذي معادلته بالشكل الأتي :-

$$Z_t = \mu + \emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + a_t$$

ولقد تم تقدير معلمات الأنموذج وكانت النتائج كما مبين في الجدول ادناه .

جدول (2)

يبين تقدير معلمات الأنموذج ومتوسط مربعات الخطأ لسلسلة أعداد المسافرين

Final Estimates of Parameters	
Type Coef SE Coef T P	متوسط مربعات الأخطاء
AR 1 0.4706 0.0991 4.75 0.000	
AR 2 0.2991 0.0999 2.99 0.004	MS = 0.02946
Constant 0.04225 0.01753 2.41 0.018	

3. اختبار دقة الأنموذج

بعد تشخيص الأنموذج وتحديد درجته وتقدير معلماته لابد من التأكد من صحة ملاءَمة الأنموذج وكفاءته إذ تم حساب معاملات الأرتباط الذاتي للبواقي كما تم حساب :

1- اختبار (Ljung-Box):-

من خلال تطبيق احصاءة (Ljung-Box) المعدلة لفحص ملائمة الأنموذج وظهربأن قيمة (Q=18.9) عند الأزاحة (K=24) لسلسلة المسافرين ومن خلال قيمة(P-Value) وهذا يدل على أن الأخطاء (a_t) ماهي إلا متغيرات عشوائية غير مترابطة وهذا يدل على أن الأنموذج جيد وملائم وكفوء لتمثيل السلسلة الزمنية وبالتالي قبول الفرضية الأتية :

$$H_0: \rho_i(a_t) = 0 \quad , \quad i = 1, 2, \dots, k$$

جدول (3)

يبين قيم احصاءة *(Ljung-Box)* المعدلة لسلسلة اعداد المسافرين

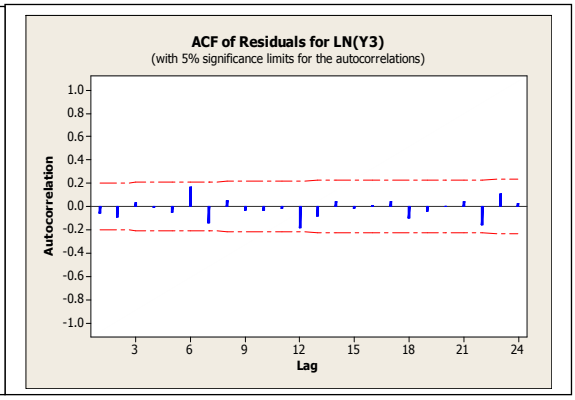
Modified Box-Pierce (Ljung-Box) Chi-Square statistic				
Lag	12	24	36	48
Chi-Square	10.7	18.9	35.0	42.5
DF	9	21	33	45
P-Value	0.296	0.590	0.375	0.580

2-اختبار البواقي :-

من خلال رسم معاملات الأرتباط الذاتي والجزئي للبواقي المستخرجة كما في الشكلين (7) و (8) ، نلاحظ أن قيم معاملات الأرتباط الذاتي للبواقي جميعها تقع ضمن حدود الثقة مما يعني أن سلسلة البواقي هي عبارة عن متغيرات عشوائية غير مترابطة وأن الأنموذج المستعمل ملائم وجيد في تمثيل بيانات المسافرين.

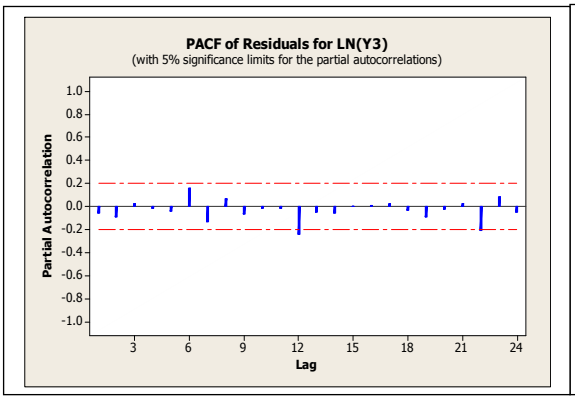
الشكل (7)

يمثل (ACF) للبواقي لسلسلة المسافرين



الشكل (8)

يمثل (PACF) للبواقي لسلسلة المسافرين



4.التنبؤ :

بعد اجتياز الأنموذج لاختبارات التشخيص يصبح بالأمكان استعمال الأنموذج للتنبؤ والجدول ادناه يمثل القيم التقديرية للأشهر الثمانية الاولى من عام (2015)لبيانات المسافرين.

جدول (4)

يبين القيم التقديرية والحدين الأدنى والأعلى مع القيم الحقيقة لسلسلة المسافرين

الشهر /2015	الحد الأدنى	القيم التقديرية	الحد الأعلى	القيم الحقيقية
كانون الثاني	11.52	11.86	12.20	11.87
شباط	11.65	12.02	12.39	12.18
آذار	11.65	12.06	12.48	12.18
نيسان	11.47	11.90	12.33	12.13
أيار	11.64	12.08	12.53	12.15
حزيران	12.04	12.49	12.95	12.15
تموز	11.88	12.34	12.80	12.18
اب	12.07	12.54	13.01	12.23

نماذج الشبكات العصبية لتحليل السلسلة الزمنية لأعداد المسافرين لمطار بغداد الدولي

استُعملت منهجية الشبكات العصبية الأصطناعية في تحليل السلاسل الزمنية و بناء أنموذج للتنبؤ في حركة المسافرين لمطار بغداد الدولي حيث تم أخذ اللوغاريتم الطبيعي لكل من سلسلة أعداد المسافرين ، وتم استخدام الشبكات العصبية المتعددة الطبقات(Multilayer Perceptron) ذات التغذية الأمامية (Feed Forward) وبطبقة مخفية واحدة (One hidden layer) لغرض التنبؤ بالقيم المستقبلية للسلسلة الزمنية من خلال إستعمال دوال التنشيط -غيرالخطية للبيانات السابقة إذ استعمل البرنامج الجاهز (SPSS 19) في نمذجة الشبكات العصبية الأصطناعية، ولتحديد معمارية الشبكة تم إختيار متغيرات الأدخال (مدخلات الشبكة) لبيانات السلسلة من y_{t-1} إلى y_{t-12} ، وقد تم استخدام دالتي تنشيط غير خطية هي الدالة اللوجستية للعقد المخفية والمسؤولة عن العمليات الرياضية وحساب الأوزان وتعديلها ، أما القيمة التقديرية ($\hat{y}_t$) فتمثل مخرجات الشبكة العصبية.

وللقيام بعملية التدريب تم تجزئة بيانات السلسلة الزمنية لأعداد المسافرين والمتكونة من (108) مشاهدة لكل سلسلة إلى (70%) بيانات التدريب و(30%) للإختبار ويمكن تغيير هذه النسب بما يلائم بيانات الأدخال وتحديد معدل للتعلم (η=0.4) ومعامل الزخم (α=0.9) وبقيم أولية عشوائية للأوزان إذ يُعد مجموع هذهِ الأوزان المضروبة مع قيم متغيرات الأدخال كمدخلات لتركيبة خطية وتؤخذ هذهِ المدخلات إلى دالة التنشيط اللوجستية في عقدتي الطبقة المخفية الأ إن ذلك يتطلب تهيئة البيانات بصورة أولية من خلال تحويلها إلى إحدى الصيغ إذ تم إستعمال الصيغة المعيارية لتحقيق مدى هذه الدالة الذي يكون ما بين (0,1) لإجراء العمليات الحسابية للتقدير فضلاً عن وجود الحد الثابت (bias) أو التحيز لكل من العقد المخفية ، وبنفس الأسلوب تم تطبيق العلاقة الخطية Linear activation function التي تجمع ما بين الطبقة المخفية وطبقة الأخراج .

وكما ذكر في الجانب النظري من ان الباحث يواجه مشكلتين رئيسيتين هما تحديد عدد عقد الأدخال وتحديد عدد العقد المخفية لذلك سنستخدم الطرائق الأحصائية لتحديد عدد عقد الأدخال والعقد المخفية

1. تحديد عقد الأدخال Determination of Inputs Nods

لتحديد عدد المتغيرات المزاحة زمنياً تم الأعتماد على الطرائق الأتية :

الطريقة الأولى :

الطريقة المقترحة من الباحث (Kihoro) واخرين اذ يتم بأخذ كل متغير مزاح زمنياً مع (Y_t) وبيان مقدار تأثيره على (Y_t) من خلال حساب $|Inf\ (L)|$ واستخراج مجموع مربعات خطأ التدريب وعليه سيكون الأنموذج هو $NN(Y_{t-k};1)\quad k=1,2,...,12$ وكما في الجدول (5) لسلسلة أعداد المسافرين كما وتم بيان المتغيرات الطبيعية المهمة وتسلسلها كما في الشكل (9)لسلسلة أعداد المسافرين .

ومن خلال ملاحظة قيم $|Inf|$ ومجموع مربعات خطأ التدريب في الجدول (5)يمكن ان تحدد الأزاحات المعنوية وهي (12,2,1) اي ان الانموذج هو: $Y_t = f(Y_{t-1}, Y_{t-2}, Y_{t-12})$

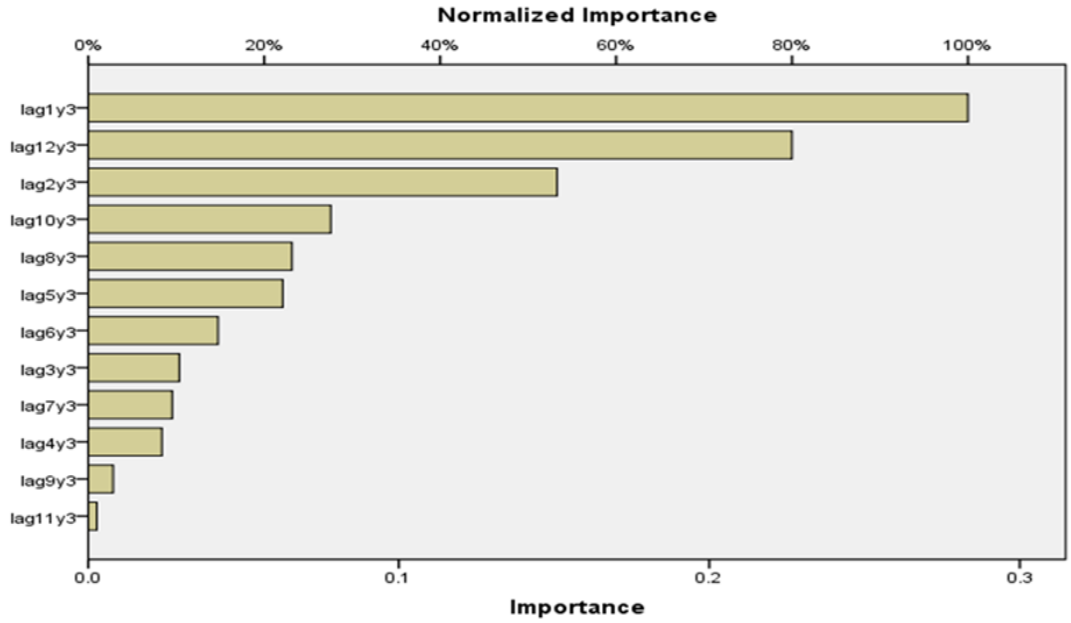
جدول (5)

يبين مقدار $|Inf|$ ومجموع مربعات خطأ التدريب للازاحات (1,...,12) لسلسلة المسافرين

Lag	S.S.Et	/inf/
1	7.581	4.752858331
2	13.33893707	5.223757991
3	12.79618161	4.423793287
4	13.31030648	5.636969828
5	18.14969953	4.271605717
6	20.69896892	4.249249794
7	16.97327228	4.108368612
8	16.45886604	4.694648958
9	14.14649564	4.954552927
10	16.67660483	3.903486631
11	19.46417847	2.474694562
12	10.58237406	4.940567112

الشكل (9)

يبين المتغيرات الطبيعية المهمة وتسلسلها حسب الاهمية لسلسلة المسافرين



كما في الشكل (9) يؤكد اختيار هذه المدخلات للشبكة العصبية الخاصة بالسلسلة الزمنية لأعداد المسافرين .

الطريقة الثانية :

الأعتماد على معاملات الأرتباط الذاتي المعنوية لسلسلة أعداد المسافرين المستقرة بعد ان اخذ لها الفرق وجد ان معاملات الأرتباط الذاتي لسلسلة أعداد المسافرين معنوية فقط عند الأزاحة الأولى والثانية والثالثة كما في الشكل (5) .

❖ **وعليه فاعن عدد مدخلات أنموذج الشبكة العصبية لسلسلة أعداد المسافرين سيكون :-**

$$Y_t = f(Y_{t-1}, Y_{t-2}, Y_{t-3})$$

الطريقة الثالثة :

ان بيانات السلسلة الزمنية المدروسة تمثل بيانات لسلسلة أعداد المسافرين شهرياً فاعن المتغيرات المزاحة هي (Y_{t-k} , $k=1,2,....,12$) اي ان أنموذج الشبكة العصبية يحوي على (12) عقدة ادخال والذي هو :- $Y_t = f(Y_{t-1}, Y_{t-2}, ..., Y_{t-11}, Y_{t-12})$

2.تحديد عدد العقد المخفيةDetermination of Hidden Nodes

الطريقة الأولى :

ان عدد العقد المخفية تعتمد على عدد عقد الأدخال ولهذا فاعن أنموذج الشبكة العصبية لسلسلة أعداد المسافرين سيكون حسب عدد عقد الأدخال كما في الجدول (6) .

جدول (6)

يمثل عدد العقد المخفية حسب عدد عقد الأدخال وأنموذج الشبكة العصبية لسلسلة المسافرين

الأنموذج	عدد عقد الأدخال	عدد العقد المخفية
NN (1,2,12;3)	3	3
NN (1,2,3;3)	3	3
NN(1,2,….,12;12)	12	12

الطريقة الثانية :

استخدام الصيغة المقترحة من قبل الباحث (Kihoro) واخرين [24] لكل طريقة من طرائق تحديد عدد عقد الأدخال فكانت النتائج كما في الجدول ادناه .

جدول (7)

يمثل عدد العقد المخفية بموجب صيغة الباحث (Kihoro) والأنموذج الخاص لسلسلة المسافرين

الأنموذج	عدد عقد الأدخال	عدد العقد المخفية
NN (1,2,12;2)	3	2
NN (1,2,3;2)	3	2
NN(1,2,….,12;1)	12	1

نظراً لعدم تساوي الطريقتين في اختيار عدد العقد المخفية لسلسلة أعداد المسافرين لذا سيتم اخذ عدد العقد (2 , 3) بالنسبة للأنموذج الأول و الثاني و (1 ,…, 12) بالنسبة للأنموذج الثالث .

وحساب المقاييس الأحصائية (AIC, AIC_C , BIC) لاختيار افضل النماذج كما مبين في الجدول(8).

جدول (8)

يبين المقاييس الاحصائية(AIC, AICC, BIC) للأنموذج الاول والثاني والثالث لسلسلة أعداد المسافرين

model	N_p	SSE	$\hat{\sigma}^2$	AIC	AICc	BIC
NN(1,2,12;2)	11	7.471	0.106	-134.63	-130.08	-98.896
NN(1,2,12;3)	16	7.883	0.112	-120.86	-110.6	-68.889
NN(1,2,3;2)	11	6.747	0.115	-124.53	-119.81	-89.111
NN(1,2,3;3)	16	8.981	0.123	-120.96	-111.24	-68.308
NN(1,2,…,12;1)	15	6.338	0.103	-108.12	-97.456	-61.46
NN(1,2,…,12;2)	29	7.073	0.105	-92.64	-45.613	0.29582
NN(1,2,…,12;3)	43	6.604	0.097	-72.56	85.1071	65.8792
NN(1,2,…,12;4)	57	14.146	0.191	-8.4383	404.812	179.893
NN(1,2,…,12;5)	71	6.273	0.098	-6.6394	-1284.6	217.641
NN(1,2,…,12;6)	85	8.214	0.112	10.5305	-1114.1	290.22
NN(1,2,…,12;7)	99	6.114	0.088	30.7789	-607.93	350.955
NN(1,2,…,12;8)	113	5.717	0.078	40.0699	-588.32	411.892
NN(1,2,…,12;9)	127	5.411	0.091	113.052	-358.14	503.9
NN(1,2,…,12;10)	141	6.271	0.111	159.4	-306.23	585.975
NN(1,2,…,12;11)	155	7.841	0.117	166.255	-377.12	662.983
NN(1,2,…,12;12)	169	9.544	0.146	213.302	-333.94	749.773

وعند ملاحظةالجدول (8) نجد ان اقل قيمة لهذه المقاييس كانت للنماذج الأتية :-

1. **الأنموذج الأول NN(1,2,12;2) الذي يضم ثلاثة متغيرات مزاحه وهي (y_{t-12},y_{t-2},y_{t-1})وعقدتين مخفيتين .**
2. **الأنموذج الثاني NN(1,2,3;2) الذي يضم ثلاثة متغيرات مزاحه وهي (y_{t-3},y_{t-2},y_{t-1})وعقدتين مخفيتين .**
3. **الأنموذج الثالث NN(1,…,12;1) الذي يضم اثنى عشرة متغير مزاح وهي(y_{t-1} , … ,y_{t-12}) وعقدة مخفية واحدة.**

3. التنبؤ:

بعد الحصول على افضل النماذج لسلسلة أعداد المسافرين نقوم بأستخدامها لحساب التقديرات ومقارنتها مع القيم الحقيقية لثمانية الأشهر الأولى من عام 2015 ، وكما مبين في الجدول (9).

جدول (9)
يبين التقديرات الشهرية لأعداد المسافرين مع القيم الحقيقية

الشهر / 2015	تقديرات الأنموذج NN(1,2,12;2)	تقديرات الأنموذج NN(1,2,3;2)	تقديرات الأنموذج NN(1,…,12;1)	القيم الحقيقية
كانون الثاني	11.88	11.88	11.86	11.87
شباط	12.02	11.98	12.06	12.18
آذار	12.09	11.99	12.01	12.18
نيسان	12.04	11.99	12.07	12.13
أيار	12.14	11.99	12.09	12.15
حزيران	12.15	12.01	12.13	12.15
تموز	12.17	12.01	12.19	12.18
اب	12.23	12.00	12.20	12.23

وتم حساب المقاييس الأحصائية (MAD,MAE,RMSE,SEP,MAPE) للنماذج الثلاثة للحصول على الأنموذج الأكفأ الذي يمثل أنموذج الشبكات العصبية المتعددة الطبقات من خلال المقارنة بين هذه المقاييس لسلسلة أعداد المسافرين وكما مبين في الجدول (10) .

جدول (10)
يبين المقاييس الأحصائية للنماذج الثلاثة لسلسلة أعداد المسافرين

النماذج (Models)	MAD	RMSE	SEP	MAPE
NN(1,2,12;2)	**0.005**	**0.073**	**0.621**	0.380
NN(1,2,3;2)	**0.019**	**0.167**	**1.429**	1.273
NN(1,2,….,12;1)	**0.008**	**0.081**	**0.691**	0.493

ومن خلال ملاحظة الجدول اعلاه نجد إن الأنموذج NN(1,2,12;2) يمتلك اقل قيمة للمقاييس الأحصائية المستخرجة لذا يعد هذا الأنموذج هو الأفضل من بين نماذج الشبكات العصبية للتنبؤ بأعداد المسافرين .
النماذج الهجينة لتحليل السلاسل الزمنية لأعداد المسافرين لمطار بغداد الدولي .

1. الأنموذج الهجين من ARIMA و ANN Hybrid Model of (ARIMA-ANN)

المرحلة الأولى :

بعد أستخدام AR(2) لتمثيل الجزء الخطي (L) للسلسلة الزمنية وحساب التقديرات المستقبلية لها كما في الجدول (4).

ولغرض بناء الأنموذج الهجين نستعمل البواقي (a_t) المستخرجة حيث تمثل المركبة غيرالخطية (N) للسلسلة التي يتم تمثيلها بأنموذج الشبكات العصبية المتعددة الطبقات للحصول على القيم المتوقعة لها اي أن الأنموذج سيكون $N = f(a_{t-1}, a_{t-2}, \dots\dots\dots, a_{t-p})$

ولتحديد المتغيرات المزاحة $(a_{t-1}, a_{t-2}, \dots, a_{t-p})$ تم أستخدام الطريقة المقترحة من الباحث (Kihoro) واخرون من خلال حساب قيم $|Inf\ (L)|$ وتحديد العقد المخفية حيث كانت المتغيرات المزاحة المهمة هي (a_{t-8}, a_{t-12})لسلسلة أعدادالمسافرين وكانت عدد عقد الأدخال والعقد المخفية هي(8,12;2)NN حيث اصبح الأنموذج كما يلي :$N_t = f(a_{t-8}, a_{t-12})$

وقد تم أستخدام هذه النماذج لسلسلة المسافرين لحساب التقديرات لأعداد المسافرين للأشهر الثمانية الأولى من عام 2015 .

المرحلة الثانية :

في هذه المرحلة يتم جمع القيم المقدرة بواسطة أنموذج (ARIMA) مع تقديرات البواقي التي تم حسابها من قبل أنموذج (ANN) ، والتي بدورها تنتج القيمة المتوقعة النهائية للسلسلة الزمنية حسب المعادلة (4) وكما مبين في الجدول (11) .

جدول (11)

يبين القيم التقديرية للأنموذج الهجين (ARIMA-ANN) لسلسلة المسافرين

الشهر/ 2015	كانون2	شباط	آذار	نيسان	أيار	حزيران	تموز	اب
ARIMA-ANN	11.87	12.03	12.09	11.94	12.13	12.49	12.34	12.53

2.الأنموذج الهجين من ANN و ARIMA Hybrid Model of (ANN-ARIMA)

وفي هذا النوع من النماذج الهجينة ايضاً تتكون من مرحلتين ولكن عكس الأنموذج الهجين السابق وكما مبين ادناه :-

المرحلة الأولى :

بعد استخدام أنموذج الشبكات العصبية الأصطناعية (ANN) لتمثيل الجزء غير الخطي (N) للسلسلة الزمنية وحساب التقديرات المستقبلية لها كما مبين في الجدول (9) .

ولغرض بناء هذا الأنموذج نستعمل البواقي (a_t) المستخرجة من أنموذج الشبكات العصبية المتعددة الطبقات حيث تمثل المركبة الخطية (L) للسلسلة التي يتم تمثيلها بأنموذج (ARIMA) للحصول على القيم التقديرية لها .

بعد استخراج البواقي من أنموذج الشبكات العصبية لسلسلة اعداد المسافرين تم رسم سلسلة البواقي للتعرف على سلوكها وتبين أنها مستقرة ، وللمزيد من الدقة تم استخراج قيم معاملات الأرتباط الذاتي ACF والأرتباط الذاتي الجزئي PACF ورسم كل منهما على التوالي حيث لم يتم التعرف على الأنموذج .

ولتحديد رتبة الأنموذج بشكل ادق تم توفيق عدد من النماذج واختيار الأنموذج الأفضل وفق معايير المفاضلة وتبين أن أفضل أنموذج من نماذج بوكس- جنكينز هو الأنموذج(1,0,1)ARIMA كون قيم معياري (معيار معلومات بيزBIC ومعيار اكيكي AIC) لهذا الأنموذج اقل قيمة من بين جميع القيم أضافة الى قيمة متوسط مربعات الخطأ التي هي ايضاً اقل قليلة مقارنةً مع القيم الأخرى .

وبعد تشخيص الأنموذج وتحديد درجته وتقديره لابد من التاكد من صحة ملاءَمة الأنموذج وكفاءته وتم ذلك من خلال تطبيق احصاءة (Ljung-Box) لفحص ملائمة الأنموذج وظهر بأنها معنوية عند الأزاحة K=12 ومن خلال قيمة P-Value وهذا يدل على أن الأخطاء غير مرتبطة ببعضها البعض وهذا دليل على أن الأنموذج جيد وملائم ، فضلاً عن استخراج معاملات الأرتباط الذاتي والجزئي للبواقي (الأخطاء) للأنموذج المقدر ورسمها ولوحظ أن قيم معاملات الأرتباط الذاتي للبواقي جميعها تقع ضمن حدود الثقة مما يعني أن سلسلة البواقي عشوائية وأن الأنموذج المستعمل جيد وملائم .

المرحلة الثانية : في هذه المرحلة يتم جمع القيم المقدرة بواسطة أنموذج (ANN) مع تقديرات البواقي التي تم حسابها من قبل أنموذج (ARIMA) ، والتي بدورها تنتج القيمة المتوقعة النهائية للأنموذج الهجين حسب المعادلة (4) وكما مبين في الجدول (12) .

جدول (12)

يبين القيم التقديرية للأنموذج الهجين (ARIMA- ANN) لسلسلة المسافرين

الشهر/ 2015	كانون2	شباط	آذار	نيسان	أيار	حزيران	تموز	اب
ARIMA-ANN	11.83	12.06	12.11	12.05	12.18	12.17	12.18	12.23

المقارنة بين أساليب التقدير وبناء النماذج لأعداد المسافرين لمطار بغداد الدولي

لقد تمت المقارنة بين الأساليب الثلاثة (أنموذج بوكس جنكينز ، وأنموذج الشبكات العصبية ، والأنموذج الهجين) المستعملة في الدراسة لتقدير نماذج التنبؤ بالأعتماد على بعض المقاييس الأحصائية (MAD, RMSE , SEP , MAPE , MAE) والجدول (13) يوضح هذه المقارنة لسلسلة اعداد المسافرين وقد تم إعتماد الأنموذج الهجين (ANN-ARIMA)لحساب التنبؤات المستقبلية لأعداد المسافرين وذلك لإمتلاكهِ أقل قيمة للمقاييس الأحصائية من بين النماذج .

جدول (13)

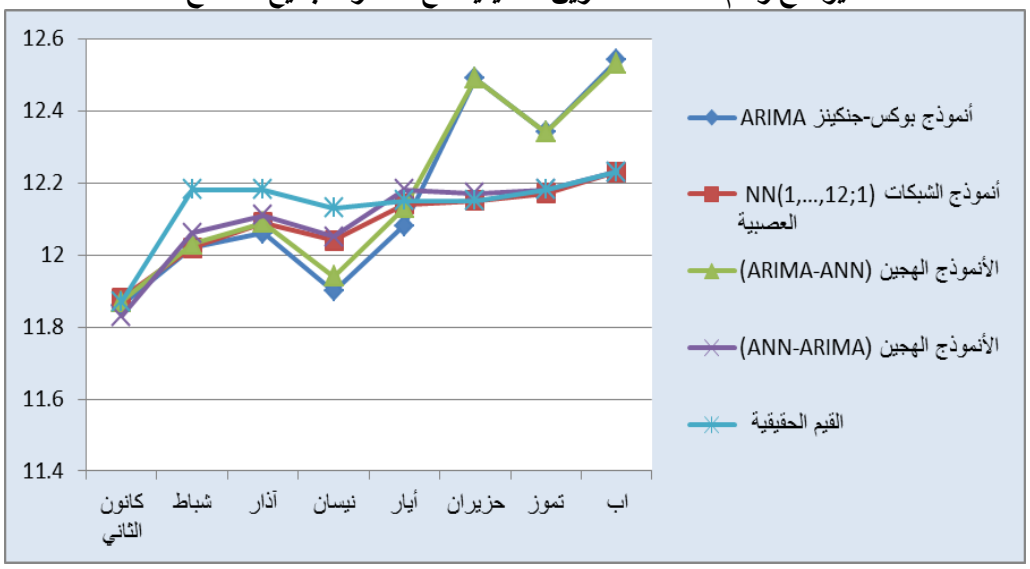
يبين القيم التقديرية للنماذج الأربعة مع القيم الحقيقة وجملة من المقاييس الأحصائية لسلسلة أعداد المسافرين

الشهر / 2015	بوكس-جنكينز ARIMA	الشبكات العصبية NN(1,2,12;2)	الهجين (ARIMA-ANN)	الهجين (ANN-ARIMA)	القيم الحقيقية
كانون الثاني	11.86	11.88	11.87	11.83	11.87
شباط	12.02	12.02	12.03	12.06	12.18
آذار	12.06	12.09	12.09	12.11	12.18
نيسان	11.90	12.04	11.94	12.05	12.13
أيار	12.08	12.14	12.13	12.18	12.15
حزيران	12.49	12.15	12.49	12.17	12.15
تموز	12.34	12.17	12.34	12.18	12.18
اب	12.54	12.23	12.53	12.23	12.23
المقاييس					
MAPE	1.438	0.380	1.283	0.371	
SEP	1.753	0.621	1.654	0.512	
RMSE	0.205	0.073	0.193	0.060	
MAD	0.022	0.005	0.020	0.004	

ايضاً بالأعتماد على قيم الجدول (13) تم رسم القيم الحقيقية مع القيم التقديرية للأشهر الثمانية الأولى من عام 2015 لسلسلة اعداد المسافرين وكما مبين في الشكل (10) .

الشكل (10)

يوضح رسم اعداد المسافرين الحقيقية مع المقدرة لجميع النماذج



التنبؤ بأعداد المسافرين لمطار بغداد الدولي بأستخدام الأنموذج الهجين(ANN-ARIMA)

بعد تقدم الأنموذج الهجين (ANN-ARIMA) على باقي النماذج وذلك لأمتلاكه اقل المقاييس الأحصائيةلذا تم اعتباره الأنموذج الأكفأ من بين هذه النماذج حيث تم أستعماله للتنبؤ المستقبلي للفترة الزمنية لسلسلة اعداد المسافرين، والجدول (14) يمثل القيم المستقبلية لسلسلة أعداد المسافرين للفترة الزمنية من أيلول/2015 ولغاية كانون الأول/2016 قبل وبعد التحويل اللوغارتمي .

جدول (14)

يبين القيم التنبؤية لسلسلة اعداد المسافرين لمطار بغداد الدولي

السنة	الشهر	اعداد المسافرين	اعداد المسافرين بعد التحويل
2015	أيلول	12.18	194983
	تشرين الأول	12.19	196859
	تشرين الثاني	12.13	184980
	كانون الأول	12.15	188239
2016	كانون الثاني	12.13	185967
	شباط	12.17	193138
	آذار	12.13	185046
	نيسان	12.14	187305

	أيار	12.16	191649
	حزيران	12.19	196829
	تموز	12.21	200643
	آب	12.23	203877
	أيلول	12.20	199520
	تشرين الأول	12.18	194726
	تشرين الثاني	12.19	196637
	كانون الأول	12.17	193213

الاستنتاجات والتوصيات

1. الأستنتاجات Conclusions

من خلال الدراسة النظرية والتطبيقية للموضوع توصل الباحث الى جملة من الأستنتاجات وكما يلي:-

1- أن السلسلة الزمنية لأعداد المسافرين تؤلف سلسلة زمنية غير مستقرة وذلك من خلال ملاحظة رسم هذه السلسلة ودالتي الأرتباط الذاتي والجزئي المستخرجة لهذه السلسلة.

2- عند حساب معاملات الأرتباط الذاتي والجزئي بعد استقرار السلسلة وجد أن دالة الأرتباط الذاتي تتناقص اسياً نحو الصفر في حين دالة الأرتباط الجزئي أنقطعت بعد الأزاحة الثانية ، مما يشير الى أن الأنموذج المقترح لتمثيل السلسلة الزمنية هو أنموذج الأنحدار الذاتي من الدرجة الثانية AR(2) وهذا ماأكدته نتائج دالة الأرتباط الذاتي الموسعة للعينة .

3- عند أستخدام المقاييس الأحصائية(AIC, AIC_C , BIC) لأختيار أفضل أنموذج للشبكات العصبية المتعددة الطبقات كانت النماذج NN(1,2,12;2) , NN(1,2,3;2) , NN(1,...,12;1) لسلسلة المسافرين، حيث تم أختيار أفضل هذه النماذج بالأعتماد على جملة من المقاييس الأحصائية وقد تبين أن الأنموذج NN(1,2,12;2) هو الأفضل .

4- تفوق الطريقة المقترحة من الباحث (Kihoro) لتجديد عدد عقد الادخال لامتلاكها اقل قيمة من للمقاييس الاحصائية المستخدمة (MAD, RMSEM, SEP, MAPE).

5- بعد مقارنة نماذج السلاسل الزمنية غير الهجينة والهجينة لأعداد المسافرين وبالأعتماد على جملة من المقاييس الأحصائية المستخدمة تفوق الأنموذج الهجين (ANN-ARIMA) على النماذج الأخرى من خلال أمتلاكه أقل قيمة لهذه المقاييس حيث تم استخدامه لحساب التنبؤات المستقبلية لأعداد المسافرين للفترة الزمنية من ايلول / 2015 ولغاية كانون الثاني / 2016 .

2. التوصيات Recommendation

1- أستخدام الأنموذج الهجين (ANN-ARIMA) من قبل المنشأة العامة للطيران المدني العراقي للتنبؤبأعداد المسافرين وذلك لما يتميز به من دقة في النتائج.

2- أجراء دراسة لتحليل السلسلة الزمنية بأعداد المسافرين غير المستقرة بأخذ الفروق الكسرية لها و أستخدام الأنموذج الهجين (ARFIMA-ANN) لها ومقارنته مع النماذج غير الهجينة (ANN,ARFIMA) .

3- تغيير نوع الشبكة من الشبكة العصبية الأمامية (FNN) الى الشبكة العصبية الأسترجاعية كشبكة (Elmans) أو شبكة (Jordan) في دراسات اخرى بغية زيادة دقة التنبؤات .

المصادر

1. العباسي،عبد الحميد ،(2013) ،"مقدمة في الشبكات العصبيةالاصطناعية وتطبيقاتها في العلوم الاجتماعية بأستخدامSPSS"،قسم الإحصاء الحيويوالسكاني/معهد الدراسات والبحوث الإحصائية .
2. فاندل ، والتر،(1983)،"السلاسل الزمنية من الوجهة التطبيقية ونماذج بوكس- جنكينز" ، تعريب عبد المرضي عزام ، المملكة العربية السعودية .
3. وزارة النقل/ المنشأة العامة للطيران المدني/قسم التخطيط والمتابعة، بيانات لأعداد المسافرين الشهرية للفترة الزمنية (2006-2014) لمطار بغداد الدولي ، بيانات غير منشورة .
4. Box G, E.P & Jenkins, G. M.,(1976), "Time series analysis forcasting and control sanfrancisco Helden-day" .
5. Dr.Andranik Macrdechian, (2003)," Designing a neural network for forecasting financial and economic time series " .
6. Fengxia Zheng and Shouming Zhong,(2011),"Time series forecasting using a hybrid RBF neural network and AR model based on binomial smoothing".
7. G. Zhang, B. E. Patuwo, Michael Y. Hu. , (2001), " The Use of Neural Networks in Forecasting", Review of Economic Sciences, 6, TEI of Epirus, pp. 161-176.
8. I. Kaastra & M. Boyd , (1996),"Designing a neural network for forecasting financial and economic time series", Neurocomputing, 10,215–236.

9. **Julian Faraway , Chris Chatfield , (1998), " Time Series Forecasting with neural networks : acomparative study using the airline data " Appl. Statist. 47, part 2 , pp. 231-250.**
10. **Kihoro, J.M, Otieno, R.O. , Wafwa, (2004), "SEASONAL Time Series Forecasting ;Acomparatime Study of ARIMA and ANN models", African Journal of Science and Technology (AJST) science and Engineering series.**
11. **Makridakis, S., Wheelwright, SC, and Hyndman, R. (1998), "Forecasting Methods and Applications", 3rd ed., John Wiley & Sons, New York.**
12. **Mirza Cilimkovic , (2011), " Neural Networks and Back Propagation Algorithm", Institute of Technology Blanchardstown , Ireland.**
13. **N. Merh, et al.,(2010), "A comparison between Hybrid Approaches of ANN and ARIMA for Indian stock trend forecasting," Business Intelligence Journal, vol. 3, pp. 23-43 .**
14. **Šterba Ján and Hilovsda Katarína ,(2010),"The Implementation of Hybrid ARIMA-Neural Network Prediction Model for Aggregate Water Consummation Prediction ", Faculty of Mechanical Engineering – Slovak University of Technology in Bratislava.**
15. **William W.S. Wei, (2006) "Time Series Analysis Univariate and Multivariate Methods".**
16. **Zhang, G.-P. , (2003),"Time series forecasting using a hybrid ARIMA and neural network model", Neurocomputing, 50,159–175 .**

.....

.....

.....