AL KUT JOURNAL OF ECONOMIC AND ADMINISTRATIVE SCIENCES
Publisher: College of Economics and Management - Wasit University

مقارنة كفاءة الأنموذج الرمادي GM(1,N) مع انماذج (ARIMA) للتنبؤ بالسلاسل الزمنية مع التطبيق

"Comparing the efficiency of the GM(1,N) gray model with ARIMA models for time series forecasting with application"

محمد مؤيد شكر [1] أ.م.د سعد صبر محمد[2]

كلية الإدارة والاقتصاد/جامعة واسط

المستخلص

يتطلب استعمال الأساليب التقليدية للتحليل الإحصائي توفر كميات كبيرة من البيانات، مما يبرز الحاجة إلى تطوير منهجيات جديدة لدراسة الأنظمة التي تعاني من نقص المعلومات، أو حالات عدم اليقين غالبًا ما تواجه هذه الأنظمة تحديات تتعلق بتباين البيانات، وعدم معرفة التوزيع الإحصائي لها خاصة في حالة العينات الصغيرة أو السلاسل الزمنية القصيرة، وفي هذا السياق برزت النظرية الرمادية أداة حديثة لدراسة الأنظمة غير المؤكدة حيث تعتمد على تقنيات مثل (التوليد الرمادي) لتنعيم البيانات و(النمذجة الرمادية)، لتبسيط وتحسين دقة النتائج.

تتناول هذه الدراسة مقارنة بين نموذجين رئيسين: الأول هو الأنموذج الرمادي GM(1,N) الذي يعتمد على المشتقة الأولى (الدرجة التفاضلية) لتمثيل الاتجاهات الزمنية مع إمكانية التعامل مع متغيرات متعددة. يتميز هذا الأنموذج بقدرته العالية على التنبؤ خصوصًا عند التعامل مع عينات صغيرة مما يجعله أداة فعالة لتحليل الاتجاهات المستقبلية في السلاسل الزمنية. أما النموذج الثاني فهو النموذج التقليدي ARIMA يعدُّ من أكثر النماذج شيوعًا في التنبؤ بالسلاسل الزمنية لذلك قامت الدراسة إلى إجراء مقارنة دقيقة بين الأنموذجين لتحديد الأنموذج الأكثر كفاءة في التنبؤ بالقيم المستقبلية.

وقُدرت معلمات الأنموذجين باستعمال طريقة المربعات الصغرى الاعتيادية(OLS) وقيُّمت كفاءتهما باستعمال مقاييس مثل RMSEوMAD فضلاً عن R^2لتحديد الأفضلية وكما تم اعتماد نسبة التباين (Q) لتحديد مدى ملاءمة البيانات لتكون رمادية ولقياس دقة التنبؤ .

طُبقت النماذج على بيانات إنتاج الغاز الطبيعي في العراق للمدة بين 1997 و2022 مع مراعاة العوامل المؤثرة مثل الغاز المستهلك والغاز المحروق واستعملت عينة بحجم 26 مشاهدة لتحليل البيانات واستعين ببرامج إحصائية متقدمة مثل

MATLAB (V-2021) وMinitab (V-2021) وEviews12 لاستخراج النتائج. أظهرت الدراسة تفوق الأنموذج الرمادي GM(1,N) على الأنموذج التقليدي ARIMA مما يبرز فعاليته في تحسين دقة التنبؤ في ظل البيانات المحدودة وغير المستقرة.

الكلمات المفتاحية: النظام الرمادي، الأنموذج الرمادي GM(1,N)، انموذج ARIMA، تحليل البيانات الصغيرة.

Abstract

The application of traditional statistical analysis methods often requires large datasets, emphasizing the need for developing new approaches to study systems characterized by limited information or uncertainty. Such systems frequently face challenges related to data heterogeneity and unknown statistical distributions, particularly in cases of small samples or short time series. In this context, grey theory has emerged as a modern tool for studying uncertain systems, utilizing techniques such as grey generation for data smoothing and grey modeling for simplifying structures and improving accuracy.

This study compares two key statistical models. The first is the GM(1,N) grey model, which relies on the first derivative (differential degree) to represent temporal dynamics while accommodating multiple variables. This model is known for its high predictive accuracy, especially when dealing with small samples, making it an effective tool for analyzing future trends in time series. The second model is the traditional ARIMA model which remains one of the most widely used models for time series forecasting. The study aims to conduct a rigorous comparison between these two models to identify which one performs better in predicting future values.

The parameters of both models were estimated using the Ordinary Least Squares (OLS) method, and their efficiency was evaluated using metrics such as RMSE, MAD, and R^2 additionally the Q variance ratio was employed to assess the suitability of the data for grey modeling and to measure prediction accuracy.

The models were applied to natural gas production data in Iraq for the period 1997–2022, considering influential factors such as gas consumption and gas flaring. A dataset of 26 observations was used for the analysis, and advanced statistical software such as MATLAB (V-2021),Minitab (V-2021),and Eviews12 was utilized for result extraction. The findings demonstrated that the GM(1,N) grey model outperformed the traditional ARIMA model,

highlighting its effectiveness in improving predictive accuracy under conditions of limited and unstable data.

Keywords: gray system, GM(1,N) model, ARIMA model, small data analysis.

1-1 المقدمة Introduction

في السنوات الأخيرة أصبحت الأنظمة الرمادية تحظى بشعبية متزايدة، بوصفها بديلاً للأنظمة التقليدية، فأسس البروفيسور دينج جولونج هذا النظام في عام 1982 في الصين. وقد أسهمت أبحاثه في معالجة الكثير من القضايا مما أدى إلى انتشار هذه النظرية أولاً في الصين ثم في دول أخرى. شهدت أبحاث الأنظمة الرمادية زيادة ملحوظة في التسعينيات والعقد الأول من القرن الحادي والعشرين وأصبحت تُستخدم في مجالات متنوعة مثل الاقتصاد والهندسة والعلوم البيئية والاجتماعية.

تُسمى هذه النماذج "نماذج رمادية" بسبب طبيعة عملها الغامضة وغير الدقيقة حيث في المراحل الأولى من التحليل يتم التعامل مع البيانات بطريقة غير واضحة حيث يُستخدم اللون الأسود للإشارة إلى المعلومات غير المعروفة واللون الأبيض للمعلومات المعروفة بالكامل واللون الرمادي للمعلومات المعروفة جزئيًا او فيها عدم تأكد (عدم اليقين).

2-1 مشكلة البحث Problem of the Research

تواجه الكثير من الدراسات تحديات كبيرة في التنبؤ بالاتجاهات المستقبلية بدقة مما يؤثر سلبًا على اتخاذ القرارات الاستراتيجية ان هذه التحديات تتفاقم بشكل خاص عند التعامل مع مجموعات بيانات صغيرة التي تعاني من عدم الاستقرارية حيث تكون دقة التنبؤات عادة أقل من هنا ظهرت مشكلة وهي كيفية التنبؤ في مثل هكذا بيانات.

3-1 هدف البحث Aim of the Research

يهدف البحث الى استعمال النماذج الرمادية وبالتحديد الأنموذج GM(1,N) في التنبؤ بالبيانات المتاحة التي تعاني من قلة المعلومات وعدم الاستقرارية أو الاضطرابات ومن ثم نقوم بمقارنة أداء الأنموذج الرمادي GM(1,N) مع أداء النماذج التقليدية مثل ARIMA من خلال تطبيق هذه النماذج على بيانات حقيقية تتعلق بإنتاج الغاز الطبيعي. يهدف البحث إلى تحديد أي من الأنموذجين يوفر دقة تنبؤ أعلى في ظل الظروف الصعبة المرتبطة بالبيانات المحدودة.

الجانب النظري

1-2 الأنموذج الرمادي GM(1,N):

الأنموذج الرمادي GM(1,N) هو انموذج تنبؤي يعتمد على نظرية الأنظمة الرمادية حيث يمثل انموذجًا من الدرجة الأولى يتضمن متغيرًا تابعًا واحدًا وأكثر من متغير مستقل (N) يعدُّ هذا الأنموذج مناسبًا للعينات الصغيرة أو البيانات التي تفتقر إلى المعلومات الكاملة.[13]

المبدأ الأساسي لهذا الأنموذج هو إقامة علاقة بين بيانات المدخلات وبيانات المخرجات من خلال تطوير معادلة تفاضلية بعد إجراء عملية التوليد التراكمي يقوم انموذج GM(1,N) بتقليل تأثير العشوائية وعدم الانتظام في سلسلة البيانات الأصلية مما

يؤدي إلى تحسين تسلسل البيانات تدريجيًا، ومن ثمَّ، تتميز البيانات الناتجة بقاعدة نمو متزايدة مما يساعد على توضيح قيم التطوير والتغير المستقبلية. [29]

بمجرد إنشاء المعادلة التفاضلية يتم حلها باستعمال طريقة المربعات الصغرى الاعتيادية (Ordinary Least Squares - OLS) من خلال هذه الطريقة يتم الحصول على معامل α الذي يعكس اتجاه التغيير في السلسلة $Y^{(1)}$وعمود البيانات الأصلية $Y^{(0)}$ كما يتم الحصول على معامل β الذي يعكس العلاقات المتغيرة بين البيانات ولتي تسمح بالتنبؤ بسلوك النظام في المستقبل. [31]

تمثل المعادلة التفاضلية من الدرجة الأولى الأساس الذي يُستند إليه في نمذجة $y^{(1)}$ ويمكن تلخيص خطوات هذا النموذج بشكل منهجي لتقديم رؤية شاملة حول كيفية استعماله في التنبؤ.

لتكن السلسلة الأصلية

$$Y_i^{(0)} = \left\{Y_i^{(0)}(1), Y_i^{(0)}(2), \dots, Y_i^{(0)}(m)\right\}, \quad (i = 1,2, \dots, N) \qquad (1)$$

حيث يمثل (i) هنا الى عداد المتغيرات لسلسلة بيانات النظام في عينة n، حيث يمثل (0) لـ $Y_1^{(0)}$ العنوان السلسلة الاصلية و(i=1) يمثل المتغير المعتمد وعندما (i=2,3,…,N) فهي تمثل المتغيرات المستقلة ويمكن كتابتها بصوره أدناه:[31]

$$Y_1^{(0)} = \left(y_1^{(0)}(1), y_1^{(0)}(2), \dots, y_1^{(0)}(n)\right) \qquad (2)$$

$$Y_2^{(0)} = \left(y_2^{(0)}(1), y_2^{(0)}(2), \dots, y_2^{(0)}(n)\right) \qquad (3)$$

$$Y_3^{(0)} = \left(y_3^{(0)}(1), y_3^{(0)}(2), \dots, y_3^{(0)}(n)\right) \qquad (4)$$

$$\vdots\ Y_N^{(0)} = \left(y_N^{(0)}(1), y_N^{(0)}(2), \dots, y_N^{(0)}(n)\right) \qquad (5)$$

بعد تحديد متغيرات الدراسة (المتغيرات التوضيحية والمتغير المعتمد)، يتم تطبيق المراحل الآتية للأنموذج: (عملية التوليد المتراكمة، تحديد معلمات، التنبؤ باستعمال عملية التوليد المتراكمة العكسية).

عملية التوليد المتراكمة: -

بعد إجراء عملية التوليد من الدرجة الأولى حيث يُرمز بـ ($Y_i^{(1)}$) للسلسلة التي تم الحصول عليها وفق المعادلة الآتية:[17]()

$$Y_i^{(1)} = \left\{Y_i^{(1)}(1), Y_i^{(1)}(2), \dots, Y_i^{(1)}(n)\right\} \qquad (6)$$

ويعبر عنها بصيغة اخرى

$$Y_1^{(1)} = \left(y_1^{(1)}(1), y_1^{(1)}(2), \dots, y_1^{(1)}(n)\right) \quad (7)$$

$$Y_2^{(1)} = \left(y_2^{(1)}(1), y_2^{(1)}(2), \dots, y_2^{(1)}(n)\right) \quad (8)$$

$$Y_3^{(1)} = \left(y_3^{(1)}(1), y_3^{(1)}(2), \dots, y_3^{(1)}(n)\right) \quad (9)$$

$$\vdots$$

$$Y_N^{(1)} = \left(y_N^{(1)}(1), y_N^{(1)}(2), \dots, y_N^{(1)}(n)\right) \quad (10)$$

ويعبر عنها بصيغة أخرى[21,30]

$$Y_i^{(1)} = \sum_{i=1}^{j} y_i^{(0)}(j) \qquad ,j = 1,2,\dots,n \quad (11)$$

وتحديد معلمات الأنموذج كمايأتي: -

لتكن للمعادلة التفاضلية للأنموذج GM(1,N) كما يأتي:[19]

$$\frac{dy_1^{(1)}(j)}{dt} + \alpha y_1^{(1)}(j) = \sum_{i=2}^{n} \beta_i y_i^{(1)}(j) \quad (12)$$

انطلاقا من المعادلة السابقة يمكن الحصول على المعادلة التفاضلية النهائية للأنموذج الرمادي GM(1,N):[14]

$$Y_1^{(0)}(j) + \alpha z_1^{(1)}(j) = \sum_{i=2}^{n} \beta_i y_i^{(1)}(j) \quad (13)$$

حيث $Z_1^{(1)}(k)$ يعبر عليها بالعبارة الآتية:[25]

$$Z_1^{(1)}(j) = \frac{y_1^{(1)}(j) + y_1^{(1)}(j-1)}{2} \quad (14)$$

ويمكن إعادة صياغة المعادلة (13) كما يأتي: [24]

$$Y_1^{(0)}(j) = -\alpha z_1^{(1)}(j) + \sum_{i=2}^{n} \beta_i y_i^{(1)}(j) \quad , \quad i = 2,3,\dots,N \quad ,j = 2,3,\dots,n \quad (15)$$

ويمكن التعبير عنها ايضاً كما يأتي:

$$y_1^{(0)}(2) = -\alpha z_1^{(1)}(2) + \beta_2 y_2^{(1)}(2)$$

$$y_1^{(0)}(3) = -\alpha z_1^{(1)}(3) + \beta_3 y_3^{(1)}(3)$$

$$y_1^{(0)}(n) = -\alpha z_1^{(1)}(n) + \beta_N y_N^{(1)}(n)$$

- حيث تشير (α) يمثل معامل التطوير Developed Coefficient الذي يعكس اتجاه التطور لسلسلة $Y^{(1)}$ والبيانات الأساسية $Y^{(0)}$.

- وتشير (β_i) يمثل معلمة التنسيق Coordination Parameter التي تعكس التغيرات في السلسلة الأصلية $Y^{(0)}$.

تمثل قيمة (β_i) مدى تأثير قيمة المتغير المستقل عند النقطة (i) على المتغير التابع إذا كانت (β_i) أكبر من (0)، فهذا يعني أن كل زيادة في المتغير المستقل (i) تؤثر إيجابًا على المتغير التابع أما في حالة العكس عندما تكون(β_i) أصغر من الصفر فإن أي زيادة في المتغير المستقل (i) تؤثر سلبًا على المتغير التابع حيث يتم تقدير معلمات الأنموذج الرمادي باستعمال طريقة المربعات الصغرى. [28]

يمكننا صياغة المعادلة (13) على شكل المصفوفات وكما يأتي:

$$\underline{\hat{\gamma}} = \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = (C'C)^{-1} C' \underline{Y_i} \qquad (16)$$

حيث:

$$\underline{\hat{\gamma}} = \left[\hat{\alpha}, \hat{\beta}_2, \hat{\beta}_3, \dots, \hat{\beta}_N\right]^T, Y_i = \left[y_1^{(0)}(2) y_1^{(0)}(3), \dots, y_1^{(0)}(n)\right]^T$$

حيث يمكن التعبير عن C بالمصفوفة الآتية:[15]

$$C = \begin{bmatrix} -z_1^{(1)}(2) & y_2^{(1)}(2) & \dots & y_N^{(1)}(2) \\ -z_1^{(1)}(3) & y_2^{(1)}(3) & \dots & y_N^{(1)}(3) \\ \vdots & \vdots & \vdots & \vdots \\ -z_1^{(1)}(n) & y_2^{(1)}(n) & \dots & y_N^{(1)}(n) \end{bmatrix}$$

والتنبؤ بالقيم المستقبلية وذلك اعتمادا على دالة التوليد العكسية بعد الحصول على المعلمات α و β يمكن كتابة معادلة التنبؤ الخاصة بالأنموذج الرمادي GM(1,N) على النحو الآتي:[17]

$$\hat{y}^{(1)}(j+1) = \left(y_1^{(0)}(1) - \sum_{i=2}^{n-1} \frac{\hat{\beta}_i y_i^{(1)}(j+1)}{\hat{\alpha}} \right) e^{-\alpha t} + \sum_{i=2}^{n-1} \frac{\hat{\beta}_i y_i^{(1)}(j+1)}{\hat{\alpha}} \qquad (17)$$

وبالاعتماد على المعادلة العكسية لدالة التوليد يتم الحصول على القيمة المتنبأ بها عند 1 + j وذلك من خلال العبارة الآتية:[11]

$$\hat{y}^{(0)}(j+1) = \hat{y}^{(1)}(j+1) - \hat{y}^{(1)}(j) \qquad ,j = 1,2,\dots,n \qquad (18)$$

2-2 نماذج بوكس-جنيكنز Box-Jenkins Methods:

منهجية Box-Jenkins هي طريقة منهجية لتحليل السلاسل الزمنية والتنبؤ بها. ويركز في درجة الأولى على تحديد وتقدير نماذج بيانات السلاسل الزمنية، وخاصة تلك التي يمكن التعبير عنها كنماذج المتوسط المتحرك المتكامل الذاتي (ARIMA)[3] تم اكتشاف نماذج بوكس-جينكنز (Box-Jenkins) في المدة ما بين السنوات (1970-1976) وتُرمز لها اختصارًا بـ (B-J M) تستعمل في تحليل السلاسل الزمنية في مجال الزمن وتنقسم الى قسمين الأولى هي النماذج الموسمية Seasonal Modelsوسوف لن ندخل في تفاصيلها لاعتماد النماذج من النوع الثاني، وهي النماذج اللاموسمية Non seasonal Models وتستعمل لتمثيل نوعين

من السلاسل هما:[26]

1-2-2 انماذج السلاسل الزمنية المستقرة:

تعرف هذه النماذج من السلاسل الزمنية التي تتمتع بخاصية الاستقرارية أي لا يوجد فيها اتجاه عام في بياناتها ولها وسط حسابي ثابت تذبذب حوله وتتضمن هذه النماذج ثلاثة أنواع:

1- نماذج الانحدار الذاتي (AR): Autoregressive **Methods**

يعرف الانحدار الذاتي على انه الارتباط بين المشاهدات الحالية لسلسلة الزمنية مع المشاهدات السابقة لنفس السلسلة ويرمز لدرجة الأنموذج (P) يمثل بعدد صحيح ويمكن تمثيل نماذج الانحدار الذاتي بدرجة إبطاء AR (p) كما يأتي:[26]

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t \qquad (19)$$

فمثلاً انموذج الانحدار الذاتي من الدرجة الأولى (AR $_{(1)}$) وهو:

$$Y_1 = \phi_1 Y_{t-1} + e_t \qquad (20)$$

وإذا كانت Y_t تمثل انحراف المشاهدات عن المتوسط μ فتكون المعادلة بالشكل الآتي:

$$(Y_t - \mu) = \phi_1 (Y_{t-1} - \mu) + e_t$$

$$Y_t = (1 - \phi_1)\mu + \phi_1 Y_{t-1} + e_t \qquad (21)$$

شرط السكون لمعاملة الانحدار الذاتي (ϕ_1) :

ان قيمة $\boldsymbol{\phi_1}$ يجب ان تكون أقل من الواحد الصحيح لتكون السلسلة ساكنة:

$$-1 < \phi < 1$$

حيث ان معامل الارتباط الذاتي ρ_k لهذا انموذج:

$$\rho_k = \phi_1^k \qquad (22)$$

وان معامل الارتباط الجزئي $\boldsymbol{\phi_{kk}}$ يكون مساوي لمعلمة الانحدار الذاتي AR(1) بينما معاملات الارتباط الجزئي الأخرى تساوي صفر.[5]

$$(Y_t - \mu) = \phi_1(Y_{t-1} - \mu) + e_t$$

$$Y_t = (1 - \phi_1)\mu + \phi_1 Y_{t-1} + e_t \qquad (23)$$

اما انموذج الانحدار الذاتي من الدرجة الثانية ($AR_{(2)}$) وهو:

يمكن كتابة الأنموذج على النحو الآتي:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t \qquad (24)$$

وان دالة الارتباط الذاتي لهذا الأنموذج:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \qquad (25)$$

وشرط الاستقرار في هذه الدالة هو:

$$-1 < \phi < 1$$

بينما تتضاءل معلمات الارتباط الذاتي الجزئيPACF للأنموذج AR (2) بعد التأخير الثاني [26]

$$\phi_{11} \neq 0$$

$$\phi_{22} \neq 0$$

حيث إن $\phi_{22} \neq 0$ عند التأخير الثاني $k > 2$ حيث يمكن التعبير عن كل رمز كالاتي:

Y_t : مشاهدات السلسلة الزمنية في الزمن t

ϕ : معالم النموذج

P : رتبة الانحدار الذاتي

$\boldsymbol{\mu}$ **:** متوسط المشاهدات.

2- نماذج الأوساط المتحركة Moving Average Methods (MA):

يمكن تمثيل الظاهرة المدروسة بنموذج اخر غير انموذج الانحدار الذاتي هو نموذج الأوساط المتحركة والفرق ما بين هذا الأنموذج والأنموذج السابق بأن الأوساط المتحركة تعتمد على حدود الأخطاء في أزمان مختلفة وان نماذج الأوساط المتحركة من الرتبة (q) يرمز بالرمز ((MA(q)) يمكن أن يمثل بالنموذج الرياضي الاتي:[5,23]

$$Y_t = -e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \qquad (26)$$

عندما يتم إضافة معلمات المتوسط المتحرك لمعادلة الارتباط الذاتيACF ، تصبح المعادلة كلاتي:

$$\rho_k = \frac{\left(-\theta_k + \theta_1\theta_{k+1} + \ldots + \theta_{q-k}\theta q\right)}{(1 + \theta_1^2 + \ldots + \theta_1^2)} \qquad (2-79)$$

عندما:

$K \leq q$

$\rho_k = 0$,$K > q$

وان دالة الارتباط الذاتي الجزئي PACF لا تنعدم معلماتها إلى الصفر مطلقاً (حيث تتناقص بشكل أسّي).

فمثلاً انموذج الأوساط المتحركة من الدرجة الأولى ((MA (1) وهو: [6,10]

$$Y_t = e_t - \theta_1 e_{t-1} \qquad (27)$$

ان عمل دالة الارتباط الذاتي ودالة الارتباط الذاتي الجزئي في انموذج MA(1) مشابه عمل انموذج AR(1)

وان نموذج الأوساط المتحركة من الدرجة الثانية ((MA (2) وهو:

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} \qquad (28)$$

حيث يمكن صياغة معادلة دالة الارتباط الذاتي ACF الأنموذج MA(2) كما يأتي:

$$\rho_k = \frac{-\theta^2}{(1 + \theta_1^2 + \theta_2^2)} \qquad (29)$$

عندما:

$k \leq 2$

$\rho_k = 0$, $k > 2$

أما معادلة دالة الارتباط الجزئيPACF كما يأتي:

$$\phi_{kk} = \frac{-\theta_1^k(1-\theta_1^2)}{\left(1+\theta_1^{2^{(k+1)}}\right)} \qquad (30)$$

ويمكن التعبير عن الرموز أعلاه:

Y_t : مشاهدات السلسلة في الزمن t

θ : معالم الأنموذج

q : رتبة الأوساط المتحركة

3- نماذج الانحدار الذاتي والأوساط المتحركة

Autoregressive Moving Average Methods ARMA

أحياناً من الممكن تمثيل الظاهرة بأنموذج انحدار ذاتي مختلط من الرتبة (p) واوساط متحركة من الرتبة (q) للحصول على ما يسمى بالأنموذج المختلط يرمز له بـ ((ARMA(p,q) وكما مبين أدناه:[32،20]

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \qquad (31)$$

فمثلاً انموذج ARMA(1,1) هو

$$Y_t = \phi_1 Y_{t-1} + e_t - \theta_1 e_{t-1} \qquad (32)$$

وهذا ناتج عن عملية الدمج بين الأنموذجين AR(1) و MA(1) ، مع أخذ دالة الارتباط الذاتي ACF للنموذج:

$$\rho_k = \frac{(1-\phi_1)(\phi_1-\theta_1)}{1+\phi_1^2-2\phi_1\theta_1)} \qquad (33)$$

وان صيغة دالة الارتباط الجزئي PACF كما يأتي:

$$\phi_{kk} = \phi_1\, \rho_{k-1} \qquad (34)$$

حيث إن: $k \geq 1$

2-2-2 نماذج السلاسل الزمنية غير المستقرة:

وتنطبق هذه النماذج لتمثيل السلسلة بالتي يكون الاتجاه العام أحد مكوناتها مما يجعل لها عدة أوساط تتذبذب حولها البيانات وهذه النماذج يمكن تمثيلها بنماذج النوع الأول لكن بعد إزالة عدم الاستقرارية من السلاسل الاصلية باستعمال طريقة الفروق

ومن خلال اخذ عدد مناسب من الفروق للبيانات الاصلية ولفترات زمنية تمثل طول الموسم ويضاف لرموز حرف (d) [26] عند إدخال معامل الفروق d إلى انموذج $ARMA_{(p,q)}$ يتحول الأنموذج إلى انموذج ARIMA(p,d,q) وتتضمن هذه النماذج نفس نماذج السلاسل الزمنية المستقرة ولكن لتتميز بين هذا النوع من النماذج الأولى تضاف كلمة (Integrated) الى اسم النماذج للدلالة على استعمال نماذج السلاسل الزمنية المستقرة على السلاسل الزمنية غير المستقرة بعد تحويلها الى السلاسل الزمنية ولتمثيل هذا النوع من النماذج يفضل استعمال النموذج المختلط بعد إضافة الحرف (I) ليصبح رمز النموذج (ARIMA) ويصبح الأنموذج من الدرجة (p,d,q)، إن عملية تحويل السلاسل الزمنية غير المستقرة الى أخرى مستقرة سيجنبنا كثيراً من المشاكل إذ إن الخصائص الزمنية المستقرة مستقلة عن الزمن، فضلاً عن ذلك يكون من السهل تطوير نظرية الاستقرارية لتشمل السلاسل الزمنية الاحتمالية أي تكون السلاسل الجديدة التي يتم التنبؤ بها مستقرة أيضاً، اما من ناحية العملية فالحسابات الخاصة بتقدير معالم النماذج المستقرة تحتاج الى عمل أقل ويعرف الأنموذج كلاتي:[33 1]

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \qquad (35)$$

ولجعل السلسلة مستقرة في هذا الأنموذج وعلى افتراض ان Y_t تمثل بيانات السلسلة الزمنية الأصلية و Z_t تمثل بيانات السلسلة الزمنية بعد أخذ الفرق الأول فإن معادلة احتساب الفرق الأول تكون كالآتي:

$$Z_t = \nabla Y_t - Y_{t-1}$$

$$= Y_t - \beta Y_t$$

$$= (1 - \beta)\mathrm{Y_t} \qquad (36)$$

أي أن قيمة السلسلة الزمنية المستقرة $\mathrm{Z_t}$ في الزمن t تساوي الفرق بين قيمة السلسلة الأصلية $\mathrm{Y_t}$ في الزمن t

وقيمتها $\mathrm{Y_{t-1}}$ في الزمن السابق t-1.

أي ان

Z_t : تمثل بيانات السلسلة الزمنية بعد أخذ الفرق الأول $\nabla = (1 - \beta)$ ، فإن معادلة احتساب الفروق من الرتبة الثانية تكون كالآتي:[10]

$$Z_t = \nabla 2 Y_t - \nabla Y_t - \nabla Y_{t-1}$$

$$= Y_t - Y_{t-1} - Y_{t-1} + Y_{t-2}$$

$$= Y_t - 2Y_{t-1} + Y_{t-2}$$

$$= Y_t - 2\beta Y_t + \beta^2 Y_t$$

$$= (1 - 2\beta + \beta^2)Y_t$$

$$= (1 - \beta)^2 Y_t \qquad (37)$$

إذا كانت d عددًا من الفروق اللازمة للحصول على سلسلة زمنية مستقرة، فإن المعادلة العامة للسلسلة المستقرةZ_t تكون كما يأتي:

$$Z_t = \nabla^d Y_t = (1 - \beta)^d Y_t \qquad (38)$$

حيث ان :

Z_t :تمثل السلسلة الزمنية المستقرة بعد أخذ d فروق.

Y_t :تمثل السلسلة الزمنية الأصلية.

∇^d :يمثل عملية أخذ d فروق من السلسلة الأصلية.

$(1 - \beta)^d$:يمثل عملية المرحلة المتأخرة (Lag Operator) المطبقة d مرات.

ومن ثمَّ،، فإن السلسلة الزمنية Y_t تصبح مستقرة إذا تم أخذ d فروق منها للحصول على السلسلة Z_t حيث يمثل d أقل عدد من الفروق اللازمة للحصول على سلسلة زمنية مستقرة في المتوسط. [27]

3-2 مراحل تطبيق نماذج بوكس - جينكنز في التنبؤ:

1- مرحلة التعريف والتشخيص (Identification & Diagnostic):

في هذه المرحلة تحديد رتب (p,d,q) لأنموذج ARIMA؛ وذلك من خلال مراقبة دالتي الارتباط الذاتي (ACF) والارتباط الذاتي الجزئي (PACF).[7]

- إذا كان شكل الارتباط الذاتي (ACF) يقع داخل حدود فترة الثقة 95% من البداية، فإن معامل الارتباط الذاتي ρ_k لا يختلف جوهرياً عن الصفر. هذا يعني أن سلسلة البيانات مستقرة ومتكاملة من الرتبة صفر ويمكن إجراء التحليلات على القيم الأصلية للمتغير Y_t دون إجراء أي تحويلات. [10]
- أما إذا كان شكل انتشار دالة الارتباط الذاتي (ACF) لا يقع داخل حدود فترة الثقة 95% لعدد من الفجوات الزمنية، فإن ρ_k يختلف جوهرياً عن الصفر، مما يشير إلى أن السلسلة الزمنية غير مستقرة وتحتاج إلى أخذ الفروق لتصبح سلسلة مستقرة عند أخذ الفرق الأول يصبح $d = 1$ وعند أخذ الفرق الثاني يصبح $d = 2$ في انموذج ARIMA(p,d,q).

يمكن تحديد p , q من خلال النظر إلى شكل انتشار دالة الارتباط الذاتي (ACF) ودالة الارتباط الذاتي الجزئي (PACF):

- عندما تنخفض الارتباطات الذاتية بصورة أسية إلى الصفر، فهذا يعني وجود انموذج انحدار ذاتي (AR) تتحدد رتبته من عدد الارتباطات الذاتية الجزئية التي تختلف معنوياً عن الصفر.
- إذا انخفضت الارتباطات الذاتية الجزئية بصورة أسية إلى الصفر، فإن الأنموذج هو نموذج MA، وتتحدد رتبته من عدد الارتباطات الذاتية ذات الدلالة الإحصائية.
- إذا انخفضت كلا الارتباطات الذاتية والذاتية الجزئية إلى الصفر بصورة أسية فإن هذا يشير إلى انموذج ARMA.

يتم تحديد رتبة كل من AR وMA من خلال مراقبة انتشار دالة الارتباط الذاتي والارتباط الذاتي الجزئي إذا كانت دالة الارتباط الذاتي لا تتنازل بسرعة مع زيادة درجات الإبطاء فهذا يعني أن السلسلة الزمنية غير مستقرة وتحتاج إلى أخذ الفروق.

2-مرحلة التقدير والاختبار(Estimation & Testing):

بعد مرحلة تشخيص الأنموذج وتحديد رتبته يتم تقدير المعلمات هذا النموذج. التي تمثل المعلمات

$(\phi_i , i = 1 , 2 , \ . \ . \ . , p)$ في نموذج AR(p) و$(\theta_i , i = 1 , 2 , \ . \ . \ . , q)$ في نموذجMA(q)

ϕ_i وθ_i في انموذج ARMA(p,q) تُعتبر طريقة المربعات الصغرى (Ordinary Least Squares Estimation - OLS) إحدى الطرائق الكفؤة للتقدير تتمثل الفكرة الأساسية لهذه الطريقة في تقليل مجموع مربعات الأخطاء إلى أدنى حد ممكن حيث تُفضل هذه الطريقة على غيرها لسهولة تقدير المعلمات ويمكن الوصول إلى هذه الطريقة من خلال الخطوات الآتية:[4]

إذا كان الأنموذج AR(1):

$$Y_t = \phi Y_{t-1} + e_t \qquad (39)$$

$$S = \sum_{t=2}^{T} e_t^2 = \sum_{t=2}^{T} (Y_t - \phi_1 Y_{t-1})^2 \qquad (40)$$

$$\frac{\partial S}{\partial \phi_1} = \frac{1}{\hat{\sigma}^2} \sum_{t=2}^{T} (Y_t - \hat{\phi}_1 Y_{t-1})\, Y_{t-1} = 0$$

$$\hat{\phi}_1 = \sum_{t=2}^{T} \frac{Y_t Y_{t-1}}{Y_{t-1}} \qquad (41)$$

$$\hat{\sigma}^2 = \frac{1}{T-2} \sum_{t=2}^{T} (Y_t - \hat{\phi}_1 Y_{t-1})^2 \qquad (42)$$

- **إذا كان الأنموذج MA(1):**

$$S = S(\theta_1) = \sum_{t=q+1}^{T} e_t^2 \qquad \ldots minimum \qquad (43)$$

$$S = \sum_{t=q+1}^{T} (Y_t + \theta_1 e_{t-1})^2 \qquad (44)$$

$$\frac{\partial S}{\partial \theta_1} = 0 \quad for \ i = 1,2, \dots, q$$

$$\frac{\partial S}{\partial \theta_1} = 2 \sum_{t=q+1}^{T} (Y_t + \hat{\theta}_1 e_{t-1})\, e_{t-1} = 0 \qquad ; \frac{\partial S}{\partial \theta_1} = 0$$

$$\hat{\theta}_1 = \frac{-\sum_{t=q+1}^{T} Y_t e_{t-1}}{\sum_{t=q+1}^{T} e_{t-1}^2} \qquad (45)$$

$$\hat{\sigma}_e^2 = \frac{1}{T-q} \sum_{t=q+1}^{T} e_t^2 \qquad (46)$$

3- مرحلة التنبؤ(Forecasting):

يعدُّ التنبؤ من الأهداف المهمة في العلم ويأتي دوره بعد فهم الظاهرة موضوع البحث وتفسيرها وتحديد العلاقة بينها وبين الظواهر الأخرى المحيطة بها والمقصود بالتنبؤ هو قدرة الباحث على وضع صورة مستقبلية للظاهرة مستندًا إلى دراسة الواقع الحالي بعد تحديد رتب الأنموذج (p,d,q) وتقديرها يتم استعمال الأنموذج الملائم للتنبؤ حيث يتم ذلك من خلال إدخال القيم الحالية والسابقة للمتغير التابع Y_t والبواقي e_tكقيم تقديرية لحساب القيمة المستقبلية الأولى المتنبأ بها Y_{t+1} ، التي تُعرف بالتنبؤ لفترة مستقبلية واحدة.[9]

4-2 جودة الملاءمة

1- اختبار الاستقرار Dickey-Fuller Test

هو اختبار يُستخدم في تحليل السلاسل الزمنية لتحديد ما إذا كانت السلسلة الزمنية تحتوي على جذر أحادي (unit root)، مما يعني أنها غير مستقرة فهي رمادية اما إذا كانت القيمة p أقل من 0.05 فهذا يشير إلى أن البيانات مستقرة ومن ثمَّ، ليست رمادية.[2]

$$\Delta Y_t = \alpha + \beta t + \emptyset y_{t-1} + \varepsilon_t \qquad (47)$$

حيث ان:ΔY_t : هو التغير في القيمة الحالية للسلسلة الزمنية

α : هو الحد الثابت

βt : هو مصطلح الاتجاه الزمني (يمكن استعماله في حالة وجود اتجاه)

$\emptyset$: هو معامل الجذر الأحادي

y_{t-1} : هو قيمة السلسة في الزمن السابق

ε_t : هو حد الخطأ العشوائي

5-2 اختبار دقة الأنموذج:

لاختبار دقة النماذج المستعملة وذلك من خلال:

2-5-1-اختبار التباين اللاحق Posterior-variance-test

يعد اختبار التباين اللاحق أداة مهمة في تقييم جودة النماذج الرمادية حيث يساعد في تحديد مدى دقة وموثوقية التوقعات الناتجة عن الأنموذج ويمكن حسابة من خلال:[8]

حساب المتوسط للتسلسل الأصلي $\boldsymbol{Y^{(0)}}$:

$$\bar{Y}^{(0)} = \frac{1}{n}\sum_{j=1}^{n} y^{(0)}(j) \qquad (48)$$

حساب التباين للتسلسل الأصلي $\boldsymbol{Y^{(0)}}$:

$$S_1 = \sqrt{\frac{\sum_{j=1}^{n}[y^{(0)}(j)-\bar{Y}^{(0)}]^2}{n-1}} \qquad (49)$$

يتم قياس جودة الأنموذج حسب المعايير الآتية:[22,12]

جدول رقم (1) معايير قياس جودة الأنموذج

نسبة تأهيل التباين Q	
درجة دقة التنبؤ	نسبة التباين (Q)
ممتاز	$0.35 \geq Q$
كفؤ	$0.50 \geq Q > 0.35$
بالكاد مؤهل	$0.65 \geq Q > 0.50$
غير مؤهل	$Q > 0.65$

ولحساب انحراف الخطأ المتبقي:

$$S_2 = \sqrt{\frac{\sum_{j=1}^{n}[\varepsilon^{(0)}(j)-\overline{\varepsilon}^{(0)}]^2}{n-1}} \qquad (50)$$

حيث ان:

$$\varepsilon^{(0)}(j) = |y^{(0)}(j) - \hat{y}^{(0)}(j)| \quad j = 1,2,\dots,n \qquad (51)$$

وان متوسط المتبقي يحسب المعادلة الاتية:

$$\overline{\varepsilon}^{(0)} \frac{1}{n}\left[\varepsilon^{(0)}(j)\right] \qquad (52)$$

حيث يشير (Q) إلى نسبة التباين ويتم حسابه حسب الصيغة الآتية:

$$Q = \frac{S_2}{S_1} \qquad (53)$$

6-2 معايير المفاضلة بين أكثر من نموذج:

1- **الجذر المتوسط المربع الخطأ ((Root Mean Square Error (RMSE):** هو عبارة عن جذر متوسط مربعات الفرق بين القيم الفعلية والقيم المقدرة وهو مقياس للدقة وجودة المقدر وتعطى بالشكل الآتي:[31,18]

$$RMSE = \sqrt{\sum_{j=1}^{n} \frac{(y^{(0)}(j) - \hat{y}^{(0)}(j))^2}{n}} \qquad (54)$$

2- **متوسط الخطأ المطلق (Mean Absolute Deviation - MAD):** هو عبارة عن مطلق متوسط مربعات الفرق بين القيمة الفعلية والقيمة المقدرة وهو مقياس للدقة وجودة المقدر وتعطى بالشكل الآتي:[28]

$$MAD = \sum_{j=1}^{n} \left| \frac{(y^{(0)}(j) - \hat{y}^{(0)}(j))^2}{n} \right| \qquad (55)$$

3- معامل التحديد R^2 (Coefficient of Determination):

هو مقياس إحصائي يستخدم لتقييم مدى جودة الأنموذج في تفسير التغيرات في المتغير التابع (القيم الحقيقية) بناءً على المتغيرات المستقلة أو القيم المتنبأ. يعكس R^2 نسبة التغير في البيانات الفعلية التي يمكن تفسيرها بواسطة الأنموذج وكما في المعادلة الرياضية أدناه [16]

$$R^2 = \frac{SSR}{SST} \qquad (56)$$

حيث: SSR: مجموع مربعات الموضحة وهو مجموع الفروقات المربعة بين القيم المتنبأ ومتوسطها.

SST: مجموع المربعات الكلية وهو مجموع الفروقات المربعة بين القيم الفعلية ومتوسطها.

الجانب العملي

1-3 تمهيد

قمنا بتحليل بيانات السلاسل الزمنية المتعلقة بكمية إنتاج الغاز الطبيعي في العراق[1] حيث تشمل هذه البيانات عوامل مؤثرة مثل الغاز المستهلك والغاز المحروق. تقاس هذه الكميات بوحدات (مليون متر مكعب قياسي) وتمتد المدة الزمنية لهذه البيانات من عام 1997 حتى عام 2022، مما يشمل 26 مشاهدة كما هو موضح في الجدول (2) الذي تم الحصول عليه من الجهاز المركزي للإحصاء / هيئة الإحصاء ونظم المعلومات الجغرافية.

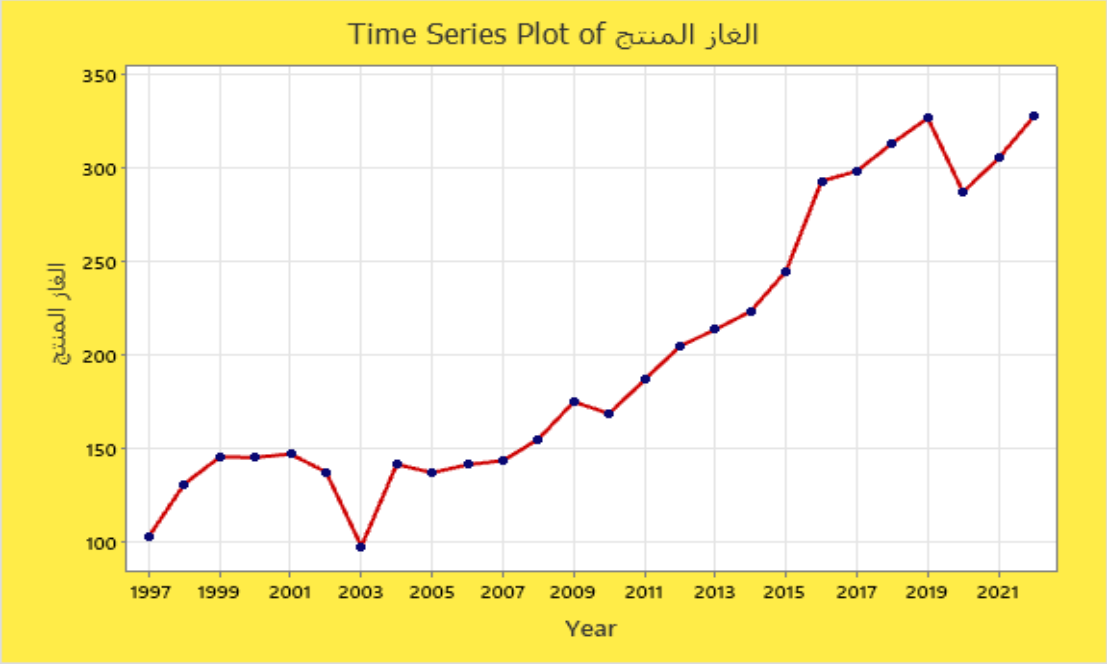
يهدف تحليل هذه البيانات إلى تحديد الاتجاهات والأنماط في إنتاج الغاز الطبيعي مع الأخذ بعين الاعتبار العوامل المؤثرة على مستويات الإنتاج مثل الغاز المستهلك والغاز المحروق وان إدخال هذه العوامل سيساعد في تقديم رؤى شاملة حول ديناميكيات سوق الغاز الطبيعي في العراق وكيف تؤثر هذه العوامل المختلفة على مستويات الإنتاج من خلال هذا التحليل سنتمكن من فهم أفضل لكيفية تفاعل هذه العوامل معًا مما يسهم في تحسين استراتيجيات الإنتاج والتخطيط المستقبلي للتنبؤ بإنتاج الغاز الطبيعي في العراق.

جدول (2): بيانات الغاز الطبيعي المنتج ولعوامل المؤثرة علية الغاز المستهلك والغاز المحروق

السنة	المنتج	المستهلك	المحروق	السنة	المنتج	المستهلك	المحروق
1997	103.28	8.474	1.854	2010	168.87	9.313	7.574
1998	131.25	9.632	3.493	2011	186.92	8.991	9.701
1999	145.64	10.009	4.555	2012	204.96	8.52	11.976
2000	145.43	10.023	4.52	2013	213.86	8.954	12.432
2001	147.23	10.451	4.272	2014	223.64	8.981	13.383
2002	137.58	10.417	3.341	2015	245.13	8.851	15.662
2003	97.81	5.542	4.239	2016	293.26	11.612	17.714
2004	141.73	7.213	6.958	2017	298.7	13.231	16.639

16.834	14.522	313.58	2018	6.64	7.083	137.23	2005
17.258	15.453	326.99	2019	7.173	6.979	141.52	2006
14.173	14.812	287.38	2020	6.998	7.372	143.7	2007
14.476	15.737	305.53	2021	6.241	9.275	155.16	2008
14.632	16.789	327.89	2022	7.38	10.14	175.2	2009

يعرض الشكل الآتي الرسم البياني للسلسلة الزمنية المتعلقة بموضوع الرسالة الخاص بالتطبيق الثاني:



شكل (2) كمية انتاج الغاز الطبيعي المنتج من 1997-2022.

يتضح من الشكل (2) أن السلسلة الزمنية تُظهر تذبذبات ملحوظة مما يشير إلى وجود تقلبات كبيرة في إنتاج الغاز الطبيعي تؤكد الإحصاءات الوصفية المتعلقة بالسلسلة هذا التوجه؛ حيث يُلاحظ أن الفارق بين أعلى وأدنى قيمة كبير نسبيًا مما يعكس تنوعًا في مستويات الإنتاج.

كما يتبين من الجدول أدناه، الذي يوضح الإحصاءات الوصفية للبيانات، تفاصيل إضافية حول هذه التذبذبات مما يعزز فهمنا للسلسلة الزمنية بشكل أكثر دقة.

جدول (3) الإحصاءات الوصفية للكمية المنتجة من الغاز الطبيعي

الوسط الحسابي	أعلى قيمة	أقل قيمة	الانحراف المعياري
199.98	327.89	97.81	74.90

زيادة عن ذلك، تشير نتائج ان اختبار Shapiro-Wilk كما هو موضح في الجدول (4) إلى أن البيانات لا تتبع التوزيع الطبيعي عند مقارنة القيم بمستوى الدلالة (α=0.05) حيث إن القيم الناتجة عن كلا الاختبارين أقل من (0.05) مما يدل على عدم تطابق البيانات مع التوزيع الطبيعي.

جدول (4) اختبار التوزيع الطبيعي

	Tests of Normality		
	Shapiro-Wilk		
	Statistic	Df	Sig.
الغاز المنتج	0.879	26	0.005

2-3 أنموذج التنبؤ الرمادي GM(1,N):

تتضمن خطوات تطبيق الأنموذج الرمادي GM(1,N) للتنبؤ بالقيم المستقبلية لإنتاج الغاز الطبيعي ما يأتي:

الخطوة الأولى: عملية التوليد التراكمي (AGO)

تم إنشاء تسلسل جديد لسلسلة البيانات الأصلية باستعمال المعادلة (11) وقد أسفرت هذه العملية عن النتائج المبينة في الجدول (5) تساعد هذه العملية في تحويل البيانات الأصلية إلى شكل تراكمي مما يسهل تحليل الاتجاهات والتنبؤات المستقبلية باستعمال الأنموذج الرمادي.

جدول (5) نتائج عملية التوليد التراكمي للمتغير التابع (الغاز المنتج) والمستقل الأول (الغاز المستهلك) والمستقل الثاني (الغاز المحروق)

ت	الغاز المنتج	الغاز المستهلك	الغاز المحروق	ت	الغاز المنتج	الغاز المستهلك	الغاز المحروق

75.238	121.923	1971.63	14	1.854	8.474	103.28	1
84.939	130.914	2158.55	15	5.347	18.106	234.53	2
96.915	139.434	2363.51	16	9.902	28.115	380.17	3
109.347	148.388	2577.37	17	14.422	38.138	525.6	4
122.73	157.369	2801.01	18	18.694	48.589	672.83	5
138.392	166.22	3046.14	19	22.035	59.006	810.41	6
156.106	177.832	3339.4	20	26.274	64.548	908.22	7
172.745	191.063	3638.1	21	33.232	71.761	1049.95	8
189.579	205.585	3951.68	22	39.872	78.844	1187.18	9
206.837	221.038	4278.67	23	47.045	85.823	1328.7	10
221.01	235.85	4566.05	24	54.043	93.195	1472.4	11
235.486	251.587	4871.58	25	60.284	102.47	1627.56	12
250.118	268.376	5199.47	26	67.664	112.61	1802.76	13

الخطوة الثانية: إيجاد قيم التسلسل الجديد للوسط المتولد حيث في هذه المرحلة، يتم حساب قيم التسلسل الجديد للوسط المتولد استنادًا إلى المعادلة (14) النتائج المستخلصة من هذه الحسابات موضحة في الجدول (6) حيث تساعد هذه القيم في تقديم فهم أفضل للاتجاهات العامة في البيانات مما يعزز دقة الأنموذج الرمادي في التنبؤ بالقيم المستقبلية.

جدول (6) قيم التسلسل الجديد للوسط المتولد

1887.195	14	———	1
2065.09	15	168.905	2
2261.03	16	307.35	3
2470.44	17	452.885	4

2689.19	18	599.215	5
2923.575	19	741.62	6
3192.77	20	859.315	7
3488.75	21	979.085	8
3794.89	22	1118.565	9
4115.175	23	1257.94	10
4422.36	24	1400.55	11
4718.815	25	1549.98	12
5035.525	26	1715.16	13

الخطوة الثالثة: مقدرات المربعات الصغرى الاعتيادية (OLS)

باستعمال قيم التسلسل الجديد للوسط المتولد التي تم الحصول عليها في الخطوة السابقة تم حساب تقديرات المربعات الصغرى الاعتيادية (OLS) استنادًا إلى المعادلة (16) وكما يأتي:

$$\underline{\hat{\gamma}} = \begin{bmatrix} \hat{\alpha} \\ \widehat{\beta_1} \\ \widehat{\beta_1} \end{bmatrix} = \begin{bmatrix} 1.93 \\ 19.33 \\ 19.29 \end{bmatrix}$$

الخطوة الرابعة: حساب قيم الاستجابة الزمنية

استنادًا إلى متجه المعاملات المقدر $\underline{\hat{\gamma}}$ يتم تقدير قيم الاستجابة الزمنية باستعمال المعادلة (17) ويوضح الجدول (7) نتائج التقدير بالنسبة للسلسلة التي تم استعمالها لأغراض التقدير التي تتكون من 26 مشاهدة. تساعد هذه القيم في فهم كيفية استجابة السلسلة الزمنية للتغيرات مما يعزز دقة الأنموذج في التنبؤ بالقيم المستقبلية.

جدول (7) قيم الاستجابة الزمنية للسلسلة التي استعملت لغرض التقدير

3644.62	21	1475.329	11	103.488	1
3958.707	22	1630.79	12	232.2708	2

4286.372	23	1806.322	13	380.1526	3
4576.735	24	1975.504	14	526.5556	4
4879.406	25	2162.745	15	674.2548	5
5194.186	26	2368.03	16	812.1734	6
		2582.232	17	910.1726	7
		2806.221	18	1052.134	8
		3051.715	19	1189.608	9
		3345.43	20	1331.374	10

الخطوة الخامسة: إجراء معكوس عملية التوليد التراكمي (IAGO)

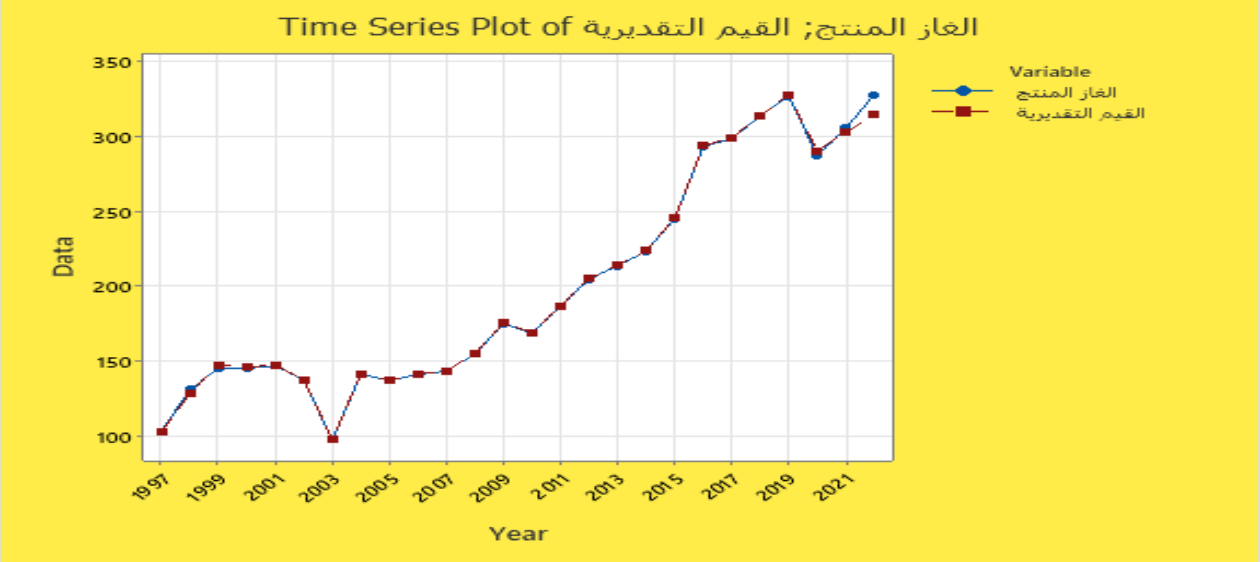
في هذه المرحلة، يتم تنفيذ معكوس عملية التوليد التراكمي (IAGO) للحصول على القيم التقديرية للسلسلة؛ وذلك باستعمال المعادلة (18) حيث تم إدراج القيم الناتجة في الجدول (8) التي تعكس التقديرات النهائية من الغاز المنتج والمستهلك والمحروق التي تسهم هذه النتائج في تقديم صورة أوضح عن التوقعات المستقبلية للسلسلة واتجاهاتها مما يساعد في تعزيز دقة التحليلات والتنبؤات.

جدول (8) القيم التقديرية الناتجة من معكوس عملية التوليد التراكمي للسلسلة التي استعملت لغرض التقدير

190.749	14	103.28	1
200.138	15	107.166	2
209.989	16	112.441	3
220.325	17	117.976	4
231.169	18	123.783	5
242.548	19	129.875	6
254.486	20	136.268	7

267.012	21	142.975	8
280.155	22	150.012	9
293.944	23	157.396	10
308.413	24	165.143	11
323.593	25	173.272	12
339.520	26	181.801	13

يوضح الرسم البياني أدناه القيم الحقيقية مقارنة بالقيم التقديرية للسلسلة موضوع الدراسة التي تشمل جميع مشاهدات العينة البالغة 26 مشاهدة. يتبين أن الأنموذج الرمادي قد ساهم في تمهيد السلسلة وجعلها أكثر انتظامًا مما أدى إلى ظهور اتجاه واضح في البيانات وان إضافة العوامل المؤثرة قد ساعدت في تعزز فعالية الأنموذج في تقديم تقديرات دقيقة وإمكانية استعماله في التنبؤ بالقيم المستقبلية.



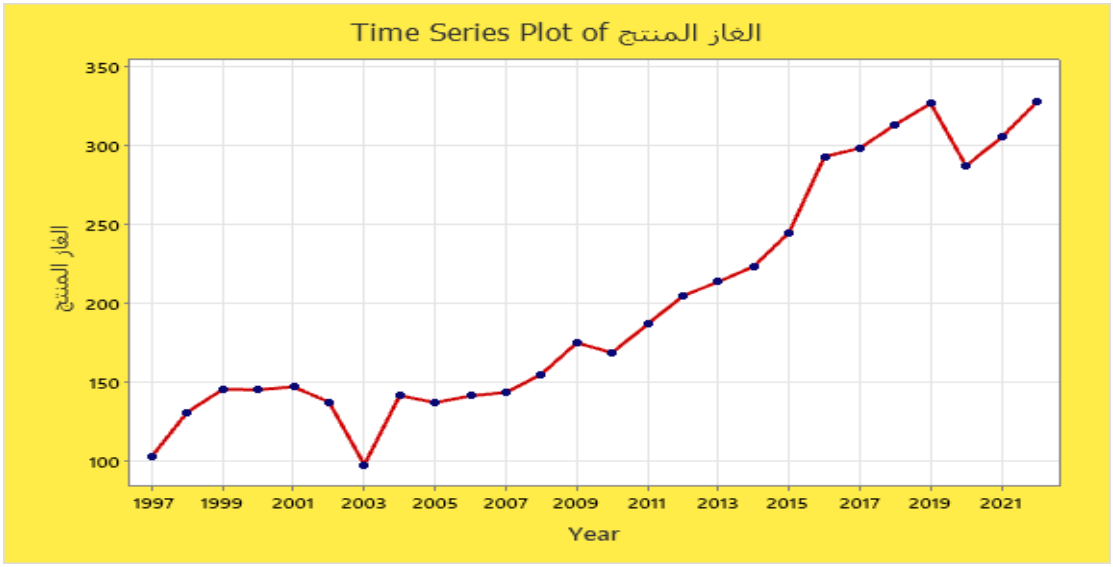
شكل (3) سلوك السلسلة المقدرة بالأنموذج الرمادي مقارنة بالسلسلة الأصلية لبيانات الغاز الطبيعي المنتج

3-3 أنموذج التنبؤ ARIMA:

جدول رقم (9) بيانات كمية الغاز المنتج

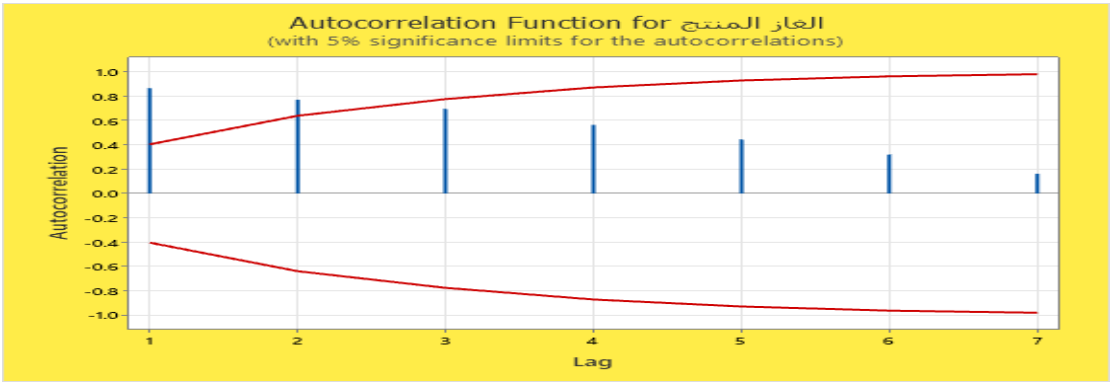
السنة	المنتج	السنة	المنتج
1997	103.28	2010	168.87
1998	131.25	2011	186.92
1999	145.64	2012	204.96
2000	145.43	2013	213.86
2001	147.23	2014	223.64
2002	137.58	2015	245.13
2003	97.81	2016	293.26
2004	141.73	2017	298.7
2005	137.23	2018	313.58
2006	141.52	2019	326.99
2007	143.7	2020	287.38
2008	155.16	2021	305.53
2009	175.2	2022	327.89

يعرض الشكل الآتي الرسم البياني للسلسلة الزمنية الخاص بالغاز المنتج:



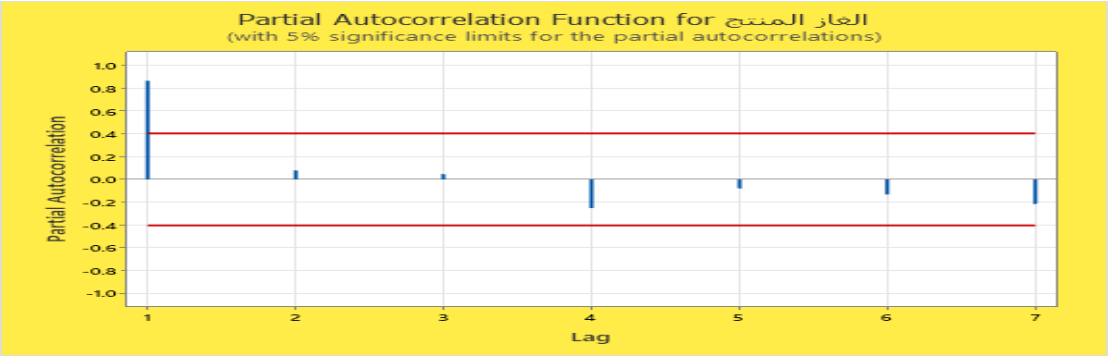
شكل (4) كمية انتاج الغاز الطبيعي المنتج من 1997-2022.

يتضح من الشكل (4) أن السلسلة الزمنية تُظهر تذبذبات ملحوظة ولكي نتأكد يجب تطبيق بعض الاختبارات لمعرفة مدى استقراريه السلسلة الزمنية لذلك نذهب الى توضيح معاملات الارتباط الذاتي للسلسلة الزمنية في البداية كما مبين في الرسم البياني أدناه:



شكل (5) معاملات الارتباط الذاتي لكمية المنتج من الغاز الطبيعي

من خلال الرسم البياني لمعاملات الارتباط الذاتي (ACF) لكمية إنتاج الغاز الطبيعي، يتبين أن هناك بعض القيم التي تتجاوز الحدود المعنوية مما يشير إلى وجود ارتباطات ذات دلالة إحصائية في بعض المُدد الزمنية، ومع ذلك تبقى جميع القيم في المُدد الزمنية اللاحقة ضمن الحدود المعنوية مما يدل على عدم وجود ارتباطات ملحوظة في تلك المُدد وكذلك تم حساب معاملات الارتباط الجزئي للسلسلة الزمنية كما مبين في الرسم البياني أدناه:



شكل (6) معاملات الارتباط الجزئي لكمية المنتج من الغاز الطبيعي

من خلال الرسم البياني لمعاملات الارتباط الجزئي (PACF) لكمية إنتاج الغاز الطبيعي يتضح أن معامل الارتباط الجزئي يسجل قيمة مرتفعة في المدة الزمنية الأولى (Lag 1) ثم ينخفض بشكل ملحوظ في المُدد الزمنية الآتية. يشير هذا إلى أن العلاقة بين القيمة الحالية والقيمة السابقة تكون قوية، لكنها تتلاشى بسرعة في المُدد اللاحقة. كما يتبين أن هناك بعض القيم التي تتجاوز الحدود المعنوية كما هو الحال في Lag 1 بينما تبقى جميع القيم في المُدد الزمنية اللاحقة ضمن الحدود المعنوية.

جدول رقم (10) اختبار جذر الوحدة

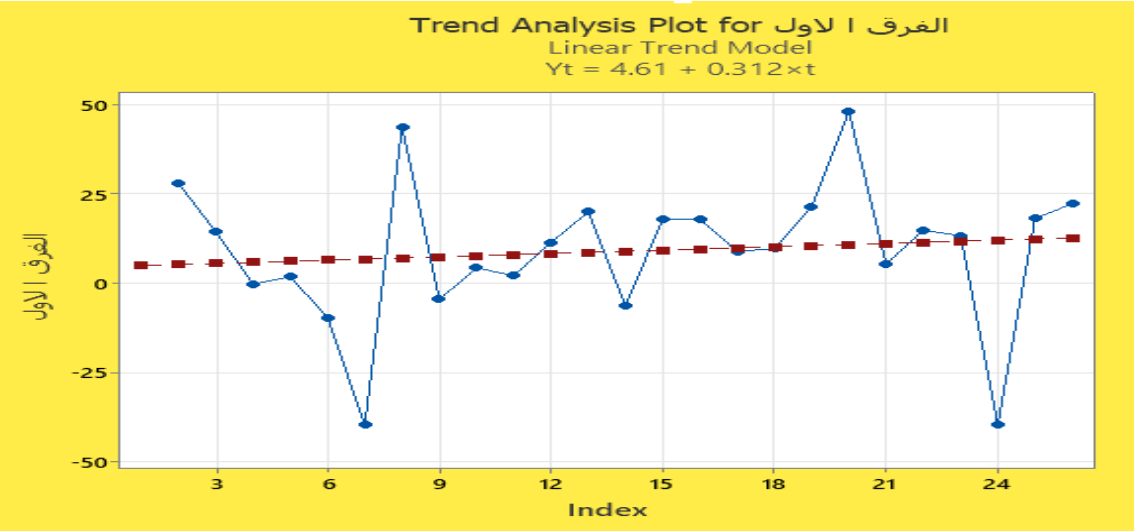
Null Hypothesis: Y has a unit root Exogenous: Constant Lag Length: 0 (Automatic - based on SIC, maxlag=5)				
Prob.*	t-Statistic			
0.9314	-0.162844	Augmented Dickey-Fuller test statistic		
	-3.724070		1% level	Test critical values:
	-2.986225		5% level	

يتبين من خلال الجدول أعلاه ان السلسلة الزمنية تعاني من جذر الوحدة لان القيمة الاحتمالية (Prob=0.9314) وهي اكبر من مستوى الدلالة (α=0.05) ، وهذا يدل على ان السلسلة الزمنية غير مستقرة؛ فإنه من الضروري أخذ الفروق للسلسلة (Differencing) لجعلها مستقرة قبل تطبيق نموذج ARIMA.

جدول رقم (11) بعد اخذ الفرق الاول

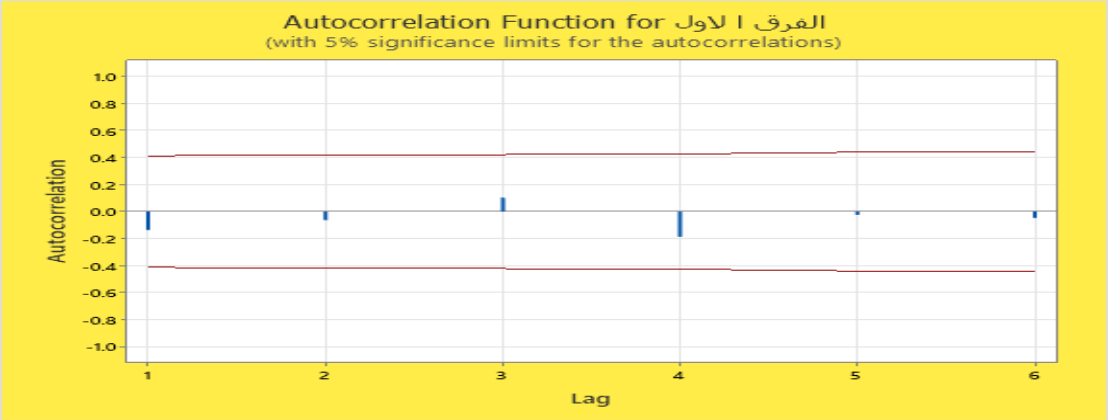
Null Hypothesis: D(Y) has a unit root Exogenous: Constant Lag Length: 0 (Automatic - based on SIC, maxlag=5)				
Prob.*	t-Statistic			
0.0002	-5.453152	Augmented Dickey-Fuller test statistic		
	-3.737853		1% level	Test critical values:
	-2.991878		5% level	
	-2.635542		10% level	

القيمة المحسوبة لـ t هي 5.453152-، وهي أقل بكثير من القيم الحرجة عند جميع مستويات الدلالة هذا يعني أن السلسلة الزمنية أصبحت مستقرة عند اخذ الفرق الاول وان القيمة الاحتمالية 0.0002 هي أيضاً أقل بكثير من مستوى الدلالة (0.05) ويمكن توضيحها أيضا من خلال الرسم البياني:



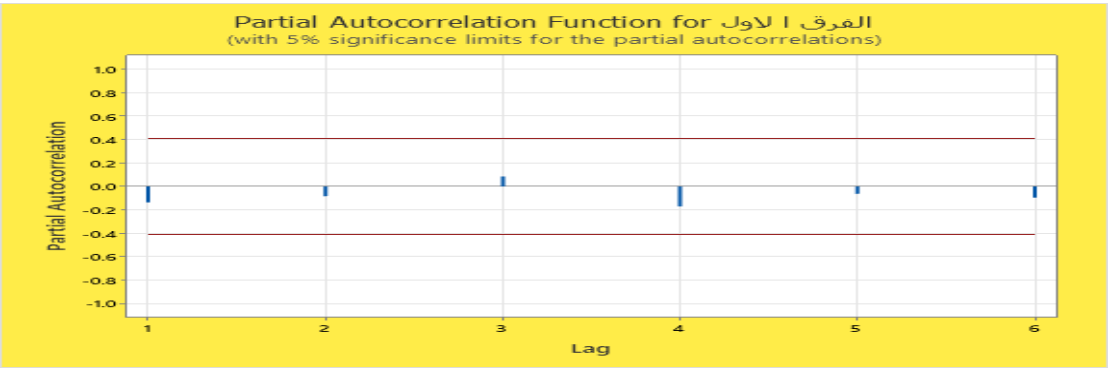
شكل (7) استقراريه السلسلة بعد اخذ الفرق الأول

يتبين بعد اخذ الفرق الأول ان اتجاه السلسلة الزمنية قد فقد عكس النماذج الرمادية التي تحافظ على اتجاه السلسلة الزمنية ويتبين من خلال تطبيق معاملات الارتباط الذاتي للسلسلة الزمنية بعد اخذ الفرق الأول ان جميع النتوء قد أصبحت داخل مدة الثقة، وكما مبين في الرسم البياني أدناه:



شكل (8) معاملات الارتباط الذاتي بعد اخذ الفرق الاول

وكذلك يتبين من خلال معاملات الارتباط الجزئي للسلسلة الزمنية بعد اخذ الفرق الأول انه أصبحت جميع النتوء داخل فترة الثقة وكما مبين في الرسم البياني أدناه:



شكل (9) معاملات الارتباط الجزئي بعد اخذ الفرق الاول

بعد تحقيق الاستقرارية في السلسلة الزمنية نختار أفضل انموذج بين نماذج ARIMA

جدول رقم (12) نموذج ARIMA (1,1,1).

Type	Coef	SE Coef	T-Value	P-Value
AR 1	0.739	0.272	2.72	0.013

MA 1	0.947	0.233	4.06	0.001
Constant	2.350	0.267	8.81	0.000

أولاً: معامل الانحدار الذاتي AR(1) يبلغ 0.739 هذا يشير إلى أن قيمة الغاز المنتج في المدة الحالية ترتبط بشكل إيجابي مع قيمتها في المدة السابقة بمعنى آخر التغييرات في الإنتاج تميل إلى الاستمرار على المدى القصير حيث أن ارتفاع الإنتاج في الماضي يرتبط بارتفاعه في الحاضر.

ثانياً: معامل المتوسط المتحرك MA (1) يبلغ 0.947. هذا يُظهر وجود ارتباط إيجابي بين قيم الغاز المنتج في المُدد الزمنية المتتالية وهذا يعني أن التغيرات في الإنتاج قد تنتقل أو تنعكس من مدة إلى أخرى.

ثالثاً: الحد الثابت في الأنموذج يبلغ 2.350. هذا يمثل المستوى المتوقع للغاز المنتج عند غياب أي تأثيرات أخرى في الأنموذج ويتبين ان جميع المعاملات ذات دلالة إحصائية معنوية مما يشير إلى أن انموذج ARIMA (1,1,1) يمثل بشكل جيد الخصائص الزمنية لسلسلة بيانات الغاز المنتج.

جدول رقم (13) نموذج ARIMA (2,1,1).

Type	Coef	SE Coef	T-Value	P-Value
AR 2	-0.006	0.249	-0.02	0.980
MA 1	0.950	0.276	3.45	0.002
Constant	2.316	0.277	8.36	0.000

يُظهر الجدول نتائج تقدير انموذج ARIMA (2,1,1) الذي يتضمن معاملات الانحدار الذاتي (AR) والمتوسط المتحرك (MA) والحد الثابت.

أولًا: معامل الانحدار الذاتي AR (2) يبلغ 0.006- هذه القيمة قريبة من الصفر وتشير إلى عدم وجود تأثير ملحوظ من القيم السابقة على القيمة الحالية للغاز المنتج حيث أن القيمة السلبية لا تعكس ارتباطًا قويًا، وكذلك نرى ان AR (2) ان القيمة الاحتمالية لـ P-Value هي أكبر من 0.05 وهي غير معنوية.

ثانيًا: معامل المتوسط المتحرك MA (1) يساوي 0.950. هذه القيمة إيجابية وذات دلالة إحصائية معنوية

(P-Value = 0.002)، مما يدل على أن هناك ارتباطًا قويًا بين القيمة الحالية للغاز المنتج والقيمة السابقة من المتوسط المتحرك ومن ثمَّ، التغييرات في الإنتاج تميل إلى أن تؤثر على القيم الآتية.

ثالثًا، الحد الثابت للأنموذج يبلغ 2.316 وهو أيضًا ذو دلالة إحصائية معنوية (P-Value = 0.000) مما يشير إلى أن هناك مستوى ثابتًا للإنتاج عند غياب أي تأثيرات أخرى.

بشكل عام يُظهر انموذج ARIMA (2,1,1) أن الارتباط مع القيم السابقة ضعيف بينما يرتبط الغاز المنتج بشكل قوي بالتغيرات السابقة عبر المتوسط المتحرك.

ولهذا سوف نذهب الى مقاييس المقارنة ونرى ايهما أفضل انموذج حتى نقوم بمقارنته مع النماذج الرمادية الأخرى.

4-3 المقارنة بين النماذج

للمقارنة بين الأنموذج الرمادي ونماذجARIMA ، اعتمد على عدة معايير تتضمن جذر متوسط مربعات الخطأ (RMSE) ومتوسط الخطأ المطلق (MAD) ومعامل التحديد R^2 ؛ فضلاً عن ذلك تم استعمال معيار نسبة التباين (Q) الذي يساعد في تحديد مدى ملاءمة النموذج ليكون نموذجًا رماديًا أو انموذجًا تقليديًا حيث حُسبت قيم هذه المعايير لجميع النماذج المعنية، كما هو موضح في الجدول (14). تسهم هذه العملية في اختيار الأنموذج الأكثر ملاءمة حقق نتائج دقيقة وموثوقة من خلال تحليل هذه المعايير، يمكن تقييم أداء النماذج بشكل موضوعي، مما يساعد في اتخاذ قرارات مستندة إلى بيانات دقيقة حول الأنموذج الأفضل للتنبؤ.

جدول (14) معايير المقارنة للنماذج المقدرة للتنبؤ بالإنتاج الغاز الطبيعي في العراق من 1997-2022

النموذج	RMSE	MAD	R^2	نسبة تأهيل التباين	
				نسبة التباين Q	درجة دقة التنبؤ
GM(1,N)	2.80	1.22	0.91	0.03	ممتاز
ARIMA(1,1,1)	19.94	13.49	0.12	غير مؤهل	
ARIMA(2,1,1)	20.39	15.36	0.03	غير مؤهل	

عند تحليل النماذج التنبؤية الخاصة بإنتاج الغاز الطبيعي في العراق يتبين أن انموذج الرمادي GM(1,N) هو الأكثر فعالية فقد حقق هذا الأنموذج أدنى قيمة لجذر متوسط مربعات الخطأ (RMSE) حيث بلغت (2.80) وأقل قيمة لمتوسط الخطأ المطلق (MAD) بلغت (1.22) فضلاً عن ذلك، سجل أعلى قيمة لمعامل التحديد R^2 بلغت (0.91) مما يعني أن (91%) من التباين في البيانات يمكن تفسيره بوساطة هذا الأنموذج حيث تعكس هذه النتائج قدرة الأنموذج الرمادي العالية على التقاط الديناميكيات والعلاقات الموجودة في البيانات مما يُشير إلى تفوق الأنموذج الرمادي في القدرة على تمثيل الظاهرة المدروسة حيث يُظهر هذا الفارق أن الأنموذج الرمادي وخاصة في مثل هذه الأحجام من العينات يعدُّ الأكثر كفاءة في تحقيق النتائج

وبناءً على هذه النتائج يتضح أن انموذج GM(1,N) يمثل الخيار الأكثر ملاءمة لإنتاج الغاز الطبيعي حيث يقدم تفسيرات دقيقة وقدرة محسنة على فهم الديناميكيات الموجودة في البيانات مقارنة بنماذج ARIMA وبعد معرفة أي نماذج هي الأفضل تم اختيار الأنموذج الرمادي GM(1,N) للتنبؤ بإنتاج الغاز الطبيعي في العراق

جدول (15) القيم المتنبأ بها

السنة	القيم المتنبأ بها
2023	333.61
2024	351.64
2025	370.68
2026	389.71
2027	408.75
2028	427.78
2029	446.82
2030	465.86

1-4 الاستنتاجات:

1- يتبين أن انموذج الرمادي GM(1,N) هو الأكثر فعالية فقد حقق هذا الأنموذج أدنى قيمة لجذر متوسط مربعات الخطأ (RMSE)حيث بلغت (2.80) وأقل قيمة لمتوسط الخطأ المطلق (MAD) بلغت (1.22) فضلاً عن ذلك سجل أعلى قيمة لمعامل التحديد R^2 بلغت (0.91) مما يعني أن (91%) من التباين في البيانات يمكن تفسيره بواسطة هذا النموذج حيث تعكس هذه النتائج قدرة الأنموذج الرمادي العالية على التقاط الاتجاهات والعلاقات الموجودة في البيانات مما يُشير إلى تفوق الأنموذج الرمادي في القدرة على تمثيل الظاهرة المدروسة حيث يُظهر هذا الفارق أن الأنموذج الرمادي وخاصة في مثل هذه الأحجام من العينات يعدُّ الأكثر كفاءة في تحقيق النتائج وبناءً على هذه النتائج يتضح أن انموذج GM(1,N) يمثل الخيار الأكثر ملاءمة لإنتاج الغاز الطبيعي حيث يقدم تفسيرات دقيقة وقدرة محسنة على فهم الاتجاهات الموجودة في البيانات مقارنة بنماذج ARIMA.

2- النماذج الرمادية تكون كفؤ في حالة احجام العينات الصغيرة.

3- يتبين بعد اخذ الفرق الأول ان اتجاه السلسلة الزمنية قد فقدت الاتجاه عكس الأنموذج الرمادي GM(1,N) التي تحافظ على اتجاه السلسلة الزمنية.

4- يتبين ان الأنموذج الرمادي GM(1,N) يعمل حتى البيانات التي لا تتبع التوزيع الطبيعي.

5- يتبين أيضا ان الأنموذج الرمادي GM(1,N) يعمل مع البيانات التي تعاني من عدم تجانس في البيانات.

التوصيات:

1- نوصي باستعمال النماذج الرمادية المحسنة، وهي النماذج الرمادية الموجية مع نماذج ARIMA لأن كليهما يعملان مع حجوم عينات كبيره.

2- تناول النماذج الرمادية GM(N,1) ذات درجات تفاضلية (N) بدلا من الدرجة الأولى مع متغير.

3- نوصي باستعمال الأنموذج الرمادي متعدد المتغيرات في الدراسات التي تعاني من نقص المعلومات مثل العيّنات الصغيرة حيث يتيح التعامل مع عدم اليقين والتباين ويوفر هذا الأنموذج إطارًا مرنًا لتوضيح العلاقات بين المتغيرات وإجراء تحليلات ذات دلالة إحصائية مما يعزز موثوقية النتائج وجودتها في التنبؤ.

المصادر العربية

1. البيانات وزارة التخطيط / الجهاز المركزي للإحصاء / هيئة الإحصاء ونظم المعلومات الجغرافية https://cosit.gov.iq/ar/?jsn_setmobile=no
2. احمد احمد ذياب& عطا غادة مصطفى (2018) التنبؤ بالارقام القياسية لاسعار المستهلك في العراق. المجلد الثاني وقائع المؤتمر العلمي الثالث جامعة كربلاء.
3. جابر اسماء غالب& عنكوش عباس دوين (2018) استعمال نماذج بوكس وجينكنز للتنبؤ بحركة اعداد الطرود البريدية الصادرة والواردة. المجلد الثاني وقائع المؤتمر العلمي الثالث جامعة كربلاء.
4. حبيب علي سلمان & رحيمه رشيد بشير& سعيد مهند عباس (2018).نمذجة حوادث الطرق في محافظة ذي قار باستعمال تحليل السلاسل الزمنية الخطية . المجلد الثاني وقائع المؤتمر العلمي الثالث جامعة كربلاء.
5. الدكتور طالب أحمد & ,كمال نجار. (2021). استعمال السلاسل الزمنية في التنبؤ-دراسة تطبيقية على طلاب التعليم الثانوي في سورية .مجلة جامعة البعث-سلسلة العلوم الاقتصادية والسياحية.(31)43 ,
6. شمال. اياد حبيب & حميد. يونس كاظم (2018). دراسة عدد حالات الطلاق في مدينة الحلة من خلال استعمال السلاسل الزمنية في الشبكات العصبية. المجلد الثاني وقائع المؤتمر العلمي الثالث جامعة كربلاء.
7. العجيأتي ، سعد صبر محمد & علي فياض عبد الله (2018). التنبؤ بمعدلات التضخم في العراق باستمعال نماذج بوكس وجنكيز التكاملية للفترة من 1971 -2015 . المجلد الثاني وقائع المؤتمر العلمي الثالث جامعة كربلاء.
8. فراس احمد محمد & ,نور سليم. (2018). مقارنة بعض طرائق تقدير انموذج GM (1, 1) بوجود بيانات مفقودة مع تطبيق عملي.Journal of Economics and Administrative Sciences, 24(103), 420-420 .

المصادر الأجنبية

9. Ahmed, B. K., Rahim, S. A., Maaroof, B. B., & Taher, H. A. (2020). Comparison Between ARIMA And Fourier ARIMA Model To Forecast The Demand Of Electricity In Sulaimani Governorate. Qalaai Zanist Journal, 5(3), 908-940.

10. Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: forecasting and control. John Wiley & Sons.

11. Cempel, C. (2014). Teoria szarych systemów–nowa metodologia analizy i oceny złożonych systemów. Przegląd możliwości. Zeszyty Naukowe Politechniki Poznańskiej. Organizacja i Zarządzanie.

12. Chiou, H. K., Tzeng, G. H., & Cheng, C. K. (2004). Grey prediction GM (1, 1) model for forecasting demand of planned spare parts in navy of Taiwan. In Proceedings World Automation Congress. Seville, Spain. Retrieved from https://ieeexplore. ieee. org/document/1439385.

13. Du, X., Wu, D., & Yan, Y. (2023). Prediction of electricity consumption based on GM (1, Nr) model in Jiangsu province, China. Energy, 262, 125439.

14. Hai N. P., Shiu T. W., Nagai M., (2015). Predicting student learning outcomes based on a combination of Taylor approximation and gray models. VNU Scientific Journal: Educational Research, Volume 31, Issue 2 (pp 70-83)

15. Hsu, L. C., & Wang, C. H. (2009). Forecasting integrated circuit output using multivariate grey model and grey relational analysis. Expert systems with applications, 36(2), 1403-1409.

16. Judd, C. M., McClelland, G. H., & Ryan, C. S. (2017). Data analysis: A model comparison approach to regression, ANOVA, and beyond. Routledge.

17. Jun, D., Pei, W., & Xihao, D. (2018). Short-term power load forecasting based on EMD-grey model. American Journal of Electrical Power and Energy Systems, 7(4), 42-49.

18. Khan, A. M., & Osińska, M. (2023). Comparing forecasting accuracy of selected grey and time series models based on energy consumption in Brazil and India. Expert Systems with Applications, 212, 118840.

19. Lao, T., Chen, X., & Zhu, J. (2021). The optimized multivariate grey prediction model based on dynamic background value and its application. Complexity, 2021(1), 6663773.

20. Ma, L., Hu, C., Lin, R., & Han, Y. (2018, December). ARIMA model forecast based on EViews software. In IOP conference series: Earth and environmental science (Vol. 208, No. 1, p. 012017). IOP Publishing.

21. Ma, X., & Liu, Z. B. (2018). The kernel-based nonlinear multivariate grey model. Applied Mathematical Modelling, 56, 217-238.

22. Mahmood, H. O. F., Ali, M. A., & Ali, Z. J. R. (2023). Forecasting performance in iraqi stock exchange for the oil price throw the gm (1, 2) model and the impacts on economic growth. 3c Empresa: investigación y pensamiento crítico, 12(1), 311-322.

23. Mills, T. C. (2019). Applied time series analysis: A practical guide to modeling and forecasting. Academic press.

24. Mondal, K., & Pramanik, S. (2015). The application of grey system theory in predicting the number of deaths of women by committing suicide-a case study. Journal of Applied Quantitative Methods, 10(1), 48-55.

25. Moonchai, S., & Rakpuang, W. (2015). A new approach to improve accuracy of grey model GMC (1, n) in time series prediction. Modelling and Simulation in Engineering, 2015(1), 126738.

26. Paolella, M. S. (2018). Linear models and time-series analysis: regression, ANOVA, ARMA and GARCH. John Wiley & Sons.

27. Sallehuddin, R., & Hj. Shamsuddin, S. M. (2009). Hybrid grey relational artificial neural network and auto regressive integrated moving average model for forecasting time-series data. Applied Artificial Intelligence, 23(5), 443-486.

28. Sapnken, F. E., Kibong, M. T., & Tamba, J. G. (2023). A novel prediction approach using wavelet transform and grey multivariate convolution model. MethodsX, 11, 102259.

29. Tang, L. W., & Lu, Y. Y. (2018). The optimization of GM (2, 1) model based on parameter estimation of grade difference format. Systems Engineering-Theory & Practice, 38(2), 502-508.

30. Tseng, F. M., Yu, H. C., & Tzeng, G. H. (2001). Applied hybrid grey model to forecast seasonal time series. Technological Forecasting and Social Change, 67(2-3), 291-302.

31. Wang, L., Lu, H. P., & Zheng, Y. (2011). Macro prediction of traffic accident in Beijing based on model GM(1,N).In ICCTP 2011:Towards Sustainable Transportation Systems (pp.1914-1924).

32. Wang, Y. W., Shen, Z. Z., & Jiang, Y. (2018). Comparison of ARIMA and GM (1, 1) models for prediction of hepatitis B in China. PloS one, 13(9), e0201987.

33. Zhao, D., Zhang, H., Cao, Q., Wang, Z., He, S., Zhou, M., & Zhang, R. (2022). The research of ARIMA, GM (1, 1), and LSTM models for prediction of TB cases in China. Plos one, 17(2), e0262734.