

استخدام طريقي بوكس - جنكيرز و التنقية المكيفة لتحديد افضل نموذج للتنبؤ بتلوث مياه نهر دجلة باستخدام السلالسل الزمنية
الفئوية.....

استخدام طريقي بوكس - جنكيرز و التنقية المكيفة لتحديد افضل نموذج للتنبؤ بتلوث مياه

نهر دجلة باستخدام السلالسل الزمنية الفئوية

م. م عقيل حميد فرحان

جامعة ديالى - كلية الادارة والاقتصاد - قسم الاحصاء

Using the Box-Genghis and Adaptive Purification methods
to determine the best model for Redicting water pollution of
the Tigris River using class time series

Asst. Ins . Aqeel Hmeed Farhan

University of Diyala

College of Administration & Economics

استخدام طيفي بوكس - جنكير و التقنية المكيفة لتحديد افضل نموذج للتنبؤ بتلوث مياه نهر دجلة باستخدام السلسل الزمنية الفنوية.....

م. م عقيل حميد فرحان

الملخص :

تعد السلسل الزمنية الفنوية ذات أهمية بالغة لوجود العديد من التطبيقات المهمة لها في مختلف المجالات التطبيقية الواقعية إلا أنه يلاحظ وجود تغافل واضح لها في أدبيات موضوع السلسل الزمنية وذلك بسبب افتقار التقنيات لمعالجة تلك السلسل الزمنية الفنوية.

كما أن التنبؤ باتجاهها في المستقبل يتتيح لنا معرفة سلوك الظاهرة وعليه فان الهدف من هذه الدراسة هو التنبؤ للسلسل الزمنية الفنوية لحالة التلوث الحاصلة لمياه نهر دجلة من خلال شبكات تصريف مياه الامطار بسبب تجاوز سكان المدن على هذه الشبكات الخاصة بتصريف مياه الامطار فقطفي محافظة واسط باستعمال طيفي بوكس جنكير وطريقة التقنية المكيفة وقد تم تقدير المعلمات باستخدام طريقة الإمكان الأعظم التي طبقت على بيانات أسبوعية لمدة سنة كاملة بواقع (51) أسبوعاً أخذت من مختبرات تلوث البيئة الخاصة بمحافظة واسط وأظهرت نتائج التطبيق أن النموذج الملائم والكافئ لتمثيل بيانات السلسلة الزمنية الفنوية هو النموذج من الدرجة (1, 2, 4)ARMA وأظهرت نتائج التنبؤ استنتاجاً أن طريقة التقنية المكيفة تظهر نتائج أدق من بوكس- جنكير Box – Jenkins method بسبب ظهور تنبؤات متناسبة مع مثيلتها في السلسلة الأصلية.

Abstract

The categorical time series is very important because there are many important applications in the various fields of applications. In fact, there is a clear disregard for them in the literature of the time series because of the lack of techniques to treat with these categorical time series .

Predicting its future direction also allows us to know the behavior of the phenomenon ,the objective of this study is to predict by the categorical time series of the state of pollution of the water of the Tigris River through rainwater drainage networks because the urban population exceeded these networks for the drainage of rain water only in Wasit Governorate, by using Box-Jenkins method and Adaptive Filtering method ,the parameters were estimated using the maximum likelihood method ,Which were applied to weekly data and for a full year of 51 weeks taken from environmental and private pollution laboratories in Wasit Governorate, The results of the application showed that the appropriate and efficient model for representing The categorical time series data is ARMA(4,2,1) ,The results of the prediction showed our conclusion that Adaptive Filtering method showed more accurate results than Box-Jenkins method because of predictions consistent with its Original Series

المقدمة

لقد أهتم الإنسان منذ أقدم العصور باستشراف المستقبل والتنبؤ بتغيراته والتنبؤ بقيمها على مستوى المؤسسات والدول فإن قراءة المستقبل تعد عملية بالغة الأهمية لارتباطها الصميمى الوثيق مع الخطط التخطيطية والتنموية المختلفة وأن من أهم العمليات التي تؤدي إلى تطور البلدان هي عملية التخطيط، والتخطيط يكون على أساس أحسن وجه إذا اعتمدت فيه الأساليب الإحصائية التي تؤدي إلى تخطيط قائم على أساس علمية متينة. ومن ابرز الأساليب الإحصائية هي السلسل الزمنية ، التي تعد من الموضوعات المهمة والمستخدمة في تفسير سلوك الظواهر عن طريق دراسة تطورها التاريخي عبر مدة زمنية قد تكون يومية أو أسبوعية أو فصلية أو شهرية... الخ لأغراض التنبؤ بما سيحدث فيمستقبل هذه الظواهر بأقل خطأ ممكن والغاية الأساسية من استعمال السلسل الزمنية يتمثل بالحصول على نموذج يمكن استخدامه لوصف المشكلة وبالتالي التنبؤ بمستقبل الظاهرة المدروسة وعليه سنقوم بدراسة نوع معين وهم من السلسل الزمنية هي السلسل الزمنية الفئوية Categorical Time Series وهي سلسل تمثل بشكل حروف أو رموز وليس أرقاماً أي تمثل أصنافاً معينة لحالة ما مثل [أسعار دور السكن تتغير تبعاً للمتغيرات التوضيحية (كمساحة حديقة الدار و موقع الدار) ومثل أسلوب التحرك عن طريق المركبات و القطارات والطائرات والسفن] ولها أهمية واضحة في التطبيقات الطبية والبيولوجية و الجيولوجية كذلك يتم التنبؤ باستعمال الشبكات العصبية التي توصف بأنها سلسلة من الحالات التي تمر بها ظاهرة معينة خلال انتقالها عبر المحور الزمني أو أي دليل آخر استناداً إلى قوانين احتمالية تدعى بالاحتمالات الانتقالية وقد تشعبت تطبيقات سلسل ماركوف وشملت مساحه واسعة ابتداء من التطبيقات الإنسانية وحتى التطبيقات الطبية مروراً بالتطبيقات العلمية التقنية المختلفة.

الدراسات السابقة

1- في عام 2005 قامت الباحثة ابتهاج الكسو بتقديم أطروحة دكتوراه لتقدير رتبة سلسلة ماركوف على بيانات سلسلة زمنية فئوية باستخدام الشبكات العصبية طبقت على بيانات جبل بطة في محافظة نينوى 2- وفي عام 2010 قدم الباحث Biswas بتقديم بحث استخدم فيه المعلومات المتباينة لتحليل السلسل الزمنية الفئوية إذ قام بتكوين نموذج انحدار ذاتي متقطع ثم لاءم بينه وبين العمليات الفئوية على وفق نموذج انحدار ذاتي والأوساط المتحركة ARMA إذ لاحظ عدم ثبات الارتباط في البيانات الفئوية لذا لجأ الباحث إلى المعلومات المتباينة (auto-mutual information).

مشكلة البحث

يعتبر استعمال السلسل الزمنية الفئوية مشكلة بحد ذاتها وكيفية التعامل مع بيانات فئوية حقيقة ويمكن تعريف مشكلة البحث بأنها ايجاد افضل طريقة للتنبؤ للسلسل الزمنية الفئوية

هدف البحث

يهدف هذا البحث الى المقارنة بين طريقي بوكس جنكير والتقيية المكيفة في السلسل الزمنية الفئوية للتنبؤ بحالة تلوث المياه في نهر دجلة للحصول على أفضل طريقة للتنبؤ من خلال تطبيقها على بيانات فئوية حقيقة تمثلت

بحالة تلوث مياه نهر دجلة بسبب التجاوز من قبل السكان على شبكات الامطار الخاصة بتصريف مياه الامطار

في محافظة واسط

الجانب النظري

سيتم استعمال السلسل الزمنية الفئوية **Categorical Time Series** التي سيتركز عليها بحثنا كما سنتناول مسألة التنبؤ بهذه السلسلة وذلك بالاعتماد على الطائق الإحصائية التقليدية للسلسل الزمنية إذ يمكن تعريف السلسل الزمنية الفئوية Categorical Time Series بانها تلك السلسل الزمنية التي يكون فضاء الحالة لها بشكل حروف أو رموز وليس أرقاماً. وعلى سبيل المثال قام الباحثون [Stoffer, et.al; 1988] بدراسة حالة النوم Sleep-State للطفل وذلك باستخدامهم جهاز (EEG) Electronic Phalographic. إن حالة النوم

في تلك الدراسة قد صنفت (كل دقيقة) إلى ست حالات هي:

أثر تناوب النوم الهدائى (qt) quiet sleep-trace alternant

القولتية العالية للنوم الهدائى (qh) quiet sleep-high voltage

النوم الانتقالي (tr) transitional sleep

القولتية المنخفضة للنوم النشيط (al) active sleep- low voltage

القولتية العالية للنوم النشيط (ah) active sleep-high voltage

الاستيقاظ (aw)

إن المشاهدات المذكورة في تلك الدراسة هي كما يأتي (تقرأ من اليسار إلى اليمين):

{ah, ah, ah, ah, ah, ah, ah, tr,}.

فلو افترضنا أن $\{X_t; t=0,1,2,...\}$ تمثل السلسلة الزمنية لحالة النوم لذلك الطفل، فكما هو واضح فإن فضاء الحالة لهذه السلسلة هو المجموعة الفئوية $S=\{qt, qh, tr, al, ah, aw\}$. وكما هو معروف فإن التحليل الإحصائي لمثل هذه السلسلة الزمنية ليس بالأمر اليسير ، لأن عناصر فضاء الحالة ليست قيماً عددياً. لذا اقترح الباحثون [Stoffer, et.al, 1988] إجراء مقياس (Scaling) فأخذوا أولاً المقياس الآتي:

$qt=1, qh=2, tr=3, al=4, ah=5, aw=6$

ثم ناقشوا بعد ذلك، ولأسباب خاصة مقياساً آخر هو:

$qt=1, qh=1, tr=2, al=3, ah=3, aw=4$

لذا يلاحظ أن إجراء المقياس لهذه المشاهدات يكون عملية غير وحيدة وأن هذه الملاحظة تمثل نقطة ضعف واضحة في البيانات الفئوية مما يؤدي بالنتيجة إلى نتائج غير دقيقة ولا يمكن اعتمادها بشدة كبيرة وذلك لأن الرقم المقاس لا يعني شيئاً في هذه الحالة وعليه سيتم استخدام طائق التنبؤ التالية :

طريقة بوكس - جنكير [Box - Jenkins method]

يقصد بنماذج بوكس - جنكير (Box and Jenkins) هي تلك المنهجية التي طبقها كل من (George Box) و (Gwilyn Jenkins) على السلسل الزمنية عام 1976 وهذه المنهجية تعتمد في صياغتها على ثلاثة أجزاء هي

استخدام طيفي بوكس - جنكير و التئية المكيفة لتحديد افضل نموذج للتنبؤ بتلوث مياه نهر دجلة باستخدام السلسل الرمزية
الثانوية.....

(Moving Average Model) (Autoregressive) (نماذج المتوسطات المتحركة) (Moving Average -Auto Regressive)
والمودع المختلط . وتعتمد هذه المنهجية على أربع مراحل لغرض التنبؤ .

1. مرحلة تحديد النموذج .Identification
2. مرحلة تقدير المعامل .Estimation
3. مرحلة اختبار دقة النموذج Diagnostic Checking
4. مرحلة التنبؤ Forecasting

إذ يتم التنبؤ بعد إكمال مرحلة اختبار دقة النموذج اذا كان النموذج غير ملائم فانه يتم تجاهله واختبار نموذج اخر وتعاد العملية من جديد أما اذا كان النموذج ملائماً حينها يتم اجراء عملية التنبؤ على أساسه أما أنواع نماذج بوكس - جنكير Box - Jenkins فهي نموذج الانحدار الذاتي (AR) ونموذج المتوسطات المتحركة (MA) والمودع المختلط (ARMA) والمودع الموسمي (SARIMA) وغيرها من نماذج بوكس - جنكز وكالاتي.

نماذج الانحدار الذاتي Auto regressive models
ان صيغة نموذج الانحدار الذاتي من الدرجة (p) الذي يرمز له بالرمز (P) وهو كما يلي

$$Z_T - M + Q_1 Z_{T-1} - Q_2 Z_{T-2} - \dots \dots - Q_a Z_{T-a}$$

حيث ان :

Z_{t-r} تمثل المشاهدات $t=1, 2, 3, \dots, p$

Z_i تمثل مجموعة الاوزان لقيم السابقة الى $i=1, 2, 3, \dots, p$

M : عدد المعالم المقدرة

نماذج الاوساط المتحركة Moving average models

ان الصيغة الاساسية لنماذج الاوساط المتحركة من الدرجة (q) ويرمز لها بالرمز (q) هي كالاتي

$$Z_T = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots \dots - \theta_q a_{t-q}$$

حيث: Z_t : هي قيمة المشاهدة في الفترة (t)

θ_t : معالم ثابتة للنموذج $t=1, 2, 3, \dots, q$ تمثل الأخطاء العشوائية

Mixed auto regressive moving average models

ويرمز لهذه النماذج بالرمز ARMA(p,q) حيث (p) تمثل درجة الانحدار الذاتي ويمثل (q) درجة الاوساط المتحركة وتكتب او توجد على وفق الصيغة التالية

$$Z_T - M - Q_1 Z_{T-1} + Q_2 Z_{T-2} + \dots \dots + Q_p Z_{T-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots \dots - \theta_q a_{t-q}$$

-مرحلة تحديد النموذج Identification

تعتمد مرحلة تحديد النموذج على أيجاد معاملات الارتباط الذاتي (ACF) والارتباط الذاتي الجزئي (PACF) واستخدام بعض المعايير التي تستخدم لتحديد درجة ونوع النموذج ومن أهمها معيار اكياكى (AIC) وكالاتي .

[14,3]: (Auto-Correlation Function) ACF

يعرف معامل الارتباط بأنه مقياس لدرجة العلاقة بين قيم المتغير نفسه عند فترات إزاحة موسمية مختلفة ويحسب معامل الارتباط الذاتي في حالة السلسل الزمنية الموسمية عند الإزاحة k حسب الصيغة الآتية

دالة الارتباط الذاتي الجزئي [Partial Auto-Correlation Function(PACF)]

يعرف معامل الارتباط الذاتي الجزئي بأنه مقياس لدرجة العلاقة بين المشاهدين Z_{t+k} وبثبوت بقية المشاهدات الأخرى $Z_{t+1}, \dots, Z_{t+k-1}$ ويمكن حسابها على وفق الصيغة الآتية .

$$\hat{\phi}_{k+1,k+1} = \frac{\hat{\rho}_{k+1} - \sum_{j=1}^k \hat{\phi}_{kj} \hat{\rho}_{k+1-j}}{\mathbf{I} - \sum_{j=1}^k \hat{\phi}_{kj} \hat{\rho}_j} \quad \dots \quad (3)$$

وان دالة الارتباط الذاتي الجزئي (PACF) لا تقل أهمية عن دالة الارتباط الذاتي فهي أيضاً أداة مهمة في تحليل السلسل الزمنية وتستخدم أيضاً في تشخيص التموذج من خلال عشوائية الباقي أخطاء الباقي الآلية يتم تحديد الأنماذج الملائمة للمشاهدات وهناك معاير أخرى لتحديد الأنماذج مثل .

اداة من أدوات التشخيص معيار معلومات اكياكى [9,6]: Akaike information criterion

$$AIC(M) \equiv n \ln(\sigma^2) + 2M \quad \dots\dots\dots(4)$$

إذ M تمثل العدد الكلي، لمعلمات النموذج (9 ، p).

ن هي عدد مشاهدات السلسلة .

مقدار تباين الخطاء . σ^2

أذ أن الأنماذج الأفضل هو الأنماذج الذي يعطي أقل قيمة من $AIC(M)$.

مرحلة التقدير [14,8,6]: Estimation

إن عملية تدبير النماذج هي المرحلة الثانية من مراحل دراسة السلسلة الزمنية وتحليلها وتأتي بعد عملية تشخيص النموذج الملائم للسلسلة الزمنية اذ توجد عدة طرائق لتقدير معلمات النموذج واهم هذه الطرائق طريقة الإمكان الأعظم (Maximum Likelihood Method) لتقدير معلمات النموذج وذلك بجعل مجموع مربعات الخطأ أقل ما يمكن وصيغة دالة الإمكان الأعظم تكون على النحو التالي .

استخدام طريقي بوكس - جنكيز و التقيية المكيفة لتحديد افضل نموذج للتبؤ بتلوث مياه نهر دجلة باستخدام السلالس الزمنية الفئوية.....

وعند اخذ اللوغاريتم سيكون :

$$LnL(\phi, \boldsymbol{\sigma}_\alpha / Z) = -(0\frac{T}{2})Ln(2\pi) - (\frac{T}{2})Ln\boldsymbol{\sigma}_\alpha^2 + (\frac{1}{2})Ln|M_t^{(P)}| - \left[\frac{S(\phi)}{2\boldsymbol{\sigma}_\alpha^2} \right] \dots \dots \dots (6)$$

حيث ان :

وباستناد الى المعادلة (7) بالنسبة ل Φ ومساواتها للصفر وتبسيطها نحصل على تقدير Φ المطلوب

اختبار دقة النموذج [14,8,6]: Diagnostic Checking

بعد التعرف على الانموذج وتقدير معالمه تأتي مرحلة مهمة هي مرحلة اختبار مطابقة النموذج ومدى كفاءة النموذج وملاءمته لتمثيل بيانات السلسلة الزمنية في تمثيل الظاهرة المدروسة باستخدام اختبارات عده منها .

أولاً: اختبار طبيعة الباقي أي مدى تقارب الباقي المقدرة إلى التوزيع الطبيعي النظري المفترض أن تتوزع

$$\cdot a_t \approx N(0, \sigma^2)$$

ثانياً: اختبار الاستقلالية يتم من خلال اختبار معاملات الارتباط الذاتي للخطأ وكالآتي .

$$H_0: P = 0 \quad \quad \quad H_1: P \neq 0$$

ونستخرج فيه الاحصاء نفسها لاختبار فرضية المتوسط $\mu = \frac{e}{se(e)}$ وهذه الاحصاء لها توزيع طبيعي

قياسي فإذا كانت $|u| < 1.96$ عند مستوى معنوية 0.05 عند ذلك تقبل فرضية العدم .

[10,3]:Forecasting التنبؤ

بعد تحديد النموذج الملائم من خلال مرحلة التشخيص وتقدير معالمه وفحص ملاءمة النموذج يتم استخدامه في التنبؤ بالقيم المستقبلية إلى ($t+L$) فترة قادمة وذلك باخذ التوقع الشرطي عند زمن الأصل (t) لأن النموذج وبعد كتابته عند الفترة ($t+L$) أي ($E_t(Z_{t+L} / Z_1, Z_2, \dots, Z_{t-1})$)، إذ باستخدام التوقع الشرطي نحصل على التنبؤات

$$\cdot \hat{Z}_t(L)$$

آذ يمكن حساب التنبؤات للنموذج المختلط الموسمي على وفق الصيغة الآتية:

حول

$$Z_t = Z_t - \mu \quad \text{and} \quad \phi_{(p+d)}(\beta) = \phi(\beta)(1-\beta)^{-1} = (1-\phi\beta - \dots - \phi_{p+d}\beta^{p+d})$$

ويسمى بمعامل الانحدار الذاتي غير المستقر من الدرجة ($p+q$).

ويمكن كتابة المعادلة على النحو الآتي:

$$Z_{t+L} = \hat{Z}_t(L) = \Phi_S Z_{t+L-S} + \Phi_{2S} \hat{Z}_{t+L+2S} + \dots + \Phi_{PS} \hat{Z}_{t+L-PS} + a_{t+L} - \Theta_S \hat{a}_{t+L-s} \\ - \dots - \Theta_{QS} \hat{a}_{t+L-QS} \quad \dots \dots \dots \quad (8)$$

حيث إن :

$$E_t[Z_{t-L}] = Z_{t-L} \quad L = 0, 1, 2, \dots$$

$$E_t[Z_{t+L}] = Z_t(L) \quad L = 1, 2, \dots$$

$$E_t[a_{t-L}] = a_{t-L} = Z_{t-L} - Z_{t-L-S}(1) \quad L = 0, 1, 2, \dots$$

$$E_t[a_{t+L}] = 0 \quad L = 1, 2, \dots$$

طريقة التنقية المكيفة : [14,8] Adapting Filtering Method

إن أول من بدأ العمل في هذا الحقل هو (Weiner, 1949) (Kolmogorov, 1941) فضلاً عن كل من (Kalman & Bucy, 1961) و (Kalman, 1960) وقد فتحوا طريقاً واسعاً في استخدام هذا الأسلوب في مختلف التطبيقات الهندسية المتعددة [14].

إن التنبؤ بطريقة التقدير المكيف (Forecasting with Adaptive Filtering) والمعروضة من قبل Makridakis & Wheelwright, 1973)، شملت خوارزميات (Algorithms) لتعديل (Revising) المعلومات للنماذج الخطية العشوائية وذلك من خلال إضافة حد لتصحيح المعلومات الأصلية والذي هو عبارة عن نسبة بين حاصل ضرب بواقي أحدث التنبؤ وقيم المشاهدات الأولية. على أن التقدير المعدلة في ذلك الوقت كانت تستخدم نماذج الانحدار الذاتي (Autoregressive) فقط [14].

ولكن منذ عام (1977) توسيع في استخدامها لتشمل نماذج أخرى منها نماذج أخطاء الأوساط المتحركة (Moving Average Errors) لسلسل مستقرة وغير مستقرة. وقد أشار (Nau & Oliver) في عام (1979) إلى أن التتبُّؤ بطريقة التقدير المكيفة المعروضة من قبل (Makridakis and Wheelwright) تعد أكثر دقة وكفاءة تقريباً من أسلوب التقدير المقدمة من قبل (Kalman) المطبقة بشكل فعال في نماذج الانحدار الذاتي. ويعرض الباحث فيما يلي التتبُّؤ بهذه الطريقة من فترة إلى أخرى وليس ثابتة.

التنقية المكيفة إلى الأنماذج الموسمية (الضربي)

[,14,8 ,5]Adaptive filtering to multiplicative model

أن نموذج (ARIMA(p, d, q)) يمكن كتابته على وفق الصيغة الآتية

$$Z_t = \phi_{1s} Z_{t-s} + \dots + \phi_{ps} Z_{t-ps} + a_t - \theta_{1s} a_{t-s} - \dots - \theta_{qs} a_{t-qs} \quad (10)$$

حيث أن : $t = 1, 2, 3, \dots, n$

استخدام طيفي بوكس - جنكير و التئية المكيفة لتحديد افضل نموذج للتنبؤ بتلوث مياه نهر دجلة باستخدام السلسل الرمزية
الثانوية.....

S : فترة الموسم .

at تمثل مركبة الخطأ العشوائي بوسط حسابي مساوٍ للصفر وتبين $\sigma^2 a_t$

$\phi_{1s}, \dots, \phi_{ps}$: تمثل معلمات AR غير المعلومة .

$\theta_{1s}, \dots, \theta_{qs}$: تمثل معلمات MA غير المعلومة .

ويمكن تقدير المعلمات على وفق طريقة المربيعات الصغرى غير الخطية non-linear least square method باستخدام ميل الانحدار السريع بمعنى استخدام الميل عند البحث عن الأصغر آذ أن الميل عند أي نقطة على سطح الاستجابة يمكن إيجادها عن طريق اشتقاق دالة متوسط مربيعات الخطاء Mse المتمثلة a_t^2 وتكون على وفق الصيغة الآتية .

$$a_t = Z_t - \phi_{1s}Z_{t-s} - \dots - \phi_{ps}Z_{t-ps} + \theta_{1s}a_{t-s} + \dots + \theta_{qs}a_{t-qs} + \dots \quad (11)$$

وبترتيب طرفي المعادلة نحصل على .

$$Z_t - \phi_{1s}Z_{t-s} - \dots - \phi_{ps}Z_{t-ps} + \theta_{1s}a_{t-s} + \dots + \theta_{qs}a_{t-qs} \quad (12) (= a_t^2)$$

وباشتقاق الـ a_t^2 بالنسبة إلى ϕ_i و θ_j نحصل على :

$$\frac{\partial a_t^2}{\partial \phi_i} = -2a_t Z_{t-is} \quad i = 1, 2, \dots, p$$

$$\frac{\partial a_t^2}{\partial \theta_j} = 2a_t a_{t-js} \quad j = 1, 2, \dots, q$$

باستعمال الأسلوب التكراري يمكن إن نستنتج أن المعلمات المكيفة المعدلة هي [46]

$$\phi_{is}^* = \phi_{is} + 2k \nabla a^2 \quad (13)$$

حيث أن :

ϕ^* : تمثل المعلمة المعدلة (المكيفة) الجديدة .

ϕ_{is} : تمثل المعلمة قبل التعديل

K : تمثل ثابتًا اختيارياً يقوم بالسيطرة على سرعة التقارب من خلال عدد مرات التكرار المستخدمة .

∇a^2 : متجه التدرج لـ a^2 ومنها نجد .

$$i = 1, 2, \dots, S \quad \phi_{it}^* = \phi_{is} + 2ka_t Z_{t-i}$$

$$\theta_{it}^* = \theta_{is} + 2ka_t a_{t-j} \quad j = 1, 2, \dots, S \quad (14)$$

وعليه أن الصيغ(14) تطبق بشكل متكرر الى الحد او النقطة التي ينعدم فيها الاختزال او التخفيض لمتوسط مرربعات الخطأ (Mse) عند هذه النقطة وعليه فإن عملية التعديل (التئية) تنتهي وأن القيمة النهائية للمعلمة تستخدم للتنبؤ .

ومن الصيغة (65-2) تكون أخطاء الباقي المعدلة الجديدة .

$$\begin{aligned} a_t &= Z_t - \emptyset_{1s}Z_{t-1} - \emptyset_{2s} - \dots - \emptyset_{ps}Z_{t-ps} + \theta_{ps}a_{t-1} + \dots + \theta_{qs}a_{t-qs} \dots \dots \dots (15) \\ \nabla a_t &= a_t^* - a_t \\ &= -(\phi_{1s}^* - \phi_{1s})Z_{t-1} - \dots - (\phi_{ps}^* - \phi_{ps})Z_{t-ps} + (\theta_{1s}^* - \theta_{1s})a_{t-1} + \dots + (\theta_{qs}^* - \theta_{qs})a_{t-qs} \end{aligned}$$

أو

بقسمة المعادلة (16) على (at) ينتج .

ومن العلاقة الأخيرة يمكن استنتاج

عامل التعديل [6,14]:Factor of Adjusting

اقترحها العالم (Whecl Wright) و (Makridakis) عام 1977 لغرض التخلص من التذبذب الحاصل في السلسلة الزمنية والوصول إلى القيم المثلثى بصورة أسرع فإننا يمكن أن نجعل قيم السلسلة الزمنية قياسية وذلك بالقسمة على العامل القياسي (Standard Factor) واقتراح الباحثان (Whecl Wright) و (Makridakis) بأن $z_t = a_t$ يجب أن تعدلا وقد قاما باقتراح عدة طرائق للتعديل من هذه الطرائق هي قسمة كل مشاهدة في السلسلة على أكبر قيمة فيها لكنهما عدلا باقتراهم هذا بعد أن لاحظا قصوراً فيه من الناحية التطبيقية بالاقتراح

الثاني الذي يقول ان تعديل القيم يجب أن يكون بقسمة كل قيمة من القيم للسلسلة على

ولكن الباحث (Chatfield) وجد عام 1978 أن هذه الطريقة تعقد التحليل أحياناً ونتيجة لها الانتقاد قدم الباحث (Makridakis) و (Wright) اقتراحًا جديداً لمعامل التعديل الذي يتم تقسيم قيم السلسلة عليه

وکما یلی:

حيث أن :

T : طول السلسلة الزمنية

W_t : عامل التعديل للفترة t

P: درجة السلسلة الزمنية

ويستعمل معامل التعديل لإزالة تأثير الاتجاه العام للسلسلة الزمنية مما يسهل المقارنة بين السلالس المختلفة.

مقاييس دقة التنبؤ:

نتيجة لاستعمال عدة طائق للتبؤ ومنها طريقة بوكس - جنكيز وطريقة الشبكات العصبية للسلالس الزمنية لابد من توافر مقاييس او معايير لاختيار الطريقة الأفضل للتبؤ .

متوسط الخطأ النسبي المطلق Mean Absolute Percentage Error (MAPE)

وتحسب على وفق الصيغة الآتية .

استخدام طريقي بوكس - جنكيز و التقيية المكيفة لتحديد افضل نموذج للتبؤ بتلوث مياه نهر دجلة باستخدام السلالس الزمنية الفئوية.....

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| * 100 \dots \dots \dots \quad (18)$$

متوسط مربعات الخطأ \ Mean Square Error

يعد متوسط مربعات الخطأ من أكثر وأدق المقاييس الإحصائية استعمالاً وذلك لأنه يعد مقياساً لمدى قرب المعلمة المقدرة من المعلمة الحقيقة وأيضاً يعد مقياساً جيداً بين مختلف طرائق التقدير ويحسب على وفق الصيغة الآتية .

$$MSE = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2 \dots \dots \dots (19)$$

الجانب التطبيقي :

تم الاعتماد على بيانات حقيقة أخذت من مختبرات دائرة البيئة في محافظة واسط والتي استخدمت في هذا البحث تألف سلسلة زمنية فئوية Categorical Time Series بواقع (51) أسبوعاً تمثل حالات التلوث لمياه نهر دجلة بسبب التجاوز الحاصل على شبكات تصريف مياه الامطار من قبل السكان في المدينة سميت فئوية لكونها بيانات ليست عددية بل عبارة عن أسماء تمثل بأسماء (ملوثة - تلوث مقبول - غير ملوثة) والتي تمثل بالحروف أي كانت فضاء الحالة $S = \{A, B, C\}$ وهي بيانات فئوية (وصفية) إذ لا يمكن التنبؤ باستخدام الطرائق الإحصائية الاعتيادية للمشاهدات الفئوية ولمعالجة مسألة التنبؤ لهذه السلسلة نحتاج أولاً إلى ترميز المشاهدات لحالات الطقس الثلاث لتحويلها إلى بيانات عددية بحيث يمكن التنبؤ لها علمًا انه تم ترميز المشاهدات إلى ست حالات مرمزه والتي سميت بـ $\{C_1, C_2, C_3, C_4, C_5, C_6\}$ أذ تم ترميزها وشكل عشوائي، وكما يلي :

وپشكل عشوائي وكما يلي :

3 = 2, C = , B 1 = A(C1) 1 **0** **الحالة الأولى**

الحالة الثانية 2, B=A(C2), C= 2, 3, = 1,

الحالة الثالثة (C3) 1 0 1 1

الحالة الرابعة 20, C=2, B=A(C4), 3=1

الحالة الخامسة (C5) = A, B = C, 3 = 2,

الحالة السادسة 2

- Jenkins . تستخدم لاظهار نتائج التنمية طريقة

أي الطلاق، أفضلاً، التتفق باستعمال السلاسل، الذي

والتي تستخدم لإظهار نتائج التبؤ طريقة Jenkins – Box) ومقارنتها مع طريقة التقنية المكيفة وذلك لمعرفة أي الطرائق أفضل للتتبؤ باستعمال السادس الزمنية الفنوية وتم الاعتماد على البرامج الحاسوبية (Minitab 15, MATLAB) لإظهار النتائج التنبؤية للظاهرة المدروسة.

التبؤ باستخدام طريقة بوكس جنكيز

لمعرف سلوك المشاهدات يجب علينا رسم شكل الانتشار للسلسلة الزمنية الفنوية للمشاهدات الأصلية ومن خلال ملاحظة الشكل رقم (١) نلاحظ أن السلسلة مستقرة في الوسط كونها حضرت بين قيم الترميز التي اختيرت من قبل الباحث .

- مرحلة تحديد النموذج Identification

لمعرفة وتحديد النموذج يتم من خلال رسم دالة الارتباط الذاتي (ACF) ودالة الارتباط الذاتي الجزئي (PACF) للمشاهدات الأصلية المرمزه للسلسلة الزمنية الفئوية و يتضح من شكل رقم (2) و (3) التي تمتلك دالتي الارتباط الذاتي والجزئي أنها تتناقص تدريجياً مع ازيداد فترات الإزاحة (k) سالكة سلوك دالة الجيب المتضائلة تدريجياً ومن خلال هذا المؤشر نستنتج أن النموذج هو نموذج (ARMA) الاعتيادية وباستخدام معيار المقاضلة (AIC) لتحديد الأنماذج الأفضل وكما مبين في الجدول رقم (1)

(المصدر اعداد الباحث اعتمادا على بيانات الدراسة باستعمال برنامج [Minitab 19])

جدول (1)

يمثل قيمة معيار (AIC) لنماذج ARIMA لتلوث مياه نهر دجلة

Case	ARIMA	M	MS	AIC
C1	(2 , 2 , 2)	4	1.057	10.8272
	(3 , 1 , 1)	4	0.98	6.9697
	(3 , 1 , 2)	5	0.9728	8.5936
	(4 , 2 , 1)	5	0.952	5.4913
C2	(2 , 2 , 2)	4	0.952	5.4913
	(3 , 1 , 1)	4	0.89	2.0568
	(3 , 1 , 2)	5	1.342	25.002
	(4 , 2 , 1)	5	0.822	0.00324
C3	(2 , 2 , 2)	4	0.578	-19.957
	(3 , 1 , 1)	4	0.6008	-17.981
	(3 , 1 , 2)	5	0.628	-13.726
	(4 , 2 , 1)	5	0.5529	-20.2215
C4	(2 , 2 , 2)	4	1.054	10.6822
	(3 , 1 , 1)	4	0.9804	6.9905
	(3 , 1 , 2)	5	0.9521	7.4967
	(4 , 2 , 1)	5	0.822	0.00324
C5	(2 , 2 , 2)	4	0.578	-19.957
	(3 , 1 , 1)	4	0.6008	-17.981
	(3 , 1 , 2)	5	0.628	-13.726
	(4 , 2 , 1)	5	0.5529	-20.2215
C6	(2 , 2 , 2)	4	1.057	10.827
	(3 , 1 , 1)	4	0.9804	6.9905
	(3 , 1 , 2)	5	0.9728	8.5936
	(4 , 2 , 1)	5	0.9521	7.4967

حيث أن M : تمثل عدد المعالم المقدرة في النموذج

استخدام طيفي بوكس - جنكير و التقييم المكيفة لتحديد أفضل نموذج للتنبؤ بتلوث مياه نهر دجلة باستخدام السلسل الرمزية
الفنوية.....

مما تقدم نستنتج أن أقل قيمة لمعايير المفضلة المبينة في الجدول رقم (1) يحملها النموذج ARIMA (4, 2, 1) وهو النموذج الملائم للبيانات .

تقدير معالم النموذج : Estimation

قدرت معالم النموذج على وفق طريقة الإمكان الأعظم (Maximum Likelihood Method) وكما موضح في الجدول رقم (2) والتي استخرجت باستخدام البرنامج الجاهز (Minitab 15) وكالآتي .

جدول (2)

النتائج الاحصائية لتلوث مياه نهر دجلة لنموذج ARIMA

C1 مقدرات الحالة الأولى				C2 مقدرات الحالة الثانية			
Type	Coef	StDev	T	Type	Coef	StDev	T
AR 1	-0.4040	0.1370	-2.95	AR 1	-1.9250	0.5170	-3.72
AR 2	-0.1997	0.1475	-1.35	AR 2	-1.3760	0.5867	-2.35
AR 3	-0.2055	0.1470	-1.40	AR 3	-0.3452	0.3939	-0.88
AR 4	-0.4221	0.1353	-3.12	AR 4	0.0806	0.1535	0.52
MA 1	0.9499	0.0700	13.56	MA 1	-0.9344	0.5007	-1.87
Number of observations: Original series 51, after differencing 49 Residuals: SS = 41.8938 (backforecasts excluded) MS = 0.9521 DF = 44				Differencing: 2 regular differences Number of observations: Original series 51, after differencing 49 Residuals: SS = 36.1676 (backforecasts excluded) MS = 0.8220 DF = 44			
C3 مقدرات الحالة الثالثة				C4 مقدرات الحالة الرابعة			
Type	Coef	StDev	T	Type	Coef	StDev	T
AR 1	-0.5478	0.1437	-3.81	AR 1	-1.9250	0.5170	-3.72
AR 2	-0.4700	0.1505	-3.12	AR 2	-1.3760	0.5867	-2.35
AR 3	-0.4551	0.1498	-3.04	AR 3	-0.3452	0.3939	-0.88
AR 4	-0.2637	0.1417	-1.86	AR 4	0.0806	0.1535	0.52
MA 1	0.9585	0.0596	16.07	MA 1	-0.9344	0.5007	-1.87
Differencing: 2 regular differences Number of observations: Original series 51, after differencing 49 Residuals: SS = 24.3275 (backforecasts excluded) MS = 0.5529 DF = 44				Differencing: 2 regular differences Number of observations: Original series 51, after differencing 49 Residuals: SS = 36.1676 (backforecasts excluded) MS = 0.8220 DF = 44			
C5 مقدرات الحالة الخامسة				C6 مقدرات الحالة السادسة			
Type	Coef	StDev	T	Type	Coef	StDev	T
AR 1	-0.5478	0.1437	-3.81	AR 1	-0.4040	0.1370	-2.95

AR 2 -0.4700 0.1505 -3.12	AR 2 -0.1997 0.1475 -1.35
AR 3 -0.4551 0.1498 -3.04	AR 3 -0.2055 0.1470 -1.40
AR 4 -0.2637 0.1417 -1.86	AR 4 -0.4221 0.1353 -3.12
MA 1 0.9585 0.0596 16.07	MA 1 0.9499 0.0700 13.56
Differencing: 2 regular differences	
Number of observations: Original series 51, after differencing 49	
Residuals: SS = 24.3275 (backforecasts excluded)	
MS = 0.5529 DF = 44	
MS = 0.9521 DF = 44	

تم تقدير معالم النموذج لكل حالة وتعد تقديرات جيدة وكفؤة والتي تتبئ بتلوثات جيدة لمستقبل الظاهرة

- اختبار دقة النموذج :Diagnostic Checking

بعد أن تم تحديد النموذج وتقدير معالمه لابد من إجراء بعض التخسيصات على البوافي لبيان مدى دقة النموذج للسلسلة المشاهدة وحسب الاختبارات التالية .

-اختبار استقلالية البوافي :

بعد ان تم اختيار الانموذج تتم عملية اختبار الانموذج وكما مبين في الشكل (4)و (5) والتي اتضح منها انها تتبع أنماط الضجة البيضاء أي أنها غير مترابطة وبذلك فهي إن كانت طبيعية فهي مستقلة .

-الاختبار الطبيعي للتوزيع البوافي .

بعد ملاحظة أن توزيع البوافي يقترب كثيراً من التوزيع الطبيعي بمتوسط صفر وتبين ثابت وذلك من خلال قبول فرضية طبيعة التوزيع فان الرسم يبين مقدار التقارب مع التوزيع الطبيعي .

-التنبؤ :Forecasting

يعد التنبؤ آخر مرحلة من مراحل تحليل السلسلة الزمنية وبعد ان تم تحديد الانموذج الملائم وايجاد معلمات الانموذج المقدرة لكل حالة مرمزه وعليه سنقوم بالتنبؤ الداخلي (Prediction) للسلسلة الزمنية لكل حالة مرمزه، الذي يتم من خلاله المقارنة واختبار دقة التنبؤ إذ تم قطع اخر ستة اسابيع من السلسلة وذلك لاستخدامها في التنبؤ الداخلي وبعد أن وجدت قيم التنبؤ الداخلي وباستعمال برنامج Minitab تم إيجاد قيم مقاييس أخطاء التنبؤ ودونت في الجدول رقم (3) كالتالي

(3) جدول

يوضح نتائج مقاييس أخطاء التنبؤ لطريقة بوكس -جنكر

الحالات	MSE	MAPE
C1	0.762484	11.1531
C2	0.384937	30.07053
C3	0.200222	9.163263
C4	0.348937	8.691941
C5	0.200222	9.88452
C6	0.762484	51.03153

استخدام طيفي بوكس - جنكز و التقييم المكيفة لتحديد أفضل نموذج للتنبؤ بتلوث مياه نهر دجلة باستخدام السلسلة الزمنية الفرعية.....

وتبين من نتائج مقاييس أخطاء التنبؤ في جدول رقم(3) نتائج جيدة وذلك بسبب قلة الأخطاء وبعد التنبؤ المستقبلي (Forecasting) للمشاهدات ظهرت نتائج التنبؤ المستقبلي ودونت في الجدول رقم (4) وكالاتي إذ يتم أيجاد القيم المستقبلية للسلسلة المدروسة باستخدام أفضل نموذج ملائم للسلسلة وكذلكالي .

الجدول (4)

يمثل قيم التنبؤ المستقبلي للحالات المرمزة C1,C2,C3,C4,C5,C6 بطريقة بوكس - جنكز

ت	قيمة التنبؤ C1	تقريب قيمة التنبؤ C1	تحويل الأرقام إلى رموز C1	ت	قيمة التنبؤ C2	تقريب قيمة التنبؤ C2	تحويل الأرقام إلى رموز C2
1	0.87920	1	ملوثة	1	2.00148	2	ملوثة
2	0.80721	1	ملوثة	2	2.00016	2	ملوثة
3	0.739625	1	ملوثة	3	2.00204	2	ملوثة
4	0.685335	1	ملوثة	4	2.00118	2	ملوثة
5	0.665747	1	ملوثة	5	2.00225	2	ملوثة
6	0.607985	1	ملوثة	6	2.00207	2	ملوثة
7	0.554120	1	ملوثة	7	2.00280	2	ملوثة
8	0.493561	1	ملوثة	8	2.00267	2	ملوثة
9	0.428126	1	ملوثة	9	22.00347	2	ملوثة
10	0.381309	1	ملوثة	10	2.00330	2	ملوثة

ت	قيمة التنبؤ C3	تقريب قيمة التنبؤ C3	تحويل الأرقام إلى رموز C3	ت	قيمة التنبؤ C4	تقريب قيمة التنبؤ C4	تحويل الأرقام إلى رموز C4
1	3.08587	3	ملوثة	1	1.99852	2	ملوثة
2	3.12314	3	ملوثة	2	1.99984	2	ملوثة
3	3.14867	3	ملوثة	3	1.99796	2	ملوثة
4	3.16351	3	ملوثة	4	1.99882	2	ملوثة
5	3.19129	3	ملوثة	5	1.99775	2	ملوثة
6	3.23384	3	ملوثة	6	1.99793	2	ملوثة
7	3.26990	3	ملوثة	7	1.99720	2	ملوثة
8	3.29951	3	ملوثة	8	1.99733	2	ملوثة
9	3.32580	3	ملوثة	9	1.99653	2	ملوثة
10	3.35631	3	ملوثة	10	1.99670	2	ملوثة

ت	قيمة التنبؤ C5	تقريب قيمة التنبؤ C5	تحويل الأرقام إلى رموز C5	ت	قيمة التنبؤ C6	تقريب قيمة التنبؤ C6	تحويل الأرقام إلى رموز C6
1	0.914132	1	ملوثة	1	2.96982	3	ملوثة
2	0.876861	1	ملوثة	2	2.99136	3	ملوثة
3	0.851328	1	ملوثة	3	2.95411	3	ملوثة
4	0.836492	1	ملوثة	4	2.97528	3	ملوثة
5	0.808713	1	ملوثة	5	2.94054	3	ملوثة
6	0.766159	1	ملوثة	6	2.95899	3	ملوثة

استخدام طيفي بوكس - جنكير و التقنية المكيفة لتحديد افضل نموذج للتنبؤ بتلوث مياه نهر دجلة باستخدام السلسلة الزمنية
الفنوية.....

7	0.730104	1	ملوثة	7	2.92681	3	ملوثة
8	0.700487	1	ملوثة	8	2.94285	3	ملوثة
9	0.674204	1	ملوثة	9	2.91297	3	ملوثة
10	0.6436688	1	ملوثة	10	2.92680	3	ملوثة

-التنبؤ باستخدام طريقة التقنية المكيفة : Adaptive Filtering Method

بعد الحصول على النموذج الملائم للسلسلة الزمنية وتقدير معلمات الأنماذج للسلسلة الزمنية تم حساب قيم التنبؤ الداخلي أي ضمن الفترة الزمنية للسلسلة وتم إيجاد قيم مقاييس أخطاء التنبؤ دونت في الجدول رقم (5) كالتالي

جدول (5)

يوضح نتائج مقاييس أخطاء التنبؤ لطريقة التقنية المكيفة

الحالات	MSE	MAPE
C1	0.640625	9.6047
C2	0.19063	29.1094
C3	0.0222	8.495
C4	0.19063	3.3854
C5	0.0222	3.495
C6	0.640625	28.3854

وحسبت قيم التنبؤ المستقبلي لطريقة التقنية المكيفة دونت في الجدول رقم (6).إذ يتم إيجاد القيم المستقبلية للسلسلة المدروسة باستخدام أفضل أنماذج ملائم للسلسلة وكالآتي .

الجدول (6)

يمثل قيم التنبؤ المستقبلي للحالات المرمزة C1, C2, C3, C4, C5, C6 بطريقة التقنية المكيفة

ت	قيم التنبؤ C1	تحويل الأرقام إلى رموزC1	ت	قيم التنبؤ C2	تحويل الأرقام إلى رموزC2	ت	قيم التنبؤ C3	تحويل الأرقام إلى رموزC3
1	3	غير ملوثة	1	2	ملوثة	1	1	تلوث مقبول
2	2	تلوث مقبول	2	3	تلوث مقبول	2	3	ملوثة
3	2	تلوث مقبول	3	3	تلوث مقبول	3	3	ملوثرة
4	2	تلوث مقبول	4	3	تلوث مقبول	4	3	ملوثرة
5	3	غير ملوثة	5	2	ملوثرة	5	1	تلوث مقبول
6	3	غير ملوثة	6	2	ملوثرة	6	1	تلوث مقبول
7	3	غير ملوثة	7	2	ملوثرة	7	1	تلوث مقبول
8	3	غير ملوثة	8	2	ملوثرة	8	1	تلوث مقبول
9	1	ملوثرة	9	1	غير ملوثة	9	2	غير ملوثة
10	3	غير ملوثة	10	2	ملوثرة	10	1	تلوث مقبول

ت	قيم التنبؤ C4	تحويل الأرقام إلى رموزC4	ت	قيم التنبؤ C5	تحويل الأرقام إلى رموزC5	ت	قيم التنبؤ C6	تحويل الأرقام إلى رموزC6
1	2	ملوثرة	1	1	ملوثرة	1	3	ملوثرة

استخدام طريقي بوكس - جنكير و التقنية المكيفة لتحديد افضل نموذج للتنبؤ بتلوث مياه نهر دجلة باستخدام السلسلة الزمنية الفئوية.....

2	1	تلوث مقبول	2	2	غير ملوثة	2	1	غير ملوثة
3	1	تلوث مقبول	3	2	غير ملوثة	3	1	غير ملوثة
4	1	تلوث مقبول	4	2	غير ملوثة	4	1	غير ملوثة
5	2	ملوثة	5	1	ملوثة	5	3	ملوثة
6	2	ملوثة	6	1	ملوثة	6	3	ملوثة
7	2	ملوثة	7	1	ملوثة	7	3	ملوثة
8	2	ملوثة	8	1	ملوثة	8	3	ملوثة
9	3	غير ملوثة	9	3	تلوث مقبول	9	2	تلوث مقبول
10	2	ملوثة	10	1	ملوثة	10	3	ملوثة

الاستنتاجات والتوصيات :

- الاستنتاجات: مما تقدم يمكن تلخيص النتائج التالية .

1.بيّنت الاختبارات الإحصائية أن السلسلة الزمنية الفئوية مستقرة في الوسط والتبان وذلك بسبب طريقة الترميز التي اختيرت من قبل الباحث فضلاً عن كونها غير مستقرة هذا ما تبين من خلال دالة الارتباط الذاتي والارتباط الذاتي الجزائي

2.تم اختيار أفضل نموذج من بين النماذج الممكنة باستخدام معايير المفاضلة (اقل قيمة لتبان النموذج ومعيار AIC) وتم فحص ملاءمة النموذج المقترن إحصائيا من تحليل دالة الارتباط الذاتي للبواقي والتوزيع الطبيعي للبواقي .

3.وُجد أن النموذج الملائم والكافٍ لتمثيل بيانات السلسلة الزمنية الفئوية هو النموذج الموسمى المضارع من الدرجة ARIMA(4,2,1)

4.يستنتج الباحث أن ‘طريقة التقنية المكيفة تظهر نتائج أدق من نماذج بوكس - جنكير Box – Jenkins method وذلك بسبب ظهور تنبؤات متباينة مع مثيلتها في السلسلة الأصلية وقدمنا لها تنبؤات مستقبلية لواقع حالة التلوث الحاصلة في مياه نهر دجلة مستقبلا .

5.كلما زاد تقسيم البيانات إلى مجاميع أكثر كانت الدقة أكبر .

6.يتضح أن طريقة التقدير هي الأفضل لكونها هي طريقة غير متحيرة .

- التوصيات : من خلال النتائج التي تم التوصل إليها نوصي بما يلي:

1. الأخذ بنتائج هذا البحث والصيغة المعتمدة للتبؤ من قبل الجهات ذات العلاقة لاعتماد الأسلوب العلمي الملائم في التبؤ .
2. تعميم هذا البحث الى دراسات مناظرة وإجراء مقارنة بينها .
3. يوصي الباحث بأخذ بيانات أكثر للحالات المرمزه بتوسيع عدد المجاميع المرمزه للبيانات
4. يوصي بإجراء عدة طرائق لتقدير المعلمات و المقارنة بينها باستخدام المحاكاة .
- 5- يوصي بالاهتمام بشبكة الامطار ومراقبتها بسبب التلوث الذي يحمل مواد سامة وقتل الاحياء في نهر دجلة

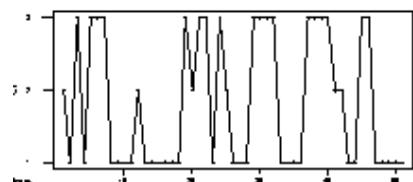
المصادر :

1. الازدي، أيمان سليمان محمد. (2002). "تقدير رتبة سلسل ماركوف مع تطبيق على سلسل الـ DNA"، رسالة ماجستير، كلية علوم الحاسوب والرياضيات، جامعة الموصل، العراق.
2. تاج ، لطفي & سرحان ، عمار (2002) "مقدمة في العمليات العشوائية " كتاب دار الطباعة والنشر
- 3 . الخضيري، محمد قدوري عبد، (1996)"دراسة مقارنة لطرائق التقدير والتبؤ لبعض نماذج بوكس - جنكير الموسمية "رسالة ماجستير إحصاء كلية الإدراة والاقتصاد ،جامعة بغداد ، العراق .
4. السعدي، اسماء ، (2002)"تقدير الاحتمالات الانتقالية لسلسل ماركوف المنتهية غير المستقرة "رسالة ماجستير ، كلية الإدراة والاقتصاد ،جامعة بغداد ، العراق.
5. السعدي ، خليل برهان الدين ، (2007)"تقلبات سعر صرف الدينار العراقي للمدة من عام 1996 ولغاية منتصف عام 2005 والتبؤ بسعره في المستقبل باستخدام سلسل ماركوف " رسالة ماجستير ، كلية الإدراة والاقتصاد ، جامعة بغداد ، العراق .
6. الصراف ، نزار مصطفى وشومان ، عبد اللطيف حسن، (2013) "السلسل الزمنية والأرقام القياسية " دار الدكتور للعلوم الإدارية والاقتصادية ،شارع المتبي ،بغداد ، العراق .
7. الكسو، ابتهاج عبد الحميد.(2005):"استخدام الشبكات العصبية في تقدير رتب سلسل ماركوف مع التطبيق على سلسلة جبل بطمة في محافظة نينوى" ، أطروحة دكتوراه ،كلية علوم الحاسوب والرياضيات ،جامعة الموصل ، العراق .

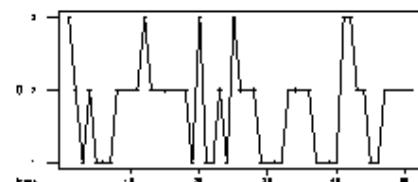
استخدام طريقي بوكس - جنكير و التنقية المكيفة لتحديد افضل نموذج للتنبؤ بتلوث مياه نهر دجلة باستخدام السلالسل الزمنية
المئوية.....

8. كنيلر ، عباس لفترة ، (2010) "بناء نموذج سلاسل زمنية للتنبؤ بإعداد المرضى المراجعين لمستشفى الزهاء في محافظة واسط بحث منشور في مجلة كلية الادارة والاقتصاد ، جامعة واسط ، العراق .
- 9-Akaike, "A new look at the statical model Identification", IEEE, transaction on automatic control, vol. (AC-19), No. (6), p. (716), 1974.
- 10-Anderson ,O.D.(1976)"Time series analysis and Forecasting "Butter Werlhs London and Boston.
- 11- Box, G.P. and Jenkins, G.M. (1976). "Time Series Analysis Forecasting and Control", Revised Edition Holden-Day Inc. San Francisco .
- 12- Lee.T .C & Judge .G.G & Zellner .A.(1970) "Estimating the Parmeters of Markov Propability Model From augregate tiam series data " (2and edition) North Holland Porbishim compauy.Nothevland Amsterdam.
- 13-Schnatter .S, Pamminger. C,(2009)"Bayesian Clustering of Categorical Time Series Using Finite Mixtures of Markov Chain Models " Department of Applied Statisyics Johannes Kepler University Linz Altenbergerstrabe 69,A-4040 Austria.
- 14-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 .

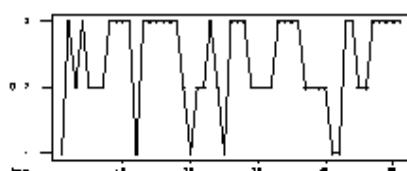
ملحق الاشكال



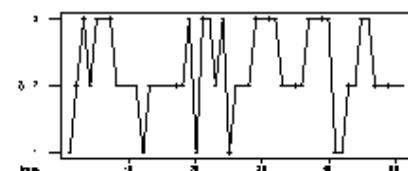
(C1)



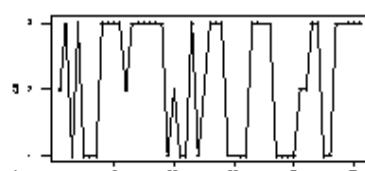
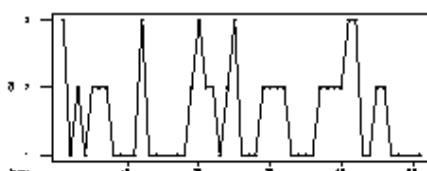
(C2)



(C3)

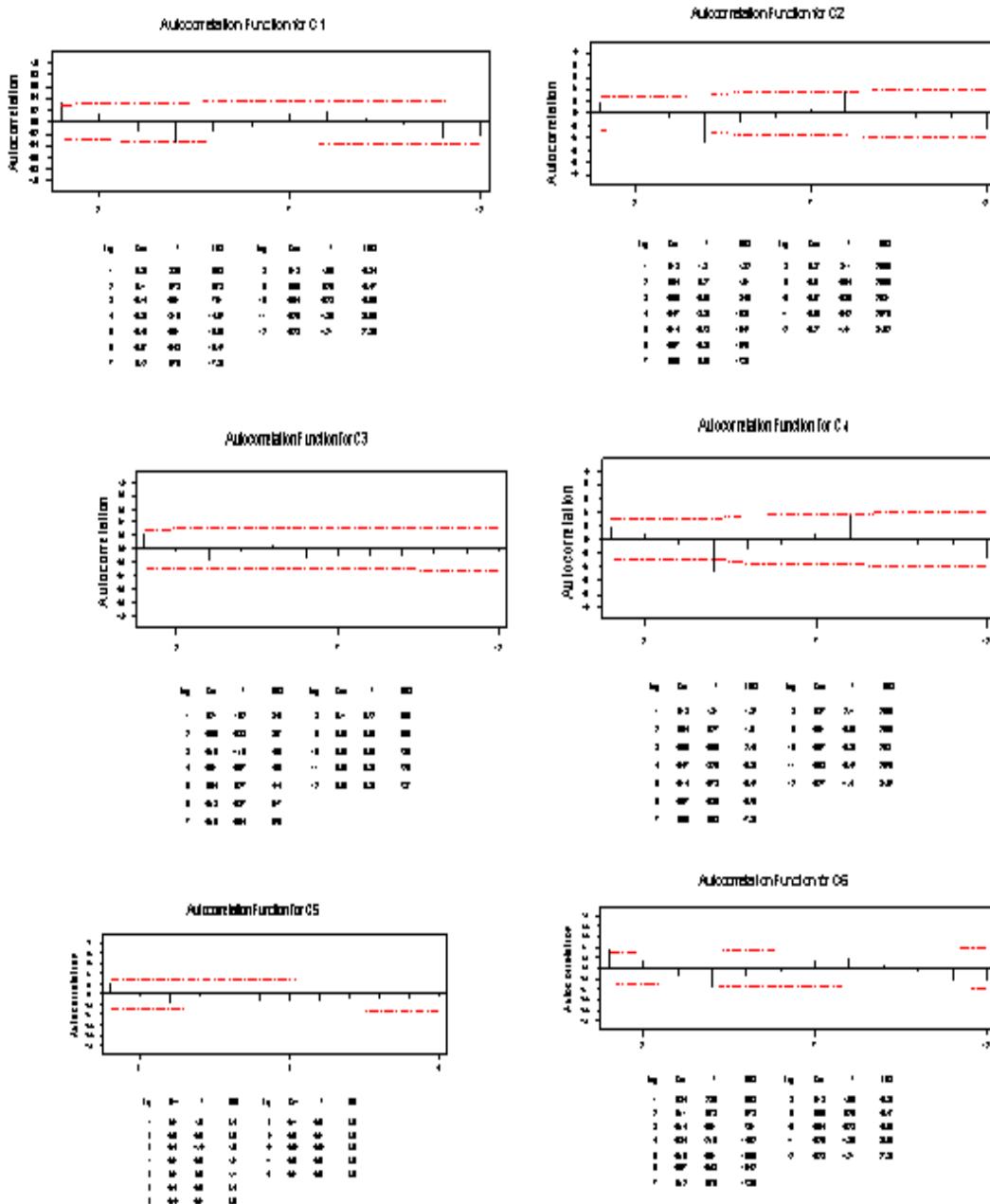


(C4)



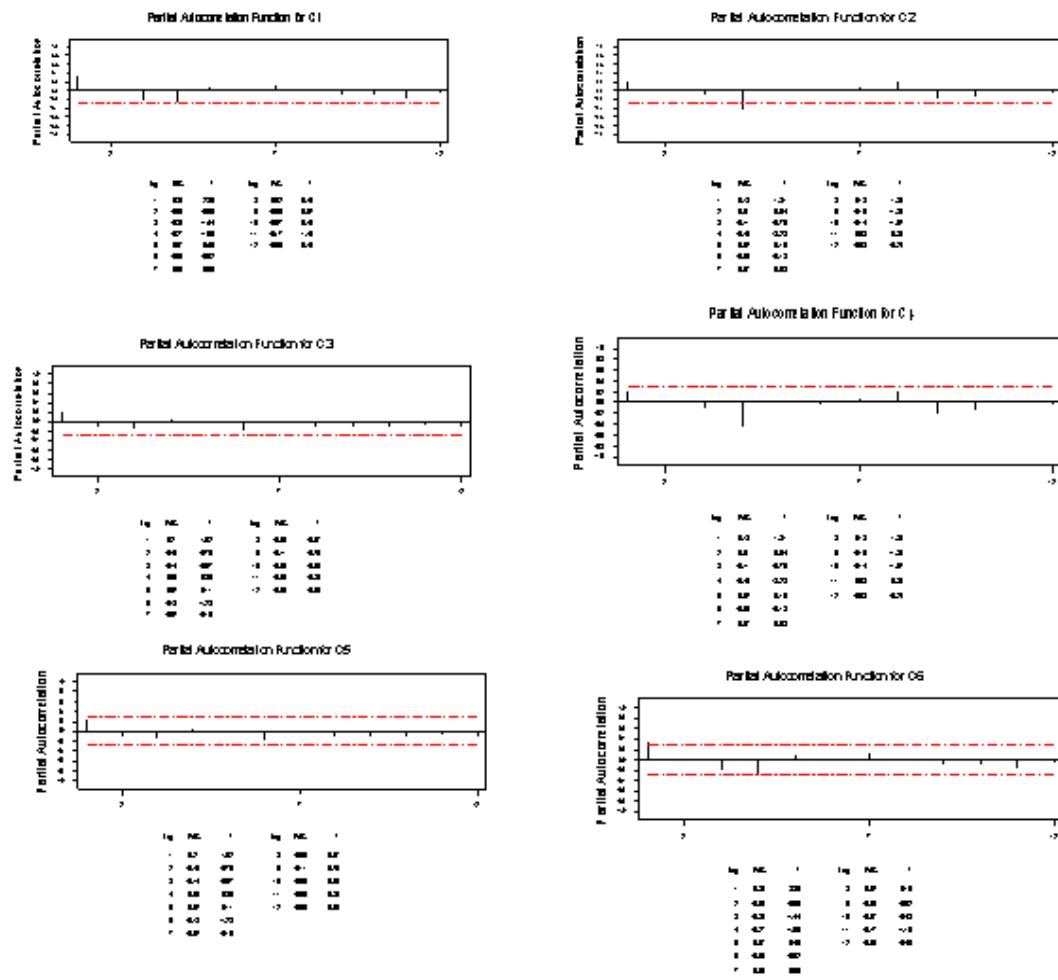
شكل (1)

يمثل رسم السلسلة للحالات الست المرمزه [Minitab 19]



(2) شكل

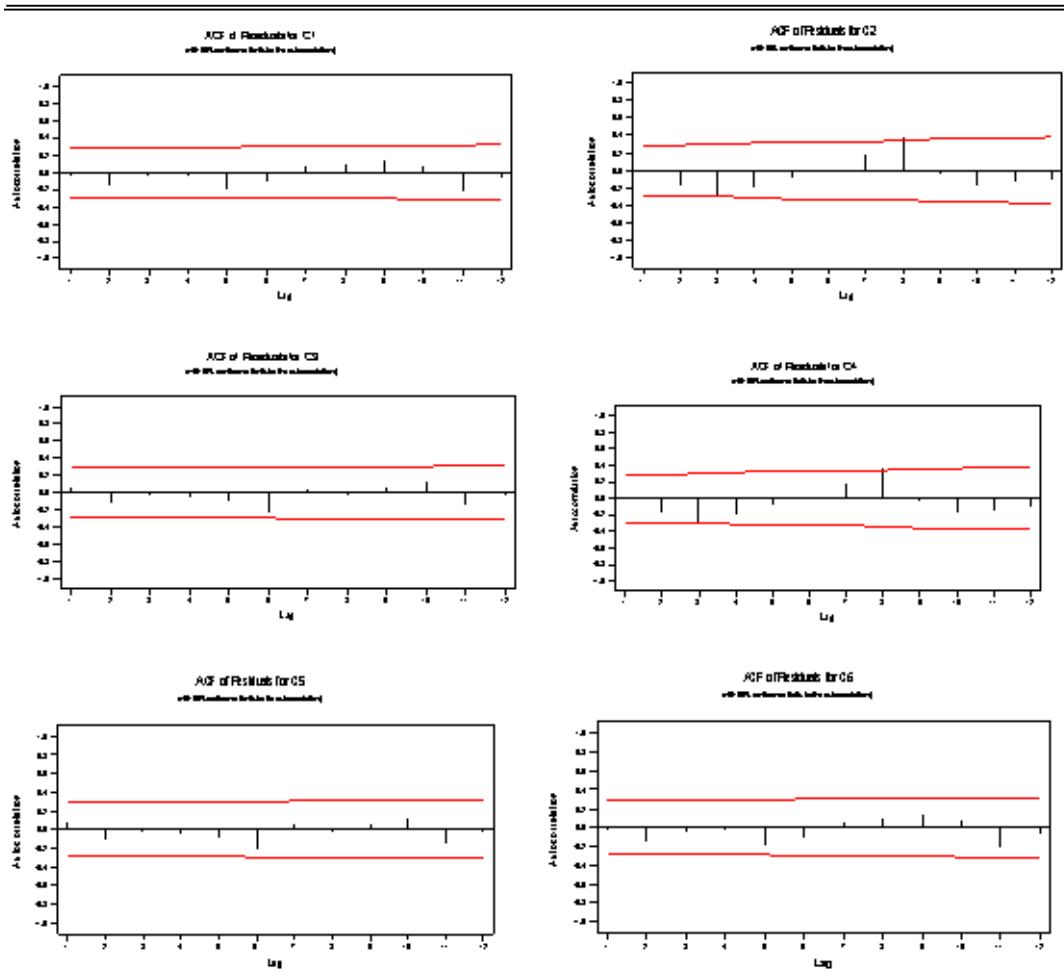
يمثل دالة الارتباط الذاتي (ACF) للسلسلة [Minitab 19]



(3) شكل

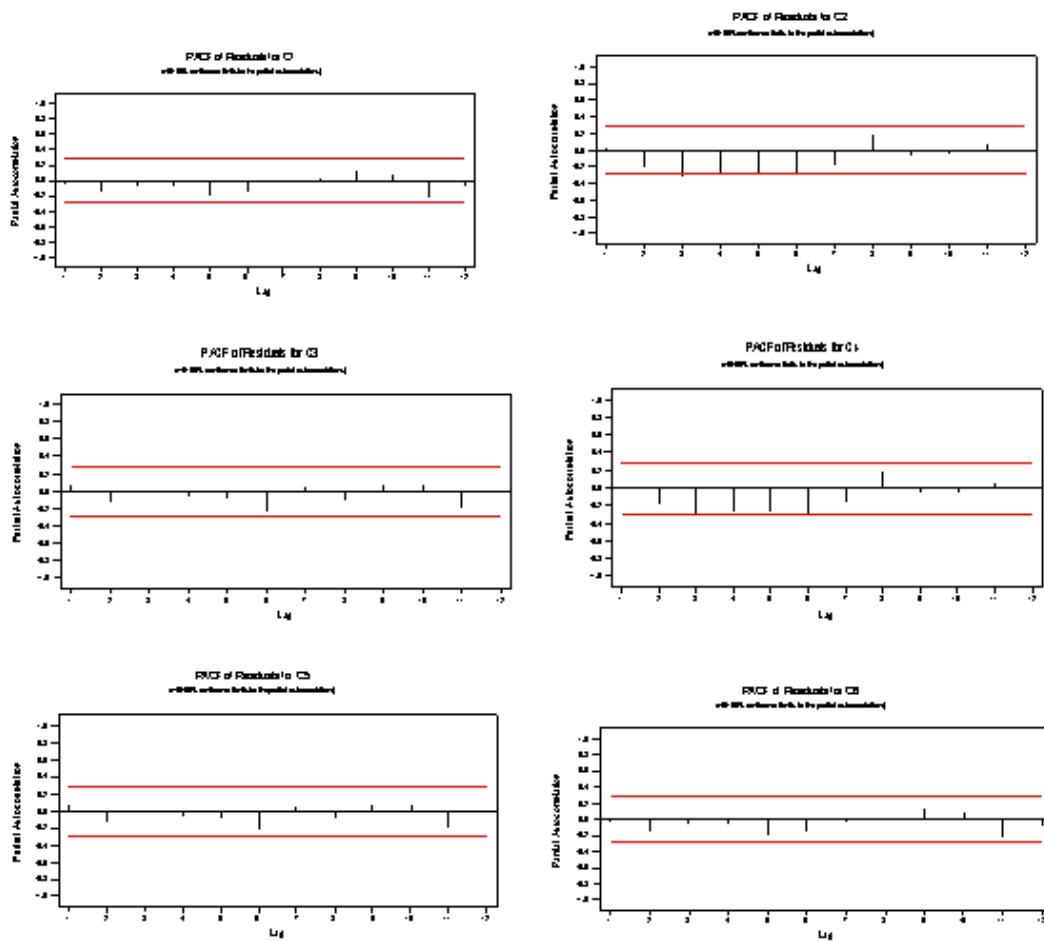
يمثل دالة الارتباط الذاتي الجزئي (PACF) [Minitab 19] للسلسلة

استخدام طيفي بوكس - جنكير و التقيية المكيفة لتحديد افضل نموذج للتبؤ بتلوث مياه نهر دجلة باستخدام السلسل الرمزية
التفويم.....



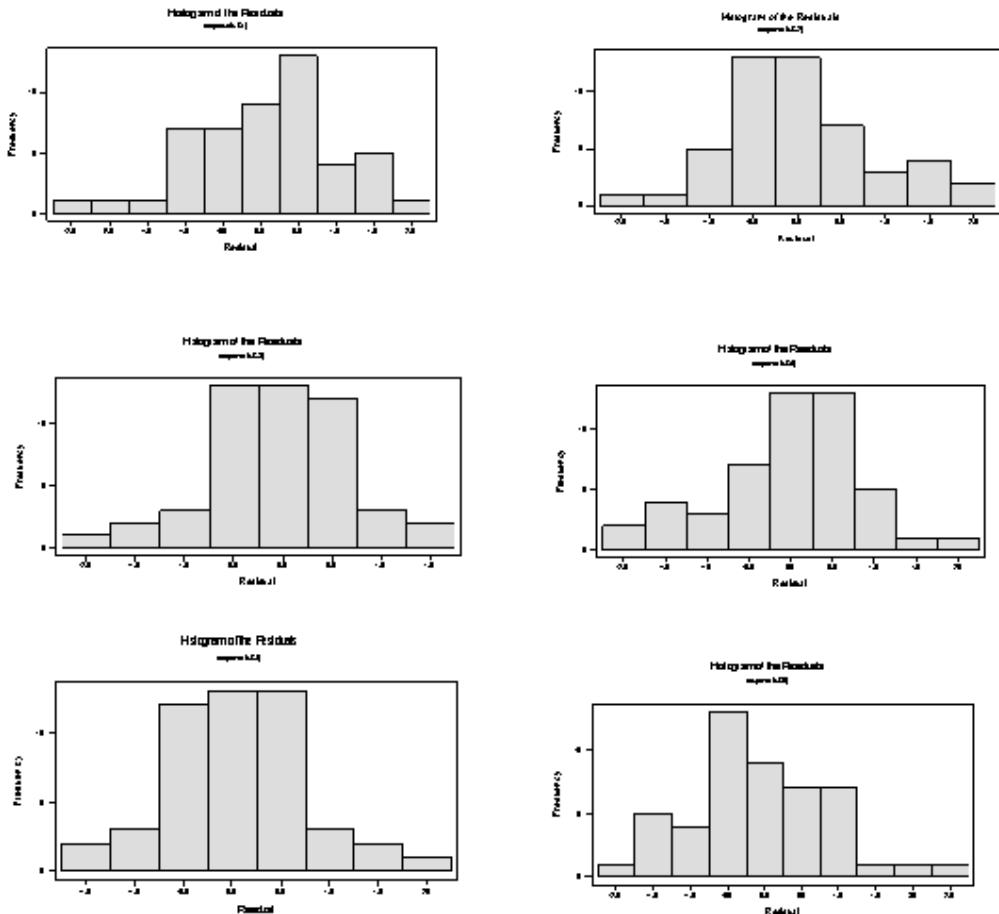
شكل (4)

يمثل دالة الارتباط الذاتي لاختبار الأنماذج [Minitab 19]



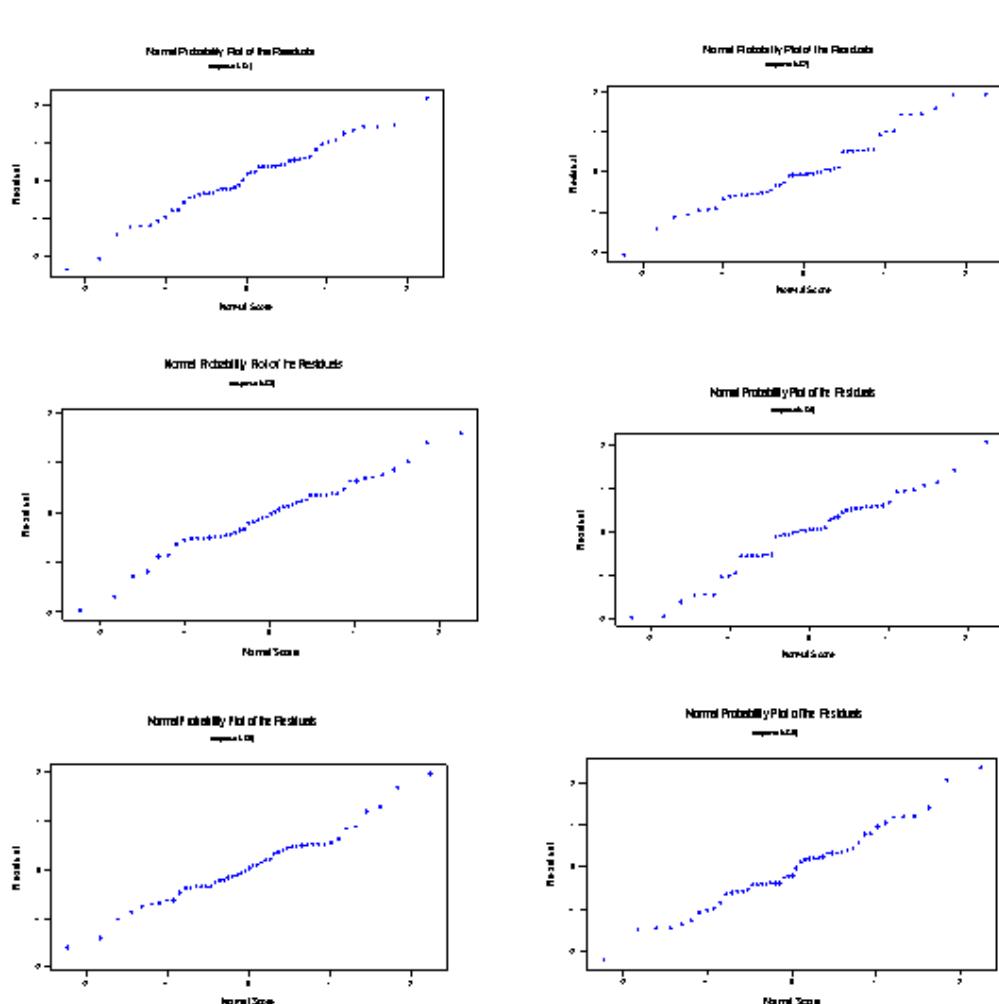
شكل رقم (5)

يمثل دالة الارتباط الذاتي الجزئي لاختبار الأنماذج



شكل (6)

يمثل توزيع الباقي لاختبار النموذج [Minitab 19]



شكل (7) يمثل احتمال انتشار الباقي لاختبار الانموذج [Minitab 19]