



مقارنة كفاءة مقدرات انموذج الصندوق الاسود ARMAX للتنبؤ بقيم الناتج المحلي الاجمالي في العراق

A Comparison of the Efficiency of Estimators for the Black Box Model ARMAX to prediction Iraq's GDP values

أ.م.د. سعد صبر محمد⁽²⁾

زهراء منذر حسن⁽¹⁾

Asst. Prof. Dr Saad Sabr Mohammed

Zahraa Munther Hassan

ssabr@uowasit.edu.iq

Zmunther523@uowasit.edu.iq

المستخلص :

إن اختيار الانموذج المناسب هو شرط اساسي للتنبؤ بالسلاسل الزمنية ، لذلك تم دراسة انموذج الصندوق الاسود والمتمثل بأنموذج الانحدار الذاتي والمتوسطات المتحركة مع مدخلات خارجية (ARMAX) واختيار الانموذج المناسب للتنبؤ على اساس خطوات محددة تبدا هذه الخطوات بالتعرف على الانموذج وذلك من خلال التحقق من استقراره السلسلة الزمنية المراد دراستها ، يليها اختيار الانموذج المناسب من خلال تحديد رتبة الانموذج بعده معايير منها ، معيار أكاكي للمعلومات (AIC) معيار المعلومات البيزي (BIC) ، ثم خطوة تقدير الانموذج بطريقة المربعات الصغرى الاعتيادية (OLS) ، وطريقة المتغيرات المساعدة الموسعة (EIV) ، تتم مقارنه هذه الطرائق باستعمال مقاييس التقييم ، بما في ذلك مقياس الجذر التربيعي لمتوسط مربعات الخطأ (RMSE) ، مقياس متوسط مطلق الخطأ النسبي (MAPE) ، والمتوسط المتماثل لمطلق الخطأ النسبي (SMAPE) . حيث اظهرت النتائج ان انموذج (3,3,1,ARMAX) أفضل انموذج مقدر بطريقة (OLS) ، لذلك تم الاعتماد عليه للتنبؤ بقيم الناتج المحلي الاجمالي في العراق حتى عام 2035.

الكلمات المفتاحية : انموذج الانحدار الذاتي والمتوسطات المتحركة مع مدخلات خارجية (ARMAX) ، معيار أكاكي للمعلومات (AIC) ، معيار المعلومات البيزي (BIC) ، طريقة المربعات الصغرى الاعتيادية (OLS) ، طريقة المتغيرات المساعدة الموسعة (EIV) ، مقياس الجذر التربيعي لمتوسط مربعات الخطأ (RMSE) ، التنبؤ .

Abstract:

The selection of an appropriate model is a fundamental requirement for forecasting time series. This study examines the black-box model, represented by the Autoregressive

Moving Average with Exogenous Inputs (ARMAX) model, and focuses on selecting the optimal model for forecasting based on a structured sequence of steps. These steps begin with model identification, which involves verifying the stationarity of the time series under investigation. Subsequently, the appropriate model is selected by determining its order using several criteria, including the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). The next step is model estimation, conducted using the Ordinary Least Squares (OLS) method and the Extended Instrumental Variables (EIV) method. These methods are compared using evaluation metrics such as the Root Mean Square Error (RMSE), the Mean Absolute Percentage Error (MAPE), and the Symmetric Mean Absolute Percentage Error (SMAPE). The results indicated that the ARMAX (3,1,3,3) model, estimated using the OLS method, is the best-performing model. Therefore, this model was employed to forecast the Gross Domestic Product (GDP) of Iraq up to the year 2035."

1-المقدمة

في الوقت الذي تؤدي فيه السلاسل الزمنية دورا مهما في تحليل البيانات والتنبؤ بالمستقبل لمختلف الظواهر الاقتصادية والهندسية والبيئية، تبرز الحاجة إلى تطوير نماذج رياضية دقيقة للتنبؤ بهذه السلاسل، لتحقيق هذا الهدف يستعمل انموذج الصندوق الاسود ARMAX، حيث تقوم هذه النماذج بالدمج بين اسلوب السلاسل الزمنية التي تعتمد على البيانات التاريخية واسلوب الانحدار بمتغيرات خارجية اضافة إلى ذلك الاخطاء السابقة.

مع تنوع التطبيقات وتعدد البيانات التي تطبق فيها هذه النماذج، تظهر الحاجة إلى تقييم كفاءة ودقة هذه النماذج في ظل ظروف مختلفة. حيث يتم في هذا البحث دراسة مقارنة لمقدرات انموذج الانحدار الذاتي والمتوسطات المتحركة مع مدخلات خارجية ARMAX للتنبؤ بالنتائج المحلي الاجمالي في العراق.

في عام 2017 م اقترح الباحثون (Hesen Liu & Lin Zhu وآخرون) [6] نهجا شاملا لتحديد نموذج دالة التحويل منخفض الترتيب يعتمد على القياس عبر الانترنت الذي يمكنه التقاط جميع اوضاع التذبذب الرئيسية للتحكم في تخميد التذبذب التكيفي والمنسق، باستعمال انموذج المتوسط المتحرك ذاتي الانحدار مع متغيرات خارجية (ARMAX) متعدد المدخلات والمخرجات (MIMO) هذه المنهجية تتكون من خمس خطوات: اختيار المدخلات، اختيار المخرجات، تحديد الانموذج، تقدير الانموذج واختبار الانموذج، في هذه الدراسة تم استعمال البيانات المحيطة والبيانات الحلقة لاختبار فعالية النهج المقترح، وقد تم توضيح هذا النهج من خلال دراسة حالة لنظام مجلس تنسيق الطاقة الكهربائية في الشمال الشرقي، والذي يتكون من 16 آلة و68 حافلة. وكانت النتائج ان النموذج القائم على القياس وباستعمال (MIMO-ARMAX) يمكنه التقاط جميع اوضاع التذبذب الرئيسية بنفس الدقة ولكن ترتيب اقل وكفاءة حسابية محسنة بالمقارنة مع انموذج مساحة حالة الفضاء الجزئي (MIMO).

وفي العام نفسه قدم الباحثون (Lu Wan & Abduselam H Beshir وآخرون) [14] دراسة تناولت المشاكل الرئيسية لأنظمة الطاقة المتجددة من حيث التوافق الكهرومغناطيسي. استعرضت هذه الدراسة تقنيات نمذجة الصندوق

الاسود المختلفة لمحاولات الطاقة المتجددة ، وعرضها وتحليلها بشكل منهجي مع التركيز بشكل خاص على استراتيجيات النمذجة السلوكية (نمذجة الصندوق الاسود) الاكثر ملائمة لتحليل التوافق الكهرومغناطيسي للأنظمة المعقدة التي تشمل محاولات الطاقة ، حيث ناقشت الدراسة ثلاث تقنيات مختلفة من نماذج الصندوق الاسود لمقارنة دقتها وإجراءات التجريب الخاصة بها .

في عام 2022 قام الباحثان (Saidon, Mohd Amin, Nasrul Amri & Mohd Saifizi وآخرون) [12] بدراسة مقارنة بين النموذجي ARX، (ARMAX) لتحديد انموذج ديناميكي لنظام تبريد متكامل يعتمد على التبريد الحراري الكهربائي (TEC) لتبريد النواذف الزجاجية ، وذلك من خلال مقاييس عدة منها مجموع مربعات الاخطاء SSE ، متوسط مربعات الخطأ MSE ، الجذر التربيعي لمتوسط مربعات الخطأ RMSE ، والانحراف المعياري . حيث اظهرت النتائج ان أنموذج ARMAX من الرتبة (4)،3،(3) اعطى نسبة تطابق تصل إلى 80.23% بين البيانات المتوقعة والبيانات الحقيقية ، مقارنة بنموذج ARX من الرتبة (4)،(4) الذي اعطى نسبة تطابق 78.14% .

2- مشكلة البحث

تكمن مشكلة هذه الدراسة في الحاجة إلى اختيار الانموذج المناسب الذي يوفر كفاءة اعلى في التنبؤ بقيم السلسلة الزمنية . يتم ذلك من خلال مقارنة مقدرات انموذج ARMAX باستعمال مقاييس دقيقة وتطبيق عملي على بيانات زمنية واقعية.

3- هدف البحث

تتمثل اهداف هذا البحث في النقاط الآتية :

1. يهدف البحث إلى تحليل ومقارنة كفاءة أنموذج الصندوق الأسود ARMAX ، في التنبؤ بقيم السلاسل الزمنية وتحديدًا للتنبؤ بقيم الناتج المحلي الاجمالي للعراق
- 2 -تسليط الضوء على بناء انموذج الصندوق الاسود ARMAX ،حيث يتم توظيف المتغير الخارجي في انموذج التنبؤ، وتحديد الانموذج الافضل من خلال استعمال معايير اختيار وطرق تقدير مختلفة .
- 3- تهدف الدراسة إلى تحديد الانموذج الأكثر دقة وفعالية في التنبؤ، مما يسهم في تحسين القدرة على التحليل والتنبؤ بالسلاسل الزمنية في مجالات متعددة.

4- نماذج الصندوق الاسود [11][15]

تصنف نماذج الصندوق الاسود على انها اساليب نمذجة احصائية تعتمد على البيانات المقاسة (المدخلات والمخرجات) فقط دون الحاجة إلى معلومات مسبقة عن ديناميكيات النظام ، اي التغيرات التي تحدث بمرور الزمن وتأثير العوامل الداخلية والخارجية عليه ، تستخدم نماذج الصندوق الاسود في مجالات متعددة مثل الهندسة ، وعلم البيانات ، والذكاء الاصطناعي ، والاقتصاد ، وغيرها .

يعبر رياضياً عن الهيكل المعمم لنماذج الصندوق الاسود الاحصائي في العلاقة الآتية :

$$A(q).Y(t) = \frac{B(q)}{F(q)} Z(t) + \frac{C(q)}{D(q)} .\varepsilon(t) \quad \dots (1)$$

حيث أن

$Y(t)$: سلسلة الإخراج

$Z(t)$: سلسلة الإدخال

$\varepsilon(t)$: الضوضاء البيضاء

وان $A(q)$ ، $B(q)$ ، $C(q)$ ، $D(q)$ ، $F(q)$ ، هي دوال متعددة الحدود تعطى على النحو التالي :

$$A(q) = 1 + a_1 q^{-1} + \dots + a_{n_a} q^{-n_a}$$

$$B(q) = 1 + b_1 q^{-1} + \dots + b_{n_b} q^{-n_b}$$

$$C(q) = 1 + c_1 q^{-1} + \dots + c_{n_c} q^{-n_c}$$

$$D(q) = 1 + d_1 q^{-1} + \dots + d_{n_d} q^{-n_d}$$

$$F(q) = 1 + f_1 q^{-1} + \dots + f_{n_f} q^{-n_f}$$

q^{-1} : تمثل مشغل الازاحة الخلفي

$A(q)$: تمثل متعدد حدود من الدرجة n_a

$B(q)$: تمثل متعدد حدود من الدرجة n_b

$C(q)$: تمثل متعدد حدود من الدرجة n_c

$D(q)$: تمثل متعدد حدود من الدرجة n_d

$F(q)$: تمثل متعدد حدود من الدرجة n_f

5- انموذج الانحدار الذاتي والمتوسطات المتحركة مع مدخلات خارجية [9] [12]

أنموذج ARMAX يشبه هيكل أنموذج ARX ، إلا انه يأخذ في الاعتبار تأثير الاضطرابات المبكرة التي قد تدخل مع المدخلات إلى النظام، حيث يتميز بقدرته على الفصل بين تأثير الاضطرابات وتأثير المدخلات ، عن طريق تضمين مصطلح متوسط الخطأ المتحرك ، مما يجعل الانموذج أكثر دقة في تمثيل النظام ، وان أنموذج ARMAX قابل للتطبيق في العديد من المجالات مثل التحكم ، العمليات ، والاقتصاد القياسي لكل من نمذجة النظام وتصميم النظام .

وأن علاقة الادخال والاخراج هي

$$A(q).Y(t) = B(q)Z(t) + C(q) \varepsilon(t)$$

يمكن وصف معلمات هيكل أنموذج ARMAX بالمعادلة الخطية الآتية :

$$\begin{aligned} Y(t) + a_1 Y(t-1) + \dots + a_{n_a} Y(t-n_a) \\ = b_1 Z(t-1) + \dots + b_{n_b} Z(t-n_b) + \varepsilon(t) + c_1 \varepsilon(t-1) + \dots \\ + c_{n_c} \varepsilon(t-n_c) \quad \dots (2) \end{aligned}$$

يمكن التعبير عن المعلمات الغير معرفة بواسطة المتجه θ

$$\theta = [a_1 \quad a_2 \dots a_{n_a} \quad b_1 \quad b_2 \dots b_{n_b} \quad c_1 \quad c_2 \dots c_{n_c}]^T$$

6- التعرف على الانموذج [7][8]

ان الخطوة الاولى والمهمة في بناء الانموذج الاحصائي هي معرفة اذا كانت كل من سلسلة الادخال وسلسلة الاخراج مستقرة ام لا في التباين والوسط ، أن العديد من السلاسل الزمنية التي تمت مواجهتها في الممارسة العملية تظهر سلوكا غير مستقر وبما أن منهجية بوكس – جينكينز مخصصة للنماذج المستقرة فقط فيجب علينا إجراء بعض التعديلات قبل ان نتمكن من نمذجة هذه السلاسل، فأذا كانت السلسلة الزمنية غير المستقرة في المتوسط فيجب اخذ الفروق الاولى، اما في التباين فيتم استعمال التحويل اللوغاريتمي وتحويل الجذر التربيعي التي تعد من اكثر التحويلات استعمالا .

7- معايير اختيار الانموذج [7]

قد يكون استعمال دوال الارتباط الذاتي لاتخاذ القرار بشأن اختيار الانموذج المناسب أمرا صعبا للغاية في الممارسة العملية، الطريقة البديلة والاكثر سهولة هي استعمال معايير احصائية لاختيار الانموذج من بين مجموعة محدودة من النماذج، حيث يتم اختيار الانموذج ذو القيمة الدنيا لهذه المعايير.

-معيار أكاي للمعلومات [1]

لتقييم جودة ملائمة الانموذج قدم Akaike معياراً للمعلومات يطلق عليه (AIC) يقابل الانموذج الافضل اقل قيمة من معيار AIC يتم تعريفه كالآتي

$$AIC = 2m - 2 \ln(\text{maximized likelihood}) \quad \dots (3)$$

وللسهولة في الاستعمال وبدلالة مقدار تباين الخطأ تكون صيغة AIC كما يلي :

$$AIC = 2m + n \ln \hat{\sigma}_\varepsilon^2 \quad \dots (4)$$

حيث ان :

m: تمثل عدد المعلمات التي سيتم تقديرها

n: عدد المشاهدات

δ_e^2 : تقدير تباين الخطأ

- معيار المعلومات البيزي [1][3]

يعرف معيار بيز احصائيا ب BIC ويطلق عليه معيار شوارتز (Schwarz) ايضا ، وان معيار BIC متسق بشكل متقارب كمعيار للاختيار ، وهذا يعني انه بالنظر إلى مجموعة من النماذج ، بما في ذلك الانموذج الصحيح ، فان احتمال قيام BIC باختيار الانموذج الصحيح يقترب من الواحد مع اقتراب حجم العينة $N \rightarrow \infty$ عكس معيار AIC الذي يميل إلى اختيار نماذج معقدة للغاية مع اقتراب $N \rightarrow \infty$ ، من ناحية اخرى وبالنسبة للعينات المحدودة غالبا ما تختار BIC نماذج بسيطة جدا نظرا لابتعادها عن التعقيد .

الشكل العام ل BIC هو :

$$BIC = -2\loglik + m. (\log n) \quad \dots (5)$$

8- تقدير معلمات الانموذج

بعد تحديد شكل الانموذج لابد من تقدير معالم الانموذج باستعمال تقنيات التقدير الاحصائي ، حيث تعد هذه المرحلة الجزء الاساسي من خطوات بناء الانموذج حيث تعتمد دقة التحليل والتنبؤات الناتجة عنها بشكل كبير على جودة تقدير الانموذج . ولتقدير نماذج الصندوق الاسود (ARX) ، (ARMAX) تم استعمال كل من طريقة المربعات الصغرى الاعتيادية (OLS) وطريقة المتغيرات المساعدة الموسعة ، والاتي شرح مفصل عنهما :

- طريقة المربعات الصغرى الاعتيادية [9][6]

تسعى طريقة المربعات الصغرى للحصول على قيم المعلمات التي توفر الوصف الاكثر دقة للبيانات، والتي يتم قياسها من حيث مدى ملائمة الانموذج للبيانات ضمن دالة الخسارة المربعة ، يتم في هذه الطريقة تقليل مجموع مربعات الخطأ (SSE) بين المشاهدات والتنبؤات .

وبالبدء تكتب معادلة الانحدار كالاتي :

$$Y(t) = \varphi^T \theta + \varepsilon(t) \quad \dots (6)$$

وان φ^T هو متجه يتضمن البيانات السابقة من المخرجات والمدخلات وكما يلي :

$$\varphi^T(\theta) = [-Y(t-1) \dots -Y(t-n_a) \quad Z(t-1) \dots Z(t-n_b) \quad \varepsilon(t-1, \theta) \dots \varepsilon(t-n_c, \theta)]$$

يعرف تقدير المربعات الصغرى على انه متجه المعلمة θ الذي يقلل من دالة الخسارة وكلائي :

$$V_n = \frac{1}{N} \sum_{t=1}^N \varepsilon^2(t) \quad \dots (7)$$

وبتعويض معادلة خطأ التنبؤ بالمعادلة (7):

$$V_n = \frac{1}{N} \sum_{t=1}^N (Y(t) - \varphi^T(t) \theta)^2$$

بعد اشتقاق دالة الخسارة ومساواتها بالصفر نحصل على :

$$[\sum_{t=1}^N \varphi(t) \varphi^T(t)] \hat{\theta} = \sum_{t=1}^N (Y(t) \varphi^T(t)) \quad \dots (8)$$

وبافتراض ان المصفوفة في المعادلة (8) هي مصفوفة غير مفردة فان :

$$\hat{\theta} = [\sum_{t=1}^N \varphi(t) \varphi^T(t)]^{-1} \sum_{t=1}^N (Y(t) \varphi^T(t)) \quad \dots (9)$$

- طريقة المتغيرات المساعدة الموسعة [5] [13]

في هذه الطريقة تم تقديم بعض التوسيعات وهي استعمال مصفوفة $m(t)$ المطورة (المعززة) اي ان $\dim \varphi(t) > \dim m(t)$ بالإضافة إلى التصفية المسبقة للبيانات ، وان لدراسة هذه التوسيعات سببين هما :

- يمكن تحسين دقة التقديرات اذا تم اختيار المتغيرات المساعدة $m(t)$ والمرشح المسبق $F(q^{-1})$ بعناية .

- قد يكون من السهل تلبية شروط الاتساق لطريقة المتغيرات المساعدة الموسعة وهي :

$$1- E m(t) \cdot F(q^{-1}) \varphi^T(t) \text{ has rank } n\theta$$

$$2- E m(t) \cdot F(q^{-1}) \varepsilon(t) = 0$$

يتم تقديم تقدير المتغيرات المساعدة الموسعة كالآتي :

$$\hat{\theta} = \arg \min_{\theta} \| [\sum_{t=1}^N m(t) \cdot F(q^{-1}) \varphi^T(t)] \theta - [\sum_{t=1}^N m(t) \cdot F(q^{-1}) Y(t)] \|_Q^2$$

حيث ان

$F(q^{-1})$: هو مرشح مسبق مستقر بشكل تدريجي .

عند تطبيق المعيار الموزون $\|X\|_Q^2 = X^T Q X$ عندما Q هي مصفوفة موجبة محددة موزونة فإن

$$\hat{\theta} = \arg \min_{\theta} [\sum_{t=1}^N m(t) \cdot F(q^{-1})\varphi^T(t)\theta - \sum_{t=1}^N m(t) \cdot F(q^{-1})Y(t)]^T Q [\sum_{t=1}^N m(t) \cdot F(q^{-1})\varphi^T(t)\theta - \sum_{t=1}^N m(t) \cdot F(q^{-1})Y(t)] \quad \dots (10)$$

لتبسيط عملية الاشتقاق لـ θ نعتد بعض الافتراضات :

$$R_N = \sum_{t=1}^N m(t) \cdot F(q^{-1})\varphi^T(t)$$

$$D_N = \sum_{t=1}^N m(t) \cdot F(q^{-1})Y(t)$$

لتصبح $\hat{\theta}$ بالشكل الآتي :

$$\hat{\theta} = [R_N \theta - D_N]^T Q [R_N \theta - D_N]$$

بعد اجراء عملية الضرب ثم الاشتقاق بالنسبة لـ θ والمساواة بالصفر يتم التوصل إلى :

$$\hat{\theta} = [R_N^T Q R_N]^{-1} R_N^T Q D_N$$

بعد التعويض عن R_N ، D_N نحصل على :

$$\hat{\theta} = \{[\sum_{t=1}^N m(t) \cdot F(q^{-1})\varphi^T(t)]^T Q [\sum_{t=1}^N m(t) \cdot F(q^{-1})\varphi^T(t)]\}^{-1} \cdot \{[\sum_{t=1}^N m(t) \cdot F(q^{-1})\varphi^T(t)]^T Q [\sum_{t=1}^N m(t) \cdot F(q^{-1})Y(t)]\} \quad \dots (11)$$

9- مقارنة طرائق التقدير

يمكن مقارنة طرائق التقدير المذكورة سابقا من خلال المقدرات لكل طريقة، حيث تتضمن عملية المقارنة مجموعة من المقاييس الاحصائية مثل الجذر التربيعي لمتوسط الخطأ (RMSE)، ومقياس متوسط الخطأ النسبي (MAPE)، ومقياس المتوسط المتماثل لمطلق الخطأ النسبي (SMAPE)، واختيار الطريقة التي تعطي اقل قيمة لهذه المقاييس باعتبارها الطريقة الأمثل للحصول على تقديرات دقيقة وموثوقة.

- مقياس الجذر التربيعي لمتوسط مربعات الخطأ [4]

كما يشير اسمه، فانه الجذر التربيعي لمتوسط مربعات الخطأ MSE إذا ان RMSE هو :

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad \dots (12)$$

y_i : القيمة الحقيقية للمخرجات

$\hat{y}_i$: القيمة التقديرية للمخرجات

- مقياس متوسط مطلق الخطأ النسبي [2]

يكتب بالصيغة الآتية :

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad \dots (13)$$

- المتوسط المتماثل لمطلق الخطأ النسبي [2]

تم اقتراح SMAPE لتعديل عيوب مقياس MAPE ويكتب بالصيغة الآتية :

$$SMAPE = \frac{100}{n} \sum_{i=1}^n \frac{2 * |\hat{y}_i - y_i|}{(|\hat{y}_i| + |y_i|)} \quad \dots (14)$$

10- التنبؤ [15]

بعد تقدير انموذج الصندوق الاسود ARMAX باستعمال الطرق المذكورة ، ، سنبدأ بمناقشة التنبؤ لهذه النماذج وكالاتي :

يتم كتابة التنبؤ بخطوة واحدة إلى الامام لأنموذج ARMAX في الوقت (t) كالاتي :

$$\begin{aligned} \hat{Y}(1) &= \hat{Y}(t+1) \\ &= -a_1 Y(t) - \dots - a_{n_a} Y(t - n_a + 1) + b_1 Z(t) + \dots + b_{n_b} Z(t - n_b + 1) \\ &\quad + \dots + c_1 \varepsilon(t) + \dots + c_{n_c} Z(t - n_c + 1) \quad \dots (15) \end{aligned}$$

الجانب التطبيقي

11- تعريف المتغيرات [10]

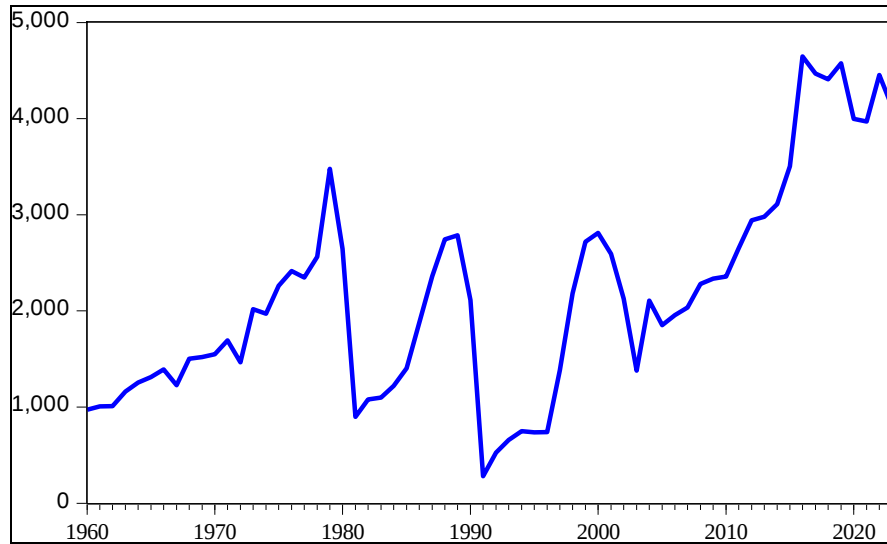
الناتج المحلي الاجمالي GDP هو القيمة الاجمالية لجميع السلع والخدمات النهائية التي تنتج داخل بلد معين ، حيث يعتبر من اهم المؤشرات الاقتصادية الرئيسية التي تشير إلى مستوى النمو الاقتصادي للعراق ، وبذلك فهو يتأثر بشكل كبير بصادرات النفط ، باعتبار العراق بلد منتج ومصدر للنفط الخام ومركز بصورة اساسية على العوائد النفطية.

تم الحصول على البيانات التي تتمثل بالناتج المحلي الاجمالي في العراق ولل سنوات (1960-2023) من البنك الدولي ومن خلال الرابط : <https://data.worldbank.org> ، اما بيانات انتاج النفط في العراق ولل سنوات (1960-2023)

تم اخذها من منظمة الدول المصدرة للبترول اوبك من خلال الرابط التالي : https://asb.opec.org/data/ASB_Data.php ، والهدف من ذلك هو دراسة تأثير انتاج النفط والذي يمثل المدخلات $Z(t)$ على الناتج المحلي الاجمالي الذي يمثل المخرجات $y(t)$.

12- التعرف على الانموذج

الشكل الآتي يوضح الرسم البياني للسلسلة $Z(t)$:



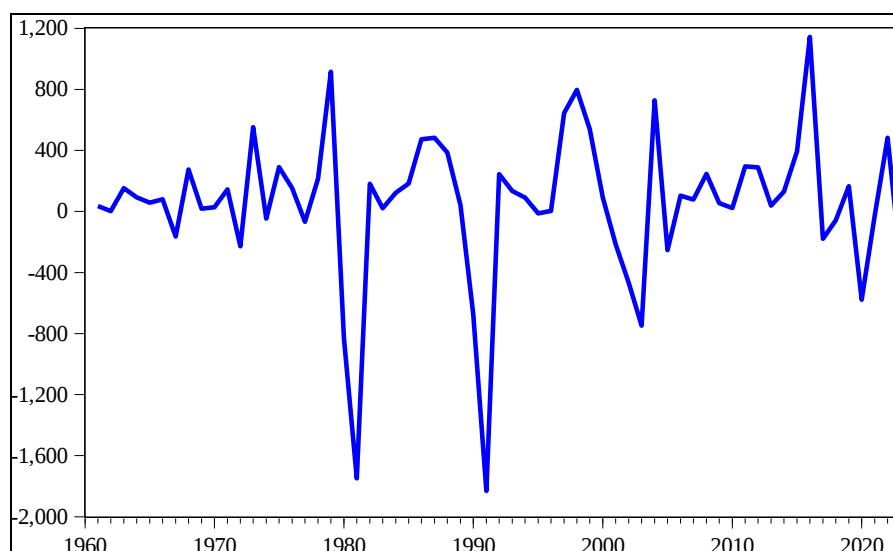
شكل (1): سلسلة $Z(t)$ للفترة من 1960 إلى 2023.

نلاحظ من الشكل (1) أن السلسلة هي متذبذبة ومتصاعدة، ويؤيد هذا الكلام الإحصاءات الوصفية المتعلقة بها، حيث إن الفرق بين أعلى قيمة وأكبر قيمة يعد كبيراً نسبياً، وكما في الجدول الآتي:

جدول(1): الإحصاءات الوصفية للسلسلة $Z(t)$

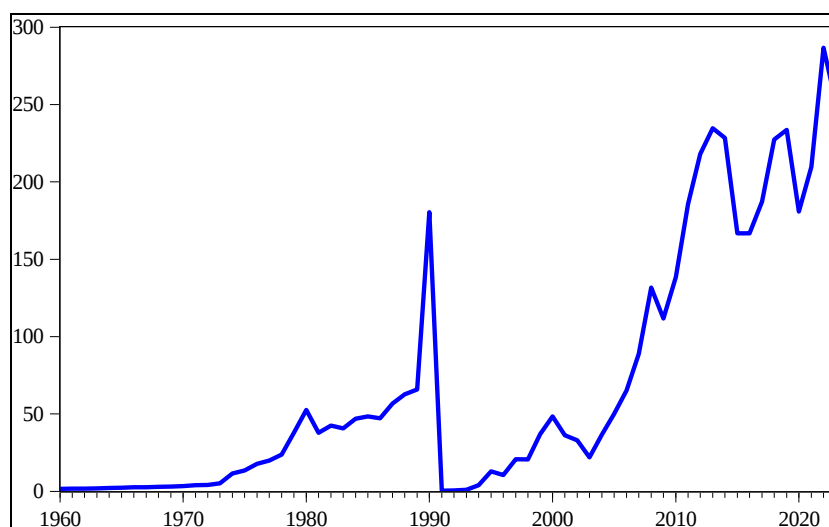
الوسط الحسابي	أعلى قيمة	أقل قيمة	الانحراف المعياري	ديكي فولر	
				المستوى	الفرق الأول
2156.125	4648	282	1110.767	0.5317	0.0000

حيث تدل النتائج أن السلسلة غير مستقرة عند المستوى، بينما تصبح مستقرة عند الفرق الأول بمعنى أنها متكاملة من الدرجة الأولى $I(1)$. وفيما يأتي رسم السلسلة $Z(t)$ بعد أخذ الفرق الأول:



شكل (2): سلسلة $Z(t)$ بعد أخذ الفرق.

بينما يوضح الشكل الآتي الرسم البياني للسلسلة $Y(t)$:



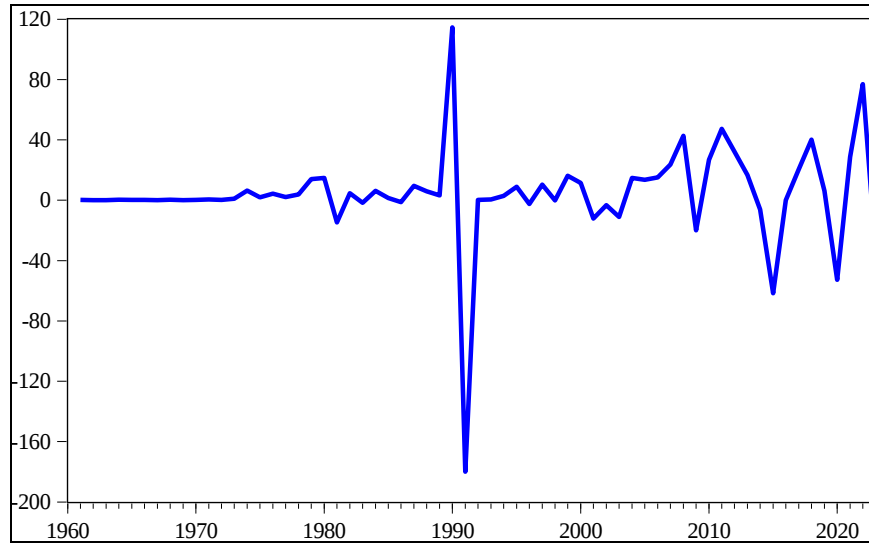
شكل (3): سلسلة $Y(t)$ للفترة من 1960 إلى 2023

نلاحظ من الشكل (3) أن السلسلة في تصاعد مستمر تقريباً، ويتضح ذلك أيضاً من الإحصاءات الوصفية المتعلقة بها، والجدول الآتي يعرض هذه المقاييس:

جدول (2): الإحصاءات الوصفية للسلسلة $Y(t)$

ديكي فولر		الانحراف المعياري	أقل قيمة	أعلى قيمة	الوسط الحسابي
المستوى	الفرق الأول				
0.0000	0.6255	82.34442	0.41	286.64	70.16656

حيث تدل النتائج أن السلسلة Y غير مستقرة عند المستوى، في حين أنها تصبح مستقرة عند الفرق الأول وهذا يعني أنها متكاملة من الدرجة الأولى $I(1)$. وفيما يأتي رسم السلسلة $Y(t)$ بعد أخذ الفرق الأول:



شكل (4) : سلسلة $Y(t)$ بعد أخذ الفرق.

13- تحديد رتبة الانموذج

لتحديد رتبة النموذج تم استعمال كل من معيار أكايك للمعلومات (AIC) ومعيار بيز للمعلومات (BIC) لكل طريقة من طرائق التقدير، وكما في الجدول الآتي:

جدول (3): نتائج معايير تحديد رتبة النموذج

Model	Order			OLS		EIV	
	a	B	q	AIC	BIC	AIC	BIC
ARMAX	1	1	1	432.22	438.60	434.18	440.56
	1	1	2	428.32	436.76	428.49	436.93

	1	1	3	424.40	434.88	425.63	436.11
	1	2	1	428.28	436.72	428.63	437.07
	1	2	2	430.12	440.67	429.51	440.06
	1	2	3	426.25	438.81	426.42	438.99
	1	3	1	424.48	434.95	425.46	435.93
	1	3	2	425.96	438.53	425.86	438.43
	1	3	3	426.53	441.19	427.39	442.05
	2	1	1	428.31	436.75	430.24	438.68
	2	1	2	429.94	440.49	429.87	440.42
	2	1	3	426.25	438.81	427.05	439.62
	2	2	1	428.58	439.13	430.00	440.56
	2	2	2	432.39	445.06	431.51	444.17
	2	2	3	427.82	442.48	428.36	443.02
	2	3	1	425.64	438.21	424.99	437.55
	2	3	2	427.95	442.61	427.04	441.70
	2	3	3	423.80	440.56	428.18	444.94
	3	1	1	424.24	434.71	424.86	435.33
	3	1	2	426.10	438.66	425.97	438.54
	3	1	3	428.66	443.32	428.33	442.99
	3	2	1	425.84	438.41	426.08	438.64
	3	2	2	427.84	442.50	427.33	441.99
	3	2	3	427.86	444.61	430.40	447.16

	3	3	1	428.03	442.69	427.00	441.67
	3	3	2	426.68	443.44	429.09	445.84
	3	3	3	423.03	433.88	429.47	448.32

نلاحظ أن أفضل انموذج مقدر بطريقة OLS هو انموذج (3,1,3) ARMAX، وأفضل نموذج بطريقة EIV هو انموذج (3,1,1) ARMAX.

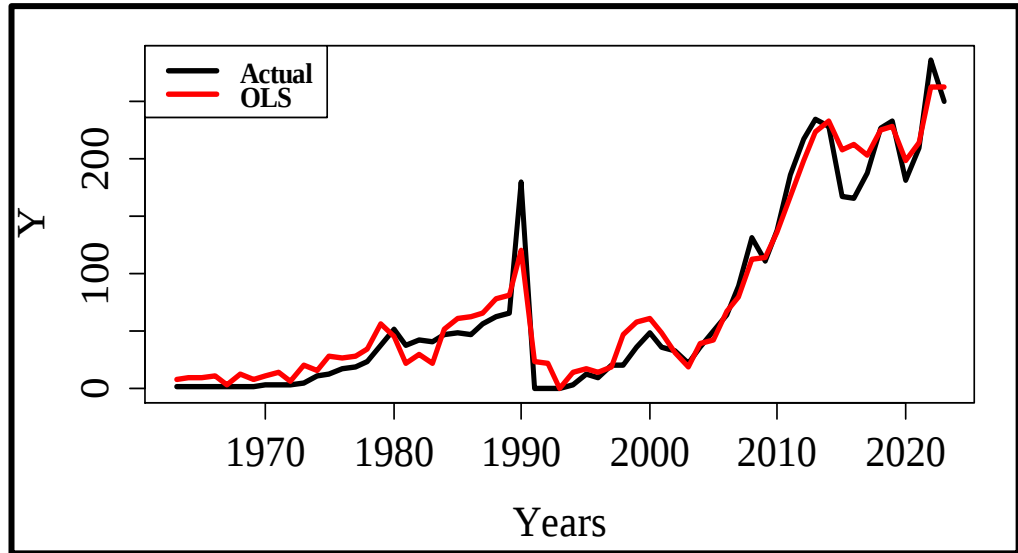
14- مقارنة طرائق التقدير

لتحديد أفضل نموذج من خلال طرائق التقدير ، سيتم إجراء مقارنة من حيث عدة معايير للمقارنة بالنسبة للنماذج المذكورة سابقاً، وكانت نتائج المقارنة كما في الجدول الآتي:

جدول (4): نتائج معايير تحديد أفضل انموذج

Methods	OLS	EIV
RMSE	30.19	30.32
MAPE	6.48	6.57
SMAPE	49.34	48.36

حيث نلاحظ أن أفضل طريقة حسب معايير المقارنة هي طريقة المربعات الصغرى الاعتيادية (OLS)، يليها طريقة العوامل المساعدة الموسعة (EIV)، ولذلك سيتم الاعتماد على طريقة المربعات الصغرى الاعتيادية في تقدير النموذج. حيث تمثل البيانات في الشكل (5) القيم الحقيقية والتقديرية لسلسلة Y.



الشكل (5): القيم الحقيقية والتقديرية لسلسلة Y.

15- التنبؤ

تم استعمال طريقة المربعات الصغرى في التنبؤ لأنموذج $ARMAX(3,1,3)$ حتى عام 2035، وكانت النتائج كما في الجدول الآتي

جدول (5): التنبؤ بالسلسلة الزمنية

Year	Forecasting
2024	236.092
2025	191.2182
2026	191.946
2027	174.938
2028	176.9088
2029	171.6953
2030	174.3801
2031	174.1547
2032	177.2141

2033	179.0924
2034	182.3337
2035	185.0925

13- الاستنتاجات والتوصيات

الاستنتاجات

استنادا إلى ما توصل اليه الجانب التطبيقي نستنتج ما يلي :

- 1- استعمل معيارن لتحديد رتبة الانموذج الافضل ، وهما (AIC و BIC)، وقد أظهرنا توافقاً على أن الرتبة الأفضل للنموذج ARMAX هي (3،1،3) لتمثيل البيانات محل الدراسة.
- 2- أظهرت طريقة المربعات الصغرى تفوقاً على طريقة المتغيرات المساعدة، حيث احتلت المرتبة الأولى وفقاً لمقاييس المقارنة (RMSE، MAPE، SMAPE)، بينما جاءت طريقة المتغيرات المساعدة في المرتبة الثانية.

التوصيات

- 1- نوصي بإجراء مقارنة بين مقدرات نماذج الصندوق الأسود ARX،ARMAX لتحليل أداء كل منهما وتحديد الانموذج الأنسب لتمثيل البيانات.
- 2- نوصي باستعمال طرائق تقدير أخرى ، مثل الأساليب البايزية والطرائق الحصينة، لتحسين عملية تقدير معالم الانموذج وزيادة دقة النتائج.
- 3- نوصي بتطبيق نموذج ARMAX متعدد المدخلات للتنبؤ بالنتائج المحلي الإجمالي، نظراً لارتباطه بالعديد من المتغيرات المؤثرة التي تتطلب تضمينها لتحسين دقة التنبؤ وموثوقيته.

المصادر

- 1- Bisgaard، S.، & Kulahci، M. (2011). Time series analysis and forecasting by example. John Wiley & Sons، Inc.
- 2- Chicco، D.، Warrens، M. J.، & Jurman، G. (2021). The coefficient of determination R-squared is more informative than SMAPE، MAE، MAPE، MSE and RMSE in regression analysis evaluation. PeerJ Computer Science، 7، 623.
- 3- Hastie، T.، Tibshirani، R.، & Friedman، J. (2009). The elements of statistical learning: Data mining، inference، and prediction. (2nd ed.). Springer.

- 4- Hodson, T. O. (2022). Root-mean-square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geoscientific Model Development*, 15(15), 5481–5487.
- 5- Ljung, L. (1999). *System identification: Theory for the user* (2nd ed.). Prentice Hall PTR.
- 6- Lin, H., Zhu, L., Pan, Z., Bai, F., Liu, Y., Lin, L., Patel, M., Fanumtas, E., & Bhatt, N. (2022). ARMAX-based transfer function model identification using wide-area measurement for adaptive and coordinated damping control. *IEEE Transactions on Smart Grid*, 13(1), 777-788.
- 7- Matroushi, S. (2011). *Hybrid computational intelligence systems based on statistical and neural networks methods for time series forecasting: The case of gold price* (Master's thesis). Lincoln University.
- 8- NCSS Statistical Software. (n.d.). *The Box-Jenkins Method* (Chapter 470). Retrieved from <https://ncss.com>.
- 9- Rachad, S., Nsiri, B., & Benassi, B. (2015). System identification of inventory system using ARX and ARMAX models. *International Journal of Control and Automation*, 8(12), 283–294.
- 10- Rasheed, S. A. (2024). *The Role of Oil Industry in The Socioeconomic Status in Iraq*. The University of Basrah, Faculty of Administration and Economics, Department of Economics.
- 11- Salah-eddine, M., Sadik, S., & Bensassi, B. (2020). Microcontroller based data acquisition and system identification of a DC servo motor using ARX, ARMAX, OE, and BJ models. *Advances in Science, Technology and Engineering Systems Journal*, 5(6), 507-513.
- 12- Saidon, M. S., Amin, N. A. M., Sulaiman, A. C., Manan, M. R., Othman, S. M., Mustafa, W. A., & Wahab, N. A. B. (2022). The ARX and ARMAX models for thermoelectric cooling on glass windows: A comparative study. *Journal of Advanced Research in Fluid Mechanics and Thermal Sciences*, 97(1), 105–118.
- 13- Söderström, T., & Stoica, P. (2002). Instrumental variable methods for system identification. *Circuits, Systems, and Signal Processing*, 21(1), 1-9.

14- Wan, L., Beshir, A. I., Wu, X., Liu, X., Grassi, F., Spadacini, G., & Pignari, S. A. (2022). Black-box modeling of converters in renewable energy systems for EMC assessment: Overview and discussion of available models. *Chinese Journal of Electrical Engineering*, 8(2), 13–25.

15- Yiu, C.-M. (2008). Statistical modelling and forecasting schemes for air-conditioning system [Doctoral dissertation, The Hong Kong Polytechnic University]. Pao Yue-kong Library, PolyU Hong Kong.