



Using Time Series Analysis to Forecast the Number of Registered Cancer Patients in the Municipality of Tobruk

Alshareef Masoud Alsunousi^{1*}, Eenas Adrees Mohammed²

^{1,2} Department of Statistics, Faculty of Economics and Political Science, University of Tobruk, Tobruk, Libya

alshareef.masoud@tu.edu.ly

نمذجة السلاسل الزمنية باستخدام ARIMA للتنبؤ بإصابات الأورام في طبرق
دراسة تحليلية للفترة 2018-2027

الشريف مسعود السنوسي^{1*}، إيناس إدريس محمد²
^{2,1} قسم الإحصاء كلية الاقتصاد والعلوم السياسية، جامعة طبرق، طبرق، ليبيا.

تاريخ الاستلام: 2025-05-27 تاريخ القبول: 2025-07-05 تاريخ النشر: 2025-07-14

Abstract:

This study aims to apply the Box–Jenkins methodology in modeling and analyzing the monthly time series data of cancer patients registered at the Oncology Department of Tobruk Medical Center during the period from January 2018 to December 2024. This time series holds significant importance in supporting healthcare decision-making and strategic planning, particularly in light of the continuous increase in cancer incidence. The analysis was conducted using ARIMA models. A range of candidate models was evaluated based on key statistical accuracy metrics, including the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), in addition to the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). Upon comparing the competing models, the results indicated that the most suitable model in terms of both goodness of fit and forecasting accuracy was ARIMA(0,1,1), as it yielded the lowest error metrics and most favorable information criteria values. The predictive power of this model was confirmed through the close alignment between the forecasted and actual values, reflecting its high accuracy in capturing the underlying dynamics of the studied time series. Based on this optimal model, statistical forecasts were generated for the expected number of patients over the future period (2025–2027), providing valuable insight for improving medical planning and healthcare service delivery.

Keywords: Cancer, ARIMA, Box–Jenkins, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Statistical Forecasting Analysis.

المخلص:

يهدف هذا البحث إلى استخدام منهجية بوكس – جنكنز (Box-Jenkins) في نمذجة وتحليل السلاسل الزمنية الشهرية لعدد مرضى الأورام المسجلين بقسم الأورام في مركز طبرق الطبي خلال الفترة الممتدة من يناير

2018 إلى ديسمبر 2024. تشمل هذه السلسلة أهمية بالغة في دعم اتخاذ القرار الصحي والتخطيط، لا سيما في ظل التزايد المستمر في حالات الإصابة بالأورام السرطانية .

تم إجراء التحليل باستخدام نماذج (ARIMA) حيث تم اختيار مجموعة من النماذج وفقا لمعايير تقييم الدقة الإحصائية وهي الجذر التربيعي لمتوسط مربعات الأخطاء (RMSE) ، متوسط الخطأ المطلق (MAE) ، متوسط النسبة المئوية للخطأ المطلق (MAPE) ، بالإضافة إلى معيار أكايكي للمعلومات (AIC) و معيار معلومات بيزيان (BIC) . وبعد مقارنة النماذج المختلفة ، أظهرت النتائج أن النموذج الأفضل من حيث الجودة والتنبؤ هو (ARIMA(0,1,1) ، و الذي تميز بأقل قيم لمقياس الخطأ وأفضل معايير المعلومات.

حيث أثبت هذا النموذج كفاءته من خلال تقارب القيم المتنبأ بها مع القيم الفعلية مما يدل على دقته العالية في تمثيل السلسلة الزمنية المدروسة . وبناء على هذا النموذج تم إجراء تنبؤات إحصائية لعدد المرضى المحتملين خلال الفترة المستقبلية (2025-2027) وهي تنبؤات يمكن أن تساهم في تحسين التخطيط الطبي والخدمات العلاجية بالمؤسسات

الكلمات الدالة: الأورام، ARIMA، بوكس-جنكنز، معيار معلومات أكايكي (AIC)، معيار معلومات (BIC) تحليل التنبؤ الإحصائي .

المقدمة:

تناول هذا البحث الجانب الصحي ولأهميته في التقليل من الإصابات وذلك لما يمثله من أهميه في الحد من معدلات الإصابة التي تشهد تزايدا ملحوظا في الفترة الأخيرة .

حيث اعتمدت هذه الدراسة على البيانات الشهرية لأعداد المصابين بمرض الأورام في الفترة من يناير 2018 إلى ديسمبر 2024 كسلسلة زمنية ، حيث تعد السلاسل الزمنية من أهم الأساليب الإحصائية التي لها القدرة على تحليل ورصد التغيرات التي تطرأ على القيم المرتبطة بعامل الزمن مما يساعدنا في فحص تلك المتغيرات واختيار النموذج المناسب للتنبؤ في المستقبل ، وكما هو الحال في تقدير أعداد المصابين بمرض الأورام حيث يعد هذا التنبؤ خطوة في دعم اتخاذ قرارات وقائية فعالة للحد من الانتشار مستقبلا .

هدف البحث:

الهدف من هذه الدراسة التنبؤ بأعداد المرضى المصابين بمرض الأورام في السنوات القادمة لي أهميته في مراكز الأورام و إدارة المستشفيات ، وكذلك الحكومة في التخطيط لبرامج صحية مناسبة من حيث توفير الأدوية العلاجية و كذلك تحديد أطباء ذوي كفاءات عالية والمناسبة لهذا العدد من المصابين .

فرضية البحث:

يشهد عدد المصابين بمرض الأورام عددا متزايدا خلال الفترة من يناير 2018 إلى ديسمبر 2024 ويعد التنبؤ بأعداد المصابين بهذا المرض في مركز طبوق الطبي قسم الأورام للأورام بمثابة مدخل أساسي للتنبؤ بعدد الإصابات في باقي المراكز الصحية الأخرى .

منهجية البحث:

تم تقسيم البحث إلى جانبين ، تناولنا الأول الجانب النظري بالأسس الخاصة بنماذج السلاسل الزمنية ومراحل بناء النموذج و طرق التقدير و التنبؤ في حين تناول الجانب التطبيقي الدراسة على بيانات واقعية لمرض الأورام للحصول على نموذج مناسب للتنبؤ بعدد المرضى في فترات لاحقة وذلك تم باستخدام برنامج SPSS

أولا: الجانب النظري

يعد تحليل السلاسل الزمنية من الموضوعات الهامة في الإحصاء التطبيقي، لما له من دور فاعل في فهم سلوك الظواهر المتغيرة عبر الزمن. ويعنى هذا النوع من التحليل بدراسة البيانات المرتبة زمنياً بهدف التعرف على العوامل المؤثرة في تطورها ضمن فترة زمنية محددة. إن الهدف الرئيس من تحليل السلاسل الزمنية يتمثل في الوصول إلى توصيف دقيق للعملية العشوائية التي تنشأ عنها هذه السلسلة ويتم ذلك من خلال تحديد مكونات السلسلة الأساسية مثل الاتجاه العام (Trend) ، الموسمية (Seasonality) ، والتقلبات العشوائية (Random fluctuations) ، ومن ثم بناء نموذج إحصائي ملائم للسلسلة. كما يستخدم النموذج الناتج لاحقاً في إجراء التنبؤات المستقبلية، وهي من التطبيقات المركزية في هذا المجال التي تتطلب قرارات مبنية على توقعات دقيقة

وبالتالي، فإن تحليل السلاسل الزمنية لا يقتصر على الوصف فقط، بل يعد أداة مهمة في دعم السياسات واتخاذ القرارات المستندة إلى البيانات.

1-1 السلسلة الزمنية The Time Series

هي عبارة عن مجموعة من المشاهدات المرتبطة مع بعضها البعض والتي تم تسجيلها عبر فترات زمنية متتالية $[X_1, X_2, \dots, X_t]$ وبذلك كل قيمة X_t تمثل قيمة المشاهدة التي وقعت في الزمن t ، حيث X_1 تمثل قيمة المشاهدة في الزمن t_1 وكذلك X_2 تمثل قيمة المشاهدة التي وقعت في الزمن t_2 وان الهدف الرئيسي من استخدام بيانات السلاسل الزمنية هو التنبؤ بالمستقبل وبذلك نستطيع تحليل الطلب في الماضي لكي نستطيع أن نعرف كيف يمكن أن يكون الطلب في المستقبل وهذا يساعد صناع القرار أنهم يأخذون قرارات مناسبة.

2-1 سكون السلسلة الزمنية Stationary Time Series

تعتبر السلسلة الزمنية مستقرة إذا كانت خصائصها الإحصائية مثل الوسط والتباين و التغيرات ثابتة مع مرور الزمن أي إن هذه الخصائص لا تتغير إذا أزحنا السلسلة إلى الأمام أو الخلف في الزمن، أي أن

- 1- ثبات الوسط $E(X_t) = \mu$
- 2- ثبات التباين $var(X_t) = E(X_t - \mu)^2 = \sigma^2$
- 3- دالة الارتباط الذاتي تعتمد علي الفجوة الزمنية (t)

3-1 الارتباط الذاتي (ACF) Autocorrelation

هو مؤشر يستخدم لقياس مدى قوة العلاقة بين القيم الحالية لمتغير معين والقيم السابقة له عند فترات تأخير مختلفة (k) وتتراوح قيمته بين $(-1, +1)$ أي أنه يقع ضمن النطاق $(-1 \leq \rho_k \leq +1)$ ويتم حسابه بالمعادلة التالية

$$\rho_k = \frac{\text{cov}(X_t, X_{t+k})}{\sqrt{\text{var}(X_t)}\sqrt{\text{var}(X_{t+k})}} = \frac{\sum_{t=1}^{n-k} (X_t - \bar{X})(X_{t+k} - \bar{X})}{\sum_{t=1}^n (X_t - \bar{X})^2}$$

حيث X_t القيم المشاهدة للسلسلة و أن التوزيع الإحصائي لمعاملات الارتباط الذاتي يتبع التوزيع الطبيعي بمتوسط صفر وتباين $\left(\frac{1}{n}\right)$

أن الرسم البياني لمعاملات الارتباط الذاتي (ρ_k) مقابل فترات الإزاحة $(k = 1, 2, 3, \dots)$ يعرف بأسم دالة الارتباط الذاتي ويرمز لها بالرمز (ACF) وتعتبر دالة الارتباط الذاتي مهمة لتحليل مدى استقرار السلسلة الزمنية بمعنى في حالة السلاسل الزمنية المستقرة فأن معاملات الارتباط الذاتي تميل الى الانحدار بسرعة نحو الصفر مع ازدياد فترات الإزاحة (k) أو تنقطع فجئه بعد عدد من فترات الإزاحة $(k = q)$ أي ان

$$\rho_k = 0 \quad \forall k > q$$

أما في حالة عدم استقرار السلسلة الزمنية بسبب وجود اتجاه عام صاعد أو نازل في المتوسط فإن دالة الارتباط (ACF) لا تنخفض بسرعة نحو الصفر و لا تنقطع، السبب تأثير القيم بنفس الاتجاه العام لمتوسط السلسلة لفترة طويلة.

وذلك يؤدي إلى ظهور معاملات ارتباط ذاتي مرتفعة عند فترات إزاحة طويلة، وتعتبر دالة الارتباط الذاتي للوفاقي Residual Autocorrelation Function (RACF) وسيلة مهمة لفحص مدى ملائمة النموذج عن طريق اختبار عشوائية الأخطاء.

وبصورة عامة فان دالة الارتباط الذاتي (ACF) في السلاسل الزمنية المستقرة تنخفض بسرعة وتقترب من الصفر مع زيادة درجات الإبطاء. ويمكن كتابة دالة الارتباط الذاتي على الصورة

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\gamma_k}{\sigma^2}, k = 0, \mp 1, \mp 2, \dots$$

4-1: دالة الارتباط الذاتي الجزئي (PACF) Partial Autocorrelation Function ,

يعتبر معامل الارتباط الذاتي الجزئي مؤشرا لقياس العلاقة بين X_t , X_{t+k} لنفس السلسلة الزمنية وذلك مع فرض استقرار السلسلة . حيث أن هذا المعامل يمثل الحد الأخير من نموذج الانحدار الذاتي من الدرجة $AR(P)$.

ويمكن حساب قيم معامل الارتباط الذاتي الجزئي باستخدام دالة الارتباط الذاتي (ACF) من الصيغة التالية

$$\hat{\varphi}_{K+L, K+1} = \frac{\sum_{j=1}^k \hat{\varphi}_{kj} \hat{\rho}_{k+1-j}}{1 - \sum_{j=1}^k \hat{\varphi}_{kj} \hat{\rho}_j}$$

تستخدم دالة الارتباط الذاتي الجزئي (PACF) بشكل أساسي في تحليل السلاسل الزمنية وكذلك تساعد في تحدي النموذج المناسب من بين نماذج العمليات العشوائية المستقرة كما تساعد على تحديد درجة النموذج ، وتستخدم أيضا لفحص مدى ملائمة النموذج لبيانات من خلال اختبار عشوائية أخطاء البواقي.

5-1: نماذج بوكس - جنكنز Box & Jenkins في تحليل السلاسل الزمنية

1-5-1: نموذج الانحدار الذاتي (AR) Autoregressive Model .

يطلق على النموذج الانحدار الذاتي من الرتبة (P) ويرمز له بالرمز (P) AR وفق الصيغة التالية

$$X_t = \alpha + \theta_1 X_{t-1} + \theta_2 X_{t-2} + \dots + \theta_P X_{t-P} + a_t$$

حيث X_t قيم مشاهدات السلسلة

α ثابت ، P تمثل رتبة الانحدار الذاتي ، $-1 < \theta < 1$ ، $\theta_1, \theta_2, \dots, \theta_P$ معالم نموذج الانحدار الذاتي الأخطاء العشوائية التي تتوزع طبيعياً بوسط حسابي صفر و تباين σ^2 ثابت .

2-5-1: نموذج الوسط المتحرك (MA) Moving Average Model .

يسمى هذا النموذج باسم نموذج الوسط المتحرك من الرتبة (q) ويرمز له بالرمز (q) MA إذا تحققت الصيغة التالية

$$X_t = \alpha - \phi_1 a_{t-1} + \phi_2 a_{t-2} + \dots + \phi_q a_{t-q} + a_t$$

حيث

α ثابت ، q تمثل رتب الأوساط المتحركة ، $-1 < \phi < 1$ ، $\phi_1, \phi_2, \dots, \phi_P$ معالم نموذج الوسط المتحرك.

3-5-1: نماذج الانحدار الذاتي و الأوساط المتحركة (النماذج المختلطة)

Mixed Autoregressive Moving Average , (ARMA)

هناك ظواهر لا نستطيع التعبير فيها عن السلسلة الزمنية بصيغة الانحدار الذاتي أو الوسط المتحرك فقط ولكن يمكن التعبير عنها بواسطة نموذج يجمع الاثنين معا الانحدار الذاتي و المتوسطات المتحركة يسمى $ARMA$ (P , q) و يكتب بالصيغة التالية

$$X_t = \alpha + \theta_1 X_{t-1} + \theta_2 X_{t-2} + \dots + \theta_P X_{t-P} + a_t - \phi_1 a_{t-1} - \phi_2 a_{t-2} - \dots - \phi_q a_{t-q}$$

1-5-4: نماذج الانحدار الذاتي و المتوسطات المتحركة التكاملية

(ARIMA) , Multiplicative Model

تعد نماذج المتوسطات المتحركة الذاتية المتكاملة (ARIMA) من أكثر نماذج السلاسل الزمنية استخدامًا وشيوعًا في الدراسات الإحصائية والتنبؤية، خصوصًا عندما تكون السلسلة الزمنية غير مستقرة (Non-stationary). وتكمن قوة هذه النماذج في قدرتها على تمثيل ومعالجة الاتجاهات الزمنية المعقدة من خلال دمج ثلاث مكونات رئيسية الانحدار الذاتي (AR) ، والمتوسط المتحرك (MA) ، بالإضافة إلى الفرق (التكامل) الذي يجعل السلسلة مستقرة.

ولكي تكون السلسلة الزمنية قابلة للنمذجة باستخدام نماذج (ARIMA) ، يشترط أولاً تحويلها إلى سلسلة مستقرة، أي ذات وسط وتباين ثابتين على طول الزمن. وتعد عملية أخذ الفرق (Differencing) الطريقة الأكثر شيوعًا لتحقيق هذا الشرط. يتم تطبيق الفرق بمقدار (d) من مرات على السلسلة الأصلية، مما يؤدي إلى إزالة الاتجاه العام أو الموسمية، وتحويل النموذج من الشكل التقليدي (ARMA(p, q) إلى الشكل (ARIMA(p, d, q)، ويكتب النموذج العام لنموذج ARIMA على الصورة التالية:

$$\theta_p(B) \nabla^d X_t = \alpha + \phi_q(B) \phi_q$$

حيث

$\theta_p(B)$: معامل الانحدار الذاتي

$\phi_q(B)$: معامل الوسط المتحرك

p: رتبة نموذج الانحدار الذاتي

d: درجة الفرق

q: رتبة نموذج الوسط المتحرك

∇^d : معامل الفرق عند الزمن d حيث $\nabla = (1-B)$

1-6-6 : مراحل منهجية Box – Jenkins

تعد منهجية بوكس-جنكنز (Box & Jenkins, 1976) واحدة من أكثر الطرق شهرة وفاعلية في تحليل السلاسل الزمنية وبناء نماذج تنبؤية من نوع (ARIMA). تعتمد هذه المنهجية على فهم البنية الإحصائية للسلسلة الزمنية، وتحديد النموذج الأمثل الذي يمكنه تمثيل الظاهرة المدروسة بدقة ثم استخدامه في التنبؤات المستقبلية.

1-6-1: مرحلة التعرف Identification

تعد مرحلة التعرف مهمة في بناء نماذج السلاسل الزمنية حيث يتم فيها اختبار النموذج المناسب الذي يتوافق مع السلسلة الزمنية ، وذلك يتم من خلال دراسة دالتي الارتباط الذاتي (ACF) و الارتباط الذاتي الجزئي (PACF) وتحليل الأشكال البيانية لكل منها لتحديد أفضل قيم لمعاملات النموذج (p, d, q) وهذا كله لا يتم إلا بعد التأكد من تحقيق الاستقرار في السلسلة الزمنية .

جدول (1) يوضح الخصائص النظرية لدالة ACF و PACF

| نوع النموذج | ACF | PACF |
|----------------|----------------------------|----------------------------|
| AR (p) | يتلاشى أسياً | ينقطع بعد الإزاحة p |
| MA (q) | ينقطع بعد الإزاحة q | يتلاشى أسياً |
| ARMA (p , q) | يتلاشى بعد الإزاحة (p-q) | يتلاشى بعد الإزاحة (p-q) |

1-6-2: مرحلة التقدير Estimation

يتم في هذه المرحلة تقدير معالم نموذج الانحدار الذاتي و نموذج الوسط المتحرك وذلك باستخدام بيانات السلسلة الزمنية و هذى يكون بعد تحديد النموذج المناسب .

هناك العديد من الطرق لتقدير معالم النموذج نذكر منها

1- طريقة العزوم The Method of Moments

2- طريقة الإمكان الأعظم Maximum Likelihood Method

3- طريقة المربعات الصغرى Least Squares Method

1-6-3: مرحلة التشخيص Diagnostic

تأتي مرحلة التشخيص بعد مرحلة التقدير حيث يتم فيها تشخيص مدى ملائمة النموذج المتحصل عليه لتمثيل بيانات السلسلة الزمنية والاعتماد عليه في الحصول على التنبؤات المستقبلية، هناك العديد من الاختبارات التي يمكن استخدامها لهذا الغرض منها

• اختبار معنوية معالم النموذج

يجب أن تكون المعالم ذات دالة إحصائية، وذلك يتم بالاعتماد على اختبار (de Student Test, t) فإذا كان النموذج يحتوي على معامل ليس له دالة إحصائية فالأبد من حذفه أي حذف رتبة النموذج (AR) أو (MA)

• تحليل البواقي

تتم عملية التحقق من ملائمة النموذج المقدر من خلال مقارنة دالة الارتباط الذاتي لسلسلة الأصلية مع المتولدة عن النموذج المقدر في حالة وجود اختلاف جوهري بين الدالتين ، هذا يدل على فشل تحديد النموذج بشكل صحيح وبذلك يتطلب إعادة عملية بناء النموذج و تقديره مرة أخرى ، أما إذا كان هناك تشابه بين الدالتين ننقل إلى تحليل بواقي النموذج وذلك من خلال فحص دالة الارتباط الذاتي للبواقي، ولكي يكون النموذج مناسباً لتمثيل السلسلة الزمنية يجب أن تقع جميع معاملات الارتباط الذاتي للبواقي داخل مجال الثقة .

$$[-T\alpha/2, T\alpha/2]$$

إذا تحقق الشرط السابق فهذا يدل على أن بواقي النموذج تتوزع بشكل عشوائي وذلك يعني إن النموذج ملائم ويمكن الاعتماد عليه في عملية التنبؤ المستقبلي وان الارتباطات الذاتية للبواقي تكون مستقلة و تتوزع طبيعياً بمتوسط يساوي صفر وتباين مقداره $(\frac{1}{n})$.

كذلك يمكن استخدام اختبار (Ljung-Box Q) لاختبار ما إذا كانت البواقي في النموذج مستقلة أو لا .

$$Q = n(n + 2) \sum_{i=1}^k \frac{\hat{\rho}^2(i)}{n - k} \sim \chi_{\alpha}^2(k - p - q).$$

إذا كانت قيمة الاختبار Q اقل من القيمة الحرجة لتوزيع X^2 فإننا نقبل الفرضية الصفرية وذلك يعني إن معاملات الارتباط الذاتي غير معنوية وهذا يدل على ملائمة النموذج أما إذا كانت Q اكبر من القيمة الحرجة لتوزيع X^2 ، فأنا نرفض الفرضية الصفرية وهذا يدل على إن معاملات الارتباط الذاتي ذات دلالة معنوية ويعد هذا مؤشر على عدم جودة النموذج في تمثيل السلسلة الزمنية .

• معايير التقييم

تستخدم هذه المعايير في المفاضلة بين النماذج المرشحة حيث يتم اختيار أفضل نموذج والذي يمتلك اقل قيمة لهذه المعايير وهذه المعايير هي

• الجذر التربيعي لمتوسط مربعات الأخطاء Root Mean Square Error

$$RMSE = \sqrt{\frac{\sum_{t=1}^n at}{n}}$$

• متوسط الخطاء المطلق Mean Absolute Error

$$MAE = \frac{\sum_{t=1}^n |at|}{n}$$

• متوسط الخطاء النسبي المطلق Mean Absolute Percentage Error

$$MAPE = \frac{\sum_{t=1}^n \left| \frac{at}{x_t} \right|}{n}$$

• معيار معلومات آكاكي Akaike Information Criterion, AIC

$$AIC = \ln(\sigma^2) + \frac{2(p + q)}{n}$$

• معيار معلومات بيزيان Bayesian Information Criterion , BIC

$$BIC = \ln(\widehat{\sigma}_e^2) + \frac{(p + q)}{n} \ln(n)$$

4-6-1: مرحلة التنبؤ Forecasting

تعتبر هذه المرحلة من أهم المراحل حيث يتم فيها التنبؤ بالقيم المستقبلية للسلسلة الزمنية وفق النموذج الملائم الذي تم اعتماده من المراحل السابقة حيث يعد النموذج الأمثل لتنبؤ هو الذي ينتج عنه أقل تقدير للأخطاء، ويمكن كتابته بالصيغة التالية

$$X_{t+L} = \widehat{X}_t(L) = \alpha + \theta_1 X_{t+L-1} + \theta_2 X_{t+L-2} + \dots + \theta_p X_{t+L-p} + a_{t+L} - \phi_1 a_{t+L-1} - \phi_2 a_{t+L-2} - \dots - \phi_q a_{t+L-q}$$

ثانياً: الجانب التطبيقي

1-2: المقدمة

تعتبر السلاسل الزمنية من الركائز الأساسية في التحليل الإحصائي، نظراً لما توفره من أدوات منهجية لتوصيف الأنماط الزمنية وتفسير الاتجاهات والتقلبات الموسمية، فضلاً عن استخدامها الفعال في التنبؤ بالقيم المستقبلية اعتماداً على البيانات التاريخية. وفي هذا السياق، يهدف هذا البحث إلى اختيار النموذج الإحصائي الأنسب من بين نماذج السلاسل الزمنية للتنبؤ بعدد الحالات المستقبلية لمرضى الأورام. وقد تم الحصول على البيانات الفعلية من قسم الأورام بمستشفى المركز الطبي بمدينة طبرق، عن الفترة الممتدة من عام 2018 إلى عام 2024. ويوظف هذا التحليل بغرض بناء نموذج تنبؤي قادر على تقدير معدلات الإصابة خلال الفترة 2025-2027، باستخدام منهجية بوكس - جنكنز.

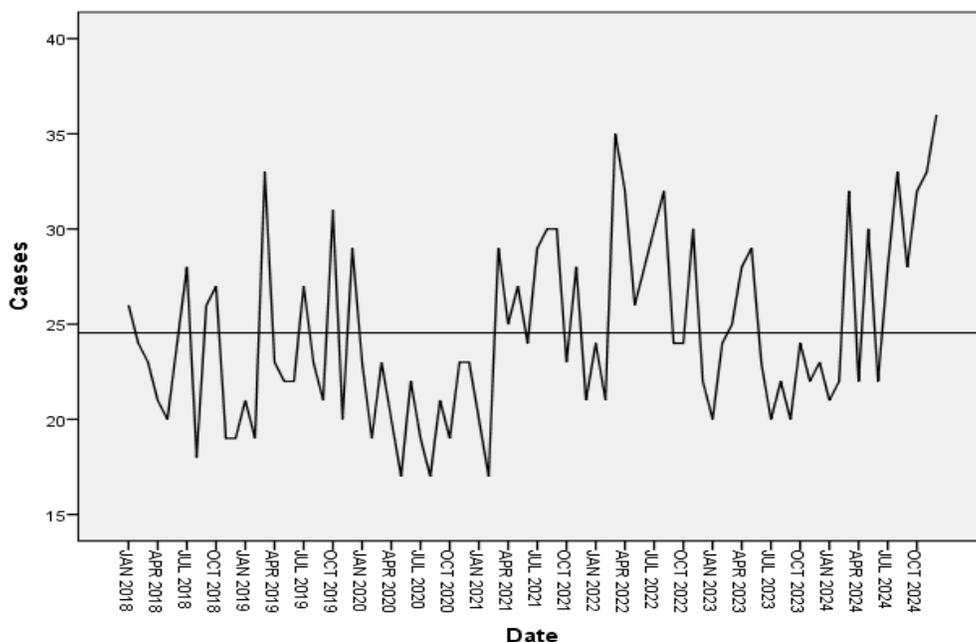
جدول (2) يوضح الإصابات الشهرية لمرضى الأورام

| الشهور | السنوات | | | | | | |
|--------|---------|------|------|------|------|------|------|
| | 2018 | 2019 | 2020 | 2021 | 2022 | 2023 | 2024 |
| يناير | 26 | 21 | 23 | 20 | 24 | 20 | 21 |
| فبراير | 24 | 19 | 19 | 17 | 21 | 24 | 22 |
| مارس | 23 | 33 | 23 | 29 | 35 | 25 | 32 |
| أبريل | 21 | 23 | 20 | 25 | 32 | 28 | 22 |
| مايو | 20 | 22 | 17 | 27 | 26 | 29 | 30 |
| يونيو | 24 | 22 | 22 | 24 | 28 | 23 | 22 |
| يوليو | 28 | 27 | 19 | 29 | 30 | 20 | 28 |
| أغسطس | 18 | 23 | 17 | 30 | 32 | 22 | 33 |
| سبتمبر | 26 | 21 | 21 | 30 | 24 | 20 | 28 |

| | | | | | | | |
|----|----|----|----|----|----|----|--------|
| 32 | 24 | 24 | 23 | 19 | 31 | 27 | أكتوبر |
| 33 | 22 | 30 | 28 | 23 | 20 | 19 | نوفمبر |
| 36 | 23 | 22 | 21 | 23 | 29 | 19 | ديسمبر |

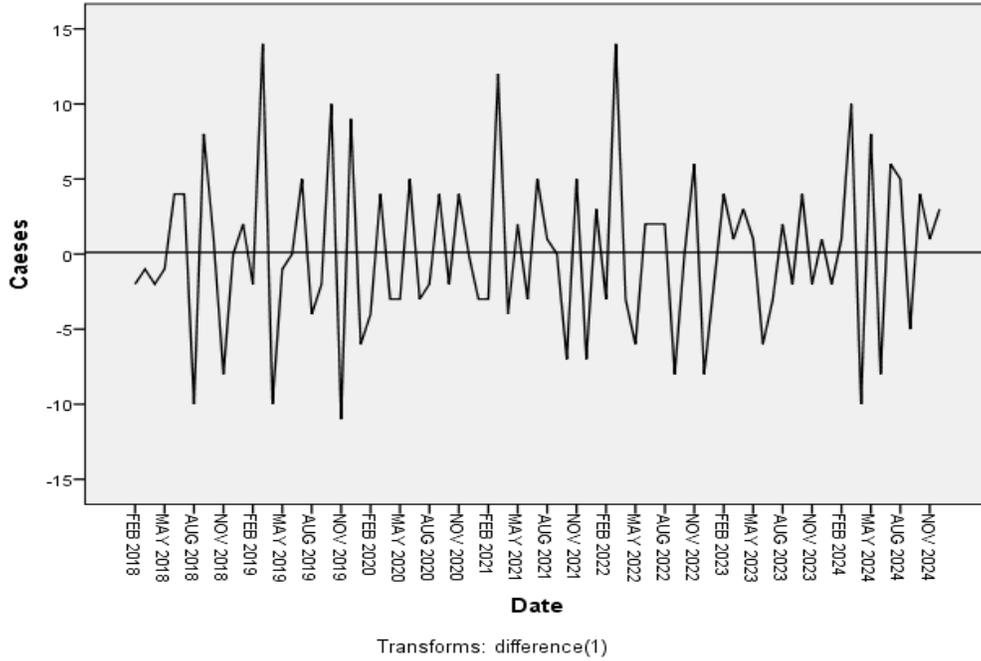
2-2: مرحلة التعرف Identification

في هذه المرحلة، يتم رسم السلسلة الزمنية للبيانات كخطوة أولية لفحص خصائصها وتحديد مدى استقرارها. وكما هو موضح في الشكل رقم (1)، تظهر السلسلة الزمنية سلوكاً غير مستقر، يتمثل في وجود اتجاه عام أو تباين غير ثابت عبر الزمن، مما يشير إلى أن البيانات لا تحقق شرط الاستقرار اللازم لنمذجة السلاسل الزمنية بشكل صحيح.



الشكل (1) يوضح السلسلة الزمنية الأصلية لمرضى الأورام

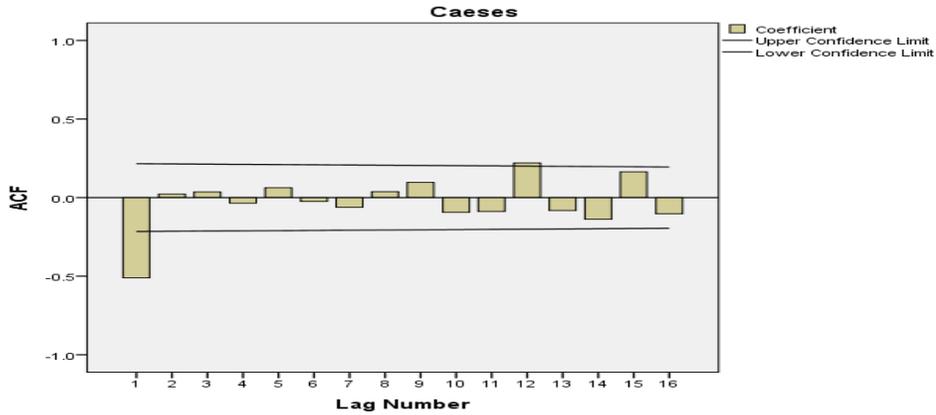
يتضح من الشكل رقم (1) أن السلسلة الزمنية الأصلية تظهر سلوكاً غير مستقر، مما يشير إلى وجود اتجاه أو تباين غير ثابت عبر الزمن. ولغرض تحويل السلسلة إلى سلسلة مستقرة، تم تطبيق الفرق الأول (First Differencing) على البيانات. كما هو موضح في الشكل رقم (2)، أدى هذا التحول إلى تحقيق الاستقرار في السلسلة الزمنية، حيث أصبحت السلسلة تظهر خصائص الاستقرار في المتوسط والتباين.



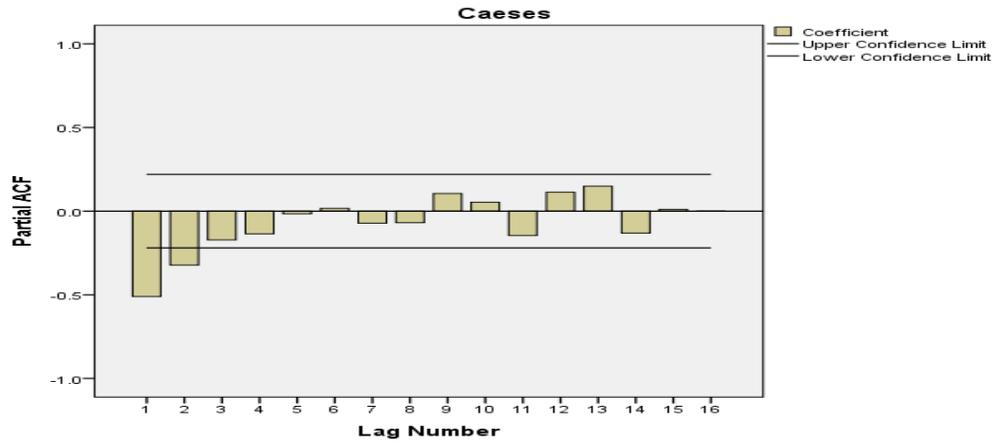
شكل (2) يوضح السلسلة الزمنية بعد أخذ الفرق الأول

2-3: تحديد رتب النموذج المناسب

بعد إجراء عملية الفرق الأول لسلسلة الزمنية و التأكد من استقرارها تأتي الخطوة المهمة في بناء النموذج (ARIMA) وهي تحديد رتب مكونات النموذج الانحدار الذاتي (AR) و الوسط المتحرك (MA) وذلك يتم من خلال تحليل سلوك كل من دالة الارتباط الذاتي (ACF) ودالة الارتباط الذاتي الجزئي (APCF) حيث تساعدنا في استنتاج القيم المناسبة لكل من (p) و (q) والتي سوف نستخدمها لاحقا في تحديد رتبة النموذج المناسب.



شكل (3) يوضح دالة الارتباط الذاتي لسلسلة الزمنية بعد أخذ الفرق الأول



شكل (4). يوضح دالة الارتباط الذاتي الجزئي لسلسلة الزمنية بعد أخذ الفرق الأول

من خلال الشكلين (3) و (4) الخاصين بدالتي الارتباط الذاتي (ACF) والارتباط الذاتي الجزئي (PACF) للسلسلة الزمنية بعد إجراء الفرق الأول، يلاحظ وجود قطع واضح بعد الإزاحة الأولى مما يشير إلى اقتراح ثلاثة نماذج (ARIMA) مناسبة هي (ARIMA(1,1,0), ARIMA(0,1,1), ARIMA(1,1,1) ، ولغرض اختيار النموذج الأمثل من بين النماذج الثلاثة تم الاعتماد على مجموعة من مؤشرات الإحصائية شملت، متوسط الخطأ المطلق (MAE)، متوسط الخطأ النسبي المطلق (MAPE)، و الجذر التربيعي لمتوسط مربعات الأخطاء (RMSE)، معيار معلومات اكايكي (AIC) بالإضافة إلى معيار معلومات بيزيان (BIC) كما هو موضح في الجدول أدناه.

جدول (3) مقارنة بين النماذج باستخدام معايير الدقة

| النموذج | RMSE | MAPE | MAE | BIC | AIC |
|--------------|-------|--------|-------|-------|-------|
| ARIMA(1,1,1) | 4.322 | 14.571 | 3.536 | 3.087 | 2.999 |
| ARIMA(0,1,1) | 4.307 | 14.643 | 3.552 | 3.027 | 2.967 |
| ARIMA(1,1,0) | 4.685 | 15.310 | 3.717 | 3.195 | 3.027 |

استنادا إلى القيم الدنيا لكل من (AIC) و (BIC)، وكذلك الأداء الأفضل في (RMSE)، فإن نموذج ARIMA(0,1,1) هو النموذج الأنسب لتمثيل السلسلة الزمنية محل الدراسة مما يجعله ملائم للاستخدام في التنبؤات المستقبلية.

2-4: تقدير معالم النماذج

ولتأكد أكثر من صحة النموذج المختار تم حساب معالم النماذج الثلاثة في الجدول أدناه.

جدول (4) يوضح معالم النماذج الثلاثة

| ARIMA | Estimate | SE | T | Sg |
|-------|-------------------|-------|--------|-------|
| | θ_1 -0.099 | 0.156 | -0.636 | 0.526 |

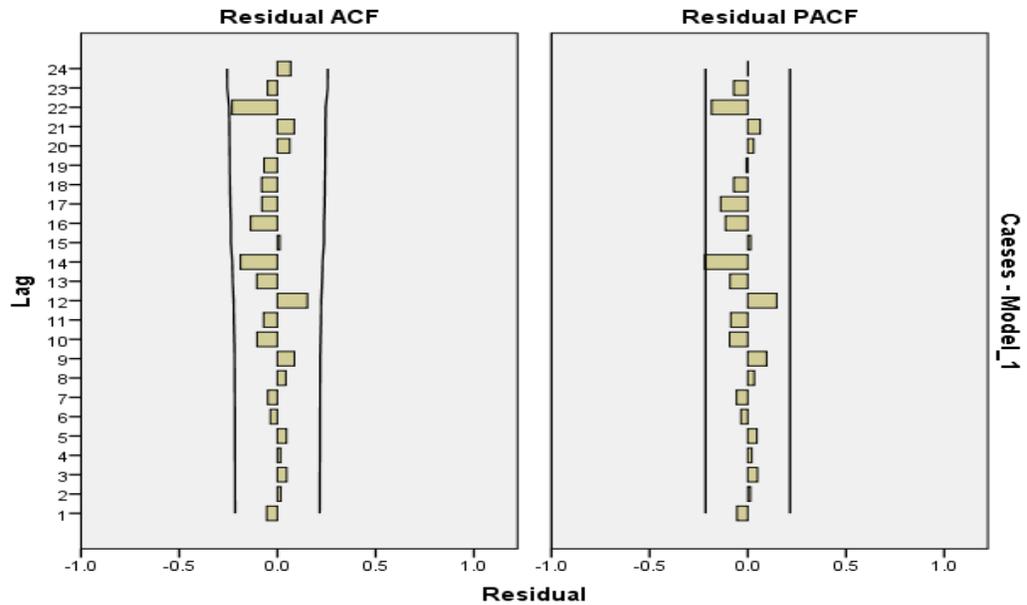
| | | | | | |
|-----------|------------|--------|-------|--------|-------|
| (1,1,1) | ϕ_1 | 0.681 | 0.122 | 5.592 | 0.000 |
| (1,1,0) | θ_1 | -0.507 | 0.096 | -5.302 | 0.000 |
| (0,1,1) | ϕ_1 | 0.733 | 0.083 | 8.861 | 0.000 |

أظهرت نتائج الجدول رقم (4) بالنسبة لنموذج الأول ARIMA(1,1,1) عدم معنوية معالم النموذج، أما النموذج الثاني ARIMA(1,1,0) يتضح معنوية معلمة النموذج ولكن قيمة (t) صغيرة جدا"، وبالنسبة لنموذج الثالث ARIMA(0,1,1) فهو ذو دلالة إحصائية مهمة عند مستوى معنوية (0.05) لمعامل المتوسط المتحرك الأولى ($\phi_1 = 0.733$) ، حيث كانت قيمة الدلالة الإحصائية (Sig) تساوي (0.000) مما يشير إلى إن هذا المعامل مختلف إحصائيا عن الصفر ، وبالتالي فإنه يعد ذو تأثير معنوي في النموذج . وبناء على ذلك يمكن اعتبار نموذج ARIMA(0.1.1) نمودجا مناسباً للتنبؤ بالسلسلة الزمنية لما أظهره من دقة إحصائية ومعنوية في تقدير المعلمات .

2-5: تحليل البواقي

تم اختبار ملائمة نموذج ARIMA(0,1,1) عن طريق تحليل بواقي النموذج وذلك من خلال رسم كل من دالة الارتباط الذاتي (ACF) ودالة الارتباط الذاتي الجزئي (PACF) للبواقي ، وكما هو موضح الشكل (5) فإن جميع معاملات الارتباط الذاتي للبواقي تقع ضمن حدود فترة الثقة مما يشير إلى إن البواقي لا تحتوي على ارتباطات زمنية ذات دلالة إحصائية .

وبناء على ذلك يمكن القول أن النموذج (0,1,1) ARIMA ملائم لتمثيل السلسلة الزمنية مما يجعله مناسباً لعملية التنبؤ المستقبلية .



شكل (5) يوضح دالتي الارتباط الذاتي و الارتباط الذاتي الجزئي للبواقي

يظهر الجدول رقم (5) نتائج اختبار Ljung-Box Q والذي يستخدم للتحقق من وجود ارتباط ذاتي بين القيم البواقي لنموذج ARIMA المقترح هذا الاختبار يقوم باختبار الفرضيتين التاليتين

فرضية العدم : لا توجد معاملات ارتباط ذاتي معنوية في البواقي حتى الرتبة k وهذا يعني أن البواقي مستقلة وتتبع عملية (White Noise).
فرضية البديلة : يوجد على الأقل ارتباط ذاتي واحد معنوي في البواقي حتى الرتبة k مما يدل على أن البواقي غير مستقلة.

$$H_0: \rho_1 = \rho_2 = \dots = \rho_K = 0$$

$$H_1: \rho_1 = \rho_2 = \dots = \rho_K \neq 0$$

أجري اختبار (Ljung-Box Q) للتحقق من استقلالية البواقي الناتجة عن النموذج، حتى الفجوة 18 (lags). كانت قيمة إحصاء (Ljung-Box Q) تساوي (13.187) بدرجات حرية قدرها (17)، مع قيمة احتمالية (Sig) قدرها (0.724). بما أن القيمة الاحتمالية أعلى من مستوى الدلالة (0.05)، فإننا لا نرفض الفرضية الصفرية القائلة بأن البواقي مستقلة إحصائياً. وعليه، تشير النتائج إلى أن بواقي النموذج لا تحتوي على ارتباط ذاتي معنوي حتى الفجوة (18)، مما يدعم صحة فرضية العشوائية في البواقي ويعزز ملائمة النموذج من ناحية الاستقلالية الزمنية.

وبالتالي فإن النموذج ARIMA(0,1,1) هو النموذج المناسب لبيانات السلسلة المدروسة.

جدول (5) إحصاءه (Ljung-Box Q)

| Model | Ljung – Box Q (18) | | |
|---------------|----------------------|----|-------|
| | Statistic | DF | Sig |
| Cases-Model-1 | 13.187 | 17 | 0.724 |

6-3: التنبؤ Forecasting

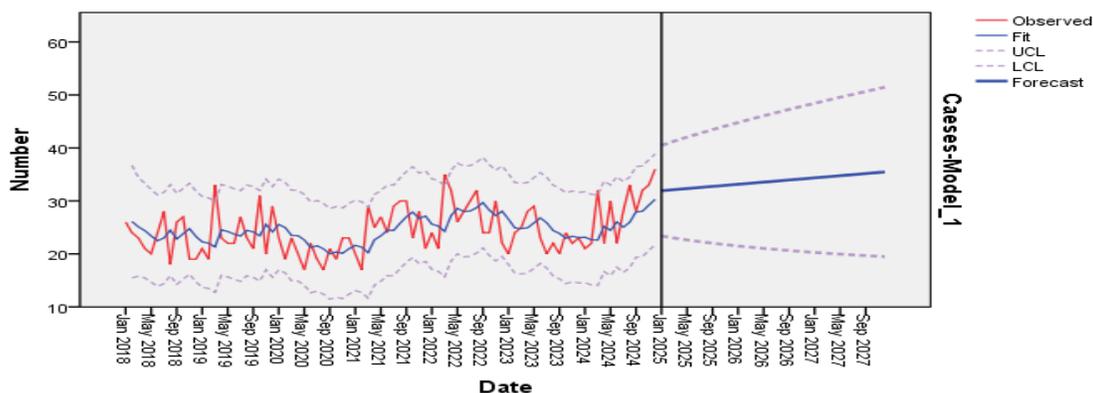
استناداً إلى النموذج الذي تم التوصل إليه في الفقرة (4-2)، تم إجراء عملية التنبؤ بأعداد المرضى المصابين بمرض الأورام خلال السنوات الممتدة من 2025 إلى 2027. وقد تم عرض النتائج المتوقعة في الجدول رقم (6)، مصحوبة بفترات الثقة التي تبين حدود التنبؤ عند دلالة إحصائية مناسبة، ما يعزز من موثوقية النتائج المتحصل عليها. ولغرض توضيح الاتجاه العام والمتغيرات البنوية في السلسلة تم تمثيل القيم المتنبأ بها بيانياً كما هو موضح في الشكل (6)، حيث أدمجت السلسلة الأصلية مع القيم المستقبلية المتوقعة ضمن نفس الإطار الزمني.

ويلاحظ من خلال الرسم البياني أن السلوك العام للسلسلة المتنبأ بها يتسم بالاتساق مع نمط السلسلة الزمنية الأصلية للفترة (2018-2024)، سواء من حيث الاتجاه أو مستوى التذبذب، مما يشير إلى كفاءة النموذج الإحصائي المستخدم في احتواء الخصائص الهيكلية للبيانات الأصلية. هذا التماثل يعزز من صلاحية النموذج لأغراض التنبؤ المستقبلي، ويمنح نتائج الدراسة درجة أعلى من الاعتمادية في دعم متخذي القرار في القطاع الصحي بشأن تخطيط الموارد والإجراءات العلاجية.

جدول (6) يوضح القيم التنبؤية لأعداد المرضى خلال ثلاث سنوات قادمة

| 2027 | | 2026 | | 2025 | | الشهور |
|----------------|---------------|------------|----------------|---------------|------------|--------|
| حدود الثقة %95 | | عدد المرضى | حدود الثقة %95 | | عدد المرضى | |
| الحدود الدنيا | الحدود العليا | | الحدود الدنيا | الحدود العليا | | |
| 48 | 20 | 34 | 45 | 21 | 33 | يناير |
| 49 | 20 | 34 | 45 | 21 | 33 | فبراير |
| 49 | 20 | 35 | 45 | 21 | 33 | مارس |
| 49 | 20 | 35 | 46 | 21 | 33 | أبريل |
| 50 | 20 | 35 | 46 | 21 | 34 | مايو |
| 50 | 20 | 35 | 46 | 21 | 34 | يونيو |
| 50 | 20 | 35 | 47 | 21 | 34 | يوليو |
| 50 | 20 | 35 | 47 | 21 | 34 | أغسطس |
| 51 | 20 | 35 | 47 | 21 | 34 | سبتمبر |
| 51 | 20 | 35 | 48 | 21 | 34 | أكتوبر |
| 51 | 20 | 35 | 48 | 20 | 34 | نوفمبر |
| 51 | 19 | 35 | 45 | 20 | 34 | ديسمبر |

الشكل (6) يوضح القيم المتنبأ بها لأعداد المرضى خلال الفترة الممتدة من (1، 1، 2025) إلى (2027,12,31)



الاستنتاجات

- 1- أظهرت نتائج التحليل أن السلسلة الزمنية المدروسة أصبحت ساكنة بعد تطبيق الفرق الأول، وهو ما تم التحقق منه من خلال الفحص لكل من دالة الارتباط الذاتي (ACF) ودالة الارتباط الذاتي الجزئي (PACF)، حيث أظهرت انقطاع واضحة بعد الإزاحة الأولى، ما يعزز فرضية السكون بعد الفرق الأول.
- 2- تم تحديد النموذج الأنسب لتمثيل البيانات باستخدام منهجية بوكس-جنكنز (Box & Jenkins) حيث أظهرت نتائج المقارنة بين النماذج المقترحة أن نموذج $ARIMA(0,1,1)$ هو الأكثر ملائمة من الناحية الإحصائية. وقد استند في هذا التحديد إلى مؤشرات الإحصائية والتي شملت كلا من معيار معلومات أكايكي (AIC)، معيار معلومات بيزيان (BIC)، والجذر التربيعي لمتوسط مربعات الأخطاء (RMSE)، حيث سجل هذا النموذج أقل القيم في جميع المؤشرات مقارنة بالنماذج الأخرى.
- 3- بعد اعتماد نموذج $ARIMA(0,1,1)$ ، تم استخدامه في التنبؤ بأعداد المرضى المصابين بالأورام السرطانية والمسجلين بقسم الأورام في بلدية طبرق للفترة الزمنية الممتدة من عام 2025 إلى 2027. وقد أظهرت القيم التنبؤية تقارباً ملحوظاً مع القيم الأصلية للسلسلة، مما يعكس دقة النموذج وكفاءته في تمثيل الظاهرة محل الدراسة، ويعزز من موثوقيته في دعم التقديرات المستقبلية واتخاذ القرار.

التوصيات

- 1- دعم البحوث التطبيقية المتقدمة حول مرضى الأورام في ليبيا وذلك بهدف الكشف عن العوامل المسببة وتحديد الأنماط الوبائية المرتبطة بانتشاره، مما يساهم في تحسين التدخلات الوقائية والعلاجية.
- 2- ضرورة اهتمام المؤسسات الصحية والحكومية المختصة بهذه الفئة من المرضى من خلال تعزيز برامج الدعم النفسي والاجتماعي وتوفير الأدوية اللازمة بشكل دائم وإنشاء مراكز علاج متخصصة في المناطق النائية والمكتظة سكانياً.
- 3- اعتماد منهجية السلاسل الزمنية ونماذج (ARIMA) Box-Jenkins كوسيلة تحليلية للتنبؤات المستقبلية في المجال الصحي نظراً لقدرتها على تقديم تقديرات دقيقة تساعد صناعات القرار في التخطيط الاستراتيجي الفعال.
- 4- نوصى بتبني نموذج $ARIMA(0,1,1)$ بشكل خاص، والذي أثبت كفاءته الإحصائية في هذا البحث، حيث أظهره دقة تنبؤية وتقارب ملحوظ بين القيم الفعلية والمتنبأ بها وفقاً لمقاييس التقييم مثل (AIC، BIC).
- 5- العمل على تطوير قواعد بيانات طبية رقمية دقيقة وشاملة للمرضى المصابين بالأورام، يتم تحديثها دورياً لتسهيل التحليل الإحصائي المستقبلي ورصد التغيرات الزمنية في معدلات الإصابة.
- 6- يجب دعم استخدام التحليلات الإحصائية التنبؤية في مراكز الأبحاث والمؤسسات الطبية لما لها من دور فعال في الحد من انتشار الأزمات الصحية قبل وقوعها، وذلك من خلال توقع أعداد المصابين وتحديد الاحتياجات الطبية مسبقاً.
- 7- توسيع نطاق استخدام النماذج الإحصائية التنبؤية لتشمل أنواعاً مختلفة من الأمراض المزمنة حيث يعزز من كفاءة القطاع الصحي.
- 8- تحفيز التعاون بين الجامعات والمراكز الصحية في تنفيذ مشاريع بحثية مشتركة، بما يعزز من كفاءة مخرجات البحث العلمي ويوسع نطاق تطبيقاته.
- 9- اتخاذ إجراءات استناداً إلى نتائج هذا البحث، خاصة في ظل تزايد حالات الإصابة، مما يستدعي تكثيف جهود التوعية والكشف المبكر والفحص الدوري للمواطنين، للحد من تفشي المرض مستقبلاً.

المراجع

- 1- أمال حيدر (2018)، "التنبؤ بسرعة الرياح الشهرية في محطة طرطوس باستخدام منهجية بوكس - جنكنز"، مجلة جامعة تشرين للبحوث والدراسات العلمية، المجلد (40)، العدد (3).
- 2- برى، عدنان عبد الرحمن، (2002)، "طرق التنبؤ الإحصائي" جامعة الملك سعود.

- 3- جبارة ، محمد (2012) " التنبؤ بالسلاسل الزمنية لمنسوب النيل الأزرق في محطة و دمدني باستخدام نماذج بوكس – جنكنز و نماذج الشبكات العصبية الاصطناعية " ، جامعة السودان للعلوم والتكنولوجيا ، السودان .
- 4- حضري ، خوله (2014) " استخدام السلاسل الزمنية من خلال منهجية بوكس –جينكنز في اتخاذ القرار الإنتاجي " ، رسالة ماجستير ، جامعة محمد خيضر –بسكرة ، الجزائر .
- 5- راضي عبد الرحيم عثمان (2018)،"استخدام منهجية بوكس – جنكنز للسلاسل الزمنية للتنبؤ بالمعدلات الشهرية لدرجات الحرارة ، ببلدية بنينا ، مجلة العلوم والدراسات الإنسانية ، العدد (52) .
- 6- ركابي ، مخلص ، محمد (2013) " استخدام منهجية بوكس و جنكنز في التنبؤ بأعداد المعتمدين من الخارج شهريا " ، جامعة أم القرى ، السعودية .
- 7- طارق بن قسيمي (2014)،"استخدام نماذج السلاسل الزمنية الموسمية للتنبؤ بمبيعات الطاقة الكهربائية دراسة حالة الشركة الوطنية للكهرباء والغاز " رسالة ماجستير ، جامعة محمد خضير ، الجزائر .
- 8- عوض الله ، إبراهيم (2015)،ط استخدام منهجية بوكس – جنكنز في التنبؤ بحالات الزواج والطلاق في قطاع غزة "رسالة ماجستير ، جامعة الأزهر – غزة ، فلسطين .
- 9- شريف شحده ، سلامة مصلح " التنبؤ بعدد مرضى السرطان باستخدام منهجية بوكس _ جنكنز في محافظة غزة " مجلة جامعة الأزهر للعلوم الإنسانية ، مجلد (20) ، العدد (1) ، 2010 .
- 10- محمد ، منعم عزيز ، (2010) ، " نماذج بوكنز _ جنكنز لتحليل السلاسل الزمنية " .
- 11- شعراوي ، سمير (2005) " مقدمة في التحليل الحديث للسلاسل الزمنية " ، الطبعة الأولى ، مركز النشر العلمي ، جامعة الملك عبد العزيز .
- 12- م.م سعدية عبد الكريم طعمه " استخدام تحليل السلاسل الزمنية للتنبؤ بإعداد المصابين الأورام الخبيثة في محافظة الأنبار " مجلة جامعة الأنبار للعلوم الاقتصادية و الإدارية ، مجلد (4) ، العدد (8) ، 2012 .

References

- 1- Box, G.E. and price , D.A. (1970), ” Distribution of Residual Autocorrelations in Autoregressive – Integrated Moving Average Time series Model “ , JASA, Vol. 55 , No.332, PP. 1509-1525.
- 2- Chris Chatfield, Time-series forecasting, Chapman&Hall/CRC, United States of America, 2002
- 3- Douglas C.Montgomery, Cheryl L.Jennings, Murat Kulahci, Introduction to time series analysis and forecasting, John Wiley & Sons INC publication, United states of America, 2008
- 4- Musa, N. A., & Alkdrawe, S. A. (2025). THE RADIATION EFFECT ON BLOOD CELLS IN CANCER PATIENTS WHO HAVE UNDERGONE EXTERNAL RADIATION THERAPY-ANALYTICAL STUDY. Bani Waleed University Journal of Humanities and Applied Sciences, 10(2), 573-581
- 5- Peter J. Brockwell & Richard A. Davis (2001), “Introduction to Time Series and Forecasting”, Second Edition, Springer-Verlag New York, Inc.
- 6- Wheel Wright, S.G. & Markedis, S. (1973), "An Examination of the Use of Adaptive filtering in Forecasting O.P.Q", vol.24, No.1, P (60-64).