https://doi.org/10.31272/jae.i151.1549
https://admics.uomustansiriyah.edu.iq
P-ISSN: 1813–6729 E-ISSN: 2707-1359

تقدير انموذج ARMAX بأستعمال طريقتي التقدير اللامعلميتين الانحدار الخطي الموضعي والتقدير المويجي

امير وليد صبري شرار
قسم الاشراف الهندسي، دائرة الاعمار الهندسي، وزارة الاعمار والإسكان والبلديات العامة، بغداد، العراق.
Email: ameerwaleed@uomustansiriyah.edu.iq, ORCID: https://orcid.org/0009-0007-1956-8455
احمد شاكر محمد طاهر المتولي
قسم الإحصاء، كلية الإدارة والاقتصاد، الجامعة المستنصرية، بغداد، العراق.
Email: ahmutwali@uomustansiriyah.edu.iq, ORCID: https://orcid.org/0000-0002-5679-0940

المستخلص

نماذج الانحدار الذاتي-أوساط متحركة مع مدخلات خارجية Autoregressive Moving Average with Exogenous Variable ويرمز لها بالمختصر ARMAX تعد من النماذج الشائعة الاستعمال في وصف السلوك الديناميكي للسلاسل الزمنية المتأثرة بمتغيرات خارجية، اذ يمثل (AR) جزء الانحدار الذاتي والمتعلق بسلسلة المدخلات، و (MA) تمثل الأوساط المتحركة والمتعلقة بسلسلة البواقي العشوائية، اما (X) فيشير الى الجزء الخارجي والمتعلق بسلسلة المدخلات. وبهذا فأن هذه النماذج تصف المكونات المحددة العشوائية Stochastic Deterministic Component

تلك النماذج يمكن تقدير معاملاتها باستعمال طرائق التقدير المعلمية علاوة عن طرائق التقدير اللامعلمية، غير أن طرائق التقدير المعلمية تستند على افتراضات قوية لغرض التوصل الى تقدير كفوء لمعاملات الانموذج، في حين طرائق التقدير اللامعلمية لا تفترض اية افتراضات على عملية التقدير، لذا يفضلها بعض الباحثين على طرائق التقدير المعلمية، القوية ومن تلك الطرائق الانحدار الخطي الموضعي (LLR) والتقدير المويجي (WAVE). وتمت المقارنة بين تلك الطريقتين بأستعمال تجارب المحاكاة.

اذ افرزت نتائج المحاكاة افضلية واضحة لطريقة الانحدار الخطي الموضعي وذلك بالاعتماد على معيار المفاضلة الجذر التربيعي لمتوسط مربعات الاخطاء RMSE الذي اظهر أدنى قيم لهذه الطريقة مقارنة بالأخرى.

معلومات البحث

تواريخ البحث:
التقديم: 2026 / 01 / 01
المراجعة: 2026 / 02 / 25
قبول النشر: 28 / 02/ 2026
نشر الكتروني: 2026 / 03 / 01
تسلسل الصفحات: 125 - 141

الكلمات المفتاحية:
نماذج ARMAX، الانحدار الذاتي، الأوساط المتحركة، الانحدار الخطي الموضعي، التقدير المويجي.

المراسلة:
أسم الباحث: امير وليد صبري
Email: ameerwaleed@uomustansiriyah.edu.iq

1. المقدمة

نماذج الانحدار الذاتي أوساط متحركة تعد من النماذج الأساسية في تحليل السلاسل الزمنية والتنبؤ بقيمها المستقبلية. تقوم فكرة هذه النماذج على ان القيم الحالية للظاهرة تعتمد على القيم السابقة لها. من نمذجة العلاقة الخطية بين القيم الحالية والقيم السابقة لها ولفترات زمنية مرتدة يمكن التعرف على سلوك الظاهرة والتنبؤ بها. في بعض الحالات تكون الظاهرة قيد الدراسة ترتبط بعلاقة مع ظاهرة أخرى أي انها تتأثر بها ويمكن الاستفادة من هذه العلاقة في تحسين أداء نماذج الانحدار الذاتي أوساط متحركة من خلال إضافة تأثير القيم الحالية والسابقة للسلسلة الزمنية المرتبطة بالسلسلة الزمنية قيد التحليل الى الانموذج. وبهذا فان الانموذج الناتج من هذه العملية سوف يصف القيم الحالية للسلسلة الزمنية موضوع البحث بدلالة القيم السابقة لتلك السلسلة والقيم الحالية والقيم السابقة للسلسلة الزمنية التي ترتبط بها بعلاقة خطية، ويطلق على هذا الانموذج بأنموذج الانحدار الذاتي أوساط متحركة مع مدخلات خارجية Auto-regressive-Moving average with exogenous variable ويكتب اختصاراً (ARMAX). تم اقتراح هذه النماذج من قبل الباحثين Box و Jenkins عام 1970، اذ وضعا اسسه النظرية والافتراضات التي يستند اليها. [1]. غير ان استعمالها الواسع كان بعد عام 1980 بعد ان وضع الباحث Baillie في هذا العام الصيغة المتكاملة لهذا الأنموذج، [2]. واستمر الباحثون في استعمال هذا الانموذج في تحليل العلاقات الثنائية. اذ عمل قسما من الباحثين على تطوير طرائق تقدير تهدف الى الحصول على تقديرات كفؤة لمعاملات انموذج ARMAX، كالعمل المقدم من قبل الباحث Bercu, [3], و الباحث Zheng, [4], والباحث Diversi واخرون, [5]. فيما استعمل قسما اخر من الباحثين انموذج ARMAX في تحليل السلاسل الزمنية لظواهر طبيعية بوجود تأثيرات خارجية, كالبحث المقدم من قبل الباحث frausto واخرون,[6], وبحث احمد طاهر واخرون [7], اللذان عملا على تقدير كميات مياه الاهوار في جنوب العراق بوجود تأثيرات خارجية متمثلة بعدة عوامل بيئية.

في هذا البحث عملنا على استعمال بعض من طرائق التقدير اللامعلمية في تقدير معاملات انموذج ARMAX وبالاعتماد على تجارب المحاكاة تمت المقارنة بين نتائج تقدير تلك الطرائق للوصول الى طريقة التقدير المثلى.

كما هو معلوم ان نماذج الانحدار الذاتي – اوساط متحركة بمتغيرات خارجية تتضمن متغيرات مرتدة زمنيا سواء كانت للسلسلة الزمنية الاصلية (سلسلة المخرجات) أو لسلسلة المتغيرات الخارجية (سلسلة المدخلات)، الامر الذي يؤدي الى حدوث مشاكل قياسية

والتي تتعارض مع افتراضات طرائق التقدير المعلمية، من هنا تظهر مشكلة البحث في البحث عن طرائق تقدير بديلة عن الطرائق المعلمية وأفضل بديل هي طرائق التقدير اللامعلمية والتي لا تضع أية افتراضات على عملية التقدير.

2. هدف البحث

يهدف البحث الى استعمال طريقتي التقدير اللامعلمية، الانحدار الخطي الموضعي ((Local Linear Regression (LLR) والتقدير المويجي ((Wavelet Estimation (WAVE) في تقدير معاملات انموذج ARMAX في محاولة للمقارنة فيما تفرزه نتائج التقدير لكلا الطريقتين بأعتماد تجارب المحاكاة للتوصل الى أفضل طريقة تقدير.

3. الجانب النظري

1.3. انموذج الانحدار الذاتي اوساط المتحركة مع مدخلات خارجية

Autoregressive Moving Average with Exogenous Variable (ARMAX)

بأفتراض وجود سلسلة زمنية تمثل بيانات ظاهر معينة والتي نرغب في تحليل سلوكها في محاولة للتنبؤ بقيمها المستقبلية بأستعمال نماذج الانحدار الذاتي أوساط متحركة. وعلى فرض ان تلك الظاهرة تتأثر بالقيم الحالية والسابقة لظاهرة او عدة ظواهر أخرى في هذه الحالة يمكن تضمين تأثير تلك الظواهر في انموذج الانحدار الذاتي أوساط متحركة لتحقيق أداء هذا الانموذج والحصول على قيم تنبؤية دقيقة، الانموذج الناتج من هذه العملية يدعى بأنموذج الانحدار الذاتي أوساط متحركة مع مدخلات خارجية وصيغته الرياضية تكتب كالأتي، بأفتراض وجود متغير خارجي واحد، [8].

$$Y_t = \sum_{i=1}^{p} \emptyset_i Y_{t-i} - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \sum_{k=1}^{r} \gamma_k X_{t-k} + \varepsilon_t \qquad (1)$$

اذان: Y_t : $t = 1,2, \dots, n$: تمثل قيم الظاهرة قيد البحث والتي هي عبارة عن سلسلة زمنية مستقرة، X_t : $t = 1,2, \dots, n$ تمثل قيم الظاهرة التي تؤثر في الظاهرة قيد البحث والتي تمثل المتغير الخارجي وهي عبارة عن سلسلة زمنية مستقرة، ε_t : $t = 1,2, \dots, n$ تمثل سلسلة الأخطاء العشوائية والتي تفترض انها تتوزع توزيع طبيعي بمتوسط يساوي الصفر وتبايـــن يساوي σ^2، q, p : تمثل رتبة الانحدار الذاتي والاوساط المتحركة على التتابع لسلسلة الظاهرة قيد البحث (Y_t)، r: درجة الانحدار الذاتي لسلسلة المدخلات (X_t)، $(\gamma_k, \theta_j, \emptyset_i)$: تمثل معلمات انموذج الانحدار الذاتي، المتوسطات المتحركة والمدخلات الخارجية على التتابع.

يمكن إعادة كتابة الانموذج (1) بدلالة متعددات الحدود وكما يلي:

$$Y_t = \frac{\Theta_q(B)}{\Phi_p(B)} \varepsilon_t + \frac{\Gamma_r(B)}{\Phi_p(B)} X_t \qquad (2)$$

اذ ان:

$$\Phi_p(B) = 1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots - \emptyset_p B^p \qquad (3)$$

$$\Theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q \qquad (4)$$

$$\Gamma_r(B) = \gamma_1 B + \gamma_2 B^2 + \cdots + \gamma_r B^r \qquad (5)$$

أذ أن: $\Phi_p(B)$: يمثل متعدد الحدود لمعلمات انموذج الانحدار الذاتي من الرتبة (p)، $\Theta_q(B)$: يمثل متعدد الحدود لمعلمات انموذج المتوسط المتحرك من الرتبة (q)، $\Gamma_r(B)$: يمثل متعدد الحدود لمعلمات انموذج المتغير الخارجـــــــي من الرتبة $(\boldsymbol{r})$، B: يمثل عامل الازاحة الى الخلف (*Backshift Operator*).

ويمكن كتابة انموذج *ARMAX(p,q,r)* بصيغة أخرى بعد إضافة رتبة أخرى للأنموذج والتي تمثل التأخر الزمني (*Delay time*) في تأثير المدخلات الخارجية (X_t) على المخرجات (Y_t)، والتي يرمز لها بالرمز (b) ليصبح مختصر الانموذج *ARMAX(p,q,r,b)* وتكتب الصيغة كالأتي: [9]

$$Y_t = \frac{\Theta_q(B)}{\Phi_p(B)} \varepsilon_t + \frac{\Gamma_r(B)}{\Phi_p(B)} X_{t-b} \qquad (6)$$

2.3. تنقية انموذج *ARIMAX*

عملية التنقية لأي سلسلة زمنية هي عبارة عن عملية تحويل تلك السلسلة الزمنية الى أخرى عشوائية نقية والتي تمثل سلسلة الأخطاء العشوائية.

عملية التنقية لسلسلة المخرجات (Y_t) وسلسلة المدخلات X_t تهدف للحصول على ابسط صورة ممكنة لتلك السلسلتين، اذ تتم عملية التنقية بإزالة أي نمط معروف ناتج عن عمليتي الانحدار الذاتي او الأوساط المتحركة. ويتم ذلك كما مبين بالمعادلتين ادناه:[7]

$$\frac{\emptyset_p(B)(1-B)^d}{\theta_q(B)} X_t = \Psi_t \qquad (7)$$

$$\frac{\emptyset_p(B)(1-B)^d}{\theta_q(B)} Y_t = \varphi_t \quad (8)$$

اذ ان $\Psi_{k,t}$: تمثل سلسلة الأخطاء العشوائية للمدخلات، φ_t: تمثل سلسلة الأخطاء العشوائية للمخرجات، وان:
$\varphi_t \sim N\left(0, \sigma^2_{\varphi_t}\right)\ , \Psi_t \sim N\left(0, \sigma^2_{\Psi_t}\right)$

3.3. السلسلة الزمنية المستقرة Stationary Time Series

عند تحليل السلاسل الزمنية يجب التأكد اولاً من استقراريتها، لأن عدم الاستقرارية تؤدي الى الحصول على نتائج غير دقيقة. الاستقرارية في السلاسل الزمنية تعبر عن السلوك الاحتمالي لمجموعة المتغيرات العشوائية عبر الزمن، بحيث تبقى خصائصها الإحصائية (المتوسط، التباين والتباين المشترك) ثابتة ولا تتغير مع مرور الزمن. والسلسلة الزمنية تكون مستقرة إذا تحققت الشروط التالية: [10]

$E(y_t) = \mu$ ثبات المتوسط، (9)

$E(y_t - \mu)^2 = Var(y_t) = \sigma^2 = \hat{\gamma}_0$ قيمة التباين ثابتة، (10)

$\gamma(k) = cov(y_t, y_{t-k}) = E[(y_t - \mu)(y_{t-k} - \mu)] = \gamma(-k)$ دالة التباين المشترك الذاتي، (11)

في حال عدم تحقق الشرطين (9) و (10) فان السلسلة الزمنية تكون غير مستقرة في المتوسط والتباين وفي حال عدم تحقق أحدهما فأنها تكون غير مستقرة بأحدهما.

يتم اولاً التأكد من استقرارية السلسلة الزمنية في التباين وفي حالة كونها غير مستقرة في التباين يتم اعتماد احدى التحويلات المعروفة منها التحويل اللوغاريتمي او الاسي لغرض تحقيق الاستقرارية في التباين. بعد التحقق من استقراريتها في التباين يجب التأكد من تحقق استقراريتها في المتوسط وهناك عدة اختبارات يمكن اعتمادها للكشف عن استقراريتها في المتوسط من عدمه سيتم التطرق اليها لاحقاً. في حال كون السلسلة الزمنية غير مستقرة في المتوسط يتم تحويلها الى سلسلة زمنية مستقرة في المتوسط بعد اخذ عدد مناسب من الفروق لقيم السلسلة الزمنية، والصيغة العامة للفروق كالأتي:

$$\nabla^d Z_t = (1-B)^d Z_t \quad (12)$$

اذ ان Z_t: السلسلة الزمنية الاصلية، B: عامل الازاحة الخلفي، d: عدد الفروق المناسبة التي يتم اخذها للبيانات الاصلية، ∇: تمثل عامل الفرق.

4.3. دالة الارتباط الذاتي Autocorrelation Function – ACF

هي دالة تستخدم في تحليل السلاسل الزمنية لمعرفة مدى ترابط القيم الحالية للسلسلة بالقيم السابقة لها (أي علاقة كل قيمة بالقيم التي سبقتها بعدد معين من الفترات الزمنية)، وتحسب هذه الدالة بقسمة التغاير الذاتي عند الازاحة k على التغاير الذاتي عند الازاحة صفر، وحسب الصيغة التالية: [11]

$$\rho_k = \frac{Cov(Y_t\, , Y_{t+k})}{\sqrt{Var(Y_t)Var(Y_{t+k})}} = \frac{E(Y_t-\mu)(Y_{t+k}-\mu)}{\sigma_Y^2} = \frac{\gamma_k}{\gamma_0}\ , k = 1,2, \dots, K \quad (13)$$

ρ_k : معامل الارتباط الذاتي، k : عدد الازاحات المحسوب عندها قيمة الارتباط الذاتي وتقدر بحدود $\left(\frac{n}{4}\right)$.

تساعد دالة الارتباط الذاتي على فهم طبيعة الاعتماد الزمني للسلسلة الزمنية، مما يمكن من التعرف على استقرارية السلسلة في المتوسط، علاوة عن تحديد نوع الانموذج ودرجته. اذ ان السلسلة التي تتناقص اسياً نحو الصفر تشير عادةً الى انموذج من نوع *AR (p)*، بينما السلسلة التي تنقطع بعد الازاحة q تميل الى انموذج *MA (q)*، كما يمكن الحكم على معنوية قيم الارتباط من خلال مقارنتها بالحدود الإحصائية التقريبية $\left(\pm \frac{1.96}{\sqrt{n}}\right)$، حيث تعد القيم الواقعة خارج هذه الحدود ذات دلالة معنوية، في حين تعتبر القيم الواقعة داخل الفترة غير معنوية.

5.3. دالة الارتباط الذاتي الجزئي Partial Autocorrelation Function – PACF

تعد دالة الارتباط الذاتي الجزئي أداة إحصائية مهمة تستخدم لتحديد رتبة انموذج الانحدار الذاتي *AR* في السلاسل الزمنية. دالة الارتباط الذاتي الجزئي تظهر مقدار العلاقة بين القيم الحالية والمتأخرة للسلسلة بعد إزالة تأثير القيم الوسطية. بمعنى اخر، توضح *PACF* مدى ارتباط القيم الحالية بالماضية بشكل مباشر فقط، وليس من خلال فترات وسيطة. تساعد دالة الارتباط الذاتي الجزئي الى جانب دالة الارتباط الذاتي في تحديد نوع ودرجة الانموذج. اذ ان السلسلة التي تتناقص اسياً نحو الصفر تشير عادةً الى انموذج من نوع *MA(q)*، بينما السلسلة التي تنقطع بعد الازاحة p تميل الى انموذج *AR(p)*، وبالاعتماد على الحدود الإحصائية $\left(\pm \frac{1.96}{\sqrt{n}}\right)$ يمكن تحديد معنوية الارتباطات الذاتية الجزئية من عدمها اذ ان القيم التي تقع داخل هذه الحدود تعتبر ذات دلالة غير معنوية في حين التي تقع خارج تلك الحدود تعتبر ذات دلالة معنوية. تحتسب دالة الارتباط الذاتي الجزئي وفق الصيغة الاتية: [12]

$$\emptyset_{kk}=\frac{\begin{vmatrix}1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_1\\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_2\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ \rho_{k-1} & \rho_{k-2} & . & \dots & \rho_1 & \rho_k\end{vmatrix}}{\begin{vmatrix}1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_{k-1}\\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_{k-2}\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ \rho_{k-1} & \rho_{k-2} & . & \dots & \rho_1 & 1\end{vmatrix}} \tag{14}$$

6.3. دالة الارتباط المتقاطع ***Cross Correlation Function-CCF***

دالة الارتباط المتقاطع يمكن استعمالها للكشف عن درجة الارتباط بين أي سلسلتين زمنيتين، وتعد أداة أساسية في تحديد الارتداد الزمني لتأثير متغيرات الزمن لأنموذج (*ARIMAX*)، وبأفتراض ان $(t, k = 0,\ \pm1,\ \pm2, \pm3, \dots)$، فأن الصيغة الرياضية لدالة الارتباط المتقاطع بين سلسلتي المدخلات (Y_t) والمخرجات (X_t) وعند الازاحة (k) تكون بالصيغة الاتية: [9]

$$\rho_{x_t y}(k)=\frac{\sigma_{x_t y}(k)}{\sigma_{x_t}\sigma_y}=\frac{E\left(X_t-\mu_{x_t}\right)\left(Y_{t+k}-\mu_y\right)}{\sqrt{E\left(X_t-\mu_{x_t}\right)^2}\sqrt{E\left(Y_{t+k}-\mu_y\right)^2}} \tag{15}$$

أذ أن: $\sigma_{x_t y}(k)$: تمثل التباين المشترك المتقاطع بين سلسلة المخرجات (Y_t) وسلسلة المدخلات (X_t) عند إزاحة مقدارها (k)، σ_y: الانحراف المعياري لسلسلة المخرجات (Y_t) ، σ_{x_t}: الانحراف المعياري لسلسلة المدخلات (X_t)، μ_y: متوسط سلسلة المخرجات (Y_t) ، μ_{x_t}: متوسط سلسلة المدخلات (X_t).

7.3. اختبار جذر الوحدة ***Unit Root Test***

تعد المعادلة المميزة لأية عملية عشوائية أداة لفحص فيما إذا كانت تلك العملية العشوائية مستقرة ام لا. اذ تتصف العملية العشوائية بأنها غير مستقرة إذا كان الواحد الصحيح يمثل جذراً للمعادلة المميزة لها، وفي حال كون بقية جذور المعادلة المميزة ذات قيمة مطلقة اقل من الواحد الصحيح فان الفرق الأول لتلك العملية العشوائية يكون مستقراً، بمعنى اخر إذا كانت بقية جذور المعادلة المميزة تقع خارج حدود الدائرة الأحادية، ويمكن توضيح اختبار جذر الوحدة بأفتراض انموذج الانحدار الذاتي *AR(1)* وفق الصيغة الآتية:[13]

$$Y_t=\emptyset_1 Y_{t-1}+\varepsilon_t \tag{16}$$

أذ أن: $\varepsilon_t \sim N(0,\sigma^2)$

اذا كانت $|\emptyset_1| < 1$: تكون السلسلة الزمنية مستقرة، واذا كانت $|\emptyset_1| > 1$: تكون السلسلة الزمنية غير مستقرة

واذا كانت $\emptyset_1 = 1$ (جذر وحدة): تكون السلسلة الزمنية غير مستقرة وتسلك عملية السير العشوائي (*Random Walk*).

هناك عدة اختبارات إحصائية يمكن اعتمادها لأختبار فيما إذا كانت العملية العشوائية لها جذر وحدة، وسنتطرق الى اهم تلك الاختبارات وهو اختبار ديكي – فلر الموسع.

1.7.3. أختبار ديكي-فولر الموسع ***Augmented Dicey – Fuller***

نظراً لكون استقرارية السلسلة الزمنية شرطاً اساسياً لضمان كفاءة التقدير اللامعلمي لأنموذج *ARMAX* قيد البحث وتجنب مشكلة الانحدار الزائف، فقد تم اعتماد اختبار ديكي – فولر الموسع كأداة تشخيصية أولية اذ هو الاختبار الذي يعد تطوير لأختبار ديكي – فلر والذي تم تطويره من قبل الباحث ديكي – فلر عام (1979) لذا تم تسميته بأختبار ديكي – فلر الموسع ويرمز له اختصاراً (*ADF*)، [14].

يستند اختبار ديكي – فلر على افتراض ان السلسلة الزمنية يمكن ان تمثل بموجب انموذج الانحدار الذاتي من الدرجة الأولى *AR (1)* وكما في الصيغة (16)، وان الفرضية المراد اختبارها تكتب كالأتي:

$H_0: \emptyset_1 = 1$ (وجود جذر وحدة اي غير مستقرة)

$H_1: |\emptyset_1| < 0$ (عدم وجود جذر وحدة اي مستقرة)

اما احصاءة الاختبار فتكون وفق الصيغة الاتية:

$$DFt=\frac{\widehat{\emptyset}_1-1}{se(\widehat{\emptyset}_1)} \tag{17}$$

اذ ان $\widehat{\emptyset}_1$: يمثل تقدير معامل الانحدار $\emptyset_1$ ، $se(\widehat{\emptyset}_1)$: تمثل الخطأ المعياري.

يعتمد اختار ديكي – فلر على فحص وجود جذر الوحدة في السلسلة الزمنية فأذا وقعت الجذور خارج دائرة الوحدة تكون السلسلة الزمنية مستقرة، وكذلك أعد الباحث ديكي – فلر جدولا خاصا بالقيم الجدولية لأحصاءة الاختبار المبينة بالصيغة (17)، أذ أن تلك الاحصاءة لا تتبع توزيع t الاحتمالي. يتم قبول فرضيه العدم إذا كانت قيمه احصاءة الاختبار DF أكبر من القيمة الجدولية مما يشير الى عدم استقراريه السلسلة الزمنية أو أنها متكاملة من الدرجة الاولى بمعنى ان السلسلة الزمنية تتبع انموذج السير العشوائي وبعكسه فأن السلسلة الزمنية تتبع انموذج الانحدار الذاتي المستقر من الدرجة الاولى AR (1)،[13].

لتضمين الاتجاه العام في هذا الاختبار، انموذج الانحدار الخطي البسيط بمتغير توضيحي يمثل الزمن t كما في الصيغة الآتية:

$$y_t = \alpha_0 + \alpha_1 t + u_t \tag{18}$$

اذ أن: الاخطاء العشوائية لهذا الانموذج عبارة عن سلسلة زمنية تتبع انموذج الانحدار الذاتي من الدرجة الاولى وكالآتي:

$$u_t = \emptyset_1 u_{t-1} + \varepsilon_t \tag{19}$$

اذ ان: ε_t: تمثل سلسلة زمنية مستقرة بمتوسط يساوي الصفر وتباين ثابت يساوي σ^2، أي أن, $Var(\varepsilon_t) = \sigma^2, E[\varepsilon_t] = 0$
ومن خلال الصيغة رقم (18) نحصل على:

$$u_t = y_t - \alpha_0 - \alpha_1 t \tag{20}$$

$$u_{t-1} = y_{t-1} - \alpha_0 - \alpha_1 (t-1) \tag{21}$$

الانموذج المختزل للسلسلة الزمنية y_t نحصل عليه بتعويض المعادلتين (20) و (21) في المعادلة (19) وكالآتي:

$$y_t - \alpha_0 - \alpha_1 t = \emptyset_1 \big(y_{t-1} - \alpha_0 - \alpha_1 (t-1)\big) + \varepsilon_t \tag{22}$$

وبأجراء العمليات الحسابية لتبسيط المعادلة (22) نحصل على الصيغة النهائية للأنموذج المختزل للسلسلة الزمنية y_t والتي تكون بالصيغة الاتية:

$$y_t = \gamma + \delta t + \emptyset_1 y_{t-1} + \varepsilon_t \tag{23}$$

اذ ان:

$$\gamma = \beta_0 (1 - \emptyset_1) + \beta_1 \emptyset_1 \tag{24}$$

$$\delta = \beta_1 (1 - \emptyset_1) \tag{25}$$

الصيغة النهائية للأنموذج المختزل للسلسلة الزمنية y_t يكون لها جذر وحده أذا كانت $(\emptyset_1 = 1)$ ويتم اختبار وجود جذر الوحدة باعتماد نفس احصاءة الاختبار المبينة في الصيغة (17)

اختبار ديكي – فلر يستند على افتراض ان حدود الخطأ العشوائي لأنموذج الانحدار الذاتي تكون نقيه (White noise) اي غير مرتبط ذاتياً، لذا طور العالم ديكي – فلر اختباره ليأخذ بنظر الاعتبار الارتباط المتسلسل من خلال تضمين عدد كافٍ من الفروق المتأخرة، والاختبار المطور سمي باختبار ديكي – فلر الموسع والذي يستند على تقدير الانموذج الاتي:

$$y_t = \gamma + \delta t + \emptyset_1 \, y_{t-1} + \sum_{j=1}^{p} \theta_i \Delta y_{t-j} + \varepsilon_t \tag{26}$$

بعد تقدير انموذج الانحدار بالعلاقة (26) يتم اختبار جذر الوحدة باعتماد نفس احصاءة اختبار ديكي – فلر والمبينة بالصيغة(17).

8.3. مراحل بناء انموذج ARMAX Model Building Stages ARMAX

تهدف عملية بناء أنموذج *ARMAX(p,q,r)* الى تحديد درجة الانموذج المثلى واجراء عملية التقدير لمعاملات هذا الانموذج ومن ثم التحقق من دقة التقدير وملائمته للبيانات تحت البحث لغرض اعتماده في عملية التقدير والتنبؤ. عملية بناء الانموذج تتضمن المراحل الآتية، [9]:

1.8.3. المرحلة الأولى: التشخيص والتنقية

1. تحقيق استقرارية سلسلتي المخرجات (Y_t) والمدخلات (X_t)، اذا كانت السلسلة الزمنية غير مستقرة في المتوسط يتم تحويلها الى سلسلة زمنية مستقرة من خلال اخذ الفروق المناسبة وبعدد (d) من الفروق، اما اذا كانت غير مستقرة في التباين فيتم تحويلها الى سلسلة مستقرة من خلال استعمال احدى التحويلات المعروفة منها التحويل اللوغاريتمي او التحويل الاسي.
2. فلترة سلسلتي المخرجات (Y_t) والمدخلات (X_t)، كما هو مبين في المعادلين رقم (*7*) و (*8*). وذلك لأن المدخلات الخارجية هي في حد ذاتها سلسلة زمنية، وكذلك في بعض الأحيان دالة الارتباط الذاتي بين سلسلتي المخرجات والمدخلات تظهر ارتباطاً غير حقيقي، هذا يؤدي الى علاقة زائفة ومبهمة التفسير بين سلسلتي المخرجات والمدخلات، ولحل هذه المشكلة تتم عملية التنقية او الفلترة.
3. تقدير دالة الارتباط المتقاطع للسلاسل الزمنية التي تم تنقيتها او فلترتها ولفترات زمنية متعددة $(t, t+k)$، وصيغته الرياضية تكتب كالاتي:

$$\rho_{\Psi_t,\varphi_t}(k) = \frac{\sigma_{\Psi_t,\varphi_t}(k)}{\sigma_{\Psi_t}\sigma_{\varphi_t}} \tag{27}$$

أذ أن: $\sigma_{\Psi_t,\varphi_t}(k)$: تمثل التباين المشترك المتقاطع بين سلسلة المخرجات (Y_t) وسلسلة المدخلات (X_t) عند إزاحة مقدارها (k)، σ_{Ψ_t}: الانحراف المعياري لسلسلة المخرجات (Y_t) التي تم تنقيتها، σ_{x_t}: الانحراف المعياري لسلسلة المدخلات (X_t) التي تم تنقيتها.

2.8.3. المرحلة الثانية: تحديد رتبة الانموذج Model Order Selection

1. تحديد رتبتي الانحدار الذاتي (p) والمتوسطات المتحركة (q)، وذلك عن طريق دالة الارتباط الذاتي *ACF* والارتباط الذاتي الجزئي *PACF*، والتي تم التطرق اليها سابقاً.
2. تحديد رتبة المدخلات الخارجية (r)، وتتم عملية تحديد الرتبة بشكل رئيسي من خلال تحليل دالة الارتباط المتقاطع *CCF* بين السلاسل الزمنية وبعد تطبيق عملية التنقية. وذلك من خلال تحديد عدد التأخيرات الزمنية المتتالية لدالة الارتباط المتقاطع *CCF* والتي تكون ذات دلالة معنوية بعد اول تأخير زمني (b).
3. تحديد رتبة التأخر الزمني (b) (*Delay time*)، ويتم تحديد الرتبة بالاعتماد على دالة الارتباط المتقاطع، اذ ان اول ارتباط متقاطع ذات دلالة إحصائية يمثل قيمة (b)، ومن الجدير بالذكر اذا كان التأخير الزمني $(b = 0)$ معنى ذلك لا يوجد تأخير زمني صافي، أي ان تأثير المدخلات الخارجية (X_t) على المخرجات (Y_t) يكون في نفس اللحظة. اما اذا كان التأخير الزمني $(b \geq 1)$، يدل ذلك على وجود تأخير زمني صافي، أي ان تأثير المدخلات الخارجية (X_t) على المخرجات (Y_t) تكون مؤجلة أي ان التأخير لا يظهر في نفس اللحظة وانما يظهر بعدة فترة زمنية.

لكون عملية تحديد رتبة انموذج ARMAX(p,q,r) تشوبها بعض التعقيدات في اختيار درجتي الانحدار الذاتي والاوساط المتحركة وكذلك رتبة التأخير الزمني لتأثير المتغير الخارجي، يمكن اختيار عدة رتب مختلفة للأنموذج وبالاعتماد على أحد معايير المفاضلة لتحديد الانموذج الامثل من بين النماذج المقترحة، ومن هذه المعايير معيار اكياكي للمعلومات *Akaike's* (*AIC*) *Information Criteria*، وصيغته الرياضية تكتب كالاتي: [13]

$$AIC = nln(\hat{\sigma}_{\alpha}^2) + 2m \tag{28}$$

أذ أن: m : تمثل عدد المعلمات في الانموذج، n: تمثل عدد المشاهدات، $\hat{\sigma}_{\alpha}^2$: مقدار تباين البواقي، وتمثل الدرجة الأفضل لهذا المعيار التي تقابل اقل قيمة.

وكذلك معيار شوارز *Schwarz Criterion* (*SC*)، وصيغته الرياضية تكتب كالاتي:

$$BIC(m) = n\,ln(\hat{\sigma}_{\alpha}^2) + m\,ln(n) \tag{29}$$

ولتصحيح التحيز في معيار (*AIC*) حيث اقترح الباحثان (*Brock Well and Davis 1993*) بأضافة $(2mn/(n-m-1))$، عندما تكون (*m*) كبيرة نسبةً الى حجم العينة (*n*) يستعمل معيار اكيكي (AIC_C) وصيغته كالأتي :

$$AIC_C = nln(\hat{\sigma}_{\alpha}^2) + \frac{2mn}{(n-m-1)} \tag{30}$$

3.8.3. المرحلة الثالثة: التقدير Estimation

المرحلة الثالثة من عملية بناء الانموذج هي التقدير، وتوجد عدة طرائق لتقدير معلمات الانموذج المختار منها الطرائق المعلمية كطريقة المربعات الصغرى (OLS)، وطريقة الإمكان الأعظم (ML)، وطريقة العزوم، ومنها الطرائق اللامعلمية والتي سيتم اعتماد طريقتين في بحثنا وهي طريقة الانحدار الخطي الموضعي (Local Linear Regression (LLR)) والتقدير المويجي (Wavelet Estimation (WAVE)) والتي سيتم التطرق اليها لاحقاً.

4.8.3. المرحلة الرابعة: دقة التشخيص Diagnostic Checking

المرحلة الرابعة تمثل دقة تشخيص الانموذج المقترح لتمثيل بيانات السلسلة الزمنية، ويتم ذلك من خلال فحص سلسلة البواقي للأنموذج المقدر ε_t والتي يجب أن تحقق ما يأتي: [15]

1. ان تكون سلسلة زمنية عشوائية تتبع التوزيع الطبيعي وبارتباطات ذاتية غير معنوية، أي انها تمثل متغيرات عشوائية غير مترابطة، ويتحقق ذلك عندما تقع تلك الارتباطات داخل حدود الثقة المبينة بالصيغة التالية ولمستوى معنوية معين:

$$-1.96\frac{1}{\sqrt{n}} \leq \hat{\rho}_k(\varepsilon) \leq 1.96\frac{1}{\sqrt{n}} \tag{31}$$

ويمكن الاعتماد على احصاءة Box and Pierce(Q) لأختبار معنوية الارتباطات الذاتية لسلسلة البواقي المقدرة أي اختبار الفرضية الآتية: [16]

$$H_0: \rho_1 = \rho_2 = \cdots = \rho_h = 0$$
$$H_1: \rho_1 \neq \rho_2 \neq \cdots \neq \rho_h \neq 0$$

واحصاء الاختبار تكون وفق الصيغة الآتية:

$$Q = n\sum_{k=1}^{h} \rho_k^2(\hat{e}) \sim \chi^2_{(h-p-q-r)} \tag{32}$$

اذ ان $n = N - d$ تمثل حجم العينة المستخدم في تشخيص الانموذج، وأن N تمثل عدد المشاهدات الحقيقية للسلسلة الزمنية، d يمثل عدد الفروق التي تم اخذها لتحقيق استقرارية السلسلة الزمنية ، h تساوي $\sqrt{n}$. وتتم عملية المقارنة من خلال مقارنة Q المحسوبة مع (χ^2) الجدولية وبدرجة حرية $(h-p-q-r)$، فأذا كانت القيمة الجدولية اكبر من القيمة Q المحسوبة تقبل فرضية العدم، أي ان الأخطاء العشوائية غير مترابطة هذا يدل على ان الانموذج جيد وملائم. اقتراح الباحثان (Box and Pierce(Q)) اجراء تعديل على صيغة احصاء الاختبار Q المبينة بالعلاقة (32) والصيغة المعدلة لصيغته تكتب كالأتي:

$$Q = n(n+2)\sum_{k=1}^{h}(n-k)^{-1}\,\rho_k^2(\hat{\varepsilon}) \sim \chi^2_{(h)} \tag{33}$$

والصيغة (32) تكون مفضلة في الاستخدام لأن مستوى المعنوية لها تكون قريبة من القيم المتوقعة.

2. ان تكون مستقلة عن سلسلة بواقي سلسلة المدخلات المفلترة Ψ_t ويتم التحقق من ذلك من خلال فحص معاملات دالة الارتباط المتقاطع بين سلسلتي البواقي ε_t و Ψ_t .

9.3. التقدير اللامعلمي لأنموذج الانحدار الذاتي والمتوسطات المتحركة مع مدخلات خارجية

Nonparametric Estimation of the ARMAX Model

طرائق التقدير المعلمية تتطلب تحقق فروض تتعلق بمتغيرات الانموذج المطلوب تقديره. غير ان هذه الفروض قد تكون غير متحققة من الناحية التطبيقية مما ينعكس ذلك على دقة التقديرات، علاوة عن ان العلاقة الدالية للأنموذج قيد البحث قد تكون غير محددة اذ قد تكون خطية او لا خطية. بناءً على ما تقدم يتم البحث عن بدائل للطرائق المعلمية وأفضل بديل هي طرائق التقدير اللامعلمية، التي لا تستند على أي فروض قوية خاصة بمتغيرات الانموذج ولا على تحديد نوع العلاقة الدالية للأنموذج،

لغرض تقدير صيغة انموذج ARMAX(p,q,r) المبنية بالمعادلة رقم (6) يتم إعادة صيغة الانموذج وفق أسلوب التقدير اللامعلمي وكالآتي:

$$y_t = f(z_t) + \varepsilon_t \tag{34}$$

اذ ان: $z_t = \left(y_{t-1}, y_{t-2}, \dots, y_{t-p}, x_t, x_{t-1}, \dots, x_{t-\mathrm{r}}, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}\right)$: تمثل متغيرات (الانحدار الذاتي - المتوسطات المتحركة مع مدخلات خارجية)، $y_{t-1}, y_{t-2}, \dots, y_{t-p}$: تمثل المتغيرات الذاتية المتأخرة (AR)، $x_t, x_{t-1}, \dots, x_{t-\mathrm{r}}$: تمثل المتغيرات التوضيحية الخارجية المتأخرة، $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-n_c}$: تمثل المتوسطات المتحركة (MA)، $\varepsilon_t \sim N(0, \sigma^2)$: هو الخطأ العشوائي المستقل والموزع طبيعياً.

الدالة (34) تمت صياغة المتغيرات المترددة زمنياً وبشكل دالة غير معلومة, اذ ان الدالة $f(.)$ هي دالة غير محددة الصيغة ويتم تحديد صيغتها من خلال البيانات والتي يتم تقديرها بالطرائق اللامعلمية . ويمكن ان نستخدم الأساليب العددية لكي نحصل على تقديرات جيدة. ولكن الانموذج (34) من الصعب تطبيقه من خلال الأساليب الجبرية بسبب وجود مشكلة زيادة الابعاد (Curse of dimensionality)، ولمعالجة هذه المشكلة نلجأ الى خاصية التجميع (Additive) هذه الخاصية تجعل الانموذج سهل ومرن بشكل جيد، ويمكن كتابة الانموذج بالصيغة التجميعية الآتية، [17]:

$$\begin{aligned} y_t = f_1(y_{t-1}) + f_2(y_{t-2}) + \dots + f_p\left(y_{t-p}\right) + g_1(x_t) + g_2(x_{t-1}) \dots + g_r(x_{t-r}) \\ + h_1(\varepsilon_{t-1}) + \dots + h_q\left(\varepsilon_{t-q}\right) + \varepsilon_t \end{aligned} \tag{35}$$

أذ أن: $f_i(.)'s$: هي دوال ذات قيمة حقيقية قابلة للقياس للمتغيرات الذاتية المتأخرة (AR)، $g_j(.)'s$: هي دوال ذات قيمة حقيقية قابلة للقياس للمتغيرات الخارجية، $h_r(.)'s$: دوال ذات قيمة حقيقية قابلة للقياس للمتوسطات المتحركة.

ويسمى الانموذج (35) بأنموذج الانحدار الذاتي والمتوسطات المتحركة التجميعي بوجود متغير خارجي، يعمل هذا الانموذج من خلال اخذ كل دالة متغير واحد فقط، وفي هذه الحالة يمكن اعتماد أساليب طرائق التقدير اللامعلمي.

1.9.3. التقدير المويجي *Wavelet Estimation (WAVE)*

ننظر في العملية الأساسية المتمثلة في تقدير دالة الانحدار f في الصيغة (34) من المشاهدات $\{(x_t, y_t, \varepsilon_t)\,; t = 1, \dots, n\}$ ، حيث $\{x_t \in [0,1]^p, y_t \in \mathbb{R}\,, \varepsilon_t \sim N(0, \sigma^2)\}$ادبيات المويجات أحادية المتغير واسعة جداً ولا يمكن تغطيتها بشكل كامل، بينما الادبيات المتعلقة بطرق المويجات في حالة متغيرات التفسير المتعددة (Multivariate) محدودة نسبياً، خصوصاً عند ارتفاع عدد المتغيرات. في حالة متعددة المتغيرات، حيث $p \geq 2$ و $z_t \in [0,1]^p$. نعتبر تقدير النماذج التجميعية كما في الصيغة (35). تعد النماذج التجميعية امتداداً طبيعياً للنماذج الخطية بهدف التقاط العلاقات الشرطية غير الخطية. تعرف المويجة $\Psi(t)$ بأنها دالة أساسية ذات دعم مدمج (Compact Support) وعزوم متلاشية (Vanishing Moments) تستخدم لتوليد نظام متعامد (Orthonormal System) في الفضاء $L^2(\mathbb{R})$ عبر عمليتي التمدد (Dilation) والانتقال (Translation) ترتبط المويجة بدالة أخرى تسمى دالة التحجيم (Scaling Function) او المويجة الاب (Father Wavelet) وترمز ب $\emptyset(t)$، وتستخدمان معاً في بناء تحليل متعدد الدقة (Multiresolution Analysis-MRA) يتيح تمثيل أي دالة قابلة للتربيع $g(t) \in L^2(\mathbb{R})$ كتوسيع مويجي بالصيغة الاتية: [18]

$$f(z) = \sum_{k=0}^{2^{j_0}-1} c_{j_0k}\,\emptyset_{j_0,k}(z) + \sum_{j=j_0}^{\infty}\sum_{k=0}^{2^j-1} d_{j,k}\,\psi_{\mathrm{j,k}}(z) \tag{36}$$

اذ ان $c_{j_0,k}$ هو معامل التقريب (approximation coefficient)

$$c_{j_0k} = \langle f(z), \emptyset_{j_0,k}\rangle = \int f(z)\,\emptyset_{j_0,k}\,dz \tag{37}$$

و $d_{j,k}$ يمثل معامل التفصيل (detail coefficient)

$$d_{j,k} = \langle f(z), \psi_{jk}(z) \rangle = \int f(z)\, \psi_{jk}(z) dz \quad (38)$$

$$\emptyset_{j_0,k}(z) = 2^{\frac{j}{2}} \emptyset\left(2^j z - k\right), \quad \psi_{j,k}(z) = 2^{\frac{j}{2}} \psi\left(2^j z - k\right)$$

وفي الحالة الإحصائية العملية، حيث تكون البيانات مأخوذة عند نقاط متقطعة Z_i متساوية التباعد، أي ان:

$$f = (f(z_1), f(z_2), \dots, f(z_n)) \quad (39)$$

فان التحويل المويجي المتقطع للدالة f يعطي بالعلاقة التالية:

$$d = Wf \quad (40)$$

اذ ان d : هو متجه ذو ابعاد $n \times 1$ يضم معاملات التحجيم المويجي المتقطعة $c_{j_0,k}$، والمعامـلات المويجة المتقطعة $d_{j,k}$، W : هي مصفوفة متعامدة من الدرجة $n \times n$ ترتبط بقاعدة المويجات المتعامدة المختارة . وترتبط هذه المعاملات المتقطعة بنظيراتها المستمرة وفق العلاقة التقريبية:

$$c_{j_0,k} \approx \sqrt{n}\alpha_{j_0,k} \,, d_{j,k} \approx \sqrt{n}\beta_{j_0,k}$$

ونظراً لتعامد W ، فان التحويل المويجي العكسي (IDWT) كما بالصيغة الاتية:

$$\hat{f} = W^T \hat{d} \quad (41)$$

يتم إعطاء مقدر انكماش الموجات البسيطة كما يلي:

$$\hat{d} \leftarrow \underset{d \in \mathbb{R}^n}{argmind} \; \frac{1}{2} \|y - W^T d\|_2^2 + \lambda \sum_{i=2}^{n} |d_t| \quad (42)$$

2.9.3. الانحدار الخطي الموضعي (Local Linear Regression (LLR))

الانحدار الخطي الموضعي هو أحد الطرائق اللامعلمية، والتي تتميز عن عدة الممهدات بمميزات، منها مرونتها في التأقلم مع التصاميم العشوائية والثابتة، وكذلك لها خاصية تقاربية قوية بين جميع مقدرات النواة (Kernel) وممهدات الشرائح التقديرية والسلاسل الزمنية المتعامدة. نفترض ان المشتقة الثانية للدالة $f(z)$ موجودة. في جوار صغير حول نقطة z, يمكن تقريب $f(s)$ موضعياً بواسطة متعددة الحدود متعددة المتغيرات عند النقطة z على النحو التالي:[19]

$$f(s) \approx f(z) + f'(z)(s - z) \approx \beta_0 + \beta_1(s - z) \quad (43)$$

اذ ان β_0 : مقدر الدالة $f(z)$, β_1 : مقدر المشتقة الأولى لدالة $f(z)$

وعليه, فان مسألة تقدير $f(z)$ تعد مسألة تقدير موضعي لأنحدار خطي، أي تقدير قيمة المقطع β_0 لذلك نأخذ في الاعتبار انحداراً خطياً موزوناً موضعياً يتمثل في إيجاد β_0 و β_1 بحيث يقللان من:

$$\sum_{I=1}^{n} \left(Y_I - \beta_0 - \beta_1(Z_I - z)\right)^2 K\left(\frac{Z_I - z}{h_n}\right) \quad (44)$$

اذ ان K دالة النواة (Kernel function)، و h هو عرض الحزمة (Bandwidth)

ليكن $\hat{\beta}_0$ و $\hat{\beta}_1$ هما حل مسألة المربعات الصغرى الموزونة في المعادلة (44) . ويعد هذا امتداداً لعمل Stone عام (1977), الذي استخدم دالة نواة بسيطة من نوع $K(x) = \frac{1}{2} I_{[|x| \le 1]}$, وقد تم دراسة هذا النهج لاحقاً من قبل كل من Cleveland (1979), Fan (in press), Lejeune (1985), Tsybakov (1986) and , Muller (1987). ويلاحظ ان $\hat{f}(z)$ هو متوسط موزون للقيم المستجيبة, ويطلق علية التمهيد الخطي (Linear Smoothing) كما ان العامل β_1 يمثل تقريباً لتقدير المشتقة الأولى $f'(z)$.
اما عن اختيار عرض الحزمة h_n, فيمكن ان يتم بشكل ذاتي من قبل الباحث, او بطريقة موضوعية بالاعتماد على البيانات. ومن أكثر الطرائق استخداماً في اختيار عرض الحزمة هي طريقة التحقق المتقاطع (Cross-Validation) كما اقترحا Stone عام 1977 والتي تختار h_n لتقليل:

$$\sum\nolimits_{I=1}^{n} \left\{Y_I - \hat{f}_{-I}(Z_I)\right\}^2 \quad (45)$$

يتم تقدير الانحدار الخطي الموضعي وحسب الصيغة التالية:

$$\begin{aligned} &f\left(x_1, \dots, x_r; y_1, \dots., y_p; e_1, \dots, e_q\right) \\ &= E\left\{\emptyset\left(Y_{i_{p+1}}\right) \middle| X_{k_1} = x_1, \dots., x_{k_r} = x_r; Y_{i_1} = y_1, \dots., Y_{i_p} = y_p; \varepsilon_{r_1} = e_1, \dots, \varepsilon_{j_q} = e_q\right\} \end{aligned} \quad (46)$$

أذ أن: e_I, ε_I : متسلسلات مستقلة ذات قيمة متوسطة تساوي صفر، وتعد متغيرات مستقلة ومتطابقة التوزيع (i.i.d)، مع وجود تباين محدود σ_e^2 و σ_ε^2 ، Y_I : القيم التابعة في الزمن I ، X_I : القيم الخارجية في الزمن I ، الدالة $\emptyset(.)$: هي دالة قياسية عشوائية وقابلة للقياس تعرف على خط الاعداد الحقيقية.

الدالة القابلة للقياس لا تفترض بشكل محدد (مثل خطي او تربيعي) وانما يتم تقديرها من البيانات باستخدام طرق لا معلميه، يفترض ان القيمة المتوقعة تكون محدودة، اي $E\left|\emptyset\left(Y_{i_{p+1}}\right)\right| < \infty$ ، تقدير الدالة $\emptyset(.)$ يسمح لنا بتقدير (التوزيعات الشرطية والعزوم الشرطية)

نفرض ان دالة الانحدار $f\left(x_1, \dots, x_r; y_1, \dots, y_p; e_1, \dots, e_q\right)$ لها تحليل تجميعي (Additive Decomposition) وحسب الصيغة (35) ولنفرض أن: [20]

$$\underline{X}_I = \left(X_{I+k_1}, \dots, x_{I+k_r}\right);\quad \underline{Y}_I = \left(Y_{I+i_1}, \dots, Y_{I+i_p}\right);\quad \underline{\varepsilon}_I = \left(\varepsilon_{I+j_1}, \dots, \varepsilon_{I+j_q}\right), \underline{Z}_I = \left(\underline{X}_I, \underline{Y}_I, \underline{\varepsilon}_I\right)$$

$$\underline{x} = (x_1, \dots, x_r);\quad \underline{y} = (y_1, \dots, y_p);\quad \underline{e} = (e_1, \dots, e_q), \underline{z} = \left(\underline{x}, \underline{y}, \underline{e}\right)$$

بأفتراض ان:

$$0 \le i_1 < \cdots < i_{p+1} \text{ ، } 0 \le j_1 < \cdots < j_{q+1} \text{ ، } 0 \le k_1 < \cdots < k_r \;.\; k_r \le i_{p+1} \le j_{q+1}$$

تمثل اعداد صحيحة فأن:

$$Y_{I+i_{p+1}} = g_1\left(Y_{I+i_1}, \dots, Y_{I+i_p}\right) + g_2\left(X_{I+k_1}, \dots, X_{I+k_r}\right) + g_3\left(e_{I+j_1}, \dots, e_{I+j_q}\right) + e_{I+i_{q+1}} \tag{47}$$

$$X_{I+k_r} = g_3\left(X_{I+k_1}, \dots, X_{I+k_{r-1}}\right) + g_4\left(\varepsilon_{I+j_1}, \dots, \varepsilon_{I+j_{q-1}}\right) + \varepsilon_{I+j_q} \tag{48}$$

يمكن ان تكون قيم k_r و j_q و i_p كبيرة، بينما q و r و p قد تكون صغيرة، يتم افتراض ان قيم التأخير $k_1, \dots, k_r$ و $j_1, \dots, j_{q+1}$ و $i_1, \dots, i_{p+1}$ معروفة مسبقاً، ثم ان دالة الانحدار في المعادلة(34) يمكن كتابتها بالصيغة التالية:

$$f\left(\underline{x}, \underline{y}, \underline{e}\right) = E\left\{\emptyset\left(Y_{i_{p+1}}\right) \middle| \underline{X}_I = \underline{x}, \underline{Y}_I = \underline{y}, \underline{\varepsilon}_I = \underline{e}\right\} \tag{49}$$

نفترض ان المشتقات الجزئية من الدرجة الثانية ل $f\left(\underline{s}\right)$ ثابتة ومستمرة عند النقطة $\underline{\boldsymbol{z}}$ يمكننا تقريب $f\left(\underline{s}\right)$ موضعيا بواسطة متعددة الحدود متعددة المتغيرات عند النقطة $\underline{z}$ على النحو التالي :

$$f\left(\underline{s}\right) \approx \beta_0 + \left(\underline{s} - \underline{z}\right)\underline{\beta_1} \tag{50}$$

Where $\beta_0 = f\left(\underline{z}\right)$, $\underline{\beta_1^T} = \partial f\left(\underline{s}\right)/\partial \underline{s}\Big|_{\underline{s} = \underline{z}}$

أذ أن: $\underline{\beta_1^T}$: يدل على تحويل $\underline{\beta_1}$ ($\underline{\beta_1}$ هو متجه العمود)، $\underline{\beta}^T = \left(\beta_0, \underline{\beta_1^T}\right)$: يعتمد على $\underline{z}$ ، نفرض $K(u)$ دالة نواة (Kernel Function) على $\mathbb{R}^d$ حيث $d = p + q + r$ ، $h = h_n$: معامل عرض النطاق (Bandwidth) للتحكم في توزيع الاوزان بالنظر الى المشاهدات $\{X_I, Y_I, \varepsilon_I\}_{I=0}^{n}$ نأخذ في الاعتبار المربعات الصغرى المرجحة متعددة المتغيرات.

$$\sum_{I=0}^{n_1}\left\{\emptyset\left(Y_{I+i_{p+1}}\right) - \beta_0 - \left(\underline{Z}_I - \underline{z}\right)\underline{\beta_1}\right\}^2 K\left(\frac{Z_I - z}{h_n}\right) \tag{51}$$

اذ ان $n_1 = n - 1 - i_{q+1}$: وهو مفترض ان يكون موجباً، $K_h\left(\underline{z}\right) = K\left(\underline{z}/h\right)/h^d$ ، h : معلمة التمهيد bandwidth التي تم تقديرها باستعمال طريقة Least Squares Cross-Validation (LSCV) التي يتم حسابها حسب الصيغة الاتية:

$$\text{LSCV(h)} = \text{n}^{-1}\sum_{t=1}^{n}\left\{Y_t - \hat{f}_{-t}(Z_t)\right\}^2 \tag{52}$$

اذ ان $\hat{f}_{-t}(Z_t)$: يمثل المقدر الممهد عند نقطة البيانات الفردية ويعتمد على ترك واحدة خارجياً Leaving-out-one وكالاتي:

$$\hat{f}_{-t}(Z_t) = \frac{\hat{f}\,(Z_t) - L_t(Z_t)Y_t}{1 - L_t(Z_t)} \tag{53}$$

اذ ان $L_t(Z_t)$: هو معامل النفوذ (leverage function) او (kernel weight) الخاصة بالنقطة Z_t , ويمثل مقدار تأثير القيمة Y_t على التقدير $\hat{f}\,(Z_t)$

$$L_t(Z_t)=\frac{w_t(z)\left[S_{n,2}(z)-(z-Z_t)S_{n,1}(z)\right]}{S_{n,0}(z)S_{n,2}(z)-\left(S_{n,1}(z)\right)^2} \tag{54}$$

$$S_{n,u}=\sum_{t=1}^{n}K\left(\frac{z-Z_t}{h}\right)(z-Z_t)^u \quad u=0,1,2 \tag{55}$$

عند تصغير المعادلة (52) بالنسبة الى β_0 و $\underline{\beta_1}$ نحصل على تقديرات أولية لــ β_0 و $\underline{\beta_1}$ على التتابع وكما نبين ادناه:

$$\underline{\hat{\beta}}(\underline{z})=\begin{pmatrix}\hat{\beta}_0(\underline{z})\\ \underline{\hat{\beta}_1}(\underline{z})\end{pmatrix}=QS_n^{-1}(\underline{z})\underline{t}_n(\underline{z}) \tag{56}$$

اذ ان:

$$Q=diag\{1,h_n^{-1},\dots.,h_n^{-1}\} \tag{57}$$

$$S_n=S_n(\underline{z})=\begin{pmatrix}s_{n,0}(\underline{z}) & \underline{s}_{n,1}^T(\underline{z})\\ \underline{s}_{n,1}(\underline{z}) & S_{n,2}(\underline{z})\end{pmatrix} \tag{58}$$

$$\underline{t}_n(\underline{z})=\begin{pmatrix}t_{n,0}(\underline{z})\\ \underline{t}_{n,1}(\underline{z})\end{pmatrix} \tag{59}$$

$$t_{n,0}(\underline{z})=\frac{1}{n_1+1}\sum_{I=0}^{n_1}\emptyset\left(Y_{I+i_{p+1}}\right)K_h(\underline{Z_I}-\underline{z}) \tag{60}$$

$$\underline{t}_{n,1}(\underline{z})=\frac{1}{n_1+1}\sum_{I=0}^{n_1}\left(\frac{\underline{Z_I}-\underline{z}}{h}\right)^T\emptyset\left(Y_{I+i_{p+1}}\right)K_h(\underline{Z_I}-\underline{z}) \tag{61}$$

$$s_{n,0}(\underline{z})=\frac{1}{n_1+1}\sum_{I=0}^{n_1}K_h(\underline{Z_I}-\underline{z}) \tag{62}$$

$$\underline{s}_{n,1}(\underline{z})=\frac{1}{n_1+1}\sum_{I=0}^{n_1}\left(\frac{\underline{Z_I}-\underline{z}}{h}\right)^T K_h(\underline{Z_I}-\underline{z}) \tag{63}$$

$$S_{n,2}(\underline{z})=\frac{1}{n_1+1}\sum_{I=0}^{n_1}\left(\frac{\underline{Z_I}-\underline{z}}{h}\right)^T\left(\frac{\underline{Z_I}-\underline{z}}{h}\right)K_h(\underline{Z_I}-\underline{z}) \tag{64}$$

وبالتالي فان تقدير $f(\underline{z})$ دالة الانحدار الموضعي هو

$$f(\underline{z})=\hat{\beta}_0(\underline{z})=\underline{e}_1^T\underline{\hat{\beta}}(\underline{z})=\underline{e}_1^TQS_n^{-1}(\underline{z})\underline{t}_n(\underline{z}) \tag{65}$$

اذ ان : $\underline{e}_1^T$ يمثل متجه الوحدة والذي يأخذ الصيغة $(1,0,\dots,0)^T$، وتكمن وظيفته الرياضية في اسـتخلاص قيمة المعلمة $\hat{\beta}_0(\underline{z})$ من قيمة المعلمات الكلي $\underline{\hat{\beta}}(\underline{z})$

3.9.3. معيار المقارنة بين طريقتي التقدير:

لغرض المقارنة بين نتائج تقدير معاملات انموذج ARMAX(p,q,r) لطريقتي التقدير اللامعلميتين موضوع البحث سيتم الاعتماد على معيار المفاضلة الجذر التربيعي لمتوسط مربعات الخطأ RMSE وفق الصيغة الآتية، [21]:

$$RMSE=\sqrt{MSE}=\sqrt{n^{-1}\sum_{t=1}^{n}(y_t-\hat{y}_t)^2} \tag{66}$$

y_t : تمثل القيمة الحقيقية للظاهرة في الزمن (t)، $\hat{y}_t$: تمثل القيمة التقديرية في الزمن (t)، n : حجم العينة

4. الجانب التجريبي .

1.4. وصف تجارب المحاكاة Description of Simulation

تم اعتماد ثلاث حجوم عينات هي (30،50،150) مشاهدة في تنفيذ تجارب المحاكاة، وذلك من خلال توليد بيانات اصطناعية تمثل المتغيرات العشوائية المستخدمة في هذا البحث، وذلك بالاعتماد على القيم الافتراضية للمعاملات. وقد جرى تكرار التجربة 500 مرة لضمان الحصول على نتائج أكثر استقراراً واتساقاً، ومن اجل تقييم كفاءة الطريقتين اللامعلميتين في التقدير تحت ظروف مختلفة من حيث حجم العينة وبدائل القيم الافتراضية.

2.4. عمليات التوليد

تتم عملية تولد المتغيرات من خلال الخطوات الاتية:

اولاً : توليد الخطأ العشوائي بواسطة التوزيع الطبيعي بمتوسط صفر وتباين واحد أي ان $\varepsilon_t \sim N(0,1)$

ثانياً : توليد المتغير الخارجي (X_t) بأعتماد انموذج الانحدار الذاتي من الدرجة الاولى AR(1) كالأتي:

$$\boldsymbol{x_t = \gamma_1 x_{t-1} + \varepsilon_t} \quad (67)$$

اذ ان : γ_1 تمثل معلمة الانحدار الذاتي من الدرجة الأولى. وقد تم توليد قيم سلسلة المدخلات (المتغير الخارجي) بأعتماد ثلاثة قيم افتراضية هي (0.6 ،0.6 ،0.3).

ثالثاً : توليد سلسلة المخرجات (Y_t)، اذ تم توليد قيمها بالاعتماد على الانموذجين التجميعيين وكما مبين في ادناه، [22]:

a. الانموذج التجميعي الأول

$$y_t = [\emptyset_1 \tan(y_{t-1})] + (x_{t-b}) + \theta \tan(\varepsilon_{t-1}) + \varepsilon_t \quad (68)$$

b. الانموذج التجميعي الثاني

$$y_t = [\emptyset_1 \tan(y_{t-1})] + [\emptyset_2 \tan(y_{t-2})] + (x_{t-b}) + [\theta exp|\varepsilon_{t-1}|] + \varepsilon_t \quad (69)$$

من الجدير بالملاحظة ان عملية توليد سلسلة المخرجات (Y_t) وفق الانموذجين (68) و (69) تتطلب توليد قيم السلسلتين الزمنيتين (Y_{t-1} و Y_{t-2}) واللتان تم توليدهما وفق انموذج الانحدار الذاتي من الدرجة الثانية. كما تم الاعتماد على سلسلة المخرجات (X_t) التي تم توليدها وفق المعادلة (67) مع الاخذ بنظر الاعتماد خيارات الابطاء الزمنية (b) وبثلاث قيم افتراضية هي (0 ،1 ،2)

رابعاً : تم الاعتماد على القيم الافتراضية لمعاملات الانموذج والمتمثلة بمعلمتي الانحدار الذاتي من الدرجة الثانية ($\emptyset_1$, $\emptyset_2$) بالنسبة لسلسلة المخرجات (Y_t)، ومعلمة الانحدار الذاتي (γ_1) بالنسبة لسلسلة المدخلات (X_t)، ومعلمة الأوساط المتحركة لسلسلة البواقي (θ)، والمبينة في الجدول (1) بالنسبة للأنموذج التجميعي الأول والجدول (2) بالنسبة للأنموذج التجميعي الثاني.

الجدول (1): يمثل الحالات التي سيتم استعمالها مع الانموذج التجميعي الاول

model	b	γ	$\emptyset_1$	θ
1	0	0.3	-0.375	0.32
		0.3	-0.749	0.654
		0.3	0.35	0.981
	1	0.6	-0.375	0.32
		0.6	-0.749	0.654
		0.6	0.35	0.981
	2	0.9	-0.375	0.32
		0.9	-0.749	0.654
		0.9	0.35	0.981

الجدول (2): يمثل الحالات التي سيتم استعمالها مع الانموذج التجميعي الثاني

model	b	γ	θ	$\emptyset_1$	$\emptyset_2$	b	γ	θ	$\emptyset_1$	$\emptyset_2$	b	γ	θ	$\emptyset_1$	$\emptyset_2$
2	0	0.3	0.32	-0.375	0.1	1	0.6	0.32	-0.375	0.1	2	0.9	0.32	-0.375	0.1
		0.3	0.32	-0.749	0.37		0.6	0.32	-0.749	0.37		0.9	0.32	-0.749	0.37
		0.3	0.32	0.35	0.7		0.6	0.32	0.35	0.7		0.9	0.32	0.35	0.7
		0.3	0.65	-0.375	0.1		0.6	0.65	-0.375	0.1		0.9	0.65	-0.375	0.1
		0.3	0.65	-0.749	0.37		0.6	0.65	-0.749	0.37		0.9	0.65	-0.749	0.37
		0.3	0.65	0.35	0.7		0.6	0.65	0.35	0.7		0.9	0.65	0.35	0.7
		0.3	0.98	-0.375	0.1		0.6	0.98	-0.375	0.1		0.9	0.98	-0.375	0.1
		0.3	0.98	-0.749	0.37		0.6	0.98	-0.749	0.37		0.9	0.98	-0.749	0.37
		0.3	0.98	0.35	0.7		0.6	0.98	0.35	0.7		0.9	0.98	0.35	0.7

5. النتائج والمناقشات

1.5. نتائج تجارب المحاكاة *Results of simulation experiments*

بالاعتماد على سلسلتي المدخلات والمخرجات التي تم تولدهما في الفقرة (2.4) وبأستعمال قيم المعاملات الافتراضية المبينة في الجدولين (1) و (2) ولحجوم العينات الثلاثة المعتمدة. تم تقدير انموذج الانحدار الذاتي بمدخلات خارجية *ARIMAX* بأستعمال طريقتي التقدير اللامعلميتين موضوع البحث وكانت نتائج التقدير مبينة في الجداول (3) و(4) و (5) و (6). ولأجل المقارنة بين ما

افرزته تلك الطريقتين من نتائج، تم الاعتماد على معيار المفاضلة جذر متوسط مربعات الأخطاء $RMSE$ للتوصل الى أفضل طريقة تقدير.

الجدول (3): يبين قيم RMSE للأنموذج التجميعي الاول عند حجوم عينات $(n = 30,50,150)$

model		المعلمات			RMSE للأنموذج التجميعي الأول عند حجم عينة $(n = 30)$		RMSE للأنموذج التجميعي الأول عند حجم عينة $(n = 50)$		RMSE للأنموذج التجميعي الأول عند حجم عينة $(n = 150)$	
	b	γ	$\emptyset_1$	θ	LLR	WAVE	LLR	WAVE	LLR	WAVE
1	0	0.3	-0.375	0.32	0.039554	0.21456	0.039838	0.265665	0.052611	0.141414
		0.3	-0.749	0.654	0.039089	0.190667	0.050224	0.265734	0.067303	0.092942
		0.3	0.35	0.981	0.0249	0.228809	0.04073	0.125706	0.049664	0.133629
		0.6	-0.375	0.32	0.040655	0.227458	0.04848	0.259163	0.052867	0.121999
		0.6	-0.749	0.654	0.032161	0.109671	0.037298	0.25118	0.053518	0.132932
		0.6	0.35	0.981	0.028262	0.182537	0.048964	0.201999	0.047544	0.12039
		0.9	-0.375	0.32	0.037447	0.159579	0.049732	0.12422	0.060786	0.063699
		0.9	-0.749	0.654	0.043231	0.392548	0.036207	0.170064	0.054062	0.095886
		0.9	0.35	0.981	0.033926	0.124694	0.036832	0.193543	0.055981	0.059309
	1	0.3	-0.375	0.32	0.024847	0.60658	0.037163	0.676914	0.050223	0.299659
		0.3	-0.749	0.654	0.030745	0.5701	0.046962	0.441803	0.052494	0.373615
		0.3	0.35	0.981	0.032031	1.095432	0.037436	0.382351	0.045891	0.516037
		0.6	-0.375	0.32	0.023827	0.518855	0.046624	0.86099	0.054556	0.507882
		0.6	-0.749	0.654	0.022015	0.742458	0.041803	0.864708	0.055002	0.499566
		0.6	0.35	0.981	0.026227	1.110822	0.051239	0.482897	0.048489	0.506599
		0.9	-0.375	0.32	0.045384	1.063618	0.047028	0.3815	0.047898	0.511015
		0.9	-0.749	0.654	0.037641	0.495688	0.047441	0.329329	0.059681	0.160887
		0.9	0.35	0.981	0.033419	0.604273	0.035227	0.87133	0.041681	0.254848
	2	0.3	-0.375	0.32	0.027225	1.102505	0.035494	0.911012	0.059933	0.215764
		0.3	-0.749	0.654	0.031339	0.780906	0.0535	0.693937	0.05478	0.545945
		0.3	0.35	0.981	0.023273	1.199664	0.041166	1.155043	0.0437	0.613032
		0.6	-0.375	0.32	0.030248	0.615978	0.03431	0.966789	0.050784	0.322535
		0.6	-0.749	0.654	0.039453	0.923593	0.038198	0.968889	0.05923	0.600615
		0.6	0.35	0.981	0.030887	1.329051	0.032534	0.953437	0.048619	0.54942
		0.9	-0.375	0.32	0.040406	0.57202	0.049582	0.899466	0.055387	0.331558
		0.9	-0.749	0.654	0.031469	0.463073	0.041576	0.939111	0.056875	0.612652
		0.9	0.35	0.981	0.02192	1.199287	0.039926	0.933698	0.067579	0.277108

الجدول (4): يبين قيم RMSE للأنموذج الثاني عند حجوم عينات $(n = 30,50,150)$ والتأخر الزمني $(b = 0)$

model		المعلمات				RMSE للأنموذج التجميعي الثاني عند حجم عينة $(n = 30)$		RMSE للأنموذج التجميعي الثاني عند حجم عينة $(n = 50)$		RMSE للأنموذج التجميعي الثاني عند حجم عينة $(n = 150)$	
	b	γ	θ	$\emptyset_1$	$\emptyset_2$	LLR	WAVE	LLR	WAVE	LLR	WAVE
2	0	0.3	0.32	-0.375	0.1	0.11813	0.320995	0.104378	0.267126	0.132888	0.16069
		0.3	0.32	-0.749	0.37	0.08814	0.199618	0.139084	0.183836	0.116116	0.12793
		0.3	0.32	0.35	0.7	0.06854	0.39232	0.079762	0.201129	0.084134	0.08845
		0.3	0.654	-0.375	0.1	0.08961	0.271205	0.15718	0.211251	0.123216	0.15341
		0.3	0.654	-0.749	0.37	0.09417	0.359509	0.100894	0.150707	0.113221	0.13284
		0.3	0.654	0.35	0.7	0.05824	0.168221	0.081914	0.292304	0.075713	0.11012
		0.3	0.981	-0.375	0.1	0.09383	0.209099	0.113958	0.549053	0.12074	0.14862
		0.3	0.981	-0.749	0.37	0.05439	0.27828	0.111671	0.214527	0.100806	0.12249
		0.3	0.981	0.35	0.7	0.07788	0.31074	0.089066	0.146585	0.071849	0.10241
		0.6	0.32	-0.375	0.1	0.08849	0.207667	0.09876	0.216205	0.088713	0.11751
		0.6	0.32	-0.749	0.37	0.06161	0.19527	0.096161	0.185702	0.087524	0.10210
		0.6	0.32	0.35	0.7	0.05939	0.37219	0.05886	0.152226	0.086206	0.12955
		0.6	0.654	-0.375	0.1	0.05735	0.204645	0.103788	0.219009	0.096549	0.11600
		0.6	0.654	-0.749	0.37	0.05733	0.347959	0.106715	0.239151	0.083872	0.10532
		0.6	0.654	0.35	0.7	0.06082	0.199593	0.048856	0.148764	0.074444	0.12739
		0.6	0.981	-0.375	0.1	0.07877	0.251089	0.103911	0.16282	0.103483	0.12878
		0.6	0.981	-0.749	0.37	0.06500	0.483413	0.089619	0.234538	0.084486	0.13090
		0.6	0.981	0.35	0.7	0.05802	0.148231	0.056906	0.269489	0.06507	0.08677
		0.9	0.32	-0.375	0.1	0.03734	0.177063	0.075942	0.176139	0.06247	0.07671
		0.9	0.32	-0.749	0.37	0.03949	0.406415	0.068468	0.17641	0.06603	0.08233

		0.9	0.32	0.35	0.7	0.03906	0.144456	0.044184	0.130152	0.056353	0.06922
		0.9	0.654	-0.375	0.1	0.04072	0.253925	0.059274	0.154991	0.076413	0.11390
		0.9	0.654	-0.749	0.37	0.06263	0.249754	0.05228	0.110657	0.055008	0.12478
		0.9	0.654	0.35	0.7	0.03915	0.112439	0.06665	0.15689	0.047993	0.06299
		0.9	0.981	-0.375	0.1	0.05652	0.254873	0.075357	0.180321	0.069242	0.09434
		0.9	0.981	-0.749	0.37	0.03507	0.179388	0.050181	0.255096	0.057658	0.08491
		0.9	0.981	0.35	0.7	0.04160	0.10751	0.042995	0.311646	0.053182	0.06701

الجدول (5): يبين قيم RMSE للأنموذج الثاني عند حجوم عينات $(n = 30,50,150)$ والتأخر الزمني $(b = 1)$

model		المعلمات				RMSE للأنموذج التجميعي الثاني عند حجم عينة $(n = 30)$		RMSE للأنموذج التجميعي الثاني عند حجم عينة $(n = 50)$		RMSE للأنموذج التجميعي الثاني عند حجم عينة $(n = 150)$	
	b	γ	θ	$\emptyset_1$	$\emptyset_2$	LLR	WAVE	LLR	WAVE	LLR	WAVE
2	1	0.3	0.32	-0.375	0.1	0.038964	0.435179	0.043077	0.872606	0.04585	0.366105
		0.3	0.32	-0.749	0.37	0.016522	1.122921	0.039764	0.867364	0.049883	0.525345
		0.3	0.32	0.35	0.7	0.020815	1.174526	0.024919	0.351231	0.034751	0.252707
		0.3	0.654	-0.375	0.1	0.041776	0.609686	0.036004	0.560689	0.047918	0.503934
		0.3	0.654	-0.749	0.37	0.022209	1.154381	0.028958	0.429438	0.045484	0.284253
		0.3	0.654	0.35	0.7	0.025253	1.180911	0.034256	0.452935	0.038959	0.319526
		0.3	0.981	-0.375	0.1	0.036663	0.422207	0.050565	0.360173	0.052127	0.462596
		0.3	0.981	-0.749	0.37	0.023792	0.469695	0.032637	0.580075	0.034607	0.280645
		0.3	0.981	0.35	0.7	0.031647	1.150695	0.031644	0.370517	0.032199	0.55229
		0.6	0.32	-0.375	0.1	0.017776	1.134903	0.036971	0.865008	0.046088	0.515174
		0.6	0.32	-0.749	0.37	0.02205	1.083993	0.041966	0.424757	0.05078	0.261701
		0.6	0.32	0.35	0.7	0.01569	0.458805	0.046995	0.375497	0.024986	0.566942
		0.6	0.654	-0.375	0.1	0.031555	1.120077	0.050181	0.8585	0.060784	0.213182
		0.6	0.654	-0.749	0.37	0.023372	1.126904	0.037005	0.50399	0.043263	0.524642
		0.6	0.654	0.35	0.7	0.025299	0.623114	0.026683	0.61144	0.02443	0.560924
		0.6	0.981	-0.375	0.1	0.038569	0.457825	0.030541	0.457235	0.057125	0.185372
		0.6	0.981	-0.749	0.37	0.021201	1.133833	0.03125	0.504003	0.040575	0.518804
		0.6	0.981	0.35	0.7	0.090758	0.514105	0.026295	0.915028	0.03993	0.347424
		0.9	0.32	-0.375	0.1	0.035031	0.265099	0.032339	0.877523	0.050833	0.172836
		0.9	0.32	-0.749	0.37	0.024514	1.116245	0.038941	0.895019	0.049931	0.278977
		0.9	0.32	0.35	0.7	0.031293	0.849482	0.045885	0.575476	0.057527	0.239358
		0.9	0.654	-0.375	0.1	0.027109	1.15345	0.045253	0.57844	0.051275	0.23431
		0.9	0.654	-0.749	0.37	0.034393	0.299812	0.043263	0.332355	0.043348	0.515862
		0.9	0.654	0.35	0.7	0.031757	0.536973	0.0316	0.569651	0.044581	0.294492
		0.9	0.981	-0.375	0.1	0.025941	1.102904	0.048092	0.318365	0.053564	0.509776
		0.9	0.981	-0.749	0.37	0.026911	0.570589	0.043434	0.350436	0.046802	0.23495
		0.9	0.981	0.35	0.7	0.027581	0.60693	0.02943	0.926973	0.026585	0.263278

الجدول (6): يبين قيم RMSE للأنموذج الثاني عند حجوم عينات $(n = 30,50,150)$ والتأخر الزمني $(b = 2)$

model		المعلمات				RMSE للأنموذج التجميعي الثاني عند حجم عينة $(n = 30)$		RMSE للأنموذج التجميعي الثاني عند حجم عينة $(n = 50)$		RMSE للأنموذج التجميعي الثاني عند حجم عينة $(n = 150)$	
	b	γ	θ	$\emptyset_1$	$\emptyset_2$	LLR	WAVE	LLR	WAVE	LLR	WAVE
2	2	0.3	0.32	-0.375	0.1	0.022346	1.23595	0.025685	1.046412	0.05479	0.406221
		0.3	0.32	-0.749	0.37	0.033423	0.857268	0.035045	0.922379	0.048396	0.226184
		0.3	0.32	0.35	0.7	0.019288	0.963637	0.022903	1.265281	0.026101	0.606678
		0.3	0.654	-0.375	0.1	0.029236	0.585513	0.033886	0.885738	0.058046	0.349009
		0.3	0.654	-0.749	0.37	0.023013	0.99691	0.0343	0.986762	0.039634	0.600309
		0.3	0.654	0.35	0.7	0.014792	1.283384	0.026336	1.077066	0.026291	0.649776
		0.3	0.981	-0.375	0.1	0.030598	0.664894	0.048005	0.423904	0.040286	0.544596
		0.3	0.981	-0.749	0.37	0.030154	1.177843	0.028062	0.696944	0.042415	0.63149
		0.3	0.981	0.35	0.7	0.023976	1.226253	0.02378	1.021246	0.030617	0.359033
		0.6	0.32	-0.375	0.1	0.032426	1.190003	0.029936	0.935224	0.054544	0.52977
		0.6	0.32	-0.749	0.37	0.027994	0.70889	0.032254	0.943205	0.049268	0.255304
		0.6	0.32	0.35	0.7	0.019891	1.269012	0.032213	0.978829	0.026361	0.732394
		0.6	0.654	-0.375	0.1	0.032184	1.529575	0.035579	0.973244	0.046717	0.440888
		0.6	0.654	-0.749	0.37	0.02906	1.31242	0.031854	0.500486	0.048041	0.549712
		0.6	0.654	0.35	0.7	0.037129	0.79067	0.022613	1.073276	0.033812	0.430525
		0.6	0.981	-0.375	0.1	0.036286	0.518035	0.043728	0.93578	0.053409	0.301117

		0.6	0.981	-0.749	0.37	0.021089	1.215492	0.023698	0.978495	0.044997	0.568579
		0.6	0.981	0.35	0.7	0.038795	1.229207	0.035716	0.438427	0.031153	0.373333
		0.9	0.32	-0.375	0.1	0.025908	0.863779	0.033739	0.551911	0.058942	0.574954
		0.9	0.32	-0.749	0.37	0.029439	1.274954	0.033732	0.757242	0.054299	0.616611
		0.9	0.32	0.35	0.7	0.022219	1.298231	0.034324	1.094518	0.031158	0.510105
		0.9	0.654	-0.375	0.1	0.023594	1.353051	0.036949	0.448223	0.043632	0.553133
		0.9	0.654	-0.749	0.37	0.037254	0.844747	0.027563	1.104864	0.051329	0.304872
		0.9	0.654	0.35	0.7	0.025954	1.308474	0.037663	0.969807	0.064908	0.626119
		0.9	0.981	-0.375	0.1	0.031064	1.208316	0.047846	0.825515	0.055824	0.259812
		0.9	0.981	-0.749	0.37	0.024691	1.256054	0.041515	0.935519	0.045816	0.545607
		0.9	0.981	0.35	0.7	0.051005	0.894947	0.030186	1.000802	0.036862	0.612114

يتبين من الجداول (3) و (4) و (5) و (6) ومن خلال ما افرزت قيم معيار المفاضلة $RMSE$ الى تفوق طريقة الانحدار الخطي الموضعي (LLR) بشكل واضح بالنسبة للأنموذج التجميعي الأول ولجميع حجوم العينات مع تقارب بسيط بين نتائج تقدير الطريقتين لأربع حالات، بالنسبة لحجم العينة الكبير $(n = 150)$ وحالة عدم وجود تباطئ في تأثير المتغير الخارجي (X_t). اما بالنسبة للأنموذج التجميعي الثاني فأشارت قيم معيار المفاضلة $RMSE$ الى افضلية طريقة الانحدار الخطي الموضعي (LLR) لجميع حجوم العينات ولحالات التباطؤ الزمني الثلاثة.

6. الاستنتاجات

1. ان طريقة الانحدار الخطي الموضعي (LLR)، بصورة عامة كانت من أفضل الطرائق.
2. تبين نتائج المحاكاة ان قيم معيار *(RMSE)* تقل كلما تزداد عدد التأخيرات الزمنية ***b***.
3. قيم معيار المفاضلة ولكلا طريقتي التقدير كانت اقل في حالة الانموذج التجميعي الأول مما يشير الى ان هذا الانموذج التجميعي افرز نتائج أفضل من الانموذج التجميعي الثاني.

7. التوصيات

1. التوسع باستعمال الطرائق اللامعلمية في تقدير انموذج ARMAX باستعمال طرق مختلفة (مثل B-Spline، Nearest Neighbor Methods...).
2. استعمال طرائق تقدير معلمية الى جانب الطرائق اللامعلمية لأغراض المقارنة.
3. استعمال خوارزميات الشبكات العصبية ومقارنتها مع طرائق التقدير المعلمية واللامعلمية.

8. مواد تكميلية

(لا توجد).

9. مساهمات المؤلفين

امير وليد صبري : تم تصميم وكتابة الجانب النظري والجانب التجريبي للبحث, احمد شاكر محمد طاهر المتولي : تفسير النتائج.

10. التمويل

(لا يوجد).

11. بيان توافر البيانات

(لا يوجد)

12. شكر وتقدير

(لا يوجد) .

13. تضارب المصالح

يُعلن المؤلفون عن عدم وجود أي تضارب في المصالح.

References

[1] Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: Forecasting and control (5th ed.). Wiley. (Reviewed in Journal of Time Series Analysis, 2016.). DOI: https://doi.org/10.1002/9781118619100. ISBN-13: 978-1118675021.

[2] Baillie, R. T. (1980). Predictions from ARMAX models. Journal of Econometrics, 12(3), 365–374. https://doi.org/10.1016/0304-4076(80)90062-7

[3] Bercu, B. (1995). Weighted estimation and tracing for ARMAX models. SIAM Journal on Control and Optimization, 33(1), 89-106. DOI: https://doi.org/10.1137/0333005 .

[4] Zheng, W. X. (2002). On least-squares identification of ARMAX models. 15th triennial world congress, Barcelona, spain, School of Quantitative Methods and Mathematical Sciences University of Western Sydney Penrith South DC NSW 1797, Australia. https://doi.org/10.3182/20020721-6-ES-1901.00476
[5] Diversi, R., Guidorzi, R., & Soverini, U. (2011). Identification of ARMAX models with noisy input and output. In Preprints of the 18th IFAC World Congress Milano (Italy). https://doi.org/10.3182/20110828-6-IT-1002.00469
[6] Frausto, H. U., Pieters, J.G., & Deltour, J. M. (2003). Modelling greenhouse temperature by means of auto regressive models. Biosystems Engineering, 84 (2), pp (147-157). https://doi.org/10.1016/S1537-5110(02)00239-8
[7] Taher, A., & Jasim, F. (2023). Forecasting of water quantities in the Iraqi marshes using MSARIMAX models. Journal of Al-Rafidain University College for Sciences, (Print ISSN: 1681-6870, Online ISSN: 2790-2293), (1), 89-99. https://doi.org/10.55562/jrucs.v54i1.579
[8] Yiu, C.-M., Jacob. (2008). Statistical modelling and forecasting schemes for air-conditioning system. (Doctoral dissertation, The Hong Kong Polytechnic University). Handle: 10397/3498.
[9] Mousa, M. A. (2020). Using the ARMAX-GARCH model in forecasting time series with an applied study (Doctoral dissertation). University of Baghdad, College of Administration and Economics.
[10] Ryan, O., Haslbeck, J. M. B., & Waldoro, L. J. (2025). Non-stationarity in time-series analysis: Modeling stochastic and deterministic trends. Multivariate Behavioral Research. https://doi.org/10.1080/00273171.2024.2436413
[11] Chatfield, C. (2004). The analysis of time series an introduction. (5th ed.), Chapman & Hall/CRC. ISBN 978-0412716409. DOI: https://doi.org/10.1201/9780203491683 . ISBN-13: 978-0203491683
[12] Wei, W. W. S. (2006). Time series analysis univariate and multivariate methods. (2nd ed.), Pearson/Addison Wesley. ISBN 978-0321322166
[13] Al-Bermani, F. A. J., & Arshid, A. J. (2019). Using time series to predict the standard numbers for residential rent in Iraq for the years 2018–2021. Journal of Al-Rafidain University College for Sciences. https://doi.org/10.29350/jaucs.2019.115
[14] Petrică, A.-C., Stancu, S., & Ghiţulescu, V. (2017). Stationarity – The central concept in time series analysis. International Journal of Emerging Research in Management and Technology, 6(1), 9-15. https://doi.org/10.23956/ijermt/V6N1/107
[15] Ibrahim, A. J. (2018). Estimation of the ARMA-GARCH model in the presence of outliers. Journal of Administration and Economics, 42(120), 243–260. https://doi.org/10.29116/admeco.2018.120.331
[16] McElroy, T., & Monsell, B. (2012). The multiple testing problem for Box-Pierce statistics. (Research Report Series, Statistics 2012-15). U.S. Census Bureau, Center for Statistical Research & Methodology. https://doi.org/10.3886/ICPSR34857
[17] Chen, R., & Tsay, R. S. (1993). Nonlinear additive ARX models. Journal of the American Statistical Association, 88(423), 955–967. https://doi.org/10.1080/01621459.1993.10476363
[18] Antoniadis, A., Bigot, J., & Sapatinas, T. (2001). Wavelet estimators in nonparametric regression: A comparative simulation study. Laboratoire IMAG-LMC, University Joseph Fourier, and Department of Mathematics and Statistics, University of Cyprus. https://doi.org/10.18637/jss.v006.i06
[19] Cai, Z., & Masry, E. (2000). Nonparametric estimation of additive nonlinear ARX time series: Local linear fitting and projections. Econometric Theory, 16(3), 465–501. https://doi.org/10.1017/S0266466600163053
[20] Ruppert, D. (1996, February 14). Local polynomial regression and its applications in environmental statistics [Article]. Cornell University.
[21] Chai, S., Qiang, L., Liu, Z., & Abedin, M. Z. (2024). Forecasting electricity price from the state-of-the-art modeling technology and the price determinants perspectives: A systematic review and recent advances. Business School, Normal University; Teesside University International Business School. https://doi.org/10.1016/j.ribaf.2023.102132
[22] Chen, R., Liu, J. S., & Tsay, R. S. (1995). Additivity tests for nonlinear autoregression. Biometrika, 82(2), 369–383.https://doi.org/10.1093/biomet/82.2.369

المصادر:

[1] بوكس، ج. إ. ب، جنكينز، ج. م، رينسيل، ج. ك، ولونغ، ج. م. (2015). تحليل السلاسل الزمنية: التنبؤ والسيطرة (الطبعة الخامسة). وايلي. (تمت مراجعته في مجلة تحليل السلاسل الزمنية عام 2016). DOI: https://doi.org/10.1002/9781118619100
[2] بايلي، ر. ت. (1980). التنبؤ بنماذج ARMAX. مجلة القياس الاقتصادي، 12(3)، الصفحات 365–374. https://doi.org/10.1016/0304-4076(80)90062-7
[3] بيركو، ب. (1995). التقدير الموزون والتعقب لنماذج ARMAX. مجلة SIAM للتحكّم والتحسين، 33(1)، 89–106.
[4] تشنغ، وي شينغ. (2002). حول التعرّف بطريقة المربعات الصغرى لنماذج ARMAX. في وقائع المؤتمر العالمي الخامس عشر للاتحاد الدولي للتحكم الآلي IFAC، برشلونة، إسبانيا. كلية الأساليب الكمية والعلوم الرياضية، جامعة ويسترن سيدني، بنريث ساوث، نيو ساوث ويلز 1797، أستراليا. https://doi.org/10.3182/20020721-6-ES-1901.00476
[5] ديفيرسي، ر.، غيدورتسي، ر.، & سوفيريني، أ. (2011). تحديد نماذج *ARMAX* ذات المدخل والمخرج الملوثين بالضوضاء. في وقائع المؤتمر العالمي الثامن عشر للاتحاد الدولي للتحكّم الآلي، ميلانو (إيطاليا). https://doi.org/10.3182/20110828-6-IT-1002.00469
[6] فراوستو، هـ. ي.، بيترز، ج. ج.، وديلتوار، ج. م. (2003). نمذجة درجة حرارة البيوت الزجاجية باستخدام نماذج الانحدار الذاتي. هندسة النظم الحيوية، 84(2)، 147–157. https://doi.org/10.1016/S1537-5110(02)00239-8
[7] احمد طاهر، فراس جاسم. (2023). " التنبؤ بكميات المياه في أنهار العراق باستخدام انموذج MSARIMAX". مجلة كلية الرافدين الجامعة للعلوم، العدد (54)، 89–99. https://doi.org/10.55562/jrucs.v54i1.579

[8] ييو، ج. م.، وجاكوب. (2008). النمذجة الإحصائية وخطط التنبؤ لأنظمة التكييف الهوائي. أطروحة دكتوراه، جامعة بوليتكنيك هونغ كونغ. https://theses.lib.polyu.edu.hk/handle/200/2728

[9] موسى (مؤمن عباس). (2020). استعمال انموذج (ARMAX-GARCH) في عملية التنبؤ للسلسلة الزمنية مع تطبيق عملي. أطروحة دكتوراه في الإحصاء جامعة بغداد-كلية الإدارة والاقتصاد.

[10] رايان، أ.، هاسلبك، ج. م. ب.، ووالدورب، ل. ج. (2025). عدم الاستقرارية في تحليل السلاسل الزمنية: نمذجة الاتجاهات العشوائية والحتمية. البحث السلوكي متعدد المتغيرات، 60(3)، 556–588. https://doi.org/10.1080/00273171.2024.2436413

[11] شاتفيلد، ك. (2004). تحليل السلاسل الزمنية: مقدمة. (الطبعة الخامسة). تشابمان وهول/ *CRC*. رقم الكتاب الدولي المعياري *(ISBN): 978-0412716409*

[12] وي، و. و. س. (2006). تحليل السلاسل الزمنية: الطرق الأحادية والمتعددة المتغيرات (الطبعة الثانية). بيرسون/أديسون ويسلي. رقم الكتاب الدولي المعياري (ISBN): 978-0321322166

[13] البيرماني (فاطمة عبد الحميد جواد) ، ارشيد (احمد جودة). (2019). استعمال السلاسل الزمنية للتنبؤ بالأرقام القياسية لإيجارات الدور السكنية في العراق للسنوات 2018-2021. مجلة كلية الرافدين الجامعة للعلوم. https://doi.org/10.29350/jaucs.2019.115

[14] بيتريكا، أ.-ك.، ستانكو، س.، وغيتيوليسكو، ف. (2017). الاستقرارية – المفهوم المركزي في تحليل السلاسل الزمنية. المجلة الدولية للبحوث الناشئة في الإدارة والتكنولوجيا، 6(1)، 9–15. https://doi.org/10.23956/ijermt/V6N1/107

[15] إبراهيم (علي جاسم). (2018). تقدير انموذج *ARMA-GARCH* بوجود القيم الشاردة. مجلة الإدارة والاقتصاد / السنة-42 العدد 120/2019. https://doi.org/10.29116/admeco.2018.120.331

[16] مكإلروي، ت.، ومونسيل، ب. (2012). مشكلة الاختبار المتعدد لإحصائيات بوكس-بيرس. (سلسلة تقارير بحثية، إحصائيات 2012-15). مكتب الإحصاء الأمريكي، مركز البحث والمنهجية الإحصائية. https://doi.org/10.3886/ICPSR34857

[17] تشن، ر.، وتساي، ر. س. (1993). نماذج *ARX* غير الخطية التجميعية. مجلة جمعية الإحصاء الأمريكية، 88(423)، 955–967. https://doi.org/10.1080/01621459.1993.10476363

[18] أنتونياديس، أ.، بيغوت، ج.، وساباتيناس، ت. (2001). تقدير المويجات في الانحدار غير المعلمي: دراسة محاكاة مقارنة. مجلة البرمجيات الإحصائية، 6(6)، 1–83. https://doi.org/10.18637/jss.v006.i06.

[19] تساي، زوون نغ غو، وماسري، إيليا. (2000). التقدير اللامعلمي للسلاسل الزمنية من نوع *ARX* اللاخطية التجميعية: الملاءمة الخطية الموضعية والإسقاطات". مجلة النظرية الاقتصادية القياسية، المجلد (16)، العدد (3)، الصفحات 465–501. https://doi.org/10.1017/S0266466600163053

[20] روبرت، د. (1996). الانحدار متعدد الحدود الموضعي وتطبيقاته في الإحصاء البيئي [محاضرة]. جامعة كورنيل.

[21] تشاي، ش.، لي، ك.، ليو، ز.، وآبيدين، م. ز. (2024). توقع أسعار الكهرباء من خلال تقنيات النمذجة الحديثة ومنظور محددات الأسعار: مراجعة منهجية وتطورات حديثة. مجلة الأبحاث في الأعمال والمالية الدولية، 67(22)، 102132. https://doi.org/10.1016/j.ribaf.2023.102132

[22] تشين، رونغ، ليو، جون س.، وتساي، روي س. (1995). الاختبارات التجميعية في الانحدار الذاتي غير الخطي. مجلة بايومتريكا ((Biometrika، المجلد 82، العدد 2، الصفحات 369–383. https://doi.org/10.1093/biomet/82.2.369

https://doi.org/10.31272/jae.i151.1549
https://admics.uomustansiriyah.edu.iq
P-ISSN: 1813–6729 E-ISSN: 2707-1359

Estimation of the ARMAX Model Using Two Nonparametric Methods: Local Linear Regression and Wavelet Estimation

Ameer Waleed Sabri Sharar
Dept. Engineering Supervision, Engineering Construction Department, Ministry of Construction, Housing, Municipalities and Public Works, Baghdad, Iraq.
Email: ameerwaleed@uomustansiriyah.edu.iq, ORCID: https://orcid.org/0009-0007-1956-8455
Ahmed Shaker Mohammed Taher Al-Mutawalli
Dept. of Statistics, College of Administration and Economics, Mustansiriya University, Baghdad, Iraq.
Email: ahmutwali@uomustansiriyah.edu.iq, ORCID: https://orcid.org/0000-0002-5679-0940

Article Information

Article History:
Received: 01 / 01 / 2025
Revised: 25 / 02 / 2026
Accepted: 28 / 02 / 2026
Available Online: 01 / 03 / 2026
Pages no: 125 – 141

Keywords:
ARMAX Models, Autoregression, Moving Average, Local Linear Regression, Wavelet Estimation.

Correspondence:
Researcher name: Ameer Waleed Sari

Email: ameerwaleed@uomustansiriyah.edu.iq

Abstract

Autoregressive Moving Average models with Exogenous inputs, denoted as ARMAX, are widely used to describe the dynamic behaviour of time series influenced by external variables. In these models, the (AR) component represents the autoregressive part of the output series, the (MA) component represents the moving average part associated with the random residuals, and the (X) component represents the exogenous component associated with the input series. Thus, these models effectively describe the stochastic and deterministic components of the data.

The parameters of these models can be estimated using both parametric and non-parametric estimation methods. However, parametric methods rely on stringent assumptions to achieve efficient estimation, whereas non-parametric methods do not impose such functional assumptions; hence, they are preferred by many researchers. Among these robust non-parametric techniques are Local Linear Regression (LLR) and Wavelet Estimation (WAVE). In this study, a comparison between these two methods was conducted using simulation experiments. The simulation results revealed a clear preference for the Local Linear Regression method, based on the Root Mean Square Error (RMSE), which yielded the lowest values among the methods.