استخدام نموذج SARIMA في التنبؤ بالسلاسل الزمنية الموسمية

Using SARIMA Models to Forecast Seasonal Time Series

إعداد

أ. سمير فرج رشيد أبو دحروج

باحث دكتوراه إحصاء ـ كلية الاقتصاد والعلوم الإدارية

جامعة الأزهر (فلسطين – غزة)

samir.abu.dahrouj@gmail.com

مجلة البحوث التجارية ـ كلية التجارة جامعة الزقازيق

المجلد الخامس والأربعين ـ العدد الرابع أكتوبر 2023

رابط المجلة: https://zcom.journals.ekb.eg/

ملخص:

يعد أسلوب تحليل السلاسل الزمنية Time Series Analysis من الأساليب الإحصائية الجديرة بالاهتمام ويعتمد أسلوب تحليل السلاسل الزمنية على تتبع الظاهرة (أو المتغير) على مدى زمني معين (عدة سنوات مثلاً)، ثم يتوقع للمستقبل بناءً على القيم المختلفة التي ظهرت في السلسلة الزمنية وعلى نمط النمو في القيم.

تناول هذا البحث التعرف على خطوات ومراحل منهجية بوكس – جنكينز الموسمية في التنبؤ على المستوى النظري والتطبيقي، لبناء أفضل نموذج يساعد على التنبؤ بالكميات المستهلكة من الغاز الطبيعي في الولايات المتحدة الأمريكية، وبمقارنة عدة نماذج مختلفة مع النموذج المقترح تبين أن النموذج الأمثل هو نموذج SARIMA (2,1,1) (2,1,1)12 لاجتيازه معظم عمليات الفحص والتشخيص بدرجة جيدة إحصائياً مقارنة بغيره من النماذج الأخرى، والحاصل على أقل القيم لمعايير التقييم (AIC, AICc, BIC)، وعليه تم استخدامه في التنبؤ بالقيم المستقبلية خلال الفترة من شهر أكتوبر 2022 حتى شهر مارس 2023، ووجد أن القيم التنبؤية تتبع نفس السلوك للسلسلة الأصلية.

كلمات مفتاحية: نموذج SARIMA، التنبؤ، السلاسل الزمنية الموسمية.

مقدمة: Introduction

يعد تحليل السلاسل الزمنية من أبرز الأساليب الإحصائية والتي تعد أحد الموضوعات المهمة والمستخدمة في تفسير سلوك الظواهر عن طريق دراسة تطورها التاريخي عبر فترة زمنية قد تكون يومية أو أسبوعية أو فصلية او شهرية أو يومية لغرض تقديرها بأقل خطأ ممكن، فضلاً عن أن تحليل السلاسل الزمنية أصبح له دور مهم في عملية اتخاذ القرارات في المجالات التطبيقية.

تعرف الموسمية في السلاسل الزمنية بأنها نمط منتظم من التغيرات يعيد نفسه بعد فترة زمنية ثابتة، كما وتعتبر من أحد المشاكل التي تواجه الباحثين عند تحليل السلاسل الزمنية والتي تؤثر سلباً على دقة النموذج الرياضي في حالة عدم معالجتها بصورة صحيحة (عوض الله، 2016).

ولغرض دراسة سلوك هذه التغيرات والوقوف على أسباب حدوثها لابد من بناء نموذج رياضي يمثلها ويجعلنا قادرين على تقديرها والتنبؤ بها مستقبلاً، فقد شهدت الأعوام الأخيرة من القرن الماضي استخداماً واسعاً لنماذج السلاسل الزمنية ولاسيما النماذج المضاعفة SARIMA (p,d,q) x $(P,Q,D)_S$ عند الازاحة الموسمية (S) (المرشدي، 2021).

ومن أهم الاساليب الإحصائية المستخدمة بشكل واسع في عمليات التنبؤ هو منهجية بوكس وجينكنز الموسمية، وتعتمد هذه المنهجية على دالة الارتباط الذاتي والذاتي الجزئي واستخدام مبدأ المتوسطات المتحركة ومبدأ الانحدار الذاتي وتشترط هذه المنهجية السكون للسلسلة الزمنية بمعنى أن يكون المتغير التابع له متوسط وتباين ثابتين خلال الفترة الزمنية موضع الدراسة.

مشكلة الدراسة: Research Problem

تتمثل مشكلة البحث في بناء نموذج تنبؤ لبيانات السلسلة الزمنية يكون له المقدرة على تصوير الواقع ودقة عالية في التنبؤات المستقبلية، بحيث يجب أن تأخذ هذه النماذج كل الاعتبارات المتعلقة ببيانات السلسلة الزمنية من حيث أنها مستقرة أم لا وكيفية اختيار النموذج الملائم من بين النماذج المختلفة، وهناك العديد من الأساليب الإحصائية التي تستخدم في بناء نماذج التنبؤ، فمنها نماذج Box & Jenkins، فعليه يتمحور التساؤل الرئيس في:

ما هو النموذج الأمثل من بين نماذج SARIMA للتنبؤ بالكميات المستهلكة من الغاز الطبيعي في الولايات المتحدة الأمريكية شهرياً؟

أهداف الدراسة: Research Objectives

- تحديد النموذج الأمثل من بين نماذج SARIMA.
- استخدام النموذج الأمثل في التنبؤ بالكميات المستهلكة شهرياً من الغاز الطبيعي مستقبلاً.

أهمية الدراسة: Importance of the Research

الأهمية التطبيقية:

تنبع الأهمية التطبيقية لهذه الدراسة من الأهمية الكبيرة التي تتمتع بها أساليب التنبؤ في تحليل بيانات السلاسل الزمنية الاقتصادية، وذلك لأنها تقدم لمنفذي ومخططي البرامج من رسم السياسات الاقتصادية المختلفة ووضع الخطط المستقبلية، ومعرفة ما تؤول إليه الظاهرة في المستقبل ووضع الحلول اللازمة للسيطرة على الظاهرة.

الأهمية الإحصائية:

تستخدم الدراسة منهجاً تحليلياً للسلاسل الزمنية المبني على منهجية Box & Jenkins الموسمية، حيث تتمثل الأهمية الإحصائية للدراسة في إيجاد نموذج قياسي يستخدم للتنبؤ بالكميات المستهلكة من الغاز الطبيعي شهرياً، والتنبؤ بالكميات المستهلكة في الأشهر القادمة من أجل التخطيط للمستقبل.

مصدر البيانات:

تمَ الحصول على بيانات الدراسة الخاصة بالكميات المستهلكة من الغاز الطبيعي في الولايات المتحدة الأمريكية شهرياً خلال الفترة الزمنية الممتدة من شهر يناير2000 حتى شهر سبتمبر 2022 من الرابط التالي: https://fred.stlouisfed.org/series/NATURALGAS

حدود الدراسة:

تمَّ تقسيم حدود الدراسة إلى حدود زمانية وموضوعية

الحدود الزمانية: الفترة الزمنية من شهر يناير2000 حتى شهر سبتمبر2022.

الحدود الموضوعية: الكميات المستهلكة من الغاز الطبيعي في الولايات المتحدة الأمريكية مليار قدم مكعب.

الدراسات السابقة

هنالك العديد من الدراسات السابقة التي تناولت الموضوع محل الدراسة، المتمثلة في الدراسات العربية، والأخرى الأجنبية، نذكر منها:

دراسة (الدوب، 2021)

هدفت الدراسة إلى التنبؤ بإجمالي الجرائم في دولة الكويت، للوصول إلى بيانات جنائية مستقبلية دقيقة، اعتمدت الدراسة على منهجية بوكس وجينكنز، بوصفها إحدى طرق تحليل السلاسل الزمنية واستخدام نموذج SARIMA للتنبؤ بإجمالي الجرائم في دولة الكويت، وذلك بالاستناد إلى بيانات إجمالي الجرائم الشهرية الصادرة عن مركز البحوث والدراسات التابع لوزارة الداخلية بالكويت، وذلك خلال الفترة من 1/2001 حتى 12/2017، للمقارنة بين نماذج SARIMA من خلال المعايير AIC،AICC،BIC .

توصلت الدراسة إلى أن النموذج الأمثل من بين نماذج SARIMS هو نموذج 12(0,1,2) SARIMA (1,1,1)، وقد استخدم للتنبؤ بإجمالي الجرائم في دولة الكويت حتى نهاية عام 2020.

دراسة (Ilmayasinta, 2021)

هدفت الدراسة إلى التنبؤ لمعرفة تقديرات السياح الأجانب الذين يأتون إلى إندونيسيا في الأشهر التالية، باستخدام نموذج الانحدار الذاتي والمتوسط المتحرك الموسمي (SARIMA)، بالاعتماد على بيانات ثانوية شهرية بعدد السياح الأجانب الذين يأتون إلى إندونيسيا عبر مطار سوكارنو هاتا للفترة من يناير 2010 إلى يونيو 2015، تم الحصول عليها من وكالة مركز الإحصاء الإندونيسي.

أظهرت نتائج تحليل البيانات ان النموذج الملائم هو $ARIMA(0,1,1)\ (0,1,1)_{12}$ للتنبؤ بقيمة البيانات للأشهر الستة القادمة، حيث تظهر نتائج التنبؤ عدد الزيادات كل شهر عن العام السابق.

دراسة (Divisekara, et al., 2020)

هدفت الدراسة نمذجة أسعار العدس الأحمر والتنبؤ بها باستخدام نموذج المتوسط المتحرك الموسمي الانحدار الذاتي المتكامل (SARIMA)، تم استخدام ثماني سنوات من البيانات الأسبوعية بدءًا من 2010 إلى 2019 والتي تضمنت 521 مشاهدة، تم الحصول عليها من شركة ساسكاتشوان، كان متوسط سعر العدس الأحمر 24.75 دولارًا لكل 100 رطل، وكانت الأسعار الأسبوعية متقلبة للغاية بمرور الوقت.

كشفت النتائج أن نموذج [52] SARIMA (2,1,2) (0,1,1) يوفر أفضل أداء في العينة وخارج العينة عند التنبؤ بأسعار العدس الأحمر، يمكن استخدام هذا النموذج من قبل كل من المزارعين والمستخدمين النهائيين في اتخاذ قرارات الإنتاج المثلى وإدارة مخاطر الأسعار الإجمالية.

دراسة (Aliyu, et al., 2021)

هدفت الدراسة نمذجة والتنبؤ بهطول الأمطار الشهرية في نيجيريا في هذه الدراسة، بالاعتماد على بيانات هطول الأمطار الشهرية النيجيرية بين 1980-2015 التي تم الحصول عليها من بوابة المناخ التابعة للبنك الدولي. تم تبني منهجية Box-Jenkin، بتطبيق نموذج SARIMA.

بينت الدراسة أن نموذج [12] SARIMA (2,0,1) (2,1,1) هو أفضل نموذج من بين النماذج الأخرى التي تناسب بيانات هطول الأمطار النيجيرية (1980-2015)، وقد استخدم للتنبؤ بهطول الأمطار الشهرية في نيجيريا من عام 2018 حتى عام 2042، وقد تم اكتشاف أن شهر أبريل هو فترة بداية هطول الأمطار في نيجيريا ونوفمبر هو فترة التراجع.

دراسة (Mamudu, et al., 2021)

هدفت الدراسة إلى التنبؤ في تصريف نهر كادونا، باستخدام بيانات التصريف الشهرية لمدة 25 عامًا (1988 إلى 2013)، تم جمع بيانات السلاسل الزمنية اليومية لتصريف نهر كادونا من وكالة الخدمات الهيدرولوجية النيجيرية من عام 1988 إلى عام 2018 في محطة جسر غوني غورا، استخدمت الدراسة 25 عامًا (1988-2013) كمجموعة تدريب أو معايرة، و5 سنوات (2014-2018) كمجموعة للتحقق، تم تحويل هذه البيانات إلى متوسطات شهرية بالمتر المكعب / ثانية.

خلصت الدراسة إلى أن نموذج SARIMA $(1, 0, 1) (0, 1, 1)_{12}$ هو الأنسب بناءً على معايير الاختيار، ويمكن أن يتنبأ بشكل كافٍ بتصريف نهر كادونا مع الحد الأدنى من الأخطاء.

دراسة (Chowdhury, et al., 2018)

هدفت الدراسة إلى فحص الرطوبة بمرور الوقت في بنغلاديش وإيجاد نموذج مناسب للتنبؤ، اعتمدت الدراسة بيانات الرطوبة من دائرة الأرصاد الجوية في بنجلاديش المسجلة في 6 محطات أرصاد جوية الأقسام للفترة من 1976 إلى 2015، وجدت الدراسة أن متوسط الرطوبة السنوية في بنغلاديش هو 78.88٪، وأن البيانات غير ثابتة ولكن تم تحويلها إلى ثابتة بعد أخذ الفرق الأول، ثم تم بناء نموذج ARIMA الموسمي باستخدام نهج Box and Jenkins.

تم تحديد نموذج ARIMA (2,0,1) $(2,1,1)_{12}$ كنموذج مناسب للتنبؤ بالرطوبة الموسمية لمدة 60 شهراً (2016-2020)

منهجية الدراسة Research Methodology

تم استخدام منهجية Box & Jenkins الموسمية للتنبؤ بالكميات المستهلكة من الغاز الطبيعي، وذلك من خلال الاطلاع على عدد من المراجع والدراسات التي تناولت هذه المنهجية في تحليل السلاسل الزمنية، حيث سيتم التعامل مع سلسلة كمية الغاز الطبيعي الشهرية، والتوصل إلى النموذج المناسب الذي يمثل هذه السلسلة لتحديد وتوفيق النموذج الأمثل للتنبؤ في المستقبل، وذلك وفق المراحل الأربعة لمنهجية Box & Jenkins الموسمية، وسيتم الاعتماد في تحليل بيانات الدراسة على البرنامج الإحصائي R.

الجانب النظري

السلسلة الزمنية الموسمية: Seasonal Time Series

يقصد بها مجموعة من القيم المشاهدة المرتبطة مع بعضها تولدت بشكل متعاقب مع استمرار الزمن وتحتوي على ظاهرة الموسمية والتي تشير إلى النمط المتماثل لحركة السلسلة الزمنية في الأشهر المتقابلة خلال السنوات المتتالية (Brock Well & Davis, 1991: 53)، أي أن السلسلة تعيد نفسها بعد فترات زمنية ثابتة (Fixed intervals) وتدعى هذه الفترة بالفترة الموسمية ونرمز لها بالرمز (S) وقد تكون (S) سنة أو فصلاً أو شهراً، أي أن:

$$\boldsymbol{f(t+s)=f(t)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (\mathbf{1})$$

ويصعب تمييز الموسمية إذا كانت مدمجة مع الاتجاه العام وهذه المشكلة يمكن تفاديها عن طريق تحديد الموسمية عندما تكون البيانات مستقرة، أي أن وجود الاتجاه العام في البيانات يعني أنها غير مستقرة وبالتالي يمكن تحويلها إلى بيانات مستقرة باستخدام الفروق.

نماذج السلاسل الزمنية الموسمية: Seasonal Time Series Models

نماذج الانحدار الذاتي الموسمية Seasonal Autoregression Models (SAR)

يقال للنموذج أنه نموذج انحدار ذاتي موسمي من الرتبة (P)، والذي يرمز له بالرمز SAR(P) إذا حقق الصيغة الآتية: (الطائى والكورانى، 2008)

$$\boldsymbol{y_t = \phi_s y_{t-s} + \phi_{2s} y_{t-2s} + \cdots + \phi_{Ps} y_{t-Ps} + \varepsilon_t} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (\mathbf{2})$$

حيث أن:

y_{t-2s}: قيم مشاهدات السلسلة الزمنية، $i = 0,1,2, \ldots\ldots, p$

s: طول الفترة الموسمية.

ϕ_{is}: معالم الانحدار الذاتي الموسمي.

p: درجة النموذج الموسمي.

$\{\varepsilon_t\}$: سلسلة الاضطرابات الهادئة حيثε_t تتبع توزيع طبيعي متوسطة صفر، وتباينه ثابتσ^2.

نماذج المتوسطات المتحركة الموسمية: Seasonal Moving Average Models (SMA)

الصيغة العامة لنموذج المتوسط المتحرك الموسمي من الدرجة (Q) هي كالتالي:

$$\boldsymbol{y_t = \varepsilon_t - \theta_1 \varepsilon_{t-s} - \theta_2 \varepsilon_{t-2s} - \cdots - \theta_Q \varepsilon_{t-Qs}} \ldots\ldots\ldots (\mathbf{3})$$

إذ أن:

θ_{is}: معالم المتوسط المتحرك الموسمي.

النماذج المختلطة (انحدار ذاتي ومتوسطات متحركة) الموسمية

Seasonal Mixed Autoregression Moving Average Models (SARMA)

الصيغة العامة للنموذج الموسمي المختلط من الرتبة (P,Q) تأخذ الشكل الآتي:

$$y_t = \varepsilon_t + \emptyset_s y_{t-s} + \cdots + \emptyset_{ps} y_{t-ps} - \theta_s \varepsilon_{t-s} - \ldots - \theta_{Qs} \varepsilon_{t-Qs} \ldots\ldots\ldots (\mathbf{4})$$

نماذج الانحدار الذاتي والمتوسطات المتحركة التكاملية الموسمية

Seasonal Auto regression Integrated Moving Average Models (SARIMA)

عند دمج النماذج الموسمية مع النماذج عير الموسمية نحصل على نموذج الانحدار الذاتي والمتوسطات المتحركة التكاملية، ويرمز لها بالرمز$SARIMA(p,d,q)(P,D,Q)_s$، والذي يكتب على الصورة التالية:(Box and Jenkins, 1976)

$$\boldsymbol{\phi_P(B^S)\nabla_S^D y_t = \theta_Q(B^S)\varepsilon_t} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (\mathbf{5})$$

حيث أن:

D.: عامل الفرق الموسمي من الدرجة $\nabla_S^D = (1 - B^S)^D$

مراحل تطبيق منهجية Box & Jenkins الموسمية

يعتبر السكون من المفاهيم العامة في السلاسل الزمنية، حيث أن منهجية Box and Jenkins الموسمية مناسبة فقط لبيانات السلاسل الزمنية التي تتسم بالسكون فلا يمكن تطبيق مراحلها في تحليل السلاسل الزمنية إلا إذا كانت السلسلة الزمنية محل الدراسة ساكنة، حيث أن أول افتراض يتم اختباره هو التحقق من سكون السلسلة (Matroushi,2011)، لذلك قبل الشروع بتطبيق المنهجية يجب تهيئة البيانات للتحقق من سكون السلسلة الزمنية، وذلك بيانياً عن طريق الرسم البياني للسلسلة الأصلية ومن خلال الرسم البياني لدالتي الارتباط الذاتي والارتباط الذاتي الجزئي، أو باختبارات إحصائية مثل: KPSS،ADF ، فإذا كانت السلسلة مستقرة فإنَّ البيانات مهيأة للتعرف، أما إذا كانت السلسلة غير مستقرة في الوسط والتباين، فإنَّه يتم معالجتها بطريقة الفروق المتتالية لإزالة الاتجاه العام في حالة عدم السكون حول الوسط، أما عدم ثبات التباين فيتم معالجته بأخذ الجذر التربيعي لبيانات السلسلة أو بأخذ تحويلة مقلوب البيانات لها، أو التحويلة اللوغاريتمية وهي أكثر التحويلات استخداماً لتثبيت التباين (Nuno, 1996).

تمر منهجية Box and Jenkins الموسمية بأربعة مراحل أساسية يتم من خلالها الوصول إلى النموذج الأمثل للتنبؤ بالسلسلة الزمنية محل الاهتمام، وتتلخص المراحل الأربعة الأساسية في الآتي:

مرحلة التعرف:

تعتبر مرحلة التعرف من أهم مراحل (Box and Jenkins, 1976)، حيث يتم من خلالها تحديد معالم النموذج $(p،d،q)$ في النماذج غير الموسمية، و(P, D, Q)في النماذج الموسمية، وذلك من خلال دالتي الارتباط الذاتي ACF والارتباط الذاتي الجزئي PACF، ومن ثم يتم مطابقة معاملات الارتباط الذاتي والجزئي مع السلوك النظري لدالتي الارتباط الذاتي والارتباط الذاتي الجزئي.

وفي هذه المرحلة عادةً يتم ترشيح عدة نماذج متقاربة يتم المقارنة بينها من خلال المعايير الإحصائية (AIC, AICc, BIC)، وللمفاضلة بين النماذج المرشحة لاختيار النموذج الأنسب الذي يمتلك أقل قيمة لهذه المعايير.

مرحلة التقدير:

ويقصد بتقدير النموذج تحديد معالمه، أي الحصول على قيم رقمية لمعلمات النموذج بناء على قيم السلسلة المشاهدة، ويتم تحديد هذه المعلمات باستخدام احدى طرق التقدير الآتية: (Abrahamand Ledolter, 1983).

- طريقة المربعات الصغرى Linear Least Squares Method.
- طريقة الإمكان الأكبر Maximum Likelihood Method.

مرحلة التشخيص:

بعد مرحلة تقدير النموذج لا بد من اختبار مدى ملائمة أو صلاحية النموذج لتمثيل بيانات السلسلة الزمنية وتوجد لذلك عدة طرق منها (حشمان،2002):

- مقارنة دالة الارتباط الذاتي للسلسلة الأصلية مع تلك المتولدة عن النموذج المقدر، فإذا لوحظ اختلاف جوهري بينهما، فإنه يكون دليلاً قطعياً على فشل عملية التحديد، وهذا يستدعي إعادة عملية بناء النموذج وتقديره من جديد.
- معاملات النموذج لا بد أن تكون ذات دلالة إحصائية أي تختلف معنوياً عن الصفر، ويستخدم لذلك اختبار ستيودنت فإذا كانت غير معنوية لا بد من استبعاد أحد رتب AR أو MA.
- تحليل البواقي: يتم التحقق من عشوائية البواقي من خلال اختبار دالة الارتباط الذاتي لبواقي النموذج المشخص ويتم ذلك باستخدام اختبار لاجونغ بوكس وبوكس بيرس.

التنبؤ:

التنبؤ هو المرحلة الأخيرة من مراحل التحليل الحديث للسلاسل الزمنية، وأنه لا يمكن الانتقال إلي هذه المرحلة إلا بعد الانتهاء من اجراء جميع الفحوص والاختبارات الإحصائية الضرورية لتشخيص النموذج الذي تم اختياره، حيث تمثل هذه المرحلة التطبيق العملي للنموذج المقترح من خلال الحصول على القيم المتوقعة للسلسلة محل الدراسة، كما أن الاهتمام لا يقتصر في مجالات الاحصاء بصفة عامة وفي مجال السلاسل الزمنية بصفة خاصة على ايجاد تنبؤات النقطة للمشاهدات المستقبلية وإنما يمتد هذا الاهتمام ليشمل بناء فترات الثقة لهذه المشاهدات(بري، 2002).

الجانب التطبيقي:

وصف البيانات

البيانات المتاحة لدينا هي بيانات شهرية بالكميات المستهلكة من الغاز الطبيعي في الولايات المتحدة الأمريكية للفترة الزمنية الممتدة من شهر يناير 2000 حتى شهر سبتمبر 2022بواقع 273 مشاهدة، وسيتم تحليل بيانات البحث باستخدام البرنامج الإحصائي R.

ولأخذ فكرة عامة عن البيانات سيتم عرض بعض المؤشرات الاحصائية البسيطة التي تدلنا على طبيعة البيانات كما يوضح الجدول (1) التالي:

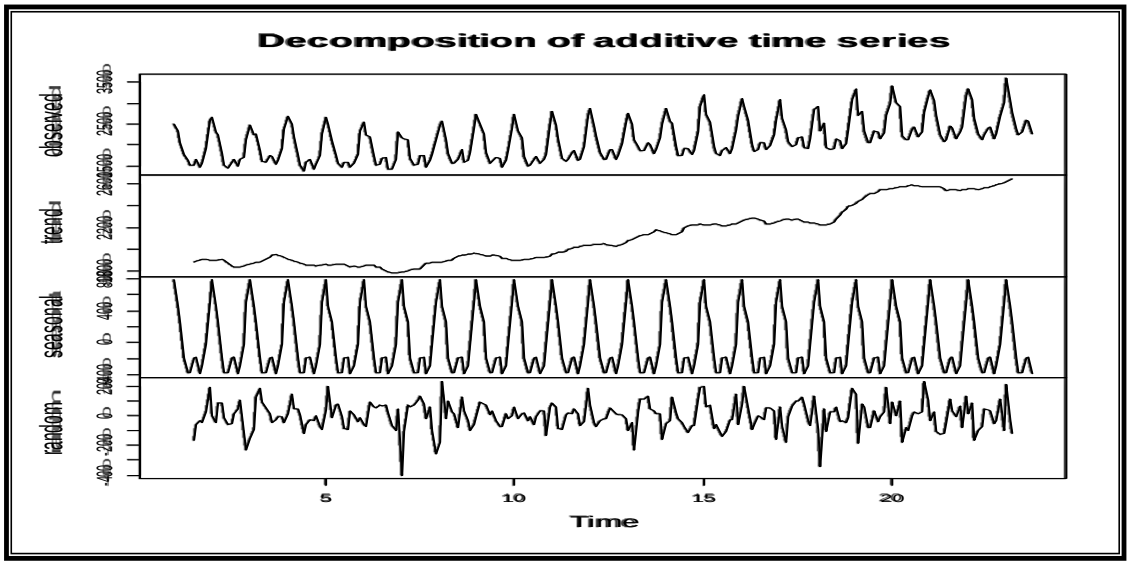
جدول (1): الاحصاءات الوصفية

الإحصاءات	حجم العينة	أقل قيمة	أكبر قيمة	الوسط الحسابي	الوسيط	الالتواء	التفرطح
القيمة	273	1368.4	3612.1	2126.2	2072.2	0.632	-0.254

تهيئة البيانات

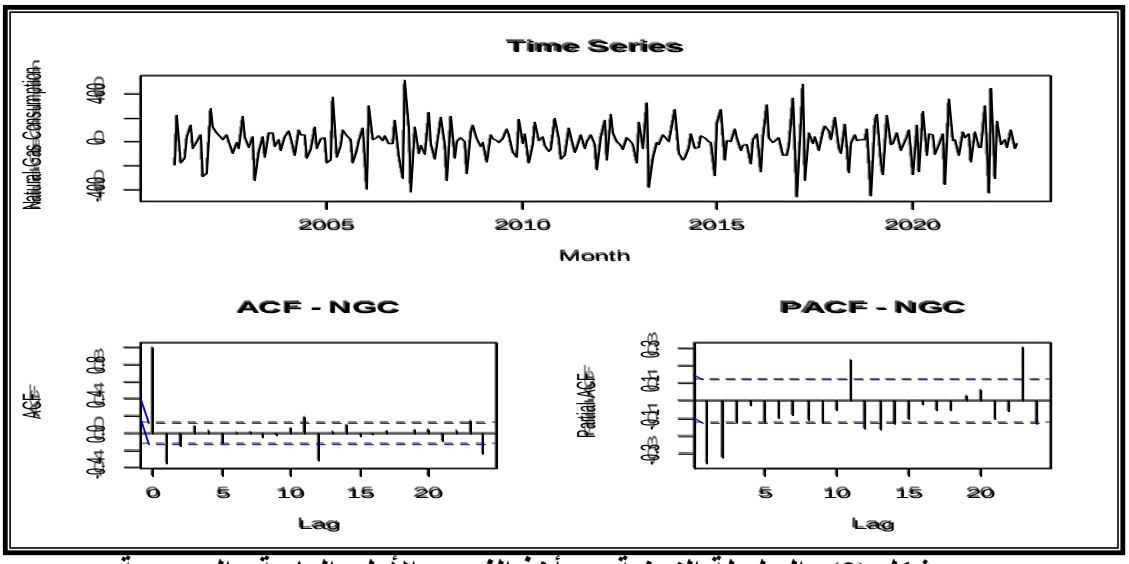
قبل البدء بتطبيق منهجية Box and Jenkinsالموسمية بمراحلها الأربع، لابدَّ من تهيئة البيانات من خلال تسكينها في الوسط والتباين، وتطبيق التحويلات اللازمة لجعلها مستقرة، ولمعرفة ذلك سيتم رسم مركبات السلسلة الزمنية ، للتعرف على خصائصها الأولية والشكل التالي يوضح ذلك:

شكل (1): مركبات السلسلة الزمنية الأصلية



يلاحظ من الشكل (1) وجود اتجاه عام متزايد مع الزمن فضلاً عن وجود تذبذبات تتكرر بانتظام وبنفس الوتيرة كل سنة، هذه التغيرات تشير على وجود مركبة الاتجاه العام والمركبة الموسمية، وهذا يعنى أنها غير مستقرة.

ولتسكين السلسلة تم استخدام الفروق العادية والفروق الموسمية لبيانات السلسلة.



شكل (2): السلسلة الزمنية بعد أخذ الفروق الأولى العادية والموسمية

من خلال شكل (2) يتضح أن السلسة استقرت بعد أخذ الفروق الأولى العادية والموسمية.

وللتأكد من صحة السكون في الوسط قمنا بإجراء اختبارين من اختبارات جذر الوحدة، هما اختبار KPSS، واختبار ADF والجدول (2) التالي يبين نتيجة الاختبارين.

جدول (2): نتيجة اختبارات جذر الوحدة.

الاختبار	p-value
KPSS	**0. 1 > 0.05**
ADF	**0.01 < 0.05**

يتضح أن قيمة p-value المقابلة لاختبار KPSS أكبر من 0.05، وأن قيمة p-value أقل من 0.05 لنتيجة اختبار ADF، وعليه فإنَّ نتائج الاختبارين تدل على استقرار السلسلة الزمنية محل الدراسة.

بناء نموذج SARIMA باستخدام منهجية Box and Jenkins الموسمية

سنطبق منهجية Box and Jenkins الموسمية للتنبؤ بالكميات المستهلكة من الغاز الطبيعي في الولايات المتحدة الأمريكية، حسب مراحلها التالية:

مرحلتي التعرف والتقدير

بعد أن تم تسكين السلسلة، تأتي مرحلة التعرف على النموذج المبدئي من خلال تحديد رتبة الانحدار الذاتي AR، والمعروفة بالرمز(p)، ورتبة المتوسطات المتحركة MA، والمعروفة بالرمز (q)، وكذلك تحديد رتبة الانحدار الذاتي الموسمي SAR، والمعروفة بالرمز(P)، ورتبة المتوسطات المتحركة الموسمية SMA، والمعروفة بالرمز (Q)، وتعتمد هذه الرتب في التقدير على أداتين أساسيتين هما: دالة الارتباط الذاتي (ACF)، ودالة الارتباط الذاتي الجزئي (PACF)، المرسومتين في الشكل (2) أعلاه:

وبالتالي يكون النموذج المبدئي هو $SARIMA(2,1,2)(2,1,2)_{12}$، وهذا النموذج يُؤخذ كنموذج أولي قابل للتعديل لاحقاً (بعد مرحلة التشخيص).

ولتوفيق أفضل نموذج للسلسلة، سنقوم بترشيح نماذج أخرى $SARIMA\ (p,d,q)\ (P,D,Q)_{12}$، ثم يتم المفاضلة بين هذه النماذج حسب أقل قيم للمعايير أدناه، كما هو مبين في جدول (3).

جدول (3): مقارنة بين نماذج SARIMA المرشحة حسب معايير التقييم

نماذج SARIMA	AIC	AICc	BIC
$SARIMA\ (2,1,2)\ (2,1,2)_{12}$	12.19643	12.19864	12.31969
$SARIMA\ (2,1,1)\ (2,1,1)_{12}$	12.19133	12.19261	12.28719
$SARIMA\ (1,1,2)\ (1,1,2)_{12}$	12.22806	12.22934	12.32393
$SARIMA\ (1,1,2)\ (2,1,1)_{12}$	12.19142	12.1927	12.28729
$SARIMA\ (1,1,0)\ (0,1,2)_{12}$	12.38721	12.38757	12.44199
$SARIMA\ (2,1,1)\ (1,1,1)_{12}$	12.2383	12.23921	12.32047
$SARIMA\ (1,1,1)\ (1,1,1)_{12}$	12.23265	12.23326	12.30113
$SARIMA\ (1,1,2)\ (1,1,1)_{12}$	12.23693	12.23784	12.3191

نلاحظ من الجدول (3) أن النموذج الأفضل هو $SARIMA\ (2,1,1)\ (2,1,1)_{12}$، وذلك لامتلاكه أقل القيم للمعايير الإحصائية (AIC, AICc, BIC)، وبذلك فإن النموذج المقترح الذي يمكن استخدامه في التنبؤ هو النموذج.$SARIMA\ (2,1,1)\ (2,1,1)_{12}$، وأن تقديرات معالم النموذج أعلاه مبينة في الجدول (4) الآتي:

جدول (4): تقديرات معالم النموذج SARIMA (2,1,1) (2,1,1)12 المقترح

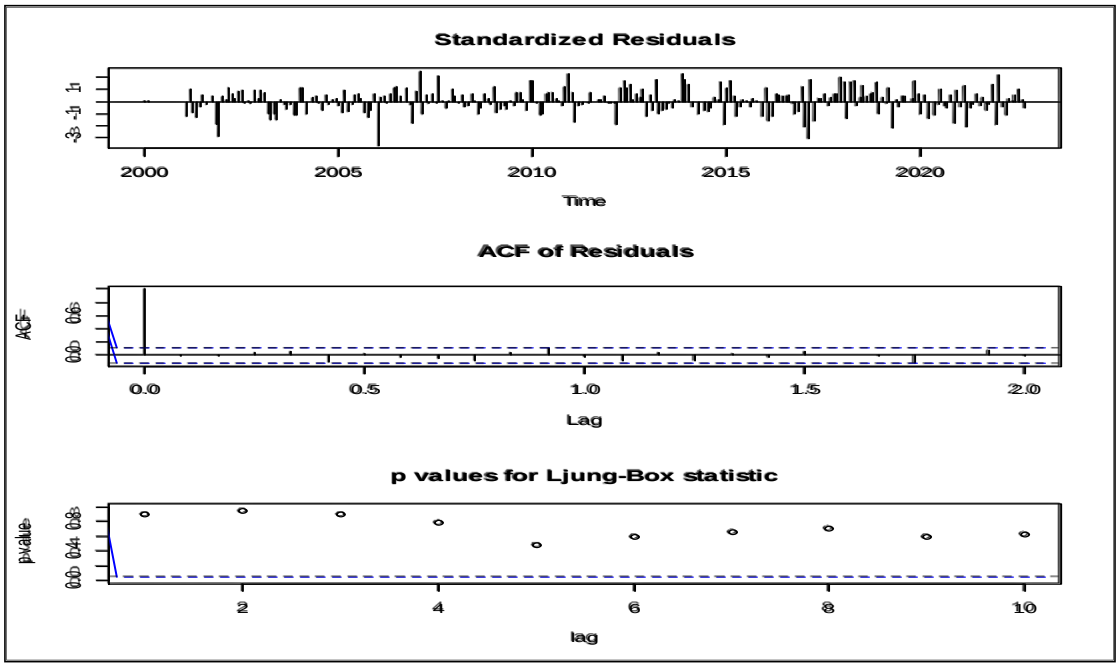
المعلمة	Estimate	SE	t.value	p.value
ar1	0.3342	0.0037	89.9642	0.0000
ar2	0.0174	0.0053	3.3095	0.0011
ma1	-0.9005	0.0287	-31.3995	0.0000
sar1	0.0430	0.0051	8.4462	0.0000
sar2	-0.2859	0.0021	-135.2057	0.0000
sma1	-0.7760	0.0238	-32.6137	0.0000

يتضح من خلال الجدول السابق أن قيمة p-value المقابلة لتقدير المعالم أقل من 0.05، وهذا دليل على معنوية معالم النموذج المقترح.

مرحلة التشخيص:

تُعد هذه المرحلة من أهم مراحل التحليل حيث يتم فيها ملائمة النموذج، وذلك من أجل تحسين النموذج وتطويره أو الابقاء عليه كما هو، حيث نقوم في هذه المرحلة بإخضاع النموذج محل الدراسة لعدد من الاختبارات لتقويم النموذج، يُعتبر تحليل البواقي جزء أساسي ومرحلة مهمة لمعرفة مدى صلاحية النموذج **$SARIMA\ (2,1,1)\ (2,1,1)_{12}$** المستخدم للتنبؤ، ومن خلال الشكل (3) التالي رسم البواقي الجزء الأول يتضح من خلاله أن البيانات تتأرجح بشكل ثابت حول الصفر، ويؤكد ذلك رسم دالة الارتباط الذاتي الجزء الثاني من الشكل نفسه حيث يتبين أن جميع معاملات الارتباط للبواقي عند كل فجوة زمنية تقع داخل حدي فترة الثقة، وهذا دليل على استقلال البواقي حيث أن معاملات الارتباط الذاتي للبواقي لا تختلف معنوياً عن الصفر.

أما الجزء الأخير من الشكل فهو يتعلق باختبار Ljung-Box حيث كانت قيم P-value عند الفجوات الزمنية جميعها أكبر من 0.05، مما يدل على أن البواقي مستقلة.



شكل (3): نتائج تحليل البواقي SARIMA (2,1,1) $(2,1,1)_{12}$

وبناء على ما سبق، يمكننا التأكيد على أن نموذج **SARIMA (2,1,1) $(2,1,1)_{12}$** اجتاز مرحلة الفحص والتشخيص ويمكن استخدامه في التنبؤ.

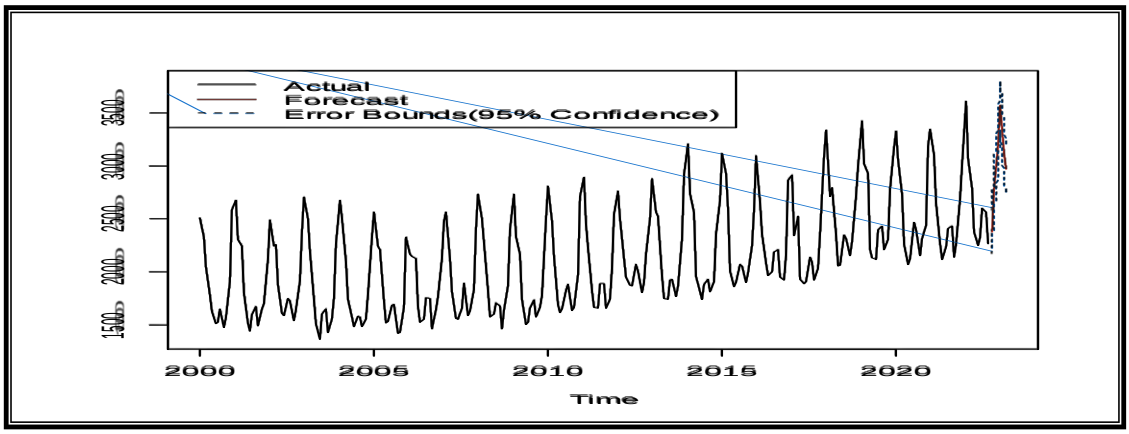
مرحلة التنبؤ:

تمَّ التوصل إلى أنَّ النموذج الأفضل للتنبؤ بالكميات المستهلكة من الغاز الطبيعي في الولايات المتحدة الأمريكية هو **SARIMA (2,1,1) $(2,1,1)_{12}$**، وذلك لاجتيازه معظم عمليات الفحص والتشخيص بدرجة جيدة إحصائياً عن غيره من النماذج الأُخرى، لذا سيتم استخدامه في التنبؤ بستة قيم مستقبلية من أكتوبر2022 حتى مارس2023 ويتضح ذلك من خلال الجدول (5) التالي:

جدول (5): القيم المتنبأ بها لنموذج SARIMA (2,1,1) (2,1,1)12

التاريخ	القيم المتنبأ بها
Oct/ 2022	2372.935
Nov/ 2022	2786.874
Dec /2022	3188.734
Jan /2023	3566.705
Feb / 2023	3115.149
Mar / 2023	2966.682

وشكل (4) التالي يوضح بيانياً القيم الفعلية من شهر يناير 2000 حتى شهر سبتمبر 2022مضافاً إليها القيم المتنبأ بها من شهر أكتوبر2022 حتى شهر مارس2023، حيث تشير التنبؤات المستقبلية أنها تتبع نفس السلوك للسلسلة الأصلية.



شكل (4): القيم الفعلية والتنبؤية للكميات المستهلكة من الغاز الطبيعي في الولايات المتحدة الأمريكية من شهر يناير 2000 حتى شهر مارس 2023

النتائج:

- السلسلة الزمنية الشهرية للكميات المستهلكة من الغاز الطبيعي في الولايات المتحدة الأمريكية تحتوي على وجود مركبة الاتجاه العام والمركبة الموسمية، وهذا يعنى أنها غير مستقرة.
- تحقق خصائص البواقي في النموذج **SARIMA (2,1,1) $(2,1,1)_{12}$** وتجاوز مرحلة الفحص والتشخيص.
- النموذج الأمثل للتنبؤ بالكميات المستهلكة من الغاز الطبيعي في الولايات المتحدة الأمريكية هو **SARIMA (2,1,1) $(2,1,1)_{12}$**
- تشير التنبؤات المستقبلية أنها تتبع نفس السلوك للسلسلة الأصلية.

التوصيات:

- التنقيب المستمر عن البيانات لتنقية بعض العوامل الأخرى التي يمكن أن تؤثر على بالكميات المستهلكة من الغاز الطبيعي في الولايات المتحدة الأمريكية.
- نوصى باستخدام أساليب تنبئية أخرى كأسلوب التمهيد الأسي وأسلوب دالة التحويل ومقارنة نتائج هذه الأساليب مع نتائج بوكس جينكنز الموسمي المقترح.
- استخدام النماذج الهجينة للتنبؤ بالكميات المستهلكة من الغاز الطبيعي.

المراجع

أولاً: المراجع العربية

بري، عدنان (2002). " طرق التنبؤ الإحصائي"، الجزء الأول، جامعة الملك سعود.

حشمان، مولود. (2010). نماذج وتقنيات التنبؤ قصير المدى. ط3، الجزائر: ديوان المطبوعات الجامعية.

الدوب، طارق. (2021). استخدام نماذج تحليل السلاسل الزمنية للتنبؤ بإجمالي الجرائم في الكويت. مجلة دراسات الخليج والجزيرة العربية، 47(181)، 248 – 235.

الطائي، فاضل والكوراني، جيهانى. (2008). التنبؤ بنماذج ARIMA الموسمية باستخدام طرائق التمهيد الأسى مع التطبيق. المجلة العراقية للعلوم الإحصائية، (12)، 171-205.

عوض الله، خلود. (2016). استخدام نماذج SARIMA و Holt - Winters في التنبؤ بالسلاسل الزمنية الموسمية (رسالة ماجستير غير منشورة)، كلية الاقتصاد والعلوم الإدارية، جامعة الأزهر، فلسطين.

المرشدي، كرار. (2021). تشخيص وتقدير نماذج السلاسل الزمنية الموسمية مع تطبيق عملي (رسالة ماجستير غير منشورة)، كلية الادارة والاقتصاد، جامعة كربلاء، العراق.

ثانياً: المراجع الأجنبية:

Abraham, B. and Ledolter, J., (1983). "Statistical Methods for Forecasting", John Wiley and Sons, New York.

Aliyu, A., Auwal, A., and Adenomon, M. (2021). Application of Sarima Models in Modelling and Forecasting Monthly Rainfall in Nigeria. Asian Journal of Probability and Statistics, 13(3), 30-43.

Box, G.E.P. and Jenkins, G.M., (1976). "Time Series Analysis; Forecasting and Control", 2nd ed., Holden-Day: San Francisco.

Brock Well, P.J., and Davis, R.A. (1991), " Time Series Theory and Methods ،2"nd ed, Springer Verlag New York Inc, New York.

Chowdhury, M., Mondal S., Islam, J. (2018). Modeling and Forecasting Humidity in bangladesh: Box – Jenkins approach. International Journal of Research – Granthaalayah, 6(4),50-60.

Divisekara, R., Jayasinghe, G., and Kumari, K.(2020). Forecasting the red lentils commodity market price using SARIMA models. Journal the Springer Nature SN Bus Econ, 1-20.

Ilmayasinta, N. (2021). Forecasting Arrival of Foreign Tourists Using Seasonal Arima Box – Jenkins. Jurnal Ilmu Matematika dan Terapan, 15(2), 223-230.

Mamudu, L., Yahaya, A., Dan'azumi, S. (2021). Application of Seasonal Autoregressive Integrated Moving Average (SARIMA) For Flows of River Kaduna, Nigerian Journal of Engineering, 28(2), 84-90.

Matroushi, S. (2011). "Hybrid computational intelligence systems based on statistical and neural networks methods for time series forecasting: the case of gold price", Lincoln University, United Kingdom.

Nuno, C. (1996). "Some Results on the Spectral Analysis of stationary Time Series" Portugal Mathematic, Vol. 53, Fasc. 2.

Abstract:

The Time Series Analysis method is one of the statistical methods worthy of attention. The time series analysis method relies on tracking the phenomenon (or variable) over a specific time period (for example, several years), and then predicts the future based on the different values that appeared in the time series and on the growth pattern. in values.

This research deals with identifying the steps and stages of the seasonal Box-Jenkins methodology in forecasting at the theoretical and applied level, to build the best model that helps predict the quantities of natural gas consumed in the United States of America, and by comparing several different models with the proposed model, it turns out that the optimal model is the SARIMA model (2,1,1) (2,1,1)12 because it passed most of the examination and diagnosis processes with a statistically good degree compared to other models, and obtained the lowest values of the evaluation criteria (AIC, AICc, BIC), and accordingly it was used in predicting future values during The period from October 2022 to March 2023, and the predictive values were found to follow the same behavior as the original series.

Keywords: SARIMA model, forecasting, seasonal time series