

أستخدام أنموذجي ARX و ARMAX للتنبؤ بالسلاسل الزمنية مع تطبيق عملي - دراسة مقارنة

ا.م.د احلام احمد جمعة

dr_ahlam_ahmed@yahoo.com

جامعة بغداد - كلية الادارة والاقتصاد - قسم الاحصاء

انمار رسول شعيث

rasool.anmar@yahoo.com

جامعة بغداد - كلية الادارة والاقتصاد - قسم الاحصاء

المستخلص

لقد اهتم الباحثون بدراسة نماذج الصندوق الاسود Black Box Models والتي تقوم بربط سلسلة المدخلات مع سلسلة المخرجات بأنموذج رياضي وفيها نوعان من النماذج هما : نماذج خطأ المعادلة وتضم أنموذج الانحدار الذاتي مع مدخلات خارجية ARX وأنموذج الانحدار الذاتي والمتوسطات المتحركة مع مدخلات خارجية ARMAX ، ونماذج خطأ الاخراج وتضم أنموذج خطأ الاخراج OE وأنموذج بوكس جنكيز BJ . وقد اخص هذا البحث فقط نماذج خطأ المعادلة لذا يهدف البحث الى التنبؤ بالسلاسل الزمنية وتحديد رتبة الأنموذج باستخدام بعض المعايير الاحصائية والهندسية ومنها معيار اكاكي للمعلومات AIC ومعيار اكاكي لخطأ التنبؤ النهائي FPE . وقد شمل هذا البحث خطوات متسلسلة والمتمثلة بتهيئة البيانات واستقراريه سلسلتي الادخال والاخراج وتحديد رتبة الأنموذج باستخدام بعض المعايير الاحصائية والهندسية وتقدير معلماته. وقد تم اختبار دقة الأنموذج وذلك باختبار معنوية الارتباط لبواقي الأنموذج المحدد بالإضافة الى اختبار معنوية الارتباط للبواقي $e(t)$ والمدخل $u(t)$. مع مقارنة التنبؤ لنماذج الصندوق الاسود المختارة باستخدام متوسط مطلق الخطأ MAE وجذر متوسط مربع الخطأ RMSE وقد تم التطبيق على بيانات حقيقية يومية وهي تصفية مياه الشرب باستخدام مقياس عكورة الماء لمشروع الكرخ في مدينة بغداد للفترة 2015/1/1 لغاية 2015/3/31.

الكلمات الرئيسية: أنموذج الانحدار الذاتي مع مدخلات خارجية، أنموذج الانحدار الذاتي والمتوسطات المتحركة مع مدخلات خارجية، أنموذج خطأ الاخراج، أنموذج بوكس - جنكيز، معيار أكاكي للمعلومات، معيار أكاكي لخطأ التنبؤ النهائي، مجموع الخطأ المطلق، جذر متوسط مربع الخطأ.

المقدمة [1] Introduction

تعتبر السلاسل الزمنية من المواضيع الاساسية التي اصبح استخدامها في مختلف العلوم، وان تحليل السلاسل الزمنية من الاساليب الاحصائية المهمة والتي تستخدم في التنبؤ بقيم الظواهر العشوائية في المستقبل. حيث ان منهجية السلاسل الزمنية تعتمد على نمذجة سلوك ظاهرة معينة خلال فترة زمنية معينة لمختلف الظواهر الاقتصادية، الاجتماعية، الفيزيائية ... الخ [3].

وتعتمد نماذج الصندوق الأسود على الانظمة الديناميكية لبيانات الادخال والاخراج حيث تقوم بربط سلسلة المدخلات مع سلسلة المخرجات بأنموذج رياضي، كما ان معلمات أنموذج الصندوق الأسود قد لا يكون لها تفسير فيزيائي. وقد تكون نماذج الصندوق الأسود خطيه وغير خطيه، حيث ان نماذجها الخطية يمكن أن تكون بشكل فترات زمنية مستمرة أو متقطعة، وتتميز نماذج الصندوق الأسود بإمكانية بناء العديد من هذه النماذج ومقارنتها لاختيار أفضل أنموذج [16].

وان تصميم الأنموذج يعتمد على البيانات المتاحة والمعلومات الأخرى في العملية، ولهذا تعتمد القدرة التمثيلية للأنموذج على قدرة تمثيل البيانات ودقه المعلومات المتاحة. والطريقة التي تستخدم فيها المعلومات للنمذجة، بمعنى يحتاج الأنموذج لعدد من الاساليب حول كيفية ارتباط المتغيرات مع بعضها البعض وهذه العلاقة التي تكون على شكل إشارات للمدخلات والمخرجات تسمى بأنموذج النظام (Model Of The System). وفي بعض الأحيان تتأثر المخرجات بإشارات أخرى فضلا عن قياس المدخلات يطلق عليها إشارات التشويش وبافتراض إن الإشارات مرتبطة مع نظام خطي (Linear System) [22].

هدف البحث Objective of the Research

يهدف البحث الى التنبؤ بالسلاسل الزمنية لكل من الأنموذجين ARX و ARMAX بعد تحديد رتبة كل أنموذج باستخدام بعض المعايير الاحصائية والهندسية ومنها معيار اكاكي للمعلومات AIC ومعيار اكاكي لخطأ التنبؤ النهائي FPE بالاضافة الى اختبار دقة الأنموذج المحدد.

الاستعراض المرجعي Review of Literature

سيتم التطرق لبعض الدراسات التي تشمل موضوع البحث حيث اهتم الباحثون بدراسة النماذج التالية (ARX, ARMAX, OE, BJ) لهذا المجال اكثر من عقدين من الزمن وازداد في السنوات الاخيرة، وفيما يلي استعراض لهذه النماذج.

- في عام 1997م درس الباحث (Ljung) [12] حول تحديد مشكلة بناء نظام النماذج، وقد تحدث عن نماذج الصندوق الاسود والصندوق الرمادي، وتطرق الى نماذج الصندوق الاسود الخطية وهي (ARX,ARMAX)، وكذلك النماذج غير الخطية للصندوق الاسود بالإضافة الى الشبكات العصبية (ANN) .
- وفي عام 2002 م تناول الباحث (Zheng)[23] تحديد مشكله المربعات الصغرى الاعتيادية (OLS) للأنموذج (ARMAX) حيث يمكن تقدير التحيز الناجم عن التشويش من خلال الاستفاده من بناء الأنموذج (ARMAX)، ولذلك فقد أوجدت طريقة المربعات الصغرى لأزاله التحيز(BELS) ليعطي أنموذج (ARMAX) غير متحيز. ولتحقيق دقة تقدير عالية بتكلفه عديده معتدلة تم استخدام الطريقة المقترحة (BELSX) وهي مفضله بسبب مدى دقتها القريبة جدا من طريقة خطأ التنبؤ (PE) الى (ARMAX).
- في عام 2003 م قام الباحثون (Frausto, Pieters & Deltour)[10] بدراسة تحقيق الى اي مدى يمكن استخدام كل من الأنموذجين (ARX,ARMAX) لوصف درجة الحرارة القياسية داخل الهواء المسببة للاحتباس الحراري المهواة بشكل طبيعي في ظروف اوربا الغربية ، واستخدمت متغيرات الإدخال كالهواء الخارجي ودرجة الحرارة ونسبة الرطوبة والشعاع الشمسي الشامل والغيوم من السماء، واتضح بان تلك النماذج ملائمة لوصف تطور درجات الحرارة والاحتباس الحراري بصوره مرضيه فيما عدا فترات التهوية، ولوحظ ان أنموذج (ARX) له اداء افضل من أنموذج (ARMAX) .
- في عام 2010 م درس الباحثون (Huusom, Poulsen, Jorgensen & Jorgensen)[11] أنموذج ARX والمقارنة مع أنموذج ARMAX و أنموذج فضاء الحالة لتأمين السيطرة على تطبيقات تنبؤ الأنموذج . وان فئة الأنموذج ترتبط مع مشكلة تقدير المعلمة لكلا النظامين (MIMO,SISO)، وقد تم تقدير معاملات الأنموذج ARMAX او فضاء الحالة من خلال اساليب التنبؤ. وبينت الدراسة بان أنموذج ARX له بساطة في نظام الديناميكية مقارنة مع نظام متعدد المتغيرات لأنموذج ARMAX .
- في عام 2011 م درس الباحثون (Diversi, Guidorzi & Soverin)[9] أنموذج ARMAX التي تستخدم في نطاق واسع وهي اداة قياسية في هندسة التحكم لوصف النظام ومراقبة التصميم لهذه النماذج ، وقد تكون غير واقعية في كثير من السياقات العلمية بسبب وجود اخطاء القياس التي تلعب دوراً هاماً في التطبيقات مثل تشخيص الاخطاء والترشيح الامثل.

2. الجانب النظري

2.1 نماذج الصندوق الاسود Black Box Models

يعد تطور النماذج الرياضية خطوة مهمة في عملية تصميم النظام، وان الاهتمام بنماذج دالة التحويل يعني الاهتمام بنماذج الصندوق الاسود، فيتم التعرف على هذه النماذج باستخدام مجموعة بيانات تم الحصول عليها من خلال اختبار الديناميكية . ولكن بنية الأنموذج المختار تنتمي لعائلة معرفة بان لديها مرونة جيدة اي انها لا تنتج المفهوم الفيزيائي [14][21]، والأنموذج الخطي العام يمثل الهيكلية الموحدة للأنظمة الخطية ، حيث يتم حساب مخرجات النظام الخطي المحدد $Y(t)$ عند الزمن t من خلال ترشيح $u(t)$ والتي تمثل مدخلات الأنموذج بمرشح خطي يرمز له $G(z)$ ويسمى بالجزء المحدد، وان المرشح الخطي للتشويش الابيض White Noise الذي يرمز له $e(t)$ هو $H(z)$ ويسمى بالجزء التصادفي وكالاتي [22] .

$$Y(t) = G(z)u(t) + H(z)e(t) \quad \dots\dots\dots (1)$$

حيث ان

$$G(z) = \frac{B(z)}{F(z) A(z)}$$

$$H(z) = \frac{C(z)}{D(z) A(z)}$$

ويمكن كتابة المعادلة رقم (1) بالشكل الاتي .

$$Y(t) = \frac{B(z)}{F(z)A(z)} u(t) + \frac{C(z)}{D(z)A(z)} e(t) \quad \dots\dots\dots (2)$$

حيث ان

$$A(z) = 1 + a_1z^{-1} + a_2z^{-2} + \dots + a_{na}z^{-na}$$

$$B(z) = b_1z^{-1} + b_2z^{-2} + \dots + b_{nb}z^{-nb}$$

$$C(z) = 1 + c_1z^{-1} + c_2z^{-2} + \dots + c_{nc}z^{-nc}$$

$$D(z) = 1 + d_1z^{-1} + d_2z^{-2} + \dots + d_{nd}z^{-nd}$$

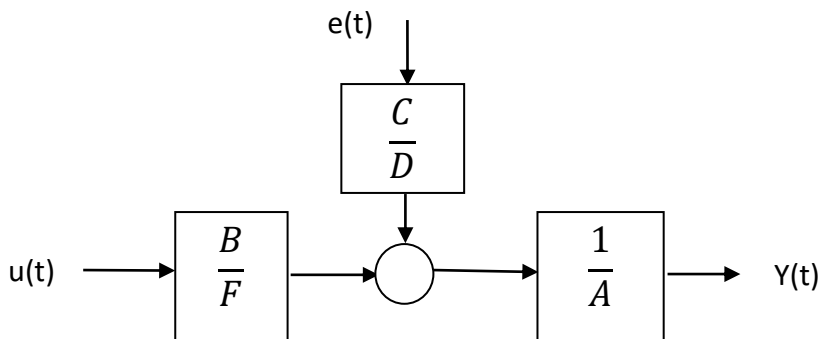
$$F(z) = 1 + f_1z^{-1} + f_2z^{-2} + \dots + f_{nf}z^{-nf}$$

وإن:

nf, nd, nc, nb, na : تمثل معاملات متعددة الحدود على التوالي

z^{-1} : تمثل مشغل الازاحة الخلفي Backward Shift [14]

ويمكن وصف هيكل الأنموذج الخطي العام من خلال الشكل أدناه [22]



الشكل (1) : يمثل هيكلية الأنموذج الخطي العام

وان التنبؤ لنماذج الصندوق الاسود هو [8]

$$\hat{Y}(t/t-1) = H_0^{-1}(z)G_0(z)u(t) + [1 - H_0^{-1}(z)]Y(t) \quad \dots\dots\dots (3)$$

وان خطأ التنبؤ لهذه النماذج هو

$$e(t) = Y(t) - \hat{Y}(t/t-1) \quad \dots\dots\dots (4)$$

وتقسم نماذج الصندوق الاسود الى قسمين هما:

اولاً: نماذج خطأ المعادلة The Equation Error models

ثانياً: نماذج خطأ المخرجات The Output Error Models

حيث اختص هذا البحث فقط على نماذج خطأ المعادلة وفيما يلي شرح لهذه النماذج

2.2 نماذج خطأ المعادلة The Equation Error Models

تعتبر نماذج خطأ المعادلة The Equation Error Models النوع الاول من نماذج الصندوق الاسود Black Box Models حيث تشمل هذه النماذج كل من أنموذج الانحدار الذاتي مع مدخلات خارجية Autoregressive With Exogenous Input Model ويرمز له ARX، وأنموذج الانحدار الذاتي والمتوسطات المتحركة مع مدخلات خارجية Autoregressive Moving Average With Exogenous Input Model ويرمز له ARMAX . وتتميز هذه النماذج عن نماذج خطأ الاخراج بانها تحتوي على المرشح الخطي الشائع ((1/ A(z)) في أنموذج العمليات المحددة وأنموذج العمليات التصادفية ، اي ان كلاً من أنموذج ARX وأنموذج ARMAX تشترك بالمرشح الخطي A(z)، ويعتبر أنموذج ARX حاله خاصة من أنموذج ARMAX عندما [22]C(z)=1، وفيما يلي شرح نماذج خطأ المعادلة :

2.2.1 أنموذج الانحدار الذاتي مع مدخلات خارجية

Autoregressive With Exogenous Input Model

يعتبر أنموذج الانحدار الذاتي مع مدخلات خارجية Autoregressive With Exogenous Input Model والذي يرمز له ARX من النماذج التي لها تطبيقات واسعة وهو ابسط أنموذج لأنه يصف مسار الادخال والايخراج، وذلك لسهولة تطبيقه في حساب تقدير المعلمات لذلك فانه يعطي اعلى تقدير لأنموذج ونتيجة فعالة اي يكون ملائماً لعدد كبير من العمليات الحقيقية. وان هذا الانموذج يسمى أيضا بأنموذج الانحدار المسيطر، وتشير قيمة X في الأنموذج ARX الى منهجية نماذج المتغير خارجي المنشأ(Exogenous Variable) [19] ، حيث يمكن توضيحه من خلال المعادلة الاتية :

$$Y(t) = \frac{B(z)}{A(z)}u(t) + \frac{1}{A(z)}e(t) \quad \dots\dots\dots (5)$$

او يمكن كتابته بالشكل الاتي

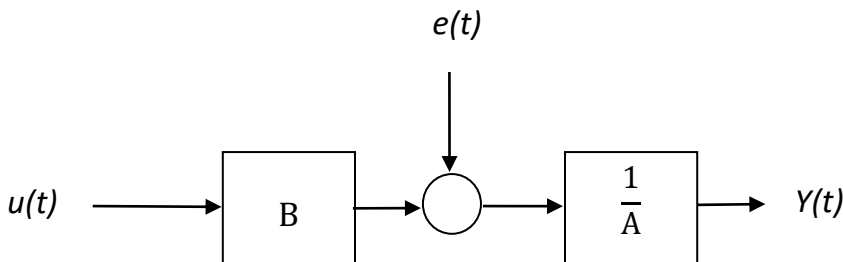
$$A(z)Y(t) = B(z)u(t) + e(t) \quad \dots\dots\dots (6)$$

حيث ان متعددات الحدود تعطى بالشكل الاتي [22]

$$A(z) = 1 + a_1z^{-1} + a_2z^{-2} + \dots + a_naz^{-na}$$

$$B(z) = b_1z^{-1} + b_2z^{-2} + \dots + b_nbz^{-nb}$$

حيث يمكن تمثيل هذا الانموذج بالشكل الاتي [14]



الشكل (2): يمثل هيكلية أنموذج الانحدار الذاتي مع مدخلات خارجية ARX

وعليه فان التنبؤ باستخدام الأنموذج ARX يكون بالشكل الاتي [22].

$$\hat{Y}(t/t - 1) = \frac{B(z)/A(z)}{1/A(z)} u(t) + [1 - A(z)]Y(t) \quad \dots\dots\dots (7)$$

ويمكن كتابة المعادلة رقم (7) بالشكل الاتي

$$\hat{Y}(t/t - 1) = B(z) u(t) + (1 - A(z))Y(t) \quad \dots\dots\dots (8)$$

وان خطأ التنبؤ لأنموذج ARX كالآتي

$$e(t) = Y(t) - \hat{Y}(t/t - 1) = Y(t) - [B(z)u(t) + (1 - A(z))Y(t)]$$

$$e(t) = A(z)Y(t) - B(z)u(t) \quad \dots\dots\dots (9)$$

2.2.2 أنموذج الانحدار الذاتي والمتوسطات المتحركة مع مدخلات خارجية Autoregressive Moving Average With Exogenous Input Model

يعتبر أنموذج الانحدار الذاتي والمتوسطات المتحركة مع مدخلات خارجية Autoregressive Moving Average With Exogenous Input Model والذي يرمز له ARMAX من نماذج خطأ المعادلة، ويعد أنموذجاً موسعاً لأنموذج ARX لاحتوائه على المرشح الخطي C(z) ، فعندما تكون C(z)=1 اي بمعنى ان nc=0 فسيتحول أنموذج ARMAX الى أنموذج ARX ، وفضلا عن ذلك فإن A(z)

يمثل المقام لمتعددة الحدود الخاص بمركبة التشويش $e(t)$ ويمكن تمثيل الأنموذج كالاتي [22].

$$Y(t) = \frac{B(z)}{A(z)} u(t) + \frac{C(z)}{A(z)} e(t) \quad \dots\dots\dots (10)$$

ويمكن كتابة الانموذج بالشكل الاتي

$$A(z)Y(t) = B(z)u(t) + C(z)e(t)$$

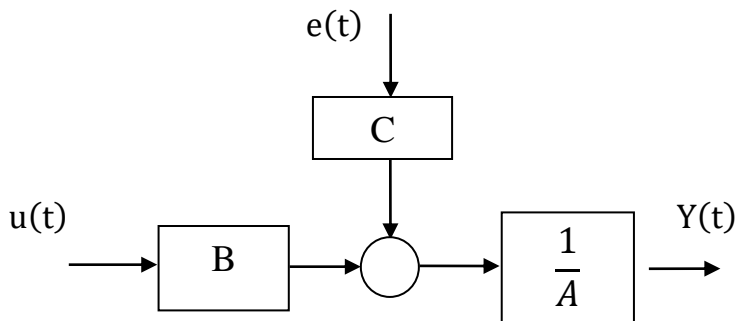
وان متعددات الحدود تكتب بالشكل الاتي

$$A(z) = 1 + a_1z^{-1} + a_2z^{-2} + \dots + a_{na}z^{-na}$$

$$B(z) = b_1z^{-1} + b_2z^{-2} + \dots + b_{nb}z^{-nb}$$

$$C(z) = 1 + c_1z^{-1} + c_2z^{-2} + \dots + c_{nc}z^{-nc}$$

والشكل التالي يوضح بنية الانموذج [14].



الشكل (4): يمثل هيكلية أنموذج الانحدار الذاتي والمتوسطات المتحركة مع مدخلات

خارجية ARMAX

اما التنبؤ باستخدام أنموذج ARMAX يمكن كتابته بالشكل الاتي

$$\hat{Y}(t/t - 1) = \frac{B(z)}{C(z)} u(t) + \left(1 - \frac{A(z)}{C(z)}\right) Y(t) \quad \dots\dots\dots (11)$$

وان خطأ التنبؤ لأنموذج ARMAX هو

$$e(t) = \frac{A(z)}{C(z)} Y(t) - \frac{B(z)}{C(z)} u(t) \quad \dots\dots\dots (12)$$

2.3 تهيئة البيانات Data Preparing

تعتبر تهيئة البيانات هي الخطوة الاولى لمعرفة هل ان سلسلتي الادخال $u(t)$ والايخراج $Y(t)$ مستقرة ام لا في الوسط والتباين ، ولمعرفة الاستقرارية يتم رسم السلسلتين، ويتم الانتقال الى الخطوة الثانية اذا كانت السلسلة مستقرة، أما اذا كانت السلسلة غير مستقرة فيتم استخدام طرق التحويل الخاصة باستقراره السلسلة الزمنية [4].

2.4 تحديد رتبة الانموذج Model Rank of Identification

تعتمد عملية تحديد رتبة الأنموذج على عدد من المعايير التي تستخدم في تحديد واختيار رتبة الأنموذج وسنطرق الى بعض المعايير الاحصائية والهندسية وهي

2.4.1 دالة الكلفة Cost Function

تعرف دالة الكلفة Cost Function أيضا بدالة الخسارة (Loss Function) أو دالة المخاطرة (Risk Function) ، ويتميز سلوك تلك الدالة بان قيمتها تقل بزيادة رتبة الأنموذج، أي أنها تتناقص كلما زادت رتبة الأنموذج، ويتوقف التناقص لهذه الدالة عند نقطة معينة وهذا يعني أن زيادة رتبة الأنموذج تكون عديمة الفائدة وهذه من مزايا دالة الكلفة . وتعتبر الأساس في اختيار رتبة [7][17]

$$V = \frac{1}{2} \sum_{t=1}^n e^2(t) \quad \dots\dots\dots (13)$$

حيث ان :

V : تمثل دالة الكلفة

n : تمثل عدد المشاهدات

$e(t)$: تمثل مقدار الخطأ ويساوي

$$e(t) = Y(t) - \hat{Y}(t)$$

2.4.2 معيار أكاي للمعلومات Akaike's Information Criteria

عرف من قبل العالم Akaike عام (1973-1974) والذي يرمز له اختصار AIC وهو يساوي ضعف عدد المعلمات مطروحاً منه دالة الامكان الاعظم للمشاهدات.

والأنموذج الذي يعطي اقل معلمات هو الأنموذج الا فضل، وتستخدم معظم البرمجيات الجاهزة مثل Matlab دالة الكلفة في حساب قيمة هذه المعيار، حيث يعطي هذا المعيار لرتبة الانموذج الصحيحة تقدير غير متسق ويميل تقديره لان يكون فوق الرتبة الحقيقية [17][24]، ويعبر عنه بالشكل الاتي [5].

$$AIC = \ln V + 2 \frac{m}{n} \dots\dots\dots (14)$$

حيث ان

m: تمثل عدد المعلمات في الأنموذج

V: دالة الكلفة والموضحة في المعادلة رقم (13)

n: تمثل عدد المشاهدات

2.4.3 معيار أكاي لخطأ التنبؤ النهائي

Akaik's Final Prediction Error Criteria

عرف من قبل العالم Akaike عام 1969 ويرمز له اختصارا FPE وانه تباين خطأ التنبؤ النهائي. يعتبر معيار أكاي للمعلومات واحد من معايير تحديد رتبة الأنموذج والتي تقابل اقل قيمة من هذا المعيار وهي الرتبة الاكثر ملائمة للمشاهدات، وان لهذا المعيار مميزات تكون مشابهة لمميزات معيار AIC من حيث زيادة عدد المعلمات وصولا الى اقل تباين وعليه يمكن حساب معيار أكاي لخطأ التنبؤ النهائي من خلال المعادلة الرياضية الاتية [13][15].

$$FPE = \left\{ \frac{1 + \frac{m}{n}}{1 - \frac{m}{n}} \right\} V \dots\dots\dots (15)$$

حيث ان

V: تمثل دالة الكلفة

n: تمثل عدد المشاهدات

m: تمثل عدد المعلمات في الأنموذج

2.5 تقدير معاملات الأنموذج

ولتقدير معاملات الأنموذج هناك عدة طرائق وقد اختص هذا البحث فقط بطريقة المربعات الصغرى الاعتيادية (OLS) وفيما يلي شرح لهذه الطريقة.

2.5.1 طريقة المربعات الصغرى الاعتيادية

ان طريقة المربعات الصغرى الاعتيادية والذي يرمز لها OLS ، هي طريقة لإيجاد تقديرات لمعاملات الأنموذج ويمكن كتابة مقدر المربعات الصغرى كما يأتي [18].

$$\hat{\theta}_n = [\sum_{t=1}^n \varphi(t) \varphi'(t)]^{-1} \sum_{t=1}^n \varphi(t)Y(t) \quad \dots\dots\dots (16)$$

حيث ان

$$\varphi'(t) = [u(t-1), \dots, u(t-m)]$$

وان

m: تمثل معاملات الانموذج

2.6 اختبار دقة الأنموذج Model Accuracy Test

توجد عدد من الاختبارات الخاصة بدقة الانموذج وهذه الاختبارات تعتمد على التشويش الابيض White Noise وذلك من خلال اختبار معنوية $R^n(\tau)$ والتي تمثل الارتباطات [25].

وتقارن إحصاءه الاختبار مع توزيع مربع كاي الجدولية، فاذا كانت الإحصاءة أصغر من قيمة مربع كاي بدرجة حرية d ومستوى معنوية α حيث ترفض الفرضية البديلة اي ان الانموذج المشخص ملائم للبيانات ومن هذه الاختبارات هي:

2.6.1 اختبار الاستقلالية

يعتبر اختبار الاستقلالية واحد من اهم الاختبارات الاساسية ، ولحساب الارتباط \hat{R}_{ue}^n في حالة المدخلات u وحد الخطأ e عند الارتداد الزمني τ حيث يتم اختبار الفرضية الاتية .

$$\left. \begin{aligned} H_0 : \hat{R}_{ue}^n(\tau) &= 0 \\ H_1 : \hat{R}_{ue}^n(\tau) &\neq 0 \end{aligned} \right\} \dots\dots\dots (17)$$

وان

$$\widehat{R}_{ue}^n(\tau) = \frac{1}{n} \sum_{t=1}^n u(t - \tau) e(t) \dots\dots\dots (18)$$

وبجعل

$$\varphi(t) = [u(t - 1) \dots u(t - m)]'$$

$$R_n = \frac{1}{n} \sum_{t=1}^n \varphi(t)\varphi'(t)$$

وهكذا فان

$$\xi_n^m = \frac{1}{\delta^2_m} \|\sum_{t=1}^n \varphi(t) e(t)\| R_n^{-1} \dots\dots\dots (19)$$

في هذه الحالة سوف يكون لها توزيع مربع كاي بدرجة حرية d وتقارن إحصاء الاختبار مع توزيع مربع كاي الجدولية، فاذا كانت الإحصاء أصغر من قيمة مربع كاي بدرجة حرية d ومستوى معنوية α حيث ترفض الفرضية البديلة اي ان الانموذج المشخص ملائم للبيانات [25].

2.6.2 اختبار البياض

يعتمد اختبار البياض على الارتباطات الذاتية للأخطاء العشوائية اي ان

$$\widehat{R}_e^n(\tau) = \frac{1}{n} \sum_{t=1}^n e(t)e(t - \tau) \dots\dots\dots (20)$$

لاختبار الفرضية الاتية

$$\left. \begin{aligned} H_0 : \widehat{R}_e^n(\tau) &= 0 \\ H_1 : \widehat{R}_e^n(\tau) &\neq 0 \end{aligned} \right\} \dots\dots\dots (21)$$

وان $e(t)$ تمثل سلسلة الخطأ White Noise وان إحصاءه اختبار البياض $\eta_{n,m}$ تساوي

$$\eta_{n,m} = \frac{n}{(\widehat{R}_e^n(0))^2} \sum_{\tau=1}^m (\widehat{R}_e^n(\tau))^2 \dots\dots\dots (22)$$

والتي تكون مقارنة الى توزيع مربع كاي بدرجة حرية d ومستوى معنويه α ، وان استقلالية الاخطاء يمكن اختبارها من خلال مقارنة الاحصاء $\eta_{n,m}$ مع قيمة مربع كاي الجدولية، حيث يتم من خلالها القبول او الرفض [6].

❖ وتوجد معايير اخرى لاختبار الدقة وهي [20]

Mean Absolute Error متوسط مطلق الخطأ

$$MAE = \frac{1}{n} \sum_{t=1}^n |e(t)| \quad \dots\dots (23)$$

Root Mean Square Error جذر متوسط مربع الخطأ

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y(t) - \hat{Y}(t))^2} \quad \dots\dots (24)$$

حيث ان

$e(t)$: تمثل مقدار الخطأ وتساوي

$$e(t) = Y(t) - \hat{Y}(t)$$

$Y(t)$: تمثل القيمة الحقيقية للفترة الزمنية t

$\hat{Y}(t)$: تمثل القيمة التنبؤية للفترة الزمنية t

n: تمثل عدد المشاهدات

4. الجانب التطبيقي 4.1 المقدمة

Introduction

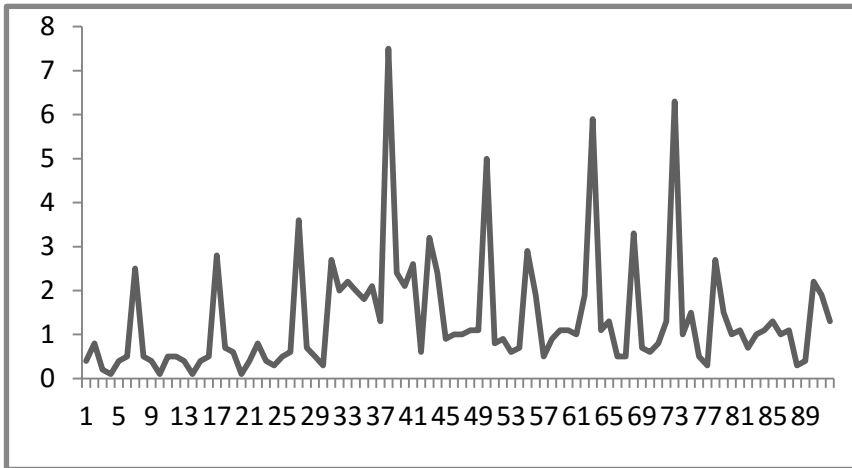
يعتبر الماء الركن الاساسي لتهيئة الظروف الملائمة للحياة اي ان الماء ضروري لبقاء الحياة ، حيث تحتوي الكرة الارضية على (75%) من الماء وان جسم الانسان يحتوي على (58% - 65%) من وزنه ماء. وقال تعالى بمحكم كتابه الكريم في سورة الانبياء (وَجَعَلْنَا مِنَ الْمَاءِ كُلَّ شَيْءٍ حَيٍّ أَفَلَا يُؤْمِنُونَ)، بحيث تتم تنقية الماء من خلال وجود محطات لمعالجة المياه والتي تكون ذات كفاءة عالية ولأن الطرائق التقليدية لا تقضي

على الملوثات الصناعية، فأن محطات التصفية تهدف الى جعل الماء الخام صالح للشرب ذو لون مقبول وخالي من الاحياء المجهرية وقيمة العكورة تكون ضمن الحد المطلوب [2]. وان الدراسة الحالية تهتم بدراسة مقياس عكورة الماء الخام والذي يحتوي على الشوائب والاجسام الصغيرة والذي يمثل مدخلات الانموذج ويرمز له $u(t)$ وهو المتغير الاول للسلسلة الزمنية، وتكون اعلى قيمه لعكورة الماء الخام هي (45) في يوم 2015/3/9 واقل قيمه لعكورة الماء الخام هي (8) في يوم 2015/1/1. اما المتغير الثاني للسلسلة الزمنية فيكون مقياس عكورة الماء الصالح للشرب والذي يكون خالي من الشوائب والاجسام الصغيرة حيث يمثل مخرجات الانموذج والذي يرمز له $Y(t)$ ، وتكون اعلى قيمة لعكورة الماء الصالح للشرب هي (7.5) في يوم 2015/2/7 واقل قيمه لعكورة الماء الصالح للشرب هي (0.1) في يوم 2015/1/4.

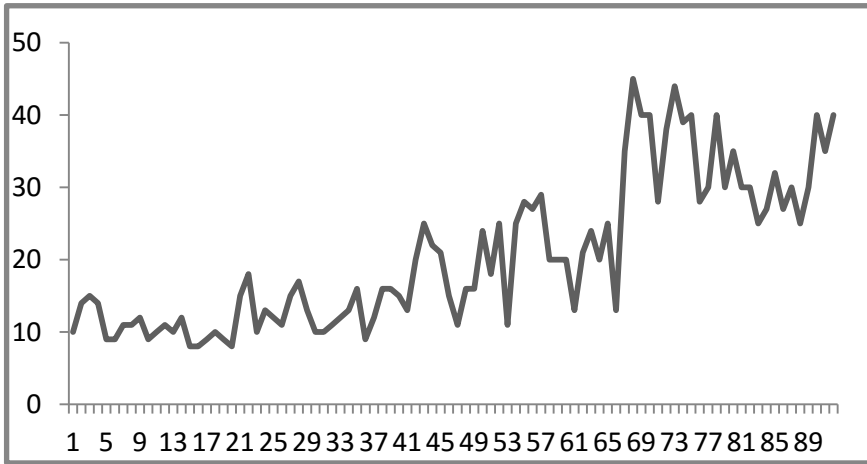
وتم الحصول على البيانات من اكبر مشاريع تصفية المياه في محافظة بغداد وهو مشروع الكرخ، حيث كانت البيانات يومية وتبدأ من 2015/1/1 ولغاية 2015/3/31 وتم تحليل البيانات من خلال كتابة برنامج بلغة MATLAB 2013

4.2 تهيئة وتحليل البيانات

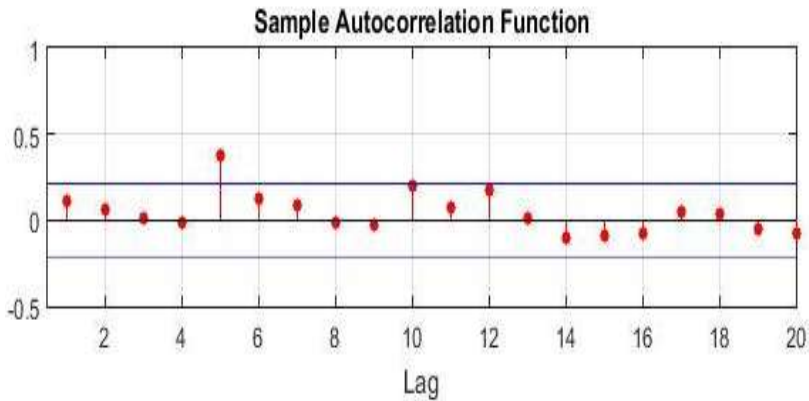
❖ رسم السلسلة الزمنية مع الاستقرارية



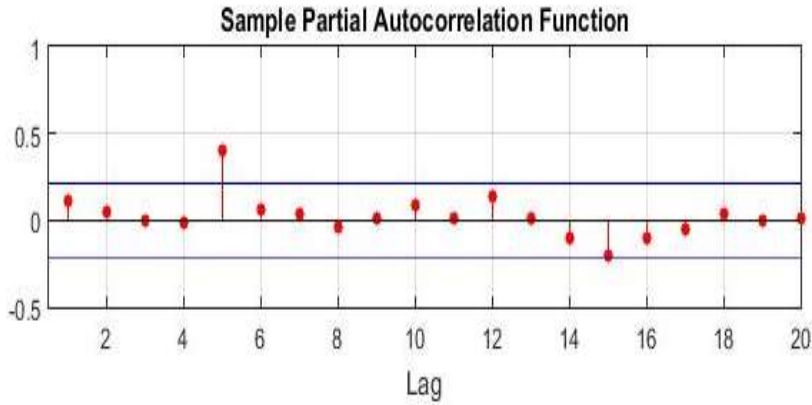
الشكل (1) : يمثل مقياس عكورة الماء الصالح للشرب لسلسلة الاخراج $Y(t)$



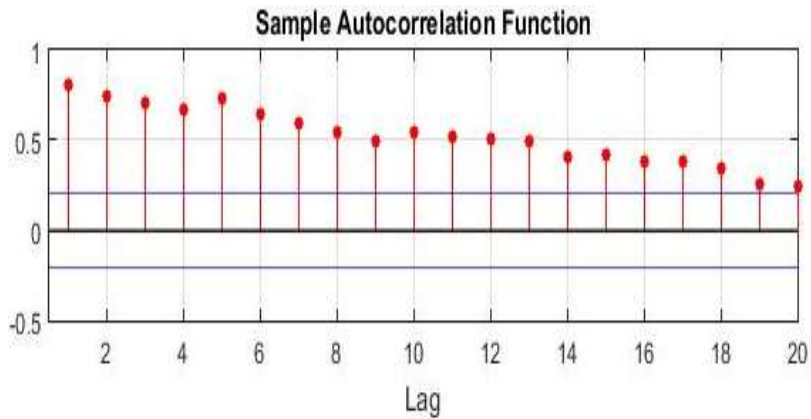
الشكل (2) : يمثل مقياس عكورة الماء الخام لسلسلة الادخال $u(t)$



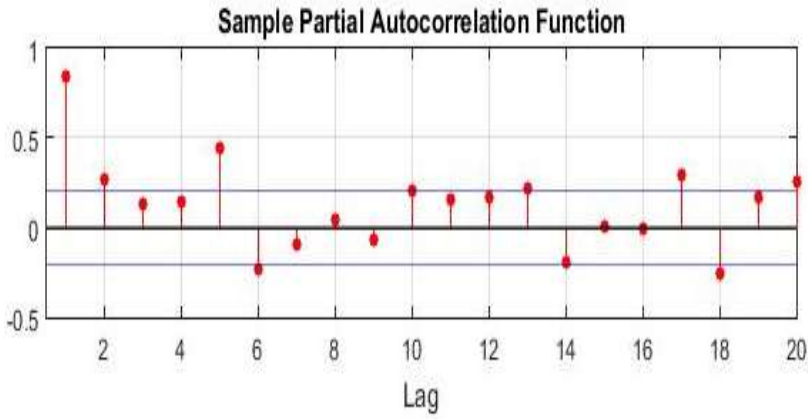
الشكل (3) : يمثل دالة الارتباط الذاتي لسلسلة الاخراج $Y(t)$



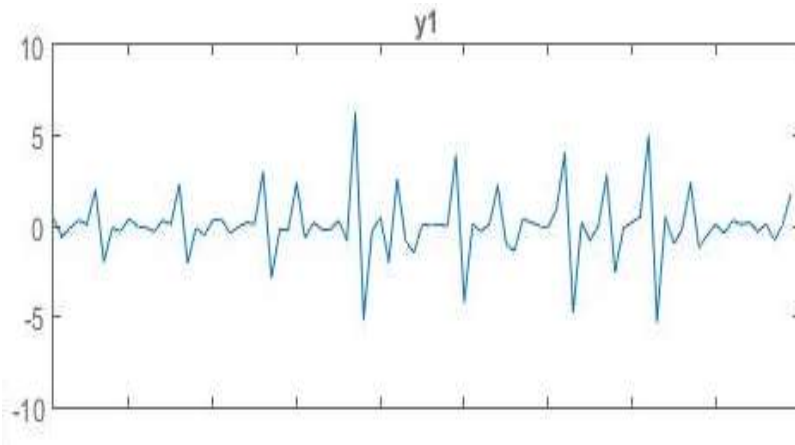
الشكل (4) : يمثل دالة الارتباط الذاتي الجزئي لسلسلة الاخراج $Y(t)$



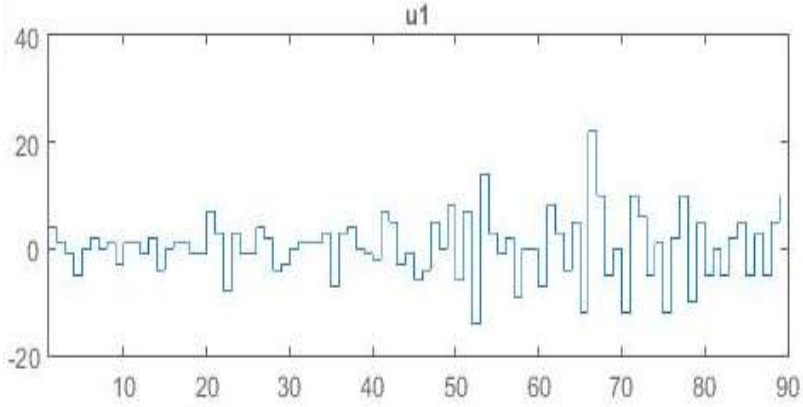
الشكل (5) : يمثل دالة الارتباط الذاتي لسلسلة الادخال $u(t)$



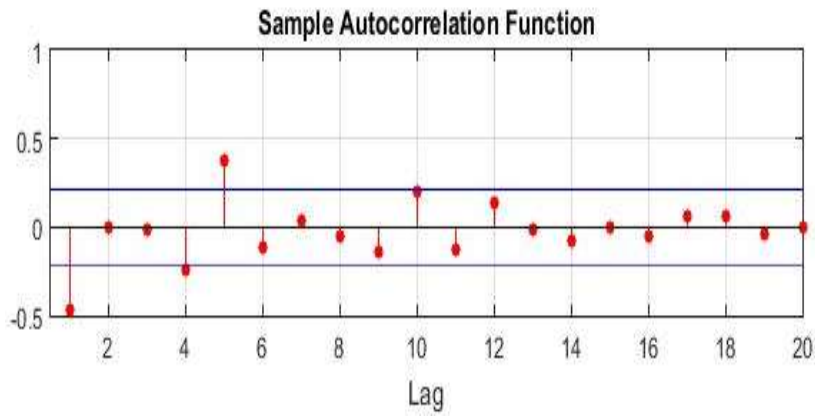
الشكل (6) : يمثل دالة الارتباط الذاتي الجزئي لسلسلة الادخال $u(t)$



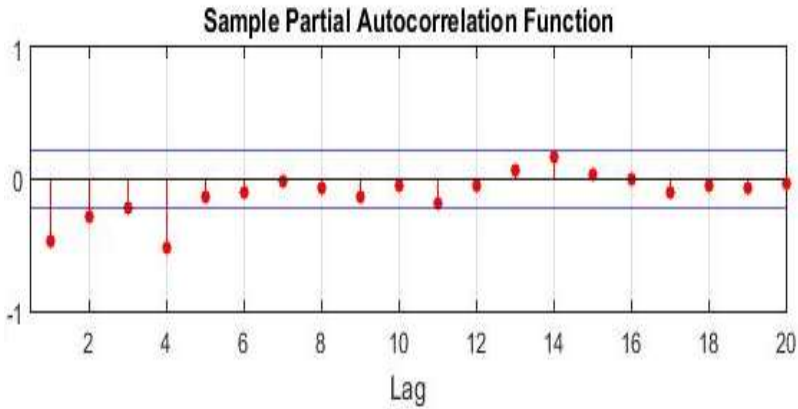
الشكل (7) : يمثل الفرق الاول لسلسلة الاخراج $Y(t)$



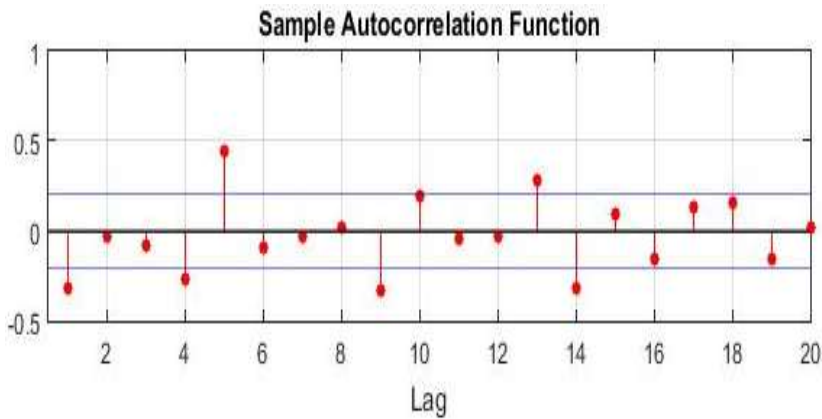
الشكل (8) : يمثل الفرق الاول لسلسلة الادخال $u(t)$



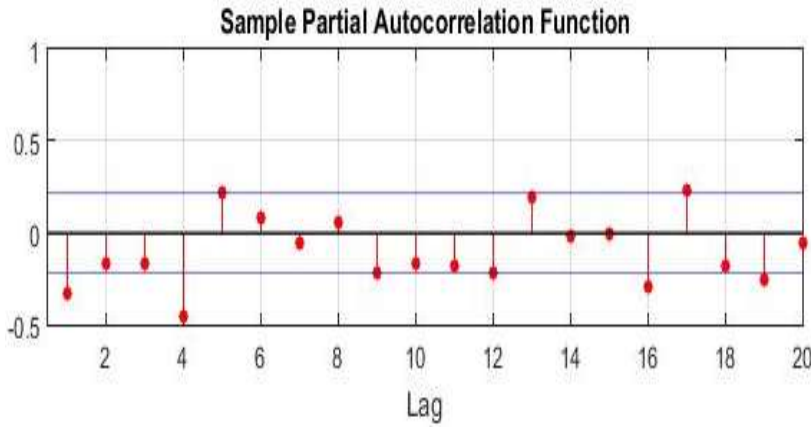
الشكل (9) : يمثل داله الارتباط الذاتي لسلسلة الاخراج $Y(t)$ بعد اخذ الفرق الاول



الشكل (10) : يمثل دالة الارتباط الذاتي الجزئي لسلسلة الاخراج $Y(t)$ بعد اخذ الفرق الاول



الشكل (11) : يمثل دالة الارتباط الذاتي لسلسلة الادخال $u(t)$ بعد اخذ الفرق الاول



الشكل (12) : يمثل دالة الارتباط الذاتي الجزئي لسلسلة الادخال $u(t)$ بعد اخذ الفرق الاول

تم رسم السلسلة الزمنية وتبين من الأشكال (1) و (2) و (3) و (4) و (5) و (6) ان سلسلتي الادخال والايخراج غير مستقرتين لذلك تم اخذ الفرق الاول لاستقرار السلسلتين ورسم سلسلتي الادخال والايخراج كما في الاشكال (7) و (8) و (9) و (10) و (11) و (12) الذي يمثل استقرارية السلسلتين

❖ تحديد الرتبة

تم تحديد رتبة الانموذج والذي تمثل عدد معاملات الانموذج لكل من انموذج ARX وانموذج ARMAX باستخدام كل من معيار AIC و FPE وتبينت النتائج ان رتبة الانموذج (ARX, ARMAX) هي:

• **ARX(5,4,1)**

$$AIC = 0.4523$$

$$FPE = 1.5438$$

• **ARMAX(1,2,1,1)**

$$AIC = 0.5093$$

$$FPE = 1.6571$$

$$ISSN (1681-6870)$$

❖ وقد تم تقدير معلمات الأنموذج ARX وأنموذج ARMAX بطريقة
المربعات الصغرى الاعتيادية

وتم التوصل الى مايلي

♦ ان صيغة أنموذج ARX (5,4,1) هي

$$A(z) = 1 + 0.8417 z^{-1} + 0.5873 z^{-2} + 0.526 z^{-3} + 0.4464 z^{-4} + 0.2484 z^{-5}$$

$$B(z) = 0.0069 z^{-1} - 0.1236 z^{-2} - 0.0479 z^{-3} - 0.0504 z^{-4}$$

♦ اما الصيغة التقديرية لأنموذج ARMAX(1,2,1,1) فهي

$$A(z) = 1 + 0.0417 z^{-1}$$

$$B(z) = 0.0417 z^{-1} - 0.128 z^{-2}$$

$$C(z) = 1 - 0.7672 z^{-1}$$

حيث ان قيمة ARX(5,4,1)

$$RMSE = 0.8717$$

$$MAE = 1.301$$

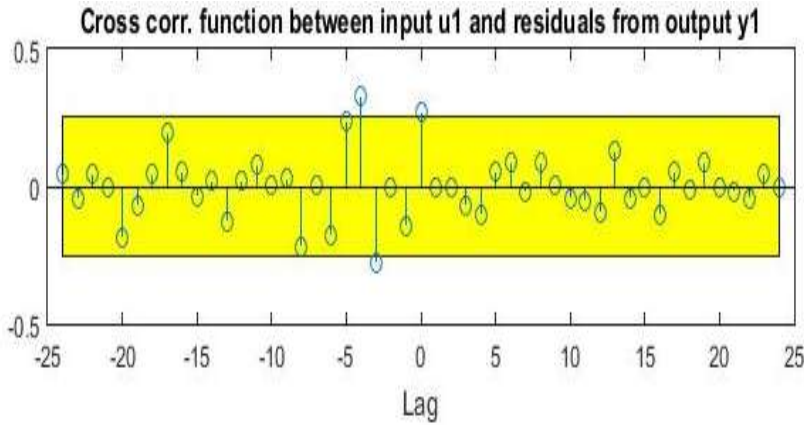
وقيمة ARMAX(1,2,1,1)

$$RMSE = 0.8234$$

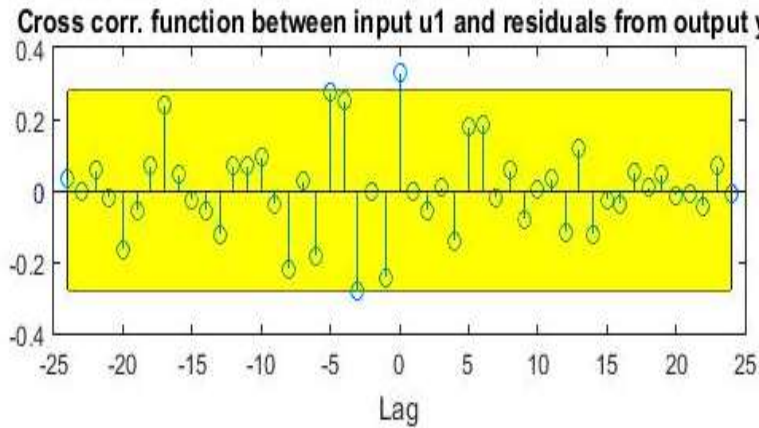
$$MAE = 1.2889$$

ولاختبار مدى دقة الأنموذج تم استخدام اختباري الاستقلالية والبياض

♦ اختبار الاستقلالية :



الشكل (13) : يمثل الارتباط المتقاطع بين الخطأ والمدخل لأنموذج ARX(5,4,1)

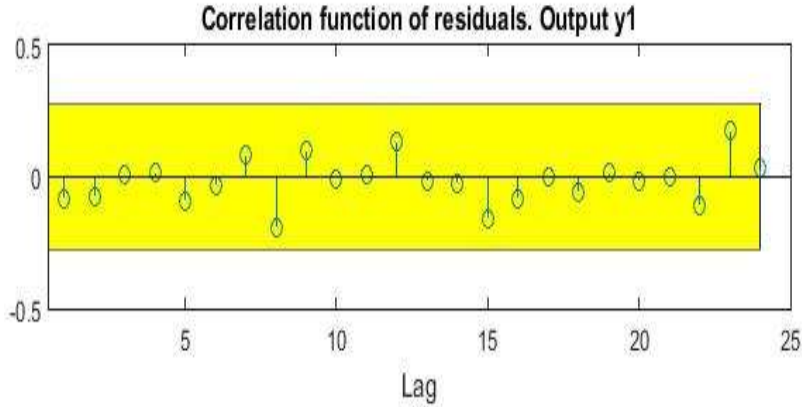


الشكل (14) : يمثل الارتباط المتقاطع بين الخطأ والمدخل

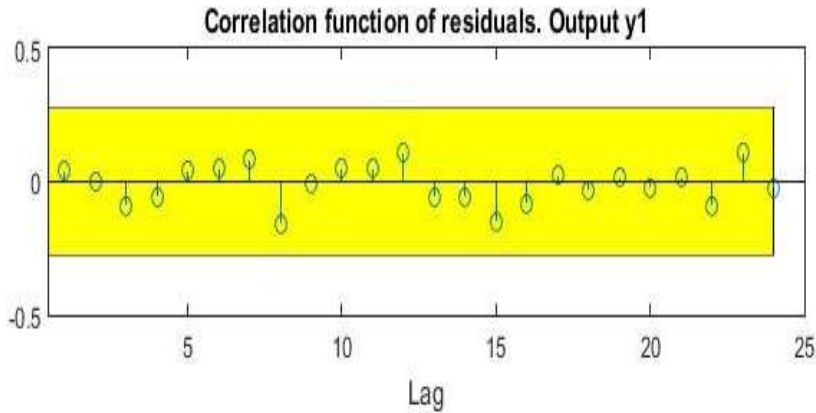
لأنموذج ARMAX(1,2,1,1)

يتضح من الاشكال (13) و (14) الخاصة بنماذج خطأ المعادلة وبصورة عامة لا يوجد ارتباط معنوي بين قيم الخطأ والمدخل وهذا يشير الى صحة افتراض بياض الخطأ .White Noise

♦ اختبار البياض



الشكل (15) : يمثل الارتباط الذاتي للخطأ لأنموذج ARX(5,4,1)



الشكل (16) يمثل الارتباط الذاتي للخطأ لأنموذج ARMAX(1,2,1,1)

ومن الاشكال (15) و (16) يتضح بعدم وجود ارتباط معنوي بين قيم الاخطاء مع بعضها البعض ولكل من النماذج الاربعة وهذا يعني عدم رفض الفرضية H_0 بان الارتباطات ليست لها فروقات وان النماذج لها صفة White Noise .

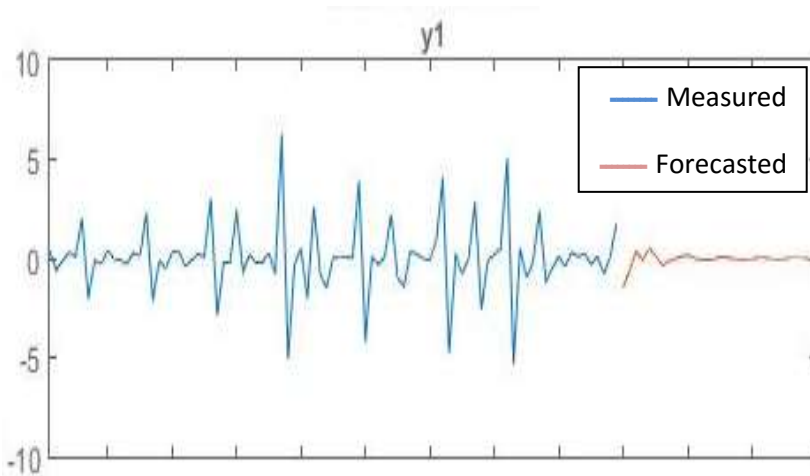
❖ التنبؤ

جرى التنبؤ بالقيم المستقبلية للمتغير $Y(t)$ باستعمال النماذج الخمسة التي تم معرفتها من خلال تحديد رتب انموذج ((ARX(5,4,1), ARMAX(1,2,1,1)) والجدول رقم (1) يبين القيم التنبؤية للنماذج وللفترة من 2015/4/1 ولغاية 2015/4/30

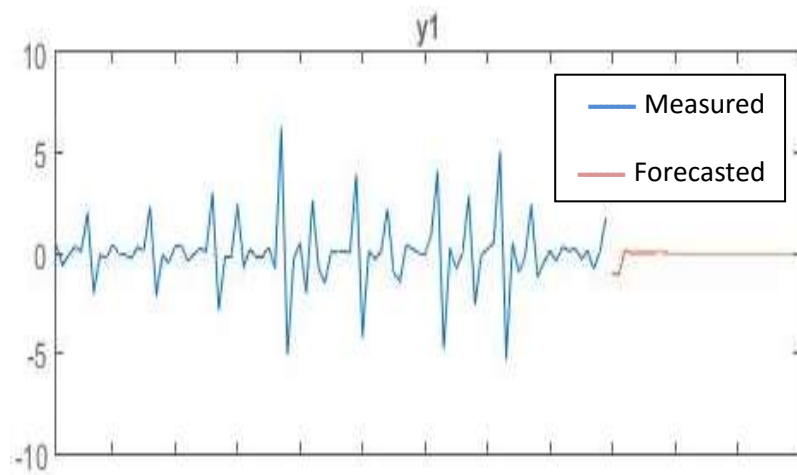
الجدول (1) : يمثل القيم التنبؤية لمقياس عكورة الماء الصالح للشرب

ARX(5,4,1)	ARMAX(1,2,1,1)
-1.47150275	-0.99509244
-0.59698561	-1.12665286
0.36058896	0.17415416
-0.07430867	-0.02692016
0.49459440	0.00416123
0.15853178	-0.00064323
-0.36845415	0.00009943
-0.14604783	-0.00001537
-0.02211930	0.00000238
0.09189104	-0.00000037
0.17534454	0.00000006
0.00481887	-0.00000001
-0.09907707	0.00000000
-0.07002432	0.00000000
-0.01267161	0.00000000
0.05452007	0.00000000
0.05581690	0.00000000
-0.00123822	0.00000000
-0.03476942	0.00000000
-0.03000541	0.00000000
-0.00159701	0.00000000
0.02289685	0.00000000
0.01911867	0.00000000
-0.00088721	0.00000000
-0.01380944	0.00000000
-0.01151451	0.00000000
0.00054181	0.00000000
0.00900528	0.00000000
0.00691035	0.00000000
-0.00067189	0.00000000

حيث تم رسم القيم التنبؤية لمخرجات الانموذج $Y(t)$ وكما في الشكلين (17) و (18) ولكل من الانموذجين ($ARX(5,4,1)$, $ARMAX(1,2,1,1)$) على التوالي.



الشكل (17) : يمثل التنبؤات للمتغير $Y(t)$ لأنموذج $ARX(5,4,1)$



الشكل (18) : يمثل التنبؤات للمتغير $Y(t)$ لأنموذج $ARMAX(1,2,1,1)$

5. الاستنتاجات و التوصيات

أ. الاستنتاجات

Conclusion

من خلال النتائج التي تم التوصل اليها في الجانب النظري والجانب التطبيقي من هذا البحث فقد تم التوصل الى الاستنتاجات التالية:

1. ان بناء نماذج الصندوق الاسود تقدم نتائج جيدة وذلك بالاعتماد الى حد كبير على مقاييس ومعايير محددة لتحديد رتب الانموذج وكذلك يعتمد على الخبرة .
2. من خلال مراحل التطبيق العملي لوحظ ان السلسلتين استقرت بعد اخذ الفرق الاول لها .
3. استخدام بعض المعايير الاحصائية للتوصل الى رتبة الانموذج الملائم وذلك بالاعتماد على المعيارين التاليين (FPE , AIC) في تحديد رتبة الانموذج وقد تم التوصل الى النماذج الملائمة الاتية ARMAX (5,4,1), ARX (1,2,1,1).
4. في اختبار دقة الأنموذج وبصورة عامة ظهرت النتائج بعدم وجود علاقة بين الخطأ والمدخل ولكل أنموذج من نماذج الصندوق الاسود . بالإضافة الى الاستقرارية بين اخطاء كل الأنموذج حيث ظهرت النتائج بانها ضمن حدي الثقة البالغة (F0.2070) .

ب. التوصيات

Recommendations

يمكن ان نوصي بما يلي

1. دراسة نماذج الصندوق الاسود باستعمال الطرائق المعلمية والطرائق اللامعلمية واجراء مقارنة بينهما لمعرفة افضل طريقة لتقدير معالم الانموذج.
2. دراسة بناء نماذج الصندوق الاسود باستعمال البيانات الموسمية بدلا من البيانات الغير موسمية.
3. دراسة نماذج الصندوق الاسود متعددة المدخلات ومخرج واحد للنتبؤ بالمياه الصالحة للشرب وذلك لوجود عدد كبير من الشواذب والاجسام الصغيرة التي تمثل المدخلات.

References

المصادر

(Arabic References)

اولاً: المصادر العربية

- [1] العقابي، أنمار رسول شعبيث، 2016، " بناء افضل أنموذج صندوق اسود باستعمال بعض الطرائق المعلمية مع تطبيق عملي"، رسالة ماجستير غير منشورة، كلية الادارة والاقتصاد، جامعة بغداد.
- [2] عزيز، حسين وحيد، عمران، حنان عبد الكريم وميثم، كفاية حسن، 2015، " تلوث المياه في جدول الكفل"، مجلة كلية التربية الاساسية للعلوم التربوية والانسانية، العدد 19، ص ص (594-605).
- [3] فاندل، والتر، 1992، " السلاسل الزمنية من الوجه التطبيقية ونماذج بوكس جنكنز"، تعريب عبد المرضي حامد عزام، دار المريخ للنشر، الرياض، المملكة العربية السعودية.
- [4] محمد، منعم عزيز، 2011، " التحليل والتنبؤ باستخدام السلاسل الزمنية"، طبعت على نفقة جامعة السليمانية .

(Foreign References)

ثانياً: المصادر الاجنبية

- [5] Babazadeh, Mehrdad, Jedermann, Reiner & Lang, Walter, 2008, "Comparative study of the Best Estimators in a New Modeling Technique Using Wireless Sensor Networks", 8th Wses international conference on simulation, Modeling and optimization (SMO'08) stantander, Cantabria, Spain, September, pp(23-25) .
- [6] Bodeux , J B & Golinval, J. C., 2001, "Application of ARMAV models to the identification and damage detection of mechanical and civil engineering", IOP Publishing Ltd. Printed in the UK , PP (1-10).
- [7] Chiras, N., (2002), "Linear and Nonlinear Modeling of Gas Turbine Engines", Ph.D. Thesis, University of Glamorgan Limassol Gyprus.
- [8] Dekker, Arnold J. Den., Bombois, Xavier & Hof, Paul M.J. Van Den, 2007, "Likelihood based uncertainty bounding in prediction error identification using ARX models: A simulation study", Proceedings of the European control conference Kos, Greece .

- [9] Diversi, Roberto, Guidorzi, Roberto & Soverini, Umberto, 2011, "Identification of ARMAX models with noisy input and output", Preprints of the 18th IFAC world congress Milano (Italy).
- [10] Frausto, H. Uchida, Pieters, J.G. & Deltour, J.M., 2003, "Modeling Greenhouse temperature using system identification by means of neural networks", Biosystems Engineering, 84 (2), pp(147-157) .
- [11] Huusom, Jakob Kjobsted, Poulsen, Niels Kjolstad, Jorgensen, Sten Bay & Jorgensen, John Bagterp, 2010, "ARX-Model based model predictive control with offset-free tracking", 20th European Symposium on computer Aided process Engineering-ESCAPE20
- [12] Ljung, Lennart, 1997, "System identification", Department of Electrical Engineering, Linkoping University.
- [13] Ljung, L., (1999), "System Identification Theory for user", 2nd ed., Prentice Hall Upper Saddle River N.J. London UK .
- [14] Ljung, Lennart, 2001, "Black-box Models from Input-output Measurements", IEEE Instrumentation and Measurement Technology Conference Budapest, Hungary, May, pp (21–23) .
- [15] Ljung, L., (2004), "System Identification Toolbox for use with MATLAB", 6.0 Mathworks Inc .
- [16] Mohler, I., Andrijic, Z. Ujevic, Bolf, N. & Galinec, G., 2013, "Distillation End point Estimation in Diesel Fuel production", Chem. Biochem. Eng., Q 27 (2),PP (125-132) .
- [17] Nelles, O., (2001), "Nonlinear System Identification from Classical Approach to Neural Network and Fuzzy Models", Springer Verlag Belin, Heidelberg Germany
- [18] Park, Herie, Martaj, Nadia, Ruellan, Marie, Bennacer, Rachid & Monmasson, Eric, 2013, "Modeling of a Building System and its Parameter Identification", J. Electr. Eng. Technol. Vol. 8, No. 5: pp (975-98) .

- [19] Rabbani, Muhammad Junaid, Hussain, Kashan, Khan, Asim-ur-Rehman & Ali, Abdullah, 2013, "Model Identification and Validation for a Heating System using MATLAB System Identification Toolbox", journal citation and DOI. Published under license by IOP Publishing Ltd.
- [20] Sahu , Pradeep Kumar & Kumar , Rajesh, 2014, "The Evaluation of Forecasting Methods for Sales of Sterilized Flavored Milk in Chhattisgarh", International Journal of Engineering Trends and Technology (IJETT), Volume 8, Number 2 .
- [21] Sjöberg, Jonas, Zhang, Qinghua, Ljung, Lennart, Benveniste, Albert ,Delyon, Bernard, Glorennec, Pierre-Yves, Hjalmarsson, HÅkan & Juditsky, Anatoli ,1995, "Nonlinear Black-box Modeling in System Identification: a Unified Overview ", Copyright Elsevier Science Ltd Prnted in Great Britain, All rights reserved, Auromarica. Vol.31, No.12, pp (1691-1724).
- [22] S. Mahmoud, Magdi & Xia, Yuanqing, 2012, "Applied Control Systems Design", Springer-Verlag London Limited.
- [23] Zheng, Wei Xing, 2002, "ON Least-Squares identification of ARMAX Models", 15th triennial world congress, Barcelona, Spain.
- [24] Wei, W.W.S., (1990), "Time Series Analysis Univariate and Multivariate Methods", Addison Wesley Publishing Company Inc. The advanced Book Program California U.S.A.

(Electronic References)**ثالثاً : المصادر الالكترونية**

- [25] <http://worldcomp-proceedings.com/proc/p2012/ICA7739.pdf>

Use of Prototypical ARX and ARMAX Prediction Time Series with Practical Application (A Comparative Study)

Assist Prof. Dr. Ahlam Ahmed Juma

dr_ahlam_ahmed@yahoo.com

University of Baghdad – College of Administration and
Economics - Department of Statistics

Anmar Rasool

rasool.anmar@yahoo.com

University of Baghdad – College of Administration and
Economics - Department of Statistics

Abstract: *The researchers interested in studying the black box models, which are linking a series of inputs with outputs series athlete models and the two models are: Error equation models and includes a model of self-gradient with external input ARX and model autoregressive and moving averages with external ARMAX inputs, models output error. The model includes directing an error OE and model Box Genghis BJ. This research has singled out only error equation models so research aims to predict the time and chained to determine the rank of the model using some statistical and engineering standards, including the standard Akaike information AIC and the standard error of the Akaike final prediction FPE.*

This research has included the sequential steps of creating data and Stability chains input and output and determine the rank of the model using some statistical and engineering standards and estimate its parameters. The test model accuracy by testing leftover moral link to the specified model in addition to the moral link test of residuals $e(t)$ and the entrance $u(t)$. Compared with the prediction models for black box selected using the average absolute error MAE and root mean square error RMSE has been the application on real data, a daily drinking water filter using Ekorh scale water project in the Karkh district of Baghdad for the period 01.01.2015 until 03.31.2015.

Keywords: Autoregressive with exogenous input model, Autoregressive moving Average with Exogenous input model, Output Error Model, The Box-Jenkins Model, Akaike's In Formation Criteria, Akaike's Final Prediction Error Criteria, sum of absolute error, Root Mean Square Error