

استعمال النموذج المختلط التكامل الموسمي للتنبؤ بالتغيرات السكانية في الوحدات الإدارية لحافظة ديالى

Using the seasonal integrative mixed model to predict population changes in the
administrative units of Diyala Governorate

م. م هبة هذال خليل⁽³⁾

وزارة التخطيط ونظم المعلومات الجغرافية
/ مديرية احصاء ديالى

Hibahathal03@gmail.com

م. م سارة عادل مظلوم⁽²⁾

الجامعة التقنية الوسطى / معهد تقني
الصويرة / قسم تقنيات المحاسبة

sarah_adel@mtu.edu.iq

م. م امل هادي رشيد⁽¹⁾

جامعة ديالى / كلية الادارة والاقتصاد /
قسم الاحصاء

amal@uodiyala.edu.iq

المستخلص

تعد المؤشرات السكانية من اهم المؤشرات في اي بلد من بلدان العالم، وتعد التغيرات السكانية حالة طبيعية نتيجة للزيادة والنقصان الحاصلة في عدد السكان وتتأثر الحركة السكانية ومعدلات النمو السكاني بعده عوامل اذ ان الدول التي تعاني من اضطرابات في الاقتصاد والحروب الدائمة نلاحظ الانخفاض في اعداد السكان نتيجة للظروف التي يعاني منها البلد مقارنة بالدول المستقرة من حيث الأمان ، حيث ان عدد السكان مستقر نوعا ما، و ان الهرم السكاني يكون في حالته الطبيعية من حيث الشكل؛ اذ يأخذ شكل الهرم ، يهدف البحث إلى التنبؤ بالنمو السكاني واثار التغيرات الحاصلة في اعداد السكان من حيث الزيادة والنقصان من خلال الوحدات الادارية ليكون عدد التنبؤ اكثر دقة من حيث التغير السكاني الحاصل في المحافظة ، وتركزت مشكله البحث في معرفة مدى التغير الحاصل في التغيرات الديموغرافية من حيث اعداد السكان ، واستعملت السلاسل الزمنية للتنبؤ ، وتوصلت الدراسة إلى افضل نموذج يمثل بيانات الحضر، هو الانموذج $SARIMA(0,1,1)(0,1,1)^{90}$ وافضل أنموذج يمثل الريف هو الانموذج $SARIMA(1,0,0)(2,1,1)^{90}$ وبعد أخذ السلسلتين تبين أنها غير مستقرة ومن أخذ الفرق الاول تم استقرارها وبعد اختبار الانموذج وتبين انه افضل انموذج لتمثيل بيانات الدراسة تم التنبؤ لها واوضحت النتائج انها مثالية بسبب قلة مقدار الخطأ وكونها قريبه من مثيلتها للبيانات الأصلية للدراسة واوضحت النتائج التنبؤية، انه توجد زيادة وفروق معنويه كبيرة من حيث ازباد اعداد السكان في الحضر بالنسب للريف

الكلمات المفتاحية:

التغيرات السكانية ، التكامل ، التنمية ، المجتمعات المعاصرة ، الموسمي ، المختلط

Abstract

Population indicators are among the most important indicators in any country in the world. Population changes are a natural phenomenon resulting from the increase and decrease in the number of inhabitants. Population movement and growth rates are influenced by several factors. Countries experiencing economic disturbances and ongoing wars often see a decline in population numbers due to the conditions affecting the country, in contrast to stable countries in terms of security, where the population remains relatively stable and the population pyramid appears in its natural shape, resembling a pyramid.

The research aims to predict population growth and the impact of changes in population numbers regarding increase and decrease through administrative units to make the predictions more accurate concerning demographic changes in the province. The research problem focused on understanding the extent of changes in demographic shifts in terms of population numbers. Time series were used for forecasting, and the study concluded that the best model representing urban data is the SARIMA(0,1,1)(0,1,1)⁹⁰ model, while the best model representing rural data is the SAMIMA(1,0,0)(2,1)⁹⁰ model.

After analyzing the two series, it was found that they were unstable, and taking the first difference stabilized them. After testing the model, it was determined to be the best model for representing the study data. The predictions showed ideal results due to the low error margin and their closeness to the original data of the study. The predictive results indicated a significant increase and meaningful differences in the population numbers in urban areas compared to rural areas.

Keywords :Population changes, integrative, development, contemporary societies, seasonal, mixed.

1. المقدمة

تعد العلاقة بين السكان والتنمية من أكثر العلاقات إشكالية في المجتمعات المعاصرة ومن هنا ظهرت أهمية على التركيز على التغيرات الديمغرافية للسكان ودمجها مع جوانب التنمية، إذ أكد على أن إهمال المتغيرات الديمغرافية الكمية والنوعية من شأنه أن يؤدي إلى اعتماد سياسات غير متكاملة وقصور واضح في برامج التنمية خاصة الهيكل العمري للسكان وعلاقته بالنمو والفقر. حيث هدفت الدراسة إلى اخذ نماذج السلاسل الزمنية للوقوف عن اهم مشاكل السكان لما لها من اهمية في التنبؤات المستقبلية ووضع الخطط والحلول اللازمة وتحليل بيانات موضوع الدراسة.

2- مشكلة البحث

يتبع العراق تقسيم مساحاته الجغرافية إلى الوحدات السكانية والمتمثلة بالمحافظات وهذه الوحدات السكانية تتكون من بلوكات خاصة وتكون في الحضر والريف، وتمحورت الدراسة حول مشكلة التنبؤ بأعداد السكان وفق التقسيمات الادارية للحضر والريف في محافظة ديالى وذلك من خلال استعمال السلاسل الزمنية (نماذج بوكس - جنكيز).

ومن هنا جاءت اهمية دراسة التغيرات الحاصلة في اعداد السكان لكونهم العامل الاساس في التنمية والتطوير. حيث تمحورت الدراسة حول بالتنبؤ بأعداد السكان لمحافظة ديالى وليكون التنبؤ بشكل أعمق، تم اخذ البيانات المستعملة بأعداد السكان حسب الوحدات الادارية ولكل جانب من الحضر والريف

3-هدف البحث

تهدف الدراسة إلى إيجاد أفضل انموذج يمثل بيانات الدراسة للوقوف عن الاسباب الحقيقية لأعداد السكان والتغيرات الحاصلة لكونهم العامل الاساس في التنمية والتطوير ومن ثم التنبؤ لها

4-السلاسل الزمنية

يطلق على تسلسل المتغيرات العشوائية $\{Y_t; t = 0, \pm 1, \pm 2, \pm 3, \dots\}$ عملية عشوائية ويعمل كنموذج لسلسلة زمنية مشاهدات. ومن المعروف أن البنية الاحتمالية الكاملة لمثل هذه العملية يتم تحديدها من خلال مجموعة التوزيعات لجميع المجموعات المحدودة من Z ، ويمكن مشاهدتها خلال فترات زمنية متعاقبة متساوية المسافات (يومية ، اسبوعية ، شهرية ، ربع سنوية ، و سنوية) تحتوي السلسلة الزمنية على متغير مستقل (independent) هو الزمن (t) والمتغير التابع (dependent) وهو قيمة الظاهرة المدروسة (Z_t). وإن المكونات (العناصر) الأربع الخاصة بالسلسلة الزمنية والتي تتأثر بالعوامل الاقتصادية والبيئية والاجتماعية والسياسية [6].

4-1 أستقرارية السلسلة الزمنية (Check Stationary of Time Series)

من اجل الوصول إلى استدلال احصائي حول بناء العملية العشوائية وعلى اساس المشاهدات لتلك العمليات ووضع افتراضات حول السلسلة الزمنية يجب ان تكون السلسلة مستقرة ، والفكرة الاساسية لاستقراريه السلسلة هي ان القوانين الاحتمالية التي تحكم سلوك العملية لا تتغير بمرور الزمن لي وجه التحديد، يقال أن العملية $\{Z_t\}$ تكون مستقرة تمامًا إذا كان التوزيع المشترك لـ $Z_{t1}, Z_{t2}, \dots, Z_{tn}$ هو نفس التوزيع المشترك $Z_{t1-k}, Z_{t2-k}, \dots, Z_{tn-k}$ لجميع اختيارات النقاط الزمنية $(t1, t2, \dots, tn)$ وجميع اختيارات الفارق الزمني k [3].

الشرط الأول للاستقرارية: ان يكون المتوسط ثابت، أي ان السلسلة لم تظهر اي اتجاه للتزايد أو التناقص مع الزمن، فإنه يمكن افتراض ان المتوسط ثابت لكل فترة زمنية وافترض إن القيمة المشاهدة عند كل فترة زمنية تمثل هذا المتوسط فيكون الوسط الحسابي مستقر إذا تحقق الشرط الاتي:

$$E(Z_t) = \mu_t = \mu \text{ is constant for all } t \quad \dots \dots (1)$$

الشرط الثاني عند ثبات التباين تكون قيمة التباين يتحقق شرط ثبات التباين اذا تحقق:

$$Var(Z_t) = E[(Z_t - \mu)^2] = \sigma^2 \dots\dots\dots (2)$$

أذ إن تباین السلسلة الزمنية ثابتاً ولا يعتمد على قيم t وتقديره:

الشرط الثالث دالة التباين المشترك الذاتي Autocovariance هي دالة فرق الزمن المشترك

$$cov(Z_t, Z_s) = E[(z_t - \mu)(z_s - \mu)] = \gamma|t - s| \dots\dots\dots (3)$$

وهي فقط مسافة الزمنية بين المتغيرين العشوائيين ولا تعتمد على الزمن t .

ويكون الاستقرار ضعيفاً نظراً لأن استقرارية التباين المشترك ل $s=t$ فإن العملية العشوائية تكون مستقرة بصورة ضعيفة عندما يكون الوسط والتباين مستقر [6]

2-1-4 دالة الارتباط الذاتي (ACF) Autocorrelation Function

ويقاس معامل الارتباط الذاتي P_k قوة العلاقة (الارتباط) بين القيم المشاهدة Z_t و Z_{t+k} من المشاهدات أو بين المتغيرات العشوائية ، ، فإن معامل الارتباط الذاتي سيكون كما يأتي :

$$P(Z_1, Z_{t+k}) = \frac{E(Z_1, Z_{t+k}) - E(Z_1)E(Z_{t+k})}{\sigma_{Zt}\sigma_{Zt-k}} \dots\dots\dots (4)$$

وبإضافة نموذج التشويش الأبيض White - Noise Model للسلسلة الزمنية Z_t التي لها متوسط μ وتباين σ^2 ودالة تباين مشترك γ_k سيكون معامل الارتباط الذاتي P_k كما يأتي [2]:

$$P_k = \frac{\gamma_k}{\gamma_0} = \frac{\gamma_k}{\sigma_{ZZ}} \dots\dots\dots (5)$$

ولإجراء الاختبار الخاص بمعرفة معنوية الارتباط الذاتي يتم استعمال اختبار جذر الوحدة (Unit Root Test) وهو أهم الأساليب المستعملة لفحص استقرارية السلسلة الزمنية ويتحقق شرط الاستقرار عندما تكون جذور الوحدة للسلسلة تقع داخل دائرة الوحدة، ومن أهم الطرق المستعملة للكشف عن استقرار السلسلة الزمنية هو اختبار ديكي _ فولر الموسع (ADF) وفيه أقترح العالمان ديكي _ فولر الموسع (Augment Dickey Fuller) عام (1981) فرضية وجود جذر الوحدة (الفرضية الصفرية) الآتية :

$$H_0: |\lambda_1| = 0 \quad VS \quad H_1: |\lambda_1| \neq 0 \quad \dots\dots\dots (6)$$

حيث تشير الفرضية الصفرية (فرضية العدم) إلى وجود جذر الوحدة بالسلسلة الزمنية مما يدل على أن السلسلة غير مستقرة والفرضية البديلة تكون عكس الفرضية الصفرية مما يدل على أن السلسلة تكون مستقرة، ومن أجل الحصول على نتائج اختبار الفرضية فإنه سوف يتم الاعتماد على طريقة المربعات الصغرى لتقدير معامل النموذج الذي تكون معادلته كالتالي [2]:

$$\Delta Z_t = \lambda_1 Z_{t-1} + \sum_{i=1}^n \theta_i \Delta Z_{t-i} + \varepsilon_i \quad \dots (7)$$

وأن المعلمة $|\lambda_1|$ يمكن الحصول عليها من المعادلة التالية :

$$|\lambda_1| = \theta - 1 \quad \dots (8)$$

وللحصول على قيمة t_{λ_1} المحسوبة في معادلة رقم (12) من خلال قسمة المعلمة $\hat{\lambda}_1$ على الانحراف المعياري كما في أدناه :

$$t_{\lambda_1} = \frac{\hat{\lambda}_1}{\hat{\sigma}_{\hat{\lambda}_1}} \quad \dots (9)$$

حيث أن : $\hat{\lambda}_1$: معلمة مقدرة بواسطة طريقة المربعات الصغرى . $\hat{\sigma}_{\hat{\lambda}_1}$: تباين المعلمة المقدرة . t_{λ_1} : القيمة المحسوبة.

وعند مقارنة t_{λ_1} المحسوبة مع قيمة t_{table} الجدولية (التي يتم الحصول عليها من جدول Dickey-Fuller) فإذا كانت المحسوبة أكبر من الجدولية فهذا يعني أن السلسلة الزمنية تكون سلسلة غير مستقرة فيتم عدم رفض الفرضية الصفرية أما إذا كانت المحسوبة أصغر من الجدولية أن السلسلة الزمنية أي تكون سلسلة مستقرة.

3-1-4 الارتباط الذاتي الجزئي PACF (Partial Autocorrelation Function)

تمثل العلاقة بين قيم متتالية لمتغير ما خلال فترتين زمنيتين مختلفتين ، مع افتراض ثبات الفترات الأخرى ، ويرمز لدالة الارتباط الذاتي الجزئي بالرمز P_{kk} .

ويمكن صياغة المعادلة الرياضية لمعامل الارتباط الذاتي الجزئي P_{kk} باتباع الصيغ الرياضية الآتية:

عند $K=1$ فإن

$$P_{kk} = \frac{P_k - \sum_{j=1}^{k-1} P_{k-1} P_{j-1}}{1 - \sum_{j=1}^{k-1} P_{k-j} P_j} \quad k = 2, 3, 4 \dots \dots \dots (10)$$

حيث

$$P_{kk} \cong \frac{1}{n} \quad k > 0 - 1$$

لقيم n الكبيرة فإن r_{kk} يكون لها تقريبا توزيع طبيعي وبالتالي نستطيع القيام بالاختبار التالي:

$$H_0 : \phi_{kk} = 0$$

$$H_1 : \phi_{kk} \neq 0$$

وذلك باستعمال الصيغة التالية

$$\frac{|P_{kk}|}{n^{-1/2}} = \sqrt{n}|P_{kk}| \quad \dots \dots \dots (11)$$

ان المعلمات $\phi_{KK}; K=1$ تسمى بدوال الارتباط الذاتي الجزئي (PACF) والتي تعتبر أداة أخرى مهمة اضافة مع (ACF) تساعدنا في دراسة ومعرفة استقرارية السلسلة الزمنية وتحديد درجة أنموذج (AR) و الأنموذج (MA) و الأنموذج (ARIMA) الملائم لتمثيل بيانات السلسلة الزمنية [5].

5- نماذج بوكس- جنيكنز Models Box of Jenkins

5-1 أنموذج الانحدار الذاتي والمتوسطات المتحركة التكاملية

(Autoregressive Integrated Moving Average Model) (ARIMA)

معظم السلاسل الزمنية التي تتكون في التطبيقات العملية في معظم مجالات المعرفة تكون غير ساكنة لذلك يجب اخذ الفروق للسلسلة للحصول على استقرارية السلسلة، لذلك سنفترض ان (d) هو الحد الأدنى من الفروق لتحويلها إلى سلسلة مستقرة أو بإجراء التحويلات، وهو يعتبر من أكثر النماذج استعمالاً في التنبؤ ويرمز له (ARIMA p,d,q) بحيث (p,q) رتب النماذج AR(p),MA(q) عدد الفروق. وصيغته العامة كالآتي [7]:

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} - a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (12)$$

5-2- الانماذج الموسمي (Seasonal models)

عند دراسة السلاسل الزمنية الربع سنوية أو الشهرية في كثير من الاحيان يشاهد نمط الموسمية اذ يلاحظ تكرار حدوث ارتفاع أو انخفاض عند الشهر نفسه أو ربع السنة تقريبا في الاعوام المتتالية، عند محاولة بناء أنموذج لمثل هذا النوع من السلاسل فإنه من الطبيعي دراسة الارتباط بين مشاهدات نفس الموسم في السنوات المتتالية فضلا عن دراسة الارتباط بين مشاهدات المواسم المتتالية.

5-2-1 نماذج الانحدار الذاتي الموسمية (Seasonal autoregressive models)

يمكن تعميم المعادلة المذكورة انفا للأنموذج SAR(p) بالصيغة الآتية :

$$Z_t = \phi_1 Z_{t-s} + \phi_2 Z_{t-2s} + \dots + \phi_p Z_{t-ps} + a_t \quad \dots (13)$$

اذ يكون رسم دالة الارتباط الذاتي (ACF) لا نموذج الانحدار الذاتي الموسمي والارتباط الذاتي الجزئي (PACF) فإنه خصائصها تشبه خصائص دالة الارتباط الذاتي لا نموذج الانحدار العادي ودالة الانحدار الذاتي الجزئي للنماذج غير الموسمية فيما عدا ذلك يلاحظ ظهور عدة قيم لمعاملات الارتباط الذاتي واضحة عند مضاعفات (S) [4].

5-2-2 نماذج الأوساط المتحركة الموسمية (Seasonal moving average models)

يمكن تعميم المعادلة للأنموذج SMA كالآتي :

$$Z_t = a_t - \theta_1 a_{t-s} - \theta_2 a_{t-2s} - \theta_q a_{t-qs} \dots (14)$$

اذ تمثل θ معلمة الاوساط المتحركة .

(a_{t-qs}) تمثل التغير العشوائي الذي حدث في نفس الموسم للسنة السابقة بطول مدة الدورة الموسمية (S) .

3-2-5 النماذج المختلطة التكاملية الموسمية

Seasonal integrated (ARIMA) models

يرمز لهذا النموذج بالرمز (SARIMA) من الرتبة $(P, D, Q)_s$ ن السلاسل الزمنية الموسمية غير المستقرة يمكن تحويلها إلى مستقرة وذلك بأخذ الفروق الموسمية (D) لها اذ ان (D=1,2 ...) اذ ان الفرق الموسمي يعتمد على مدة طول فترة الموسم وان الفرق الموسمي الأول يأخذ الصيغة الاتية [1] :

$$Z'_t = Z_t - Z_{t-s} \dots (15)$$

والانموذج المختلط الموسمي المأخوذ له الفروق الموسمية فيأخذ الصيغة الاتية :

$$\phi_p(B^s)(1-B)^d Z_t = \theta_q(B^s)e_t$$

$$\phi_p(B^s) = 1 - \phi_1 B^1 - \phi_2 B^2 - \dots - \phi_p B^{ps} \dots (16)$$

4-2-5 نماذج المتوسطات المتحرك الموسمي للانحدار الموسمي الذاتي التكاملية

Seasonal Autoregressive Integrated Moving Average Models (SARIMA(P,D,Q))

المتسلسلات الزمنية الموسمية تعطي اناط مماثلة تتكرر على فترات معينة متساوية بالبعد مثل ان يتكرر النمط كل موسم من السنة او يتكرر حسب الحالة الدراسية بوقت معين [1] .

بشكل عام يكتب أنموذج السلسلة الزمنية لنموذج الانحدار الذاتي التكاملية المتحرك

$$\text{SARIMA}(p,d,q)(P,D,Q)$$

$$\phi_p(B)\Phi_p(B^s)(1-B)^d(1-B^s)^D z_t = \delta + \theta_q(B)\Theta_q(B^s)a_t \dots (17)$$

$$a_t \sim WN(0, \sigma^2)$$

6- تحديد رتبتي الانحدار الذاتي والمتوسطات المتحركة

قام بوكس وجنيكنز بوضع أسلوباً لتحديد رتبة الأنموذج والذي يعتبر أداة من أدوات التشخيص (معرفة نوع الأنموذج) حيث يتم تحديد أنموذج (ARIMA) الملائم لها، حيث تعتبر أحد المفاتيح الفعالة للتمييز بين ال نماذج AR(p) والنماذج MA(q) والنماذج المختلطة ARMA (p,q) وتحديد رتبة كل منه [6] .

1-6 معيار أكايكي (AIC) (Akaike)

وضع العالم الياباني أكايكي في عام 1974 معيار التطوير ونمذجة السلاسل الزمنية واعتمد فيه أن يكون تقليل تباين الأنموذج مقارنة بزيادة عدد المعالم المقدرة ويستعمل هذا المعيار للتعرف على نوعية ومطابقة الأنموذج وفي تحديد الرتبة وصيغته [10]

$$AIC \approx n \log \sigma_e^2 + 2K \dots \dots \dots (18)$$

σ_e^2 يمثل متوسط مربعات الخطأ (MSE)

k: هي العدد الكلي لمعاملات الأنموذج المقدرة، وتحدد الرتبة باختيار الأنموذج الذي يقابل أقل قيمة للمعيار AIC.

2-6 معيار بيز للمعلومات (BIC) Bayesian information Criterion

معيار بيز ويحسب حسب الصيغة الآتية

$$BIC = n \log(\sigma_e^2) + m \log_n(n) \dots (19)$$

حيث أن n : تمثل حجم السلسلة الزمنية ، m : عدد معاملات النموذج .

3-6 معيار أكايكي المصحح (AICc) Akaike Information Corroected

اقترحه العالمين Hurvich and Tsai عام 1989 ويحسب حسب الصيغة الآتية:

$$AICc = n \log(\sigma_e^2) + \frac{2n(m+1)}{n-m-2} \dots \dots \dots (20)$$

ويستعمل هذا المعيار عندما تكون العينات قليلة.

7- التقدير Estimation

بعد تحديد النموذج الملائم فيتم تقدير الأنموذج باستعمال عدة طرق وسيتم استعمال طريقة الامكان الاعظم المضبوط (The maximum possible control function) حيث ان السلسلة الزمنية $Z=(Z_1, Z_2, \dots, Z_n)$ والمكونة من نموذج الانحدار الطبيعي ومن الدرجة الاولى $AR(1)$ اذن معلمة الانحدار الذاتي $|\phi| > 1$ كمية ثابتة لا تتغير بتغير الزمن وان الاخطاء العشوائية تتبع التوزيع الطبيعي [2].

$$\ln L(\phi, \theta, \sigma_a^2) = \left(-\frac{n}{2} \ln 2\pi \sigma_a^2 - \frac{S(\phi, \theta)}{2\pi \sigma_a^2}\right) \dots (21)$$

وعند العينات الكبيرة الحجم فان مصفوفة المعلومات التي تساوي $(\phi, \theta)^{-1}$ والمتوفرة بعد أن يتم حساب مصفوفة المشتقات ذات الرتبة $(p+q) \times n$ ويكون

$$a = \theta^{-1}(B)\phi(B) \dots \dots \dots (22)$$

ويتم الحصول على مقدرات الإمكان الأعظم التقريبية بعد أخذ المشتقات جزئية لها اي ان:

$$\frac{\partial a_t}{\partial \theta_i} = \theta^{-1}(B)B^i a_t \dots\dots\dots(23)$$

$$\frac{\partial a_t}{\partial \theta_j} = \theta^{-1}(B)B^j a_t \dots\dots\dots$$

وان التباينات المشتركة للتقديرات يتم حسابها عن طريق استخدام معكوس مصفوفة المعلومات

8- اختبار نماذج بوكس - جينكينز A significant test of the Box Jenkins model

بعد تقدير معالم الأنموذج الذي تم تحديده تأتي بعدها مرحلة اختبار صحة معنوية الأنموذج، وان البواقي يجب أن تتوزع بشكل عشوائي تستعمل الفرضية التالية: اذا يستعمل اختبار ليونغ_ بوكس (Ljung-Box) وهو اختبار إحصائي الهدف منه هو التأكد من عدم وجود ارتباط ذاتي للبواقي، تستعمل ايضاً الفرضية الموضوعة للاختبار للتحقق من عشوائية البيانات باستعمال مجموعة من الإزاحات [5].

$$H_0: r_1(e_t) = r_2(e_t) = \dots\dots\dots = r_k(e_t) = 0$$

$$H_1: r_1(e_t) \neq r_2(e_t) \neq \dots\dots\dots \neq$$

$$r_k(e_t) \neq 0$$

تحسب احصاء الاختبار من العلاقة كالاتي:

$$Q = m(m+2) \sum_{i=1}^k \frac{r_i^2(e)}{(m-i)} \dots\dots\dots (24)$$

إذ أن : $m=(n-d)$ ، n : يمثل عدد مشاهدات السلسلة الزمنية. ، d : عدد الفروق. k : اكبر إزاحة

$r_i(e)$: يمثل الارتباط الذاتي للبواقي عند الإزاحة (i) ، (Q) تتوزع بصورة تقريبية توزيع (χ^2) بدرجة حرية $(K-p)$.

ومن خلال قيمة الـ (p-value) اذا كانت قيمة الـ (p-value) اكبر من 0.05 تقبل فرضية العدم اي عدم وجود ارتباط ذاتي بين البواقي.

9- التنبؤ (Forecasting)

هو المرحلة الأخيرة من مراحل منهجية بوكس وجينكينز حيث هو الهدف النهائي من تحليل السلاسل الزمنية. إذ لا يمكن الوصول إلى هذه المرحلة حتى يجتاز الأنموذج المبدئي جميع الفحوصات والاختبارات التشخيصية الخاصة بأسلوب بوكس وجينكينز، ان التنبؤ بالقيم المستقبلية باستخدام الأنموذج الأفضل والذي تم تقدير معلماته، وهو الهدف المرجو من بناء الأنموذج الإحصائي، ومن خلال التنبؤ يمكننا معرفة السلوك المستقبلي لظاهرة ما بأقل خطأ ، فإذا كانت لدينا سلسلة زمنية معينة مثل $(Z_t, Z_{t-1}, \dots)$ عندها ننتبأ بالقيمة المستقبلية (Z_{t+1}) للسلسلة بأخذ التوقع الشرطي للتنبؤ في

المدة $(t+T)$ عند الزمن (t) وهو تقدير غير متحيز إلى (Y_{t+T}) لنحصل على التنبؤات اقل تباين للخطأ، وان التوقع الشرطي للتنبؤ z عند المدة t على وفق الصيغة الآتية [7]:

$$Z_{t+T} = \phi_1 E(Z_{t+T-1}) + \dots + \phi_{p+d} E(Z_{t+T-p-d}) - \theta_1 E(a_{t+T-1}) - \dots \dots \dots - \theta_{q-d} E(a_{t+T-q-d}) \dots \dots \dots (25)$$

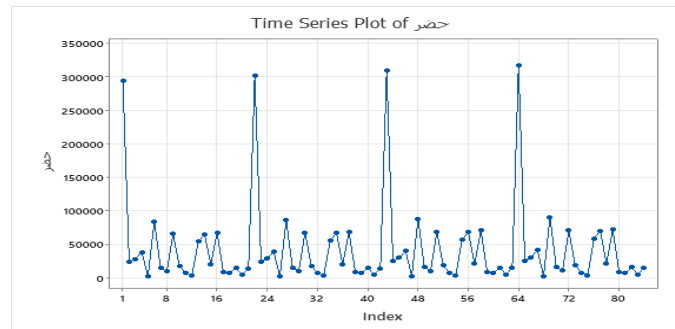
حيث Z_{t+T} : القيمة المتنبأ بها .

10- وصف البيانات وتحليل النتائج

تم جمع البيانات من خلال التقرير السنوي لهيأة الاحصاء ونظم المعلومات الجغرافية - مديرية احصاء ديالى ، حيث كان حجم العينة 21 وحدة اداري من حيث الاقضية والنواحي في المحافظة ولأربع سنوات (2020-2023) وتكونت الدراسة من 84 عينة ، وتضم اعداد السكان من (الحضر ، الريف) حيث تلخص اعداد السكان بشكل ادق كون الاختلاف بين اعداد السكان بين المنطقتين .

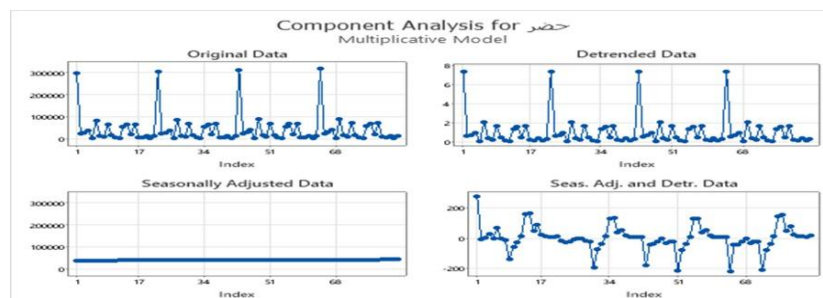
10-1 تحليل السلسلة الزمنية لا اعداد السكان بالنسبة لمتغير الحضر

نبدأ برسم السلسلة الزمنية لمعرفة سلوك السلسلة الزمنية



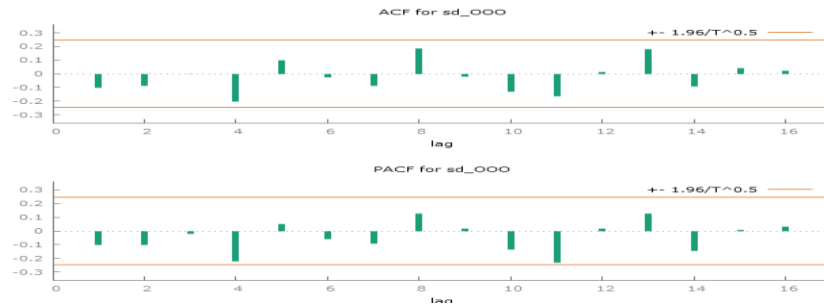
شكل (1) يمثل رسم السلسلة لتغيرات السكانية في محافظة ديالى (حضر)

خلال الشكل يمكن ملاحظة السلسلة تحتوي على الاثر الموسمي اذ ان السلسلة تتذبذب في اوقات معينة للسنوات ولتخلص من الاثر الموسمي يتم اخذ الفرق الموسمي.



شكل (2) السلسلة بعد اخذ الفرق الموسمي لها

من خلال الشكل اعلاه نلاحظ ان السلسلة تخلصت من الاثر الموسمي بعد اخذ الفرق لموسمي لها



شكل (3) دالة الارتباط الذاتي ودالة الارتباط الجزئي

نلاحظ من خلال الشكل (3) بان معاملات دالة الارتباط الذاتي مستقرة حيث تقع ضمن حدود الثقة (± 0.5) ، وكذلك بالنسبة لدالة الارتباط الجزئي والجزئي فتقع جميع معاملات ضمن حدود الثقة . وللتأكد أكثر من استقرارية السلسلة يتم فحص استقراريته من خلال جذر الوحدة باستعمال اختبار ديكي – فولر.

جدول (1) اختبار ديكي - فولر لاستقرارية السلسلة

Null hypothesis:		Data are non-stationary
Alternative hypothesis:		Data are stationary
Test Statistic	P-Value	Recommendation
-6.43843	0.000	Test statistic $\leq$ critical value of -3.90839.
		Significance level = 0.05

$H_0: \phi_1 = 0$ (الوحدة جذر وجود)

$H_1: \phi_1 \neq 0$ (عدم الوحدة جذر وجود)

من نتائج الاختبار نلاحظ ان $(p\text{-value} < 0.05)$ مما يدل على استقرارية السلسلة .

لتحديد رتبة النموذج ARIMA من خلال اختيار افضل نموذج باستعمال معايير المفاضلة (معياري اكاكي AIC ، ومعياري المصحح AICc ، ومعياري بيز).

جدول (2) اختيار النموذج Model Selection

SARIMA	AICc	AIC	BIC
$SARIMA(0,1,1)(0,1,1)^{90}$	1078.82	1078.41	1084.84
$SARIMA(1,1,0)(0,1,1)^{90}$	1080.21	1079.52	1088.09
$SARIMA(0,0,1)(0,1,1)^{90}$	1081.88	1081.19	1089.76
$SARIMA(0,1,2)(0,1,1)^{90}$	1086.94	1085.88	1096.60
$SARIMA(0,1,3)(0,1,1)^{90}$	1091.26	1089.76	1102.62
$SARIMA(1,1,1)(0,1,1)^{90}$	1111.16	1110.47	1119.05
$SARIMA(1,1,2)(0,1,1)^{90}$	1114.42	1113.36	1124.08
$SARIMA(0,1,1)(0,1,0)^{90}$	1115.54	1115.14	1121.57
$SARIMA(1,1,0)(0,1,0)^{90}$	1115.89	1115.48	1121.91
$SARIMA(0,1,2)(0,1,0)^{90}$	1115.90	1115.21	1123.78
$SARIMA(0,1,3)(0,1,0)^{90}$	1116.55	1115.50	1126.21
$SARIMA(1,1,3)(0,1,0)^{90}$	1116.93	1115.43	1128.29
$SARIMA(1,1,2)(0,1,1)^{90}$	1134.15	1132.65	1145.51
$SARIMA(1,1,1)(0,1,1)^{90}$	1137.21	1136.16	1146.87

من خلال معايير المفاضلة كان افضل نموذج هو $SARIMA(0,1,1)(0,1,1)^{90}$ ، كون قيم (معياري اكاكي المصحح AICc ومعياري اكاكي AIC و معياري بيز BIC) اقل ما يمكن بين جميع القيم المعيارية للنماذج المقترحة ولقد تم تقدير معلومات الأنموذج وفق طريقة الامكان الاعظم المضبوطة كما موضحة في الجدول وتقدير النموذج من خلال الجدول الاتي :

جدول (3) تقدير النموذج

Type	Coef	SE Coef	T- Value	P- Value
SMA 21	-0.746	0.200	-3.73	0.000
Constant	1066	261	4.09	0.000

حيث ان مقدرات النموذج كانت $P\text{-value} < 0.05$ اي ان النموذج معنوي

$$Z_t = 1066 + 0.746$$

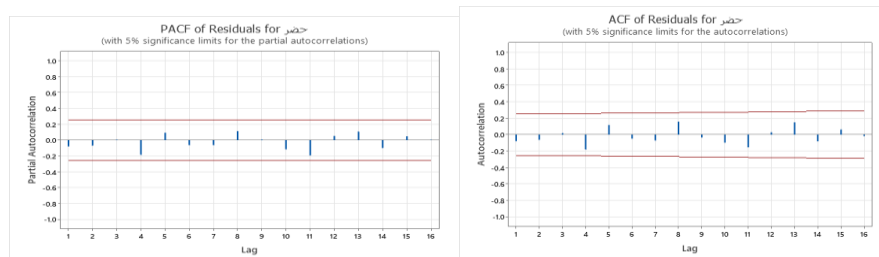
فحص ملائمة النموذج

من خلال تطبيق إحصائية (Ljung-Box) الذي يعتمد على البواقي لفحص ملائمة النموذج، وتزداد قيمة P-Value كلما تزداد عدد الإزاحات وهذا دليل على أن النموذج جيد وملائم وكفوء وهي تعني قبول الفرضية العدم (لا توجد ارتباطات ذاتية بين الأخطاء)

جدول (4) اختبار (Ljung-Box) لاختبار ملائمة النموذج

Lag	12	24	36	48
Chi-Square	9.09	33.72	39.34	84.29
DF	10	22	34	46
P-Value	0.524	0.052	0.243	0.000

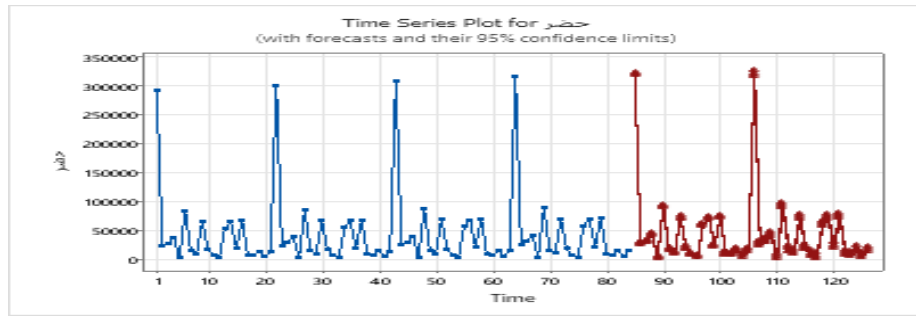
من خلال الجدول نلاحظ ان $P\text{-Value} > 0.05$ اي ان النموذج ملائم ولا يوجد ارتباطات بين المتغيرات



شكل (4) دالة الارتباط الذاتي والارتباط الجزئي للبواقي

من خلال الشكل اعلاه نلاحظ ان قيم البواقي ضمن حدود الثقة اي ان البواقي لاتعاني من ارتباطات بين المتغيرات .

بعد التعرف على السلسلة الزمنية وفحص السلسلة وتقديرها تأتي المرحلة الأخيرة من منهجية بوكس جينكنز وهي التنبؤ.



شكل (5) السلسلة المتنبا بها لأعداد سكان الحضر

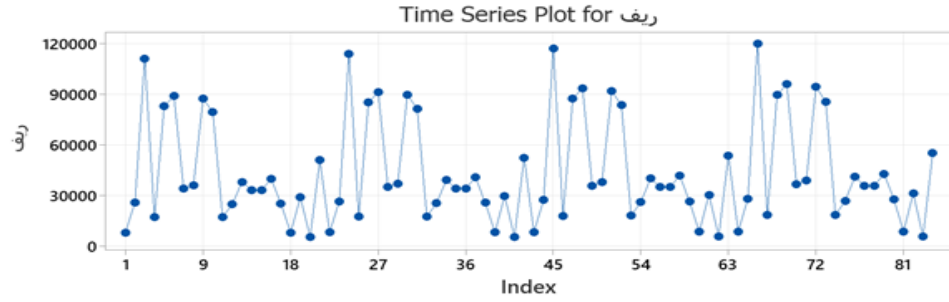
جدول (5) التنبؤ بالسلسلة الزمنية لا أعداد السكان الحضر حسب الوحدات الادارية ولسنة 2024 ، 2025

95% Limits						95% Limits					
Years	Time	Forecast	SE Foreca st	Low er	Uppe r	years	Tim e	Fore cast	SE Forec ast	Lower	Upper
2024	1	321985	1132.48	319765	324205	2025	1	323051	2278.58	318584	327518
	2	26382	1132.48	24162	28602		2	27448	2278.58	22981	31915
	3	31773	1132.48	29553	33993		3	32839	2278.58	28372	37306
	4	42525	1132.48	40305	44745		4	43591	2278.58	39124	48058
	5	3647	1132.48	1427	5867		5	4713	2278.58	246	9180
	6	92279	1132.48	90059	94499		6	93345	2278.58	88878	97812
	7	17156	1132.48	14936	19376		7	18222	2278.58	13755	22688
	8	11436	1132.48	9216	13656		8	12502	2278.58	8035	16969

	9	72470	1132.4 8	7025 0	7469 0		9	7353 6	2278. 58	69069	78003
	10	19917	1132.4 8	1769 7	2213 7		10	2098 3	2278. 58	16516	25450
	11	8557	1132.4 8	6337	1077 7		11	9623	2278. 58	5156	14090
	12	4426	1132.4 8	2205	6646		12	5491	2278. 58	1025	9958
	13	59844	1132.4 8	5762 4	6206 4		13	6091 0	2278. 58	56443	65377
	14	72030	1132.4 8	6981 0	7425 0		14	7309 6	2278. 58	68629	77562
	15	22606	1132.4 8	2038 6	2482 6		15	2367 2	2278. 58	19205	28139
	16	74050	1132.4 8	7183 0	7627 0		16	7511 6	2278. 58	70649	79583
	17	9890	1132.4 8	7670	1211 1		17	1095 6	2278. 58	6489	15423
	18	8318	1132.4 8	6098	1053 8		18	9384	2278. 58	4917	13851
	19	16633	1132.4 8	1441 3	1885 3		19	1769 9	2278. 58	13232	22166
	20	6038	1132.4 8	3818	8258		20	7104	2278. 58	2637	11571
	21	16002	1132.4 8	1378 2	1822 2		21	1706 8	2278. 58	12601	21535

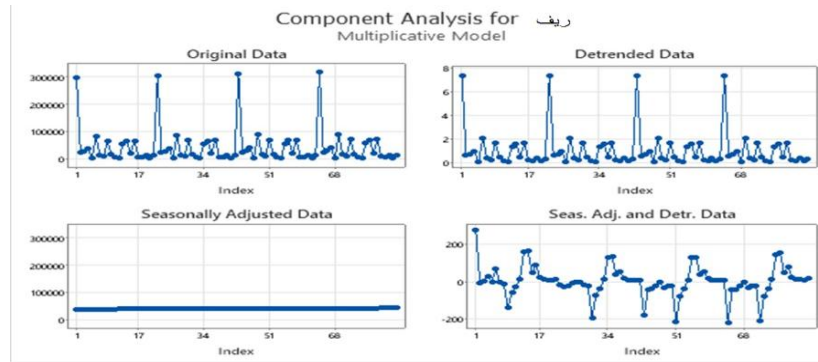
10-2 تحليل السلسلة الزمنية لمتغير الريف

نبدأ برسم السلسلة الزمنية للمتغير المعتمد للتعرف على طبيعة السلسلة



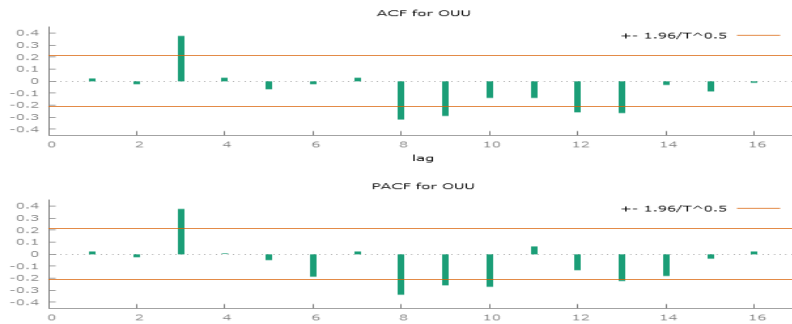
شكل (6) السلسلة لا عداد السكان للوحدات الادارية الريف

خلال الشكل يمكن ملاحظة السلسلة تحتوي على الاثر الموسمي اذ ان السلسلة تتذبذب في اوقات معينة للسنوات ولتخلص من الاثر الموسمي يتم اخذ الفرق الموسمي .



شكل (7) السلسلة بعد اخذ الفرق الموسمي لها

شكل السلسلة بعد اخذ الفرق الموسمي وازالة الاثر الموسمي منها .



شكل (8) دالة الارتباط الذاتي ودالة الارتباط الجزئي

نلاحظ من خلال الشكل (8) بان معاملات دالة الارتباط الذاتي مستقرة حيث تقع ضمن حدود الثقة (± 0.5) ، وكذلك بالنسبة لدالة الارتباط الجزئي والجزئي فتقع جميع معاملات ضمن حدود الثقة . وللتأكد أكثر من استقرارية السلسلة يتم فحص استقراريته من خلال جذر الوحدة باستعمال اختبار ديكي – فولر.

جدول (6) اختبار ديكي فولر

Null hypothesis:		Data are non-stationary
Alternative hypothesis:		Data are trend stationary
Test Statistic	P-Value	Recommendation
-5.82065	0.000	Test statistic <= critical value of -3.91477.
		Significance level = 0.05

من نتائج الاختبار نلاحظ ان ($p\text{-value} < 0.05$) مما يدل على استقرارية السلسلة .

جدول (7) اختيار النموذج

SARIMA	AICc	AIC	BIC
$SARIMA(1,0,0)(2,1,1)^{90}$	1078.82	1078.41	1084.84
$SARIMA(1,0,0)(2,1,0)^{90}$	1080.21	1079.52	1088.09
$SARIMA(0,0,1)(2,1,1)^{90}$	1081.88	1081.19	1089.76
$SARIMA(0,0,2)(2,1,1)^{90}$	1086.94	1085.88	1096.60
$SARIMA(0,0,0)(2,1,1)^{90}$	1091.26	1089.76	1102.62
$SARIMA(1,0,2)(2,1,1)^{90}$	1111.16	1110.47	1119.05
$SARIMA(0,0,0)(0,1,1)^{90}$	1114.42	1113.36	1124.08
$SARIMA(0,1,1)(0,1,1)^{90}$	1115.54	1115.14	1121.57
$SARIMA(1,1,0)(0,1,1)^{90}$	1115.89	1115.48	1121.91
$SARIMA(1,1,2)(0,0,0)^{90}$	1115.90	1115.21	1123.78
$SARIMA(0,1,1)(0,0,0)^{90}$	1116.55	1115.50	1126.21
$SARIMA(1,1,0)(0,0,0)^{90}$	1116.93	1115.43	1128.29

$SARIMA(0,1,2)(0,0,0)^{90}$	1134.15	1132.65	1145.51
$SARIMA(1,1,1)(0,0,0)^{90}$	1137.21	1136.16	1146.87

من الجدول اعلاه نحدد رتبة الانموذج ARIMA من خلال اختيار افضل نموذج باستعمال معايير المفاضلة ومن خلال معايير المفاضلة كان افضل نموذج هو $SARIMA(1,0,0)(2,1,1)^{90}$ كون قيم (معياري اكيي المصحح AICc ومعيار اكيي AIC و معيار بيز BIC) اقل ما يمكن بين جميع القيم المعيارية للنماذج المقترحة ولقد تم تقدير معلمات الأنموذج وفق طريقة الامكان الاعظم المضبوطة كما موضحة في الجدول وتقدير النموذج من خلال الجدول الآتي :

جدول رقم (8) تقدير الانموذج

Type	Coef	SE Coef	T-Value	P-Value
AR 1	-0.016	0.130	-0.12	0.0 03
SAR 21	1.9742	0.0221	89.48	0.000
SAR 42	-0.9662	0.0217	-44.48	0.000
SMA 21	0.915	0.201	4.55	0.000

من خلال اختبار Z وباعتماد على قيمة p-value نلاحظ معلمات
 $SARIMA(1,0,0)(2,1,1)^{90}$ معنوية لان قيمة p-value اقل من (0.05)

الأنموذج التقديري

$$-0.016(B)1.9742(B^{21})(1 - B)^1(1 - B^{21})^1 z_t = \delta - 0.9662(B^{21})a_t$$

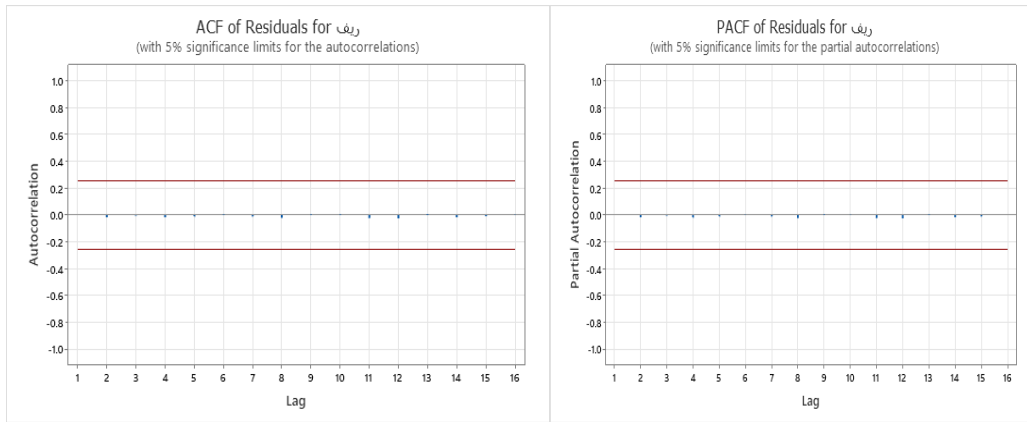
ملائمة النموذج

يستعمل اختبار Ljung –Box لاختبار ملائمة النموذج قبل البدء بالتنبؤ

جدول (9) اختبار ملائمة النموذج

Lag	12	24	36	48
Chi-Square	0.17	37.46	37.73	40.70
DF	8	20	32	44
P-Value	1.000	0.010	0.224	0.614

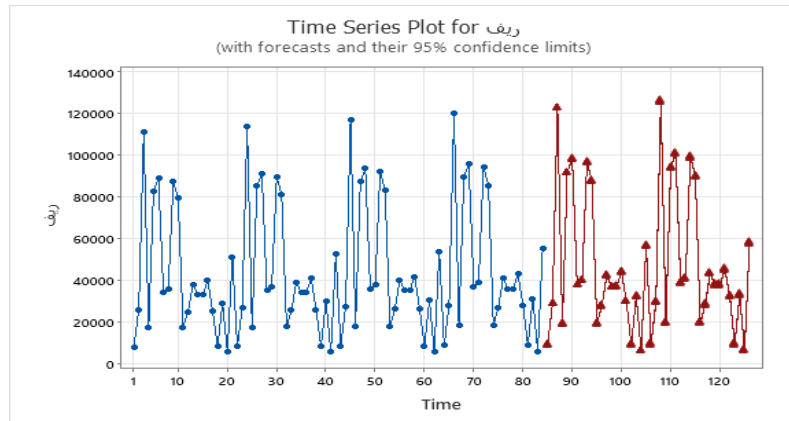
من خلال الجدول نلاحظ ان قيمة (P-Value > 0.05) مما يدل على معنوية النموذج اي لا توجد ارتباطات بين بواقي المتغيرات .



شكل (9) دالة الارتباط الذاتي والارتباط الجزئي للبواقي

من خلال الشكل اعلاه نلاحظ ان قيم البواقي ضمن حدود الثقة اي ان البواقي لاتعاني من ارتباطات بين المتغيرات .

بعد التعرف على السلسلة الزمنية وفحص السلسلة وتقديرها تاتي المرحلة الاخيرة من منهجية بوكس جينكنز وهي التنبؤ



شكل (10) القيم المتنبأ بها للسلسلة الزمنية لأعداد السكان في الريف

جدول (10) التنبؤ بالسلسلة الزمنية لا عداد السكان الريف حسب الوحدات الادارية ولسنة 2024 2025

				95% Limits						95% Limits	
years	Time	Forecast	SE Forecast	Lower	Upper	years	Time	Forecast	SE Forecast	Lower	Upper
2024	1	8919	82.986	8756	9082	2025	1	9148	189.995	8776	9520
	2	28755	82.997	28592	28918		2	29490	190.015	29118	29863
	3	122848	82.997	122686	123011		3	125993	190.015	125620	126365
	4	18974	82.997	18811	19137		4	19459	190.015	19086	19831
	5	91724	82.997	91561	91887		5	94068	190.015	93695	94440
	6	98256	82.997	98093	98419		6	100769	190.015	100396	101141
	7	37732	82.997	37570	37895		7	38702	190.015	38329	39074
	8	39795	82.997	39632	39957		8	40817	190.015	40445	41190
	9	96648	82.997	96485	96811		9	99124	190.015	98752	99497
	10	87707	82.997	87544	87870		10	89953	190.015	89580	90325
	11	19053	82.997	18890	19215		11	19544	190.015	19171	19916
	12	27555	82.997	27393	27718		12	28259	190.015	27886	28631
	13	42178	82.997	42015	42340		13	43260	190.01	42887	43632

								5		
	14	36787	82.997	36624	36949		14	37729	190.01 5	37356 38101
	15	36826	82.997	36663	36989		15	37769	190.01 5	37396 38141
	16	43975	82.997	43812	44138		16	45100	190.01 5	44728 45473
	17	29862	82.997	29700	30025		17	32358	190.01 5	31985 32730
	18	9058	82.997	8896	9221		18	9290	190.01 5	8917 9662
	19	32068	82.997	31905	32231		19	32889	190.01 5	32517 33262
	20	6085	82.997	5922	6247		20	6243	190.01 5	5870 6615
	21	56481	82.997	56318	56643		21	57941	190.01 5	57569 58314

11 – الاستنتاج

من خلال التحليل الاحصائي للسلسلتين اعداد السكان حسب الوحدات الادارية في محافظة ديالى (الحضر و الريف) نجد الاتي :

- 1- ان سلسلتي التغيرات السكانية المتمثلة بالحضر والريف كانتا غير مستقرة وتحتوي على الاثر الموسمي وقد اخذ الفرق الموسمي الاول للتخلص من هذا الاثر.
- 2- ان انموذج السلسلة الاولى والمتمثلة بالحضر كانت من النوع الموسمية التكاملية المختلط هو $SARIMA(0,1,1)(0,1,1)^{90}$
- 3- ان انموذج سلسلة الثانية المتمثلة بالريف كانت من نوع نموذج المتوسط المتحرك الموسمي للانحدار الموسمي الذاتي التكاملية المختلط هو $SAMIMA(1,0,0)(2,1,1)^{90}$
- 4- كان توفيق الأنموذج المعتمد لأعداد السكان حسب الوحدات الادارية للسلسلتين توفيقاً جيداً بعد تطبيق اختبار ليونغ بوكس Ljung-Box والذي يظهر عدم وجود فروق معنوية واختبار البواقي يشير ان سلسلة الاخطاء للأنموذج المشخص هي طبيعية.

12- التوصيات

- 1- دراسة إمكانية التغيرات السكانية والاستفادة من التعداد السكاني الذي سيجري في هذه السنة والذي بدوره سيلعب دوراً مهماً في إليه عد السكان مستقبلاً لكونه أقرب للواقع واعتماد على بياناته بكونها أكثر دقة من التوقع الحاصل في البيانات الحالية إذ اعتمدت سنة 1997 هي السنة الاساس.
- 2- دراسة امكانية استعمال البرامج الاحصائية في التنبؤ باستعمال المحاكاة.
- 3- اعتماداً على القيم التنبؤية التي تم التوصل اليها في هذا البحث نوصي بالاستفادة من النتائج واخذها بنظر الاعتبار من الجهات المختصة وفق النماذج المختارة لوضع الخطط المستقبلية.

13- المصادر

1. الدوري، أحلام احمد جمعة (2003)، (بعض الاختبارات الإحصائية لانموذج الانحدار الذاتي الطبيعي من الدرجة الأولى) أطروحة دكتوراه في الإحصاء – كلية الإدارة والاقتصاد – جامعة بغداد.
2. الزوبعي ، عبيد محمود محسن (2005)، (تشخيص وفحص مدى الملاءمة لنماذج السلاسل الزمنية المختلطة ذات الرتب الدنيا) أطروحة دكتوراه في الإحصاء – كلية الإدارة والاقتصاد – جامعة بغداد.
3. المهنا، فراس احمد محمد (2007)، (بعض طرائق التشخيص والتقدير في الانحدار الذاتي الخطي وغير الخطي ذي الرتب الدنيا) أطروحة دكتوراه في الإحصاء – كلية الإدارة والاقتصاد – جامعة بغداد.
4. Abdullah, M. M., & Masmoudi, A. (2023a). Modeling Real-life Data Sets with a Novel G Family of Continuous Probability Distributions: Statistical Properties, and Copulas. Pakistan Journal of Statistics and Operation Research, 719-746.
5. Abdullah, M., M., Mohammed, W. S., & Farhana, A. H. (2023b). The Poisson-Topp-Leone Burr Type-X II Model: Various Uncensored Applications for Statistical Modeling and Some Copulas. Statistics, Optimization & Information Computing, 11(2), 276-298.
6. Alaka S,H,(Using Buckth-Jenkins models to predict traffic accident deaths in the Holy Governorate of Karbala for the period (2010-2015)), University of Karbala - College of Administration and Economics - published master's thesis. 2017.
7. Al-Zaidi, K, N, (Traditional methods and alternative methods for testing the unit root of time series stability - a comparative study), Al-Mustansiriya College - of Administration and Economics - Doctoral dissertation, 2015.
8. Hamid, F, A & Arshad, A, (Using time series to predict residential rent index numbers in Iraq for the years 2018-2021). Journal of Al-Raiding University College for Sciences (Print ISSN: 1681-6870, Online ISSN: 2790-2293), (2), 77-96, (2019).

9. Majid, A, B, (Using time series to predict consumer price index numbers in Iraq for the period (2018-2022)), University of Baghdad - College of Administration and Economics - Higher Diploma , 2019 .
10. Mohamed, S, B, (A comparison between estimation methods for the parameters of the AR (1) autoregressive model using simulation), Technical Journal, Volume 20, Issue (2) , 2007 .