

التنبؤ بانتاج محصول البطاطس باستخدام نماذج SARIMA والشبكات العصبية

أ.د/ محمد المهدي محمد علي

أستاذ الرياضيات والاحصاء الاكتواري المتفرغ

وقائم بعمل رئيس قسم الاحصاء والرياضيات والتأمين

كلية التجارة- جامعة بورسعيد

أ.م.د/ رانيا احمد حامد

استاذ الاحصاء المساعد

كلية التجارة-جامعة بورسعيد

الباحثة

نبيللي سمير عبد الموجود شلبي

ملخص

تعتبر البطاطس من المحاصيل الخضر الأساسية في مصر وبلغ متوسط المساحة المنزرعة حوالي 200 ألف فدان وتعطى إنتاجية كلية تقدر بأكثر من 2 مليون طن سنويا والتي تنتج طوال العام في العروات الثلاث هم العروة الصيفية والعروة النيلية والعروة الشتوية.

و بتتبع إنتاجية محصول البطاطس في الفترة 1983-2014 من خلال التنبؤ بإنتاجية البطاطس مع إدخال المعالجات الإحصائية الدقيقة لدراسة وبناء نموذج للتنبؤ يعتمد على أسلوب بوكس و جينكنيز لتحليل السلاسل الزمنية الموسمية SARIMA و نموذج الشبكات العصبية الاصطناعية والمنطق المضبب وذلك لمعرفة أيهما أفضل وأكثر دقة. فضلا عن التعرف على أهم العوامل التي تؤثر على إنتاجية هذا المحصول للتنبؤ بالقيم المستقبلية.

وقد بينت الدراسة ان الشبكات العصبية ذات الانتشار العكسي BPNN وهى الافضل بين انواع النماذج العصبية من حيث السرعة والتدريب والدقة وعلى الرغم من مميزات نموذج نظام الاستلال العصبي المتكيف بقدرته على معالجة مختلف انواع البيانات الخطية وغير الخطية الي انه فشل في البيانات صغيرة الحجم.

Abstract:

Potatoes are the main vegetable crops in Egypt. The average cultivated area is about 200 thousand feddans and the total productivity is estimated at more than 2 million tons annually, which is produced throughout the year in the three lugs are the summer hook.

The potato yield was analyzed in the period 1983-2014 by predicting potato productivity with the introduction of micro-statistical treatments to study and construct a model based on the Box and Jenkins method for the analysis of the seasonal time series **SARIMA** and the model of artificial neural networks and logic. To identify the most important factors that affect the productivity of this crop to predict future values.

The study showed that the neurotransmitter networks (BPNN), the best among the neural models in terms of speed, training and accuracy. Despite the advantages of the model of the neural system adaptive ability to address various types of linear and nonlinear data to the failure in small data.

المقدمة :

يوجد العديد من الأساليب الاحصائية الممكنة لوصف السلوك الزمني الظاهرة منها أسلوب بوكس جينكينز إذ يزودنا بمنهجية نمذجة إحصائية شاملة يغطي تنوعاً واسعاً من الأنماط مثل الإستقرارية وعدم الإستقرارية والموسمية . إن إحدى المشكلات تطوير واستخدام نماذج بوكس – جينكينز هي متطلبات صياغة النموذج من حيث التوزيع الاحتمالي المفترض للبيانات ، إذ تعد في بعض الأحيان غير ملائمة لحالات محددة فقد تفشل في وصف السلوك المتغير ديناميكياً مع الزمن¹.

وهناك طرق أخرى تعتمد على الأساليب غير الخطية مثل الشبكات العصبية Neural Networks والمنطق الضبابي او الغامض Fuzzy Logic ومن مزايا هذه الشبكات تعدد استخدامها وعدم تطلبها نموذجاً قياسياً بمواصفات أو توزيعات احتمالية مفترضة للبيانات .

إن نموذج الشبكات العصبية الذي يقدم مطابقة جيدة للبيانات ربما يعطي تكهنات ضعيفة خارج العينة Out-of-Sample ، لذلك فإن التحليل الجيد للبيانات يتطلب استخدام مهارات النمذجة الإحصائية لاختيار نموذج الشبكات العصبية المناسب . كما أن العلاقة بين نماذج الشبكات العصبية ونماذج بوكس جينكينز يمكن أن تكون علاقة تكامل لا تنافس وصولاً إلى نماذج جيدة².

محصول البطاطس في مصر

¹ انتصار ابراهيم الياس " بناء نموذج موسمي لتحليل السلسلة الزمنية لمعدلات الطاقة الكهربائية المجهزة لمدينة دهوك والتنبؤ بها " كلية الإدارة والاقتصاد – جامعة نوروز – المجلد 6 العدد1(2017) ص 153
² ظافر رمضان مطر واخرون " تحليل ونمذجة السلسلة الزمنية لتدفق المياه الداخلة الي مدينة الموصل دراسة مقارنة " - المجلة العراقية للعلوم الاحصائية (18) 2010 ص 3

تعتمد مصر في جزء كبير على صادراتها على المنتجات الزراعية ولذلك تسعى الي التوسع في انتاج الزروع ذات الميزة النسبية في انتاجها والتنافسية في تصديرها. 1

و يحتل محصول البطاطس في مصر مركز الصدارة بالنسبة لمحاصيل الخضر التصديرية مما دعى في السنوات الخيرة تطوير تكنولوجيا تصنيع البطاطس وطرق تجهيزها وحفظها بدرجة كبيرة وادى ذلك الي اطالة فترة الاستفادة منها وتنوعها بما يناسب رغبات المستهلكين كما ادخلت العديد من اصناف ذات المواصفات التصنيعية العالية مما ادى الي زيادة معدل استهلاك الفرد ليصل الي 25-30 كجم للفرد سنويا سواء البطاطس الطازجة او المصنعة. 2

ويعتبر محصول البطاطس في مصر من الزروع التي تنتج طوال العام في العروات الثلاث (الصيفية والنبيلية والشتوية) والتي يمكن زراعتها خلال شهري اكتوبر ونوفمبر لأغراض تصديرية حيث يبدأ الحصول في التصدير اوائل يناير وحتى اخر مارس وخاصة في الاسواق الاوروبية كم تجدر الإشارة ان مصر تحتل المرتبة الثانية عشر بين دول العالم من حيث الانتاجية والمساحة. 3

1- مشكلة الدراسة

تتلخص مشكلة البحث في تتبع انتاجية محصول البطاطس خلال الثلاثون عاما الماضية في مصر من خلال التنبؤ بإنتاجية البطاطس مع إدخال المعالجات الإحصائية الدقيقة لدراسة وبناء نموذج دقيق لتنبؤ يمكن من خلاله اتخاذ قرارات انتاجية أفضل.

2- الهدف من الدراسة

1 المصري , منى محمود ابراهيم " دراسة اقتصادية لأثر التدخل الحكومي على السياسة السعرية لمحصول البطاطس ومستقبل صادراته في ضوء العولمة والتنافسية " مجلة مصر المعاصرة - مصر مج 93 و468, 467 لسنة 2002 ص ص 395 -445

2 " زراعة وانتاج البطاطس "نشرة مركز البحوث الزراعية ص2

3 المصري , منى محمود ابراهيم - مرجع سبق ذكره ص825

إن الهدف الرئيسي للبحث هو :- التنبؤ بالإنتاجية المستقبلية لمحصول البطاطس في مصر من خلال التعرف على البيانات الماضية وذلك لاتخاذ القرارات المناسبة التي من شأنها تحسين إنتاجية هذا المحصول

وينبثق من هذا الهدف الرئيسي هدفا آخر يتمثل في :- بناء نموذج للتنبؤ يعتمد على أسلوب السلاسل الزمنية الموسمية SARIMA وكذلك بناء نموذج الشبكات العصبية الاصطناعية وذلك لمعرفة أيهما أفضل وأكثر دقة. فضلا عن التعرف على أهم العوامل التي تؤثر على إنتاجية هذا المحصول للتنبؤ بالقيم المستقبلية.

3- أهمية الدراسة

تسعى جميع دول العالم إلى رفع مستواها الاقتصادي، الذي يؤثر بدوره ايجابا في جميع القطاعات الأخرى.

وتكمن أهمية هذه الدراسة في أنها تناقش أهم العوامل التي يمكن أن تؤثر في إنتاجية البطاطس ، ذلك أن معرفة مدى تأثير هذه العوامل والأهمية النسبية لهذا التأثير قد تساعد في عملية التنبؤ بحجم الانتاجية وبالتالي وضع الخطط المستقبلية وللإستفادة المثلي من انتاج البطاطس والحصول على نتائج أفضل وأكثر دقة للتنبؤ بالقيم المستقبلية لإنتاج هذا المحصول.

4- حدود الدراسة

أ- البرامج المستخدمة

- ❖ يستخدم R لتحليل السلاسل الزمنية والتنبؤ بالقيم المستقبلية لإنتاج محصول البطاطس
- ❖ كما يستخدم برنامج Matlab لاختبار نموذج الشبكات العصبية المناسب من خلال طريقتين هما

خوارزمية الانتشار العكسي Back-Propagation

خوارزميات نظام الاستلال العصبي الضبابي المتكيف

Adaptive Neuro Fuzzy Inference System (ANFIS)

❖ ويتم مقارنة ما تم التوصل اليه من النماذج الثلاثة للوقوف على افضلية أي منهما في عملية التنبؤ

ب- فترة الدراسة

❖ تجرى الدراسة علي بيانات سنوية لمدة ثلاثون عاما خلال الفترة (1983- 2014)

ج- مصادر الحصول علي البيانات

❖ مبنى الهيئات والحكومات بالدقي قسم البيانات الاحصائية الزراعية

❖ الابحاث والرسائل العلمية المنشورة على بنك المعرفة المصري

5- فروض البحث

اسلوب الشبكات العصبية الاصطناعية افضل للتنبؤ من اسلوب بوكس و جينكينز من حيث الدقة والسهولة وبخاصة السلاسل الزمنية الموسمية ذات القيم الشاذة .

اسلوب الاستدلال العصبي الضبابي المتكيف افضل من اسلوب الشبكات العصبية الاصطناعية واسلوب بوكس و جينكينز بما له من قدرة على معالجة مختلف انواع البيانات الخطية وغير الخطية.

منهجية البحث

ان دراسة السلاسل الزمنية لها اهمية بالغة لما تقدمه من معلومات حول العناصر الاساسية التي تتميز بها ظاهرة ما عبر الزمن ومن خلال متابعة تغيراتها وتطورها العام يمكننا بصورة جيدة من معرفة كيفية تطورهما مستقبلا كما في تحديد مختلف العوامل المؤثرة على هذه الظاهرة 1

1 فريد بختي " السلاسل الزمنية الخطية باستعمال حزمة Eviews7.0" الجمهورية الجزائرية الديمقراطية الشعبية – كلية العلوم الاقتصادية ، والتجارية وعلوم التسيير – السنة الجامعية 2014/2015 ص22

كما يعد أسلوب التحليل الذي قدمه العالمان بوكس و جينكينز في كتابهما الصادر عام (1970) احد الادوات الاحصائية التي يمكن الاعتماد عليها في تحليل السلاسل الزمنية حيث يتميز هذا الاسلوب.

1

- 1- هي منهجية شاملة تقدم حلول جيدة لجميع مراحل التحليل في شكل منظومة اكثر علمية ومنطقية لبناء النماذج وتشخيصها وتقدير معالمها والتنبؤ بالمشاهدات المستقبلية .
 - 2- تميز وثرء النماذج العشوائية التي تظهر في مجالات التطبيق المختلفة والتي تعرف بنماذج الانحدار الذاتي والمتوسطات المتحركة التكاملية او نماذج ARIMA .
 - 3- لا تفترض الاستقلال انما تستغل الارتباط الموجود بين هذه المشاهدات في عملية النمذجة والتنبؤ مما يؤدي الي تنبؤات اكثر مصداقية ودقة من الطرق التقليدية .
 - 4- تعطي فترات ثقة ذات مصداقية عالية للقيم المستقبلية مقارنة بالاساليب الاخرى .
- ولتحقيق ذلك يتطلب الامر دراسة تحليلية وافية لنماذج السلاسل الزمنية بالاعتماد على الاساليب الاحصائية والرياضية .

مركبات السلاسل الزمنية 2

تكون السلاسل الزمنية من مجموعة من المركبات التي تساعد على معرفة سلوك السلسلة وتحديد مقدار تغيراتها وادراك طبيعتها واتجاهها من اجل القيام بالتقديرات اللازمة والتنبؤات الضرورية وهذه العناصر هي:-

1- مركب الاتجاه العام 3 Long term of secular trend(T)

وهو العنصر الذي يقصد به الحركة المنتظمة عبر فترة زمنية طويلة نسبيا ويعتبر اهم العناصر المكونة للسلسلة.

¹ سمير مصطفى شعراوى " مقدمة في التحليل الحديث للسلاسل الزمنية " مركز النشر العلمى - جامعة الملك عبد العزيز جدة 21589 ص 92

² زمران كريم " نحو تقدير امثل لمعدلات الخسارة في شركات التأمين " - مجلة ابحاث اقتصادية وادارية العدد السادس عشر ديسمبر 2014 جامعة خنشلة - الجزائر ص188

³ عاشور - بدار " المفاضلة بين نموذج السلاسل الزمنية ونموذج الانحدار البسيط في التنبؤ بحجم المبيعات في المؤسسة الاقتصادية دراسة حاله (مطاحن الحزن بالمسيلة) - كلية العلوم الاقتصادية وعلوم التسيير والعلوم التجارية - جامعة محمد بوضياف بالمسيلة عام2005-2006 ص56-63

كما يقصد به تطور السلسلة في الاجل الطويل وقد يكون متزايدا اذا كانت القيمة الظاهرة تتزايد عبر الزمن او قد يكون متناقصا اذا ما اتجهت قيمة الظاهرة للتناقص و قد يأخذ شكل خط مستقيم في حالة ثبات السلسلة ويعكس الاتجاه العام تأثير العوامل طويلة الاجل على السلسلة الزمنية.

2- مركب الموسمية (S) Seasonal variation

وتعبر عن التغيرات والتذبذبات الموسمية الناتجة عن تغيرات الفصول بسبب تأثير عوامل خارجية وهى غالبا ما تتم بطريقة منتظمة في شكل دورات لا يزيد طولها عن سنة فقد تكون اسبوعية او شهرية او فصلية

3- مركب الدورية (c) Cyclical Movement

هو التغيرات التي تطرا على قيم السلسلة الزمنية بصورة منتظمة او غير منتظمة ويزيد طولها عن السنة وتتكون من دوال الجيب وشبه جيب التمام

4- مركبة العشوائية (I) Irregular variation

يقصد به التغيرات او التحركات المفاجئة في السلسلة الزمنية والتي ترجع الي عوامل عشوائية مثل الزلازل والبراكين والحروب وهى ايضا عبارة عن تذبذبات او تغيرات غير منتظمة (عشوائية) عادة ما تكون تابعة لأحداث مرتبطة باحتمالات ضعيفة جدا ومن ثم فإنها لا تقع في المدى الطويل
مراحل بوكس جينكينز¹

يمكن اعتبار منهجية بوكس وجينكينز بمثابة نظام تنبؤ كامل يمكن استخدامه مع معظم السلاسل الزمنية ولتحقيق ذلك يتطلب بناء نموذج للتنبؤ باستخدام بوكس وجينكينز مايلي :-

1- التعرف على النموذج Identification Model

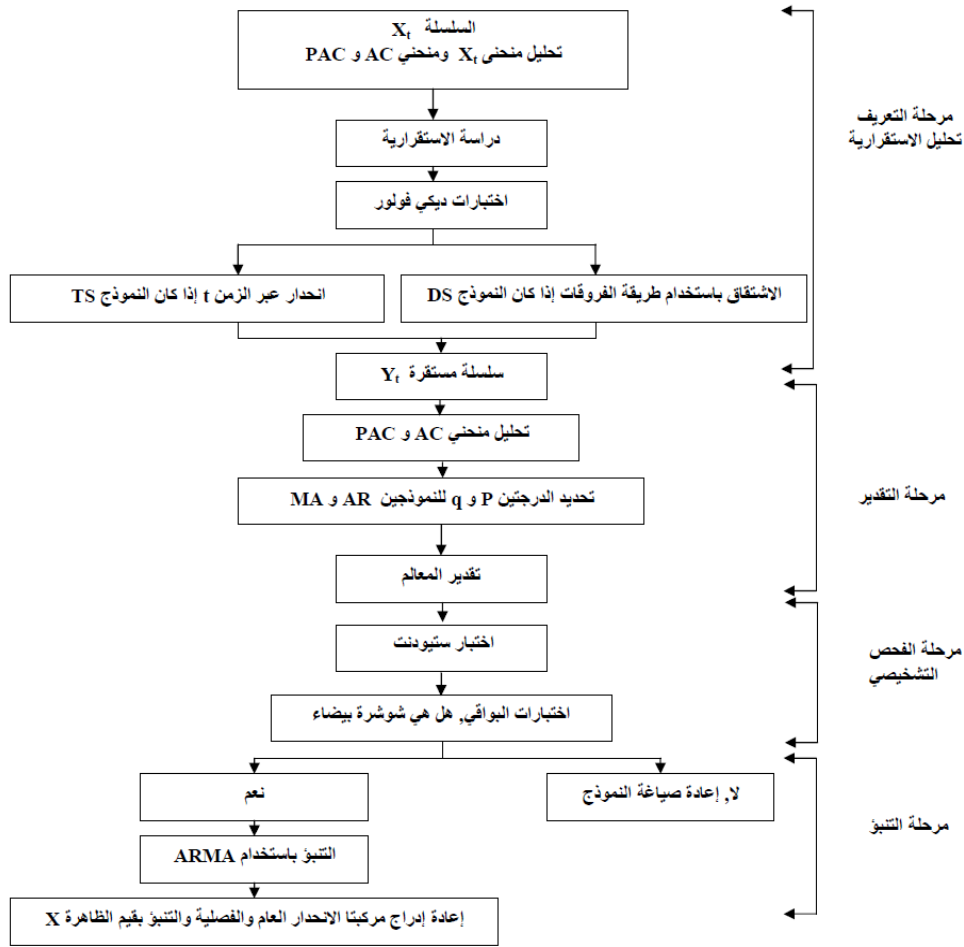
2- تقدير النموذج Model Estimation

3- تشخيص واختبار النموذج Model Diagnostic

¹رابح بلعباس " فعالية التنبؤ باستخدام النماذج الإحصائية في اتخاذ القرارات " ص9

4- التنبؤ Forecasting

عرض طريقة بوكس جينكينز:



مراحل منهجية بوكس جينكينز

source: régis bourbonnais, économétrie 6iem édition, p248

يوضح الشكل (112) مراحل منهجية بوكس وجينكينز¹

الشبكات العصبية الاصطناعية Artificial Neural Networks 2

الشبكات العصبية الاصطناعية 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 ماهي الا العناصر حسابية تسمى عصبونات او عقد)

¹ Source: regis Bourbonnais, econometrie 6iem édition , p248

² قصي حبيب الحسيني و اخرون " مقدمة في الشبكات العصبية الاصطناعية " جامعة الامير جعفر الصادق عليه السلام ص3

(Nodes, Neurons) والتي لها خاصية عصبية حيث تقوم بتخزين المعرفة العملية والمعلومات التجريبية لتجعلها متاحة للمستخدم عن طريق ضبط الأوزان

اهم مميزات الشبكات العصبية 1

الشبكات العصبية أكثر مرونة في البيئات المتغيرة لأن النظم المبرمجة وعندما تحدث تغيرات مفاجئة تصبح هذه النظم غير صالحة وعلى الرغم من ان الشبكات العصبونية تأخذ بعض الوقت لتتألم مع التغيرات المفاجئة إلا أنها تمتاز بالتأقلم مع التغيرات الحاصلة.

1- التعلم التكيفي Adaptive Learning وهي القدرة على التعلم لتنفيذ المهام بالاعتماد على عينات من المعطيات يتم ادخالها في البداية كخبرة بدائية

2- التنظيم الذاتي Self-Organization يمكن الشبكة ان تولد نمطا لتمثيل البيانات التي يتم استقبالها من مرحلة التعلم

3- عمليات في الزمن الحقيقي Real Time Operation بمعنى أنه يمكن تنفيذ حسابات وعمليات على التوازي.

4- تسامحيه الأخطاء Fault Tolerance بمعنى أن التعطيل الجزئي للشبكة يمكن ان يقود الي تخفيض الأداء ولا يؤدي الي توقف العمل

التطبيق العملي:

اولا :- مرحلة التعرف

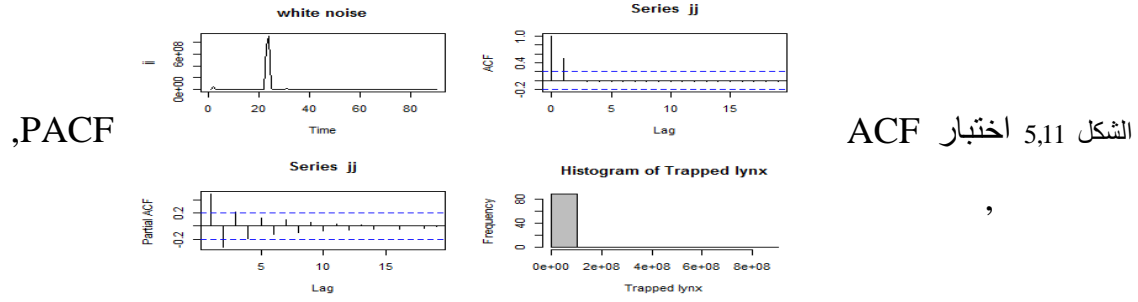
يتم ادخال بيانات السلاسل الزمنية لثلاث عروات هما العروة الصيفية والعروة النيلية والعروة الشتوية لتظهر البيانات على النحو التالي

ويتضح من الجدول السابق ان هناك سنوات تبدأ من عام 1983 : 1990 لم تتم الزراعة في العروة الشتوية حيث لذا يتم ادخالهم كسلسلة واحدة وذلك لإظهار مدى التغيرات التي تطرأ على السلسلة الزمنية وليتم تحديد طبيعتها ودراسة استقرار السلسلة ويتم التعويض عن قيم NA بالواحد الصحيح وذلك لان اللوغاريتم NA غير معرف وبالتالي تكون النتائج غير صحيحة

1 عروة نادر " استخدام الشبكات العصبونية للتنبؤ في الأسواق المالية : تطبيق لتوقع تغيرات قيم الأسهم في سوق دمشق للأوراق المالية " - المعهد العالي للعلوم التطبيقية والتكنولوجيا - الجمهورية العربية السورية ص 20

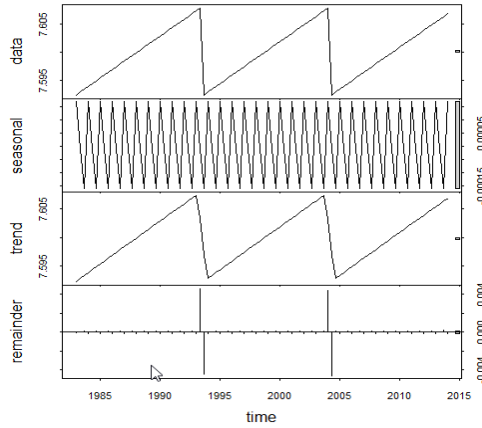
اختبارات تحديد طبيعة السلسلة الزمنية:

يصعب تحديد طبيعة استقرار السلسلة الزمنية بالرسم البيان ومن هنا نلجأ إلى استخدام مقاييس إحصائية لاختبار وجود أو عدم وجود اتجاه عام للسلسلة ومن أبسط هذه المقاييس وأكثرها استعمالاً باختبار معنوية هذا الاختلاف عن طريق الارتباط الذاتي ACF الارتباط الذاتي الجزئي PACF, كما في الشكل 5,11



ومن الشكل 5,11 نجد ان السلسلة غير مستقرة ولذا نقوم بتحليل مكونات السلسلة كما في

الشكل التالي 5,12

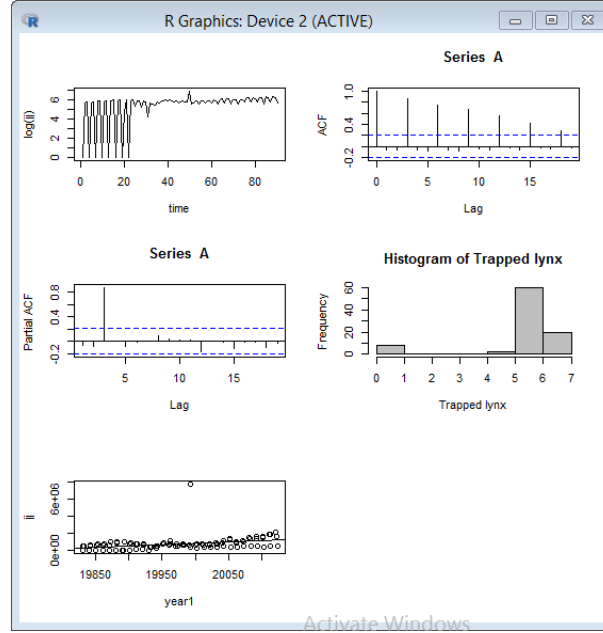


الشكل 5,12 تحليل مكونات السلسلة

يتضح من الشكل 5,12 ان هناك اتجاه عام وموسمية لذا سيتم تثبيت الاتجاه العام حيث ان قيمة ارتباط الذاتي تكون درجة الابطاء قريبة او مساوية للصفر

4- لوقي فاتح "جودة نماذج السلاسل الزمنية الموسمية المختلطة SARIMA في التنبؤ بالمبيعات دراسة حالة مؤسسة مطاحن جديع بقر ، رسالة ماجستير ، كلية العلوم الاقتصادية والتجارة وعلوم التسيير ، جامعة محمد خضير - بسكرة للعام 2013/2014 (مرجع سبق ذكره)

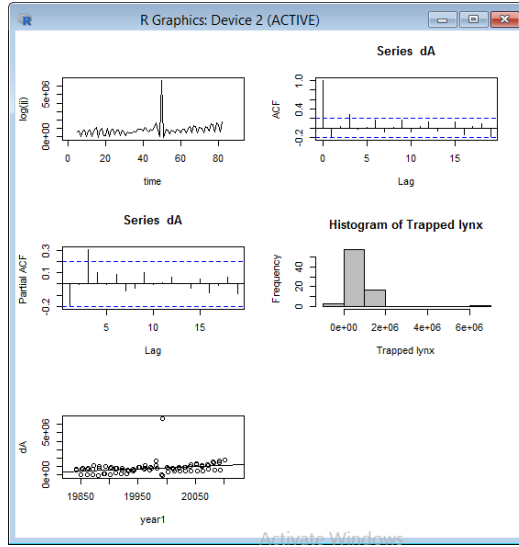
ولجعل السلسلة مستقرة يتم تثبيت التباين قبل الكشف عن مركبات السلسلة الزمنية ويتم ذلك بعدة طرق منها التحويلات اللوغاريتمية او الجذر التربيعي وسوف نأخذ بطريقة اللوغاريتمية \log_{10} ثم يتم التأكد من استقرار السلسلة بفحص وذلك باختبار معنوية هذا الاختلاف عن طريق الارتباط الذاتي ACF الارتباط الذاتي الجزئي PACF, من الشكل التالي 5,13



الشكل 5,13 اختبار ACF, PACF

يتضح من الشكل 5,13 ان السلسلة الزمنية غير مستقرة حيث نجد ACF غير مستقر PACF شبه مستقرة عدا قيمة وحيدة اما خط الانحدار يكاد يكون متطابق لذا نستخدم وقد يرجع ذلك للتأثير الموسمي كما يوضحه

5,12 نجد ان هناك تأثيرات موسمية لذا نقوم بفلتره البيانات الاصلية jj ثم نقوم بطرح البيانات الاصلية من البيانات المفلتره ويتضح ويظهر في الشكل التالي ($t1 = filter(jj, filter = c(1/6, 1/3, 1/3, 1/6), sides=2)$) وذلك للقضاء علي التأثير الموسمي كما انه مع استخدام $Lag(2)$ فانه يظهر نفس النتيجة السابقة ومع استخدام اكثر من محاولة $Lag(5)$ فانه يعطي نفس النتائج حيث تم الثابت على الفرق الثاني ويتم التأكد من ذلك عن طريق اختبار فحص الارتباط الذاتي ACF و الارتباط الجزئي PACF كما في الشكل 5,15



الشكل 5,15 اختبار ACF,PACF والمدرج التكراري وخط الانحدار

ولذا نجد ان هناك قيم شاذة لذا نتبع خطوات ARCH و GARCH المناسب لأي بيانات يمر بعدة خطوات يمكن ذكرها علي الشكل التالي

1- اول الخطوات في عملية بناء هذه النماذج هي عملية فحص البيانات والتأكد من سكونها وذلك برسم البيانات وعمل الإحصاءات الوصفية لها واستخدام احد اختبارات فحص السكون مثل ADF كما في الشكل 5,16

```
> # ADF test for dA data
> adf.test(dA, nlag = NULL, output = TRUE)
Augmented Dickey-Fuller Test
alternative: stationary

Type 1: no drift no trend
lag   ADF p.value
[1,]  0 -13.42  0.01
[2,]  1 -11.36  0.01
[3,]  2  -7.48  0.01
[4,]  3  -6.50  0.01
Type 2: with drift no trend
lag   ADF p.value
[1,]  0 -13.32  0.01
[2,]  1 -11.27  0.01
[3,]  2  -7.43  0.01
[4,]  3  -6.45  0.01
Type 3: with drift and trend
lag   ADF p.value
[1,]  0 -13.21  0.01
[2,]  1 -11.18  0.01
[3,]  2  -7.36  0.01
[4,]  3  -6.40  0.01
----
Note: in fact, p.value = 0.01 means p.value <= 0.01
```

الشكل 5,16 اختبارات فحص السكون ADF

ونجد ان $p. value \leq 0.01$ المحسوبة أي عندما تكون اقل من $p. value < 0,05$ ذلك يعني ان الدالة ساكنة

ثم القيام بعدة محاولات لاختيار نموذج SARIMA(2,2,2)(2,3,2)₃

بعد اختبار اكثر من نموذج ويتم المفاضلة بينهما حسب معيار اكايكي AIC وذلك باختيار اقل قيمة

SARIMA(2,2,2)(2,3,2)₃ = 2326.47

وتم اختيار هذه السلسلة من حيث الاقل في اختبار AIC انظر ملاحق

ثم بعد تحديد رتبة النموذج يتم اختبار Ljung.box.test لتظهر النتائج كما بالشكل 5,18

```
> library(ccgarch)
> nNA <- is.na(mod$res)==F
> ljung.box.test(mod$res[nNA])
      test stat  p-value
Lag 5   3.609629 0.6068681
Lag 10  9.406089 0.4940459
Lag 15 10.270124 0.8024131
Lag 20 10.889867 0.9490287
Lag 25 11.486470 0.9902382
Lag 30 13.410222 0.9960875
Lag 35 13.884391 0.9994594
Lag 40 13.962948 0.9999572
Lag 45 13.973561 0.9999976
Lag 50 13.993187 0.9999999
```

الشكل 5,18 اختبار Ljung.box.test

ومن هذا الاختبار Ljung.box.test نجد ان $P\text{-value} > .05$ وبذلك نجد ان هذا الاختبار لا يدل

على صحة النموذج لذا سنقوم باجراء ARCH test

ARCH test

يستخدم في اختبار عشوائية اخطاء السلسلة الزمنية أي اختبار ان الأخطاء تتبع توزيع طبيعي متمائل مستقل من خلال تمثيل T من قيم مربعات الاخطاء العشوائية لنموذج GARCH في النموذج انحدار بحد ثابت ومن ثم اختبار وجود تأثير للارتباط الذاتي وتكتب بالصيغة التالية

```
> ArchTest(c(mod$res))

ARCH LM-test; Null hypothesis: no ARCH effec

data: c(mod$res)
Chi-squared = 35.831, df = 12, p-value = 0.0003449
■
264
```

الشكل 5,19 ARCH test

فاذا كانت نتائج الاختبارات ايجابية وتدل على وجود اثر لـ ARCH ويتم استكمال باقي خطوات بناء النموذج GARCH واذا كانت النتائج سلبية و لا يوجد اثر ARCH فيتم توقيف عملية بناء

ومن الشكل 5,19 نقارن احصاءات الاختبار مع القيم الجدولية لاختبار مربع كأي بدرجة حرية (p) $\chi^2_{(p)}$ وعند مستوى معنوية α فاذا كانت القيمة المحسوبة اكبر من الجدولية نرفض الفرضية H_0 أي لا يوجد تأثير ARCH والعكس صحيح

وبالنظر في جدول نجد ان قيمة $\chi^2_{(p)}$ المحسوبة 35.831 < قيمة $\chi^2_{(0.05)}$ الجدولية 21.026 يوجد لا تأثير ARCH

1- وبعد التأكد من عدم وجود اثر لـ ARCH في السلسلة من خلال الاختبارات السابقة تبدأ عملية التعرف وتقدير النموذج المناسب للبيانات حيث يتم التعرف على النموذج الأفضل من خلال المفاضلة بين النماذج المقدره حسب معياري BIC,AIC والنموذج الذي يحصل على اقل قيمة في المعيارين بالنسبة لباقي النماذج يكون هو النموذج المناسب حيث تمثل عدد ازاحات الارتباط الذاتي p عدد المعلمات المقدره في النموذج فاذا كانت $Q_m < \chi^2_{\alpha}(m-p)$ أي ان الاخطاء عشوائية و لا يوجد تأثير ARCH والعكس صحيح

وبذلك تم التأكد من ان النتائج ايجابية لذا فان النموذج مناسب وتتم عملية التنبؤ

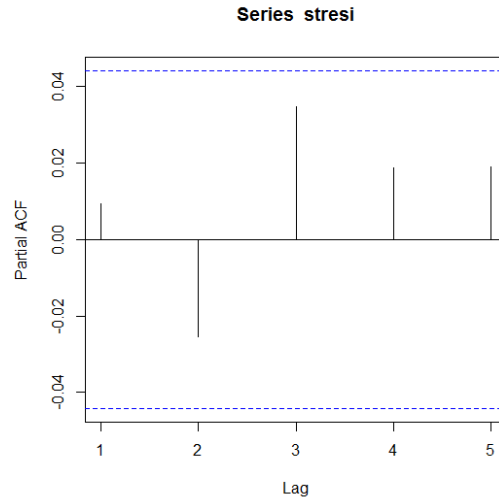
اخر الخطوات واهمها والهدف الرئيسي من عملية بناء النموذج المناسب وهي التنبؤ¹

التنبؤ يمثل المرحلة الاخيرة في بناء نماذج بوكس جينكينز و لا يمكن الانتقال الي هذه المرحلة الا بعد ان يجتاز النموذج المبدئي كافة الفحوص واختبارات التشخيصية للتأكد من صحة النموذج اذا لم يجتاز هذه الفحوص والاختبارات فإنه يجب تعديله واستبداله بنموذج اخر اكثر كفاءة

¹ خلود جمال عوض الله " استخدام نماذج SARIMA و Holt-Winters في التنبؤ بالسلاسل الزمنية الموسمية " كلية الاقتصاد والعلوم الادارية - جامعة الازهر - غزة عام 2016

على النحو المشروح في مرحلة التشخيص مع إجراء نفس خطوات مرحلة التشخيص على النموذج المعدل والتأكد من انه اجتازه بنجاح

وفي معظم حالات التنبؤ تعتبر الدقة هي مقياس الأساسي في اختيار الطريقة المناسبة ومهما كانت طريقة التنبؤ تميل للدقة وللتأكد من ذلك لابد من تقييم جودة التنبؤ بمقارنه القيم الحقيقية بالقيم المقدره هذه المقارنة تكشف لنا عن حجم الاخطاء في التنبؤ او جودة التنبؤ



الشكل 5,20 PACF

تعتمد اغلب مقاييس دقة التنبؤ على الانحرافات بين القيم الفعلية للسلسلة والقيم و المقدره

وذلك من خلال مقاييس دقة التنبؤ

متوسط الخطأ المطلق Mean Absolute Error - MEAN :

لتجاوز السلبيات في معدل الخطأ وهي تلاشي قيم الخطأ نتيجة اختلاف إشارة قيم الخطأ استعمل هذا المقياس ولكن تؤخذ القيمة المطلقة لكل خطأ فيكون المتوسط القيم دون النظر الي اشارتها كما في المعادلة (5,2)

$$MAE = \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{n} = 277825.3 \quad (5,2)$$

حيث y_t القيم المشاهدة و \hat{y}_t القيم المحسوبة و n عدد المشاهدات

انواع متوسط الخطأ المطلق

متوسط الخطأ المطلق النسبي Mean Absolute Percentage

عبارة عن متوسط القيمة المطلق للخطأ وهو عبارة عن مجموع القيم المطلقة لكل خطأ مقسمة على القيم المشاهدة ثم نقسم على عدد القيم

$$MAPE = \left[\sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{n} \right] / n \times 100 = 298.6573 \quad (5.3)$$

واخر مرحلة هي التنبؤ بالقيم المستقبلية لتكون على النحو التالي

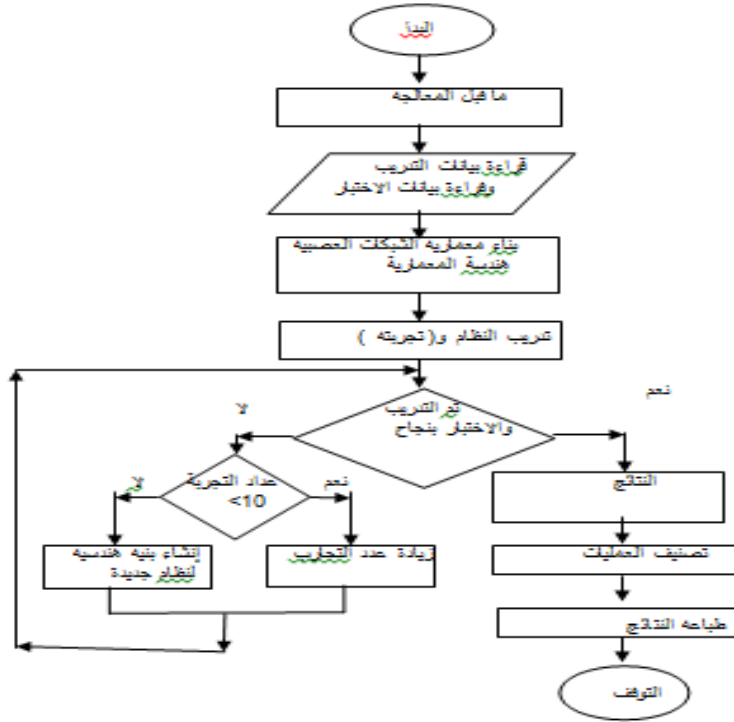
```
> predict(mod, n.ahead=6)
$pred
Time Series:
Start = 91
End = 96
Frequency = 1
[1] 791030.8 272457.0 -966337.1 954828.6 270494.8 -921389.5

$se
Time Series:
Start = 91
End = 96
Frequency = 1
[1] 774470.5 892984.7 911247.1 913049.3 1086406.5 1103652.7
```

ثانياً :- الانتشار العكسي Back-propagation¹

تعليم الانتشار العكسي Back-propagation

يمكن توضيح خطوات الانتشار العكسي كما يلي



خطوات تعليم شبكة الانتشار الخلفي للخطأ (BPNN) 1

بناء الشبكة

تمر مرحلة بناء الشبكة العصبية الاصطناعية بالمرحل التالية²

1-تحديد المدخلات :- ان اول خطوة في بناء شبكة عصبية اصطناعية هو تحديد المدخلات في

Clear the workspace

Close all the previons windows

¹ عمر صابرقاسم واسراء رستم محمد " دراسة رياضية تحليلية لخوارزمية الشبكات العصبية الاصطناعية في ملاءمة نموذج للتشخيص الطبي " مجلة الرافيدين لعلوم الحاسوب والرياضيات المجلد (10) العدد (1) 2013 (مرجع سبق ذكره)
² مناهل عبد الكريم يوسف " دراسة مقارنة بين الشبكات العصبية الاصطناعية (تميز الأرقام الانكليزية المطبوعة " كلية طب اسنان /جامعة الموصل 2009/7/15(مرجع سبق ذكره)

Load the training data (input and output data)

Load the test data

A: Build the back propagation neural Network architecture

Start training

If the training data is okay keep the BPNNT architecture

Start test data

If the test data is okay keep the BPNNT architecture. And Compute the complex variables and stop

Else go to address A

حيث $v1, v2, v3$ هي العروات الصيفية والشتوية وذلك مع استبعاد عامي 2013 و2014 وذلك للتعويض بتلك القيم ومقارنتها مع نموذج السلاسل الزمنية و الانتشار العكسي - Back propagation و نظام الاستدلال العصبي الضبابي المتكيف ANFIS

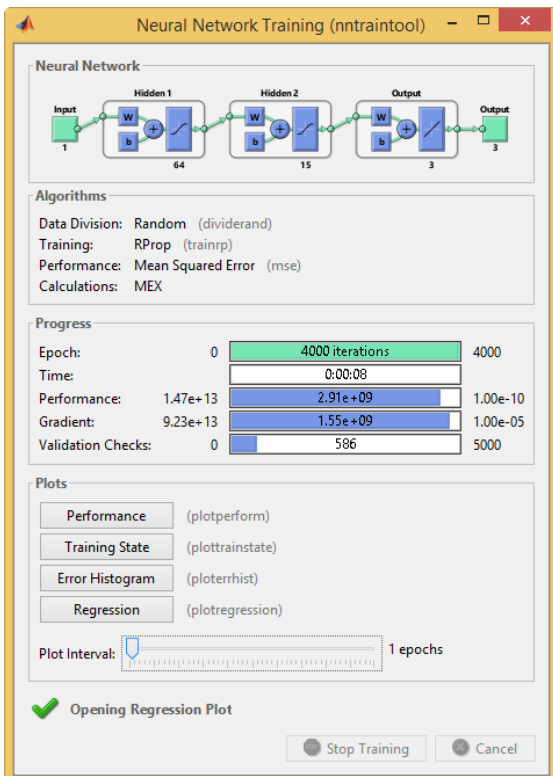
مرحلة التحليل :- بعد القيام بعملية تحليل الشبكة تبينت النتائج التالية

الخطوة الأولى : تهيئة الأوزان الابتدائية $W_N = (w_1, w_2, w_3, \dots, w_n)$ وتمثل θ قيمة حد العتبة (Threshold) حيث ان تهيئة الاوزان تاخذ قيم عشوائية صغيرة بين -0.5 و +0.5

ثم تستقبل كل وحدة ادخال $x_i, i=1, \dots, n, n=90$ اشارة دخل وتحويلها الي مصفوفة ثم تنتشر الى جميع

وحدات الطبقة الخفية وفي هذه المرحلة تم تجميع البيانات محل الدراسة وعددها (96) مشاهدة

بنسبة (75%) كمجموعة تدريب و(25%) كمجموعة اختبار وتاكد حيث يكون عندنا input واحد فقط وهم العروات وطبقتين Hidden



و(3) طبقات مخرجات وتكون السنوات حيث ينتج (6 قيم المتنباؤها وهى العروات)

الخطوة الثانية : اختيار زوج التدريب W_N, y_j حيث تمثل X_N متجه الادخال

$$X_N = (X_1, X_2, X_3, \dots, X_n)$$

وتمثل y_j الاخراج المطلوب حيث

تجمع كل وحدة خفية $z_j, j=1, \dots, p, p=15$ قيم الادخال الموزونة لحساب دالة الطبقة الخفية

Tan Sigmoid

$$z_{inj} = v_{0j} + \sum_{i=1}^n x_i v_{ij}$$

$$z_j = f(z_{inj})$$

ثم ترسل قيمة الدالة الي جميع وحدات طبقة الاخراج

الخطوة الثالثة : في الاتجاه الأمامى حساب قيمة الاخراج الحقيقي كالاتى

1 - حساب قيمة الاخراج الحقيقي من طبقة الادخال الي الطبقة المخفية

$$Y_H = f(net_{iH}), net_{iH} = \left[\sum_{i=1}^m X_{Hi} W_i - \theta_i \right]$$

n تمثل عدد العناصر في طبقة الادخال الشبكة

2- حساب قيمة الاخراج الحقيقي من الطبقة المخفية Hidden Layer الي طبقة الخرج Output

Layer

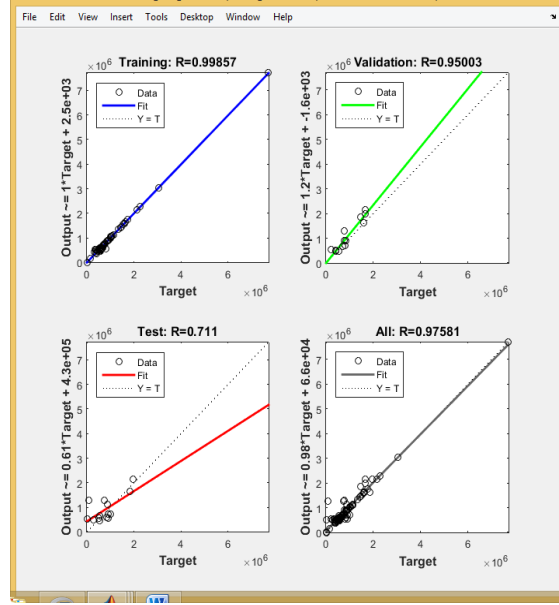
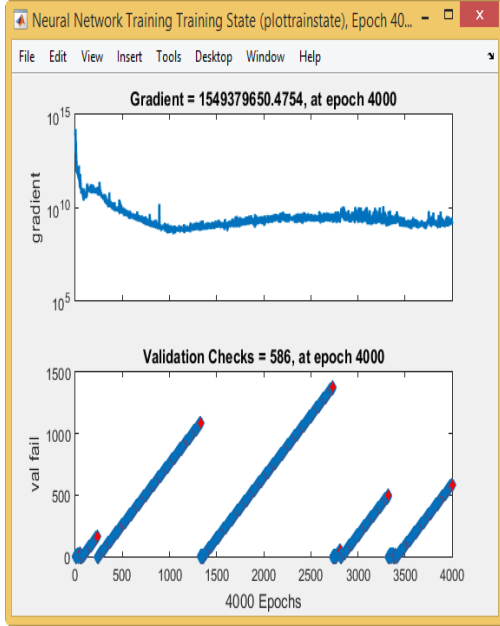
$$Y_0 = f(net_{H0}), net_{H0} = \left[\sum_{i=1}^p y_{Hi} W_i - \theta_i \right]$$

P : تمثل عدد العناصر في الطبقة المخفية للشبكة

حيث يتم ايجاد حاصل جمع كل وحدة اخراج $y_k : k = 1 \dots m, m = 6$ المضروبة بالاوزان كما يلي

$$y_k = f(y_{ink}) \quad \text{وتطبيق دالة حساب الاخراج} \quad y_{ink} = w_{0k} + \sum_{j=1}^6 z_j w_{jk}$$

يتم خلالها تحديد دالة التدريب training test: trainrp



وكذلك

ثم تحديد متوسط مربع الاخطأ Mean Squared Error (mse) performance: ليظهر في الشكل التالي

الخطوة الرابعة : حساب الخطأ ويتم من خلال الخطوات التالية $e_i = y_a - y_i \neq 0$

1- نقوم بتعديل الأوزان بين طبقة الاخراج (output Layer) والطبقة المخفية Hidden Layer

$$W_i^{new} = W_i^{old} + \Delta W_i$$

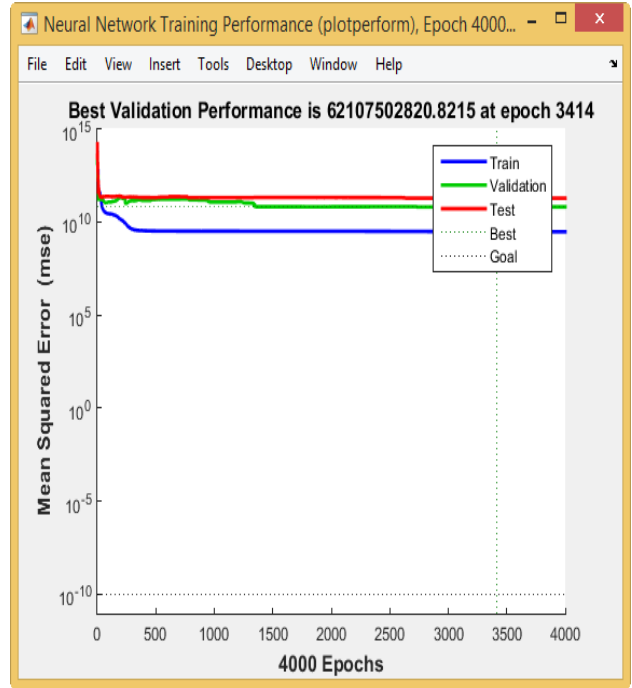
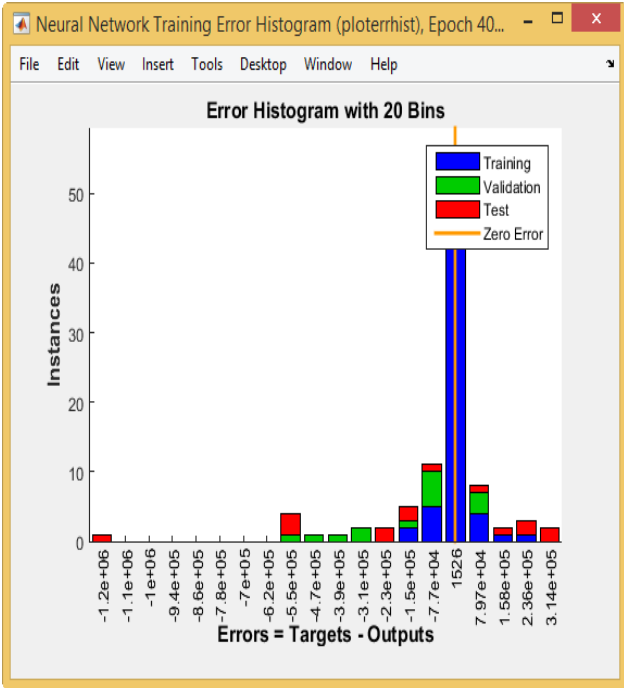
$$\Delta W_i = \alpha y_i \delta_i \dots \text{and} \dots \delta_i = y_i (1 - y_i) e_i$$

2- نقوم بتعديل الأوزان بين الطبقة المخفية Hidden Layer الي طبقة الادخال Input Layer

$$W_i^{new} = W_i^{old} + \Delta W_i$$

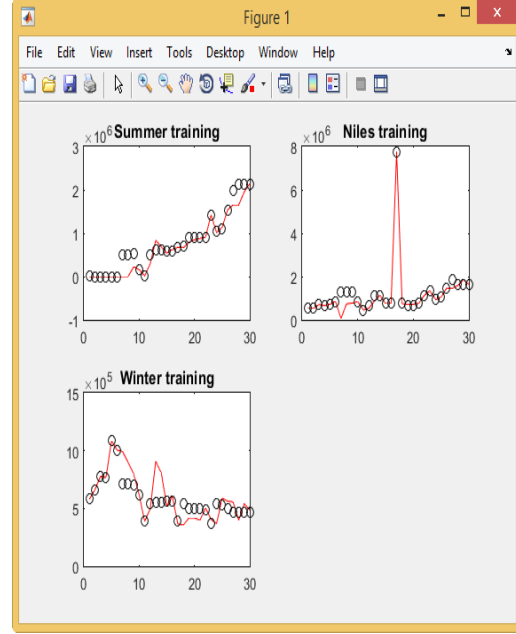
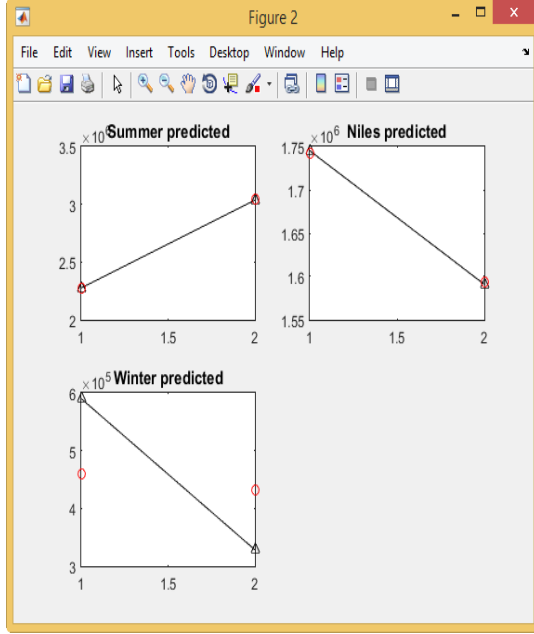
$$\Delta W_i = \alpha y_i \delta_i \dots \text{and} \dots \delta_i = y_i (1 - y_i) \sum_{j=1}^p \delta_j W_j$$

الخطوة الخامسة : تكرار الخطوات من الخطوة الثانية الي الخطوة الخامسة الي يتم الحصول الي يتم الحصول على التقارب المطلوب والذي يمثل (اقل معدل مربع الخطأ MSE)



وكذلك Error Histogram حث يحسب كما يلي

وبعد عدة محاولات تم الوقوف على افضل النتائج للتدريب ويتم التنبؤ بالسنتين 2013 و 2014
حيث يظهر الشكل التالي نتائج التدريب



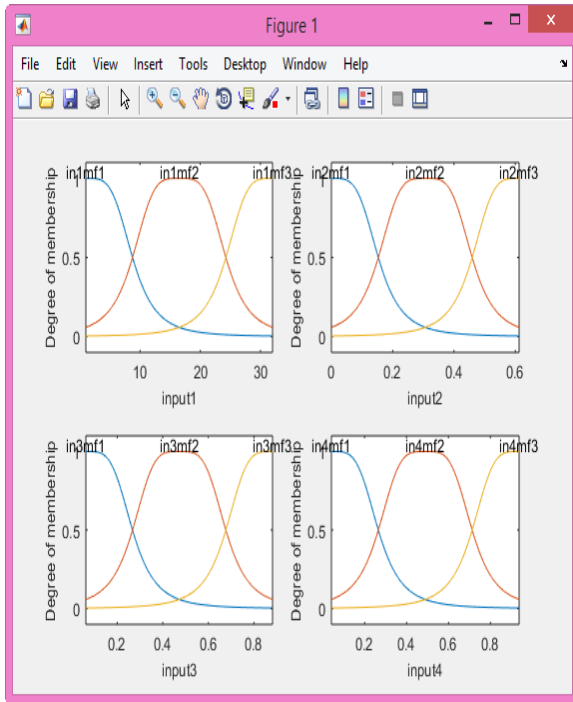
وكان نتيجة التنبؤ كما يلي $h = 27.3729$

التصنيف بواسطة نظام ANFIS

بناء نموذج الاستدلال الضبابي

الخطوة الاولى

تحديد المتطلبات الفعالة وذلك بتحديد الاهمية النسبية لكل متطلب من خلال قياس تاثير عناصر مقارنة بالمجموع وتحديد دالة العضوية على شكل جرس ويتم ادخال البيانات كنسبة مئوية حيث يتم جمع الثلاث مواسم معا والقسم على المجموع الكلي كما انها تتم على مرحلتين



المرحلة الاولى¹

ايجاد مطابقة خط انحدار بواسطة استخدام المعلومات المتاحة حول نقاط المركز للملاحظات

اي ان البيانات المدخلات تكون غير مضطربة وبعد ايجاد من هذه المرحلة اذ تعد كمتجه (α^*) يستخدم كمجموعة بيانات لتنفيذ وتستخدم طريقة المربعات الصغرى التقليدية لتقدير نموذج الانحدار

$$\alpha^* = (x'x)^{-1} x'y$$

من الشكل السابق يتضح ان الشبكة تتكون من ثلاث مداخل وخرج وحيد وقد استخدم توابع عضوية لكل دخل وهذه التوابع من النوع الجرسى²

المرحلة الثانية³

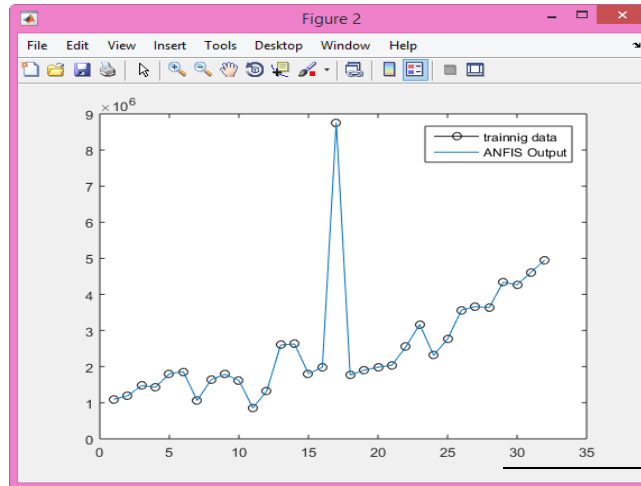
يتم تنفيذ مسألة (LP) اي استخدام طريقة معيار اقل مجموع كلي او الغموض ويعرف كمجموع

$$MinS = c_1 + c_2 + \dots + c_n$$

مع متباينات الشرطية لحد الانتماء حيث قيمة الانتماء لكل مشاهدة $Y_i = (y_i, e_i)$ هي اكبر وتساوى من حد الانتماء المفترض اذ $(h \in [0,1])$ اي ان

$$Y_{(y_i)} \geq h \dots \dots \dots \forall i = 1, \dots, N$$

$$Y_i^h \subseteq Y_i^{*h} \dots \dots \dots \bar{h}_i$$

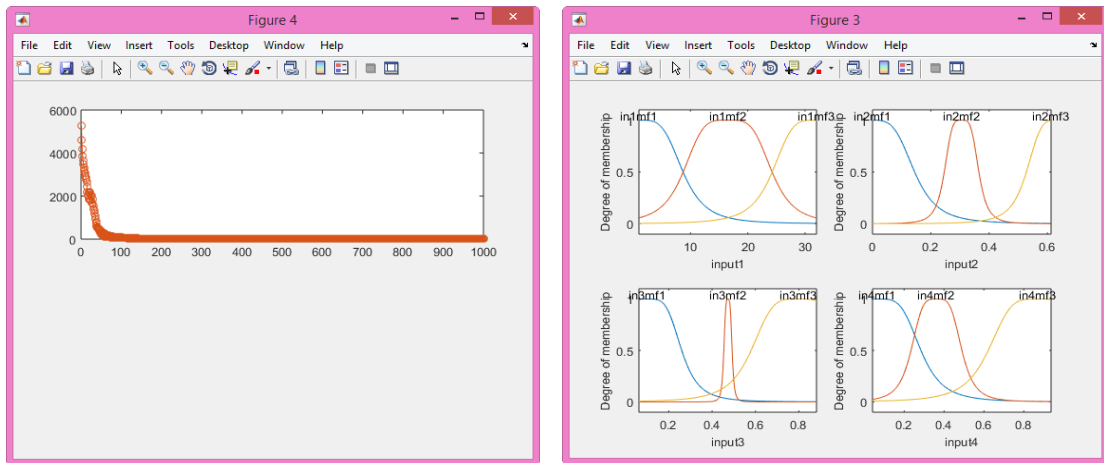


1 الياس حسن محمد " تحليل الانحدار المضطرب " المجلة العراقية للعلوم الاحصائية - كلية العلوم الحاسوب والرياضيات - جامعة الموصل العراق 2006

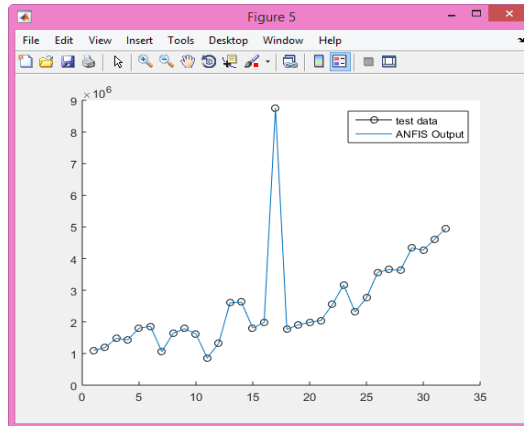
2 بلسم احمد عيد " مرجع سبق ذكره

3 الياس حسن محمد " مرجع سبق ذكره

ومن الشكل السابق نجد درجة تدريب البيانات traninning data تنطبق تمام على المخرجات output



ونجد من الشكل ان نتيجة التدريب جيدة وبالتالي نصل الي مرحلة الاختبار والفحص حيث يتم التنبؤ بعدد 6 قيم



يتضح انه اعلى نسبة

وانه من الشكل التالي

تدريب ويكاد يكون منطبق تماما حيث ان test data منطبقة تماما على مخرجات ANFIS

النتائج و التوصيات

وبمقارنه بين القيم الممتبأ بها من النماذج الثلاثة

القيم الاصلية	SARIMA(2,2,2)(2,3,2) ₃	BPNN	ANFIS
---------------	-----------------------------------	------	-------

2275942	791030	2270400	3167400
1745185	272457	1742000	2312800
589938	-966337	459900	2760500
3036233	954828	304120	3567100
1590937	270494	1593800	3659300
328275	-270494	4311000	3634200

نجد ان اهم الاستنتاجات التي تم التوصل اليها من خلال الدراسة هي

- 1- ان السلاسل الزمنية لانتاج محصول البطاطس هي سلسلة غير مستقرة وبها سنوات مفقودة وقيم شاذة.
- 2- اظهرت النتائج قدرة نموذج الشبكات العصبية الاصطناعية على التقدير والتنبؤ بانتاجية محصول البطاطس خلال العروات الثلاثة (النيلية والصيفية والشتوية) وذلك من خلال مقارنة النتائج الفعلية بالمقدرة (من الجدول السابق).
- 3- تبين افضلية الشبكات العصبية كانت من خلال امكانية ادخال تأثير الزمن كمدخل اضافي من مدخلات الشبكة الامر الذي تحسين كبير في اداء الشبكة الا ان ذلك غير ممكن في نماذج الاستدلال العصبي المتكيف لانها تعتمد علاقة فيزيائية بين المدخلات والمخرجات
- 4- بينت الدراسة ان الشبكات العصبية ذات التغذية الامامية والانتشار العكسي للخطأ وهي الافضل بين انواع النماذج العصبية من حيث السرعة والتدريب والدقة.
- 5- نجد ان الشبكات العصبية ذات الانتشار العكسي BPNN اذ تم الحصول على قيم تكاد تكون مطابقة للقيم للسنوات المقارنة لذا نجد انه افضل نموذج للتنبأ.
- 5- على الرغم من مميزات نموذج نظام الاستدلال العصبي المتكيف له القدرة على معالجة مختلف انواع البيانات الخطية وغير الخطية الي انه فشل في البيانات صغيرة الحجم.

التوصيات:

- 1- انها تحتاج الي خبره شخصيه من نوع خاص قد لا تتوفر في كثير من الباحثين.
- 2- كما تحتاج الي توافر بيانات على الاقل 60 وذلك لبناء نموذج جيد كما يمكن استخدام نموذج ARCH ,GARCH للحصول على نتائج افضل مع القيم الشاذة والمتطرفة.
- 3- يمكن استخدام اسلوب الدمج بين السلاسل الزمنية والشبكات العصبية وذلك من خلال اخذ النتائج المتحصل عليها في نموذج بوكس وجينكينز واستخدامها كمدخلات في الشبكات العصبية.
- 4- يمكن استخدام اسلوب الدمج بين السلاسل الزمنية نظام الاستدلال العصبي المتكيف وذلك من خلال اخذ النتائج المتحصل عليها في نموذج بوكس وجينكينز واستخدامها كمدخلات في نظام الاستدلال العصبي المتكيف .
- 5- يمكن استخدام اسلوب الدمج بين الشبكات العصبية نظام الاستدلال العصبي المتكيف وذلك من خلال اخذ النتائج المتحصل عليها في نموذج الشبكات العصبية واستخدامها كمدخلات في نظام الاستدلال العصبي المتكيف .

المراجع

أولاً: المراجع العربية:

- 1- انتصار ابراهيم الياس " بناء نموذج موسمي لتحليل السلسلة الزمنية لمعدلات الطاقة الكهربائية المجهزة لمدينة دهوك والتنبؤ بها " كلية الادارة والاقتصاد - جامعة نوروز - المجلد 6 العدد1(2017)
- 2- المصري , منى محمود ابراهيم " دراسة اقتصادية لأثر التدخل الحكومي على السياسة السعرية لمحصول البطاطس ومستقبل صادراته في ضوء العولمة والتنافسية " مجلة مصر المعاصرة - مصر مج 93 و468, 467 لسنة 2002
- 3- بلسم احمد عيد " دراسة مقارنة بين أداء الشبكات العصبية ونظام الاستدلال العصبي الضبابي المتكيف في تشخيص سرطان الثدي بالاعتماد علي السمات البنوية " مجلة جامعة تشرين للبحوث والدراسات العليا - سلسلة العلوم الهندسية المجلد (39) العدد(3) 2017.
- 4- جعفر محسن الخير واخرون " تصميم متحكم عصبوني ضبابي متكيف موسع لنظام التعليق النشط في نموذج نصف المركبة " مجلة جامعة البعث _ المجلد (38) العدد (55) عام 2016
- 5- خلود جمال عوض الله " استخدام نماذج SARIMA و Holt-Winters في التنبؤ بالسلاسل الزمنية الموسمية " كلية الاقتصاد والعلوم الادارية - جامعة الازهر - غزة عام 2016
- 6- رابح بلعباس " فعالية التنبؤ باستخدام النماذج الإحصائية في اتخاذ القرارات "
- 7- زرمان كريم " نحو تقدير امثل لمعدلات الخسارة في شركات التأمين " - مجلة ابحاث اقتصادية وادارية العدد السادس عشر ديسمبر 2014 جامعة خنشلة - الجزائر
- 8- سمير مصطفى شعراوى " مقدمة في التحليل الحديث للسلاسل الزمنية " مركز النشر العلمى - جامعة الملك عبد العزيز جدة 21589

- 9- ظافر رمضان مطر وآخرون " تحليل ونمذجة السلسلة الزمنية لتدفق المياه الداخلة الي مدينة الموصل دراسة مقارنة " - المجلة العراقية للعلوم الاحصائية (18) 2010
- 10- عاشور - بدار " المفاضلة بين نموذج السلاسل الزمنية ونموذج الانحدار البسيط في التنبؤ بحجم المبيعات في المؤسسة الاقتصادية دراسة حاله (مطاحن الحزن بالمسيلة) - كلية العلوم الاقتصادية وعلوم التسيير والعلوم التجارية - جامعة محمد بوضياف بالمسيلة عام 2005-2006
- 11- عبد النور , عادل "اساسيات الذكاء الاصطناعي " ط1 , دار الفيصل الثقافية , الرياض 2005
- 12- عروه نادر " استخدام الشبكات العصبونية للتنبؤ في الأسواق المالية (تطبيق لتوقع تغيرات قيم الأسهم في سوق دمشق للأوراق المالية) – ماجستير نظم المعلومات ودعم القرار - المعهد العالي للعلوم التطبيقية والتكنولوجيا- الجمهورية العربية السورية
- 13- عمر صابر قاسم , اسراء رست محمد " دراسة رياضية تحليلية لخوارزميات الشبكات العصبية الاصطناعية في ملائمة نموذج للتشخيص الطبي " كلية علوم الحاسوب والرياضيات - جامعة الموصل
- 14- عيسي , علام زكي " الشبكات العصبية البنية الهندسية , الخوارزميات , التطبيقات " شعاع للنشر والعلوم (2000)
- 15- فريد بختي " السلاسل الزمنية الخطية باستعمال حزمة Eviews7.0 " الجمهورية الجزائرية الديمقراطية الشعبية – كلية العلوم الاقتصادية ، والتجارية وعلوم التسيير – السنة الجامعية 2014/2015
- 16- قصي حبيب الحسيني و احمد عبد الامير الساعدي " مقدمة في الشبكات العصبية الاصطناعية " جامعة الامير جعفر الصادق عليه السلام
- 17- لوقي فاتح "جودة نماذج السلاسل الزمنية الموسمية المختلطة SARIMA في التنبؤ بالمبيعات دراسة حالة مؤسسة مطاحن جديع بتقوت" ، كلية العلوم الاقتصادية والتجارة وعلوم التسيير ، جامعة محمد خضير - بسكرة الجمهورية الجزائرية الديمقراطية الشعبية للعام 2013/2014 ص 61
- 18- ماهر النديوي وهيام خدام " تطوير حل مسألة الحركة المجردة العكسية المعتمد علي تقنية نظام الاستدلال العائم المتكيف عصبونياً باستخدام زاوية الاتجاه * " مجلة جامعة البعث – المجلد 38 – العدد (14)
- 19- مناهل عبد الكريم يوسف " دراسة مقارنة بين الشبكات العصبية الاصطناعية (تمييز الارقام الانكليزية المطبوعة) – مجلة التربية والعلوم – المجلد (23) , العدد(2) لسنة 2010

