

التنبؤ بأعداد الزائرين غير العراقيين في محافظة كربلاء المقدسة باستخدام اسلوب السلاسل الزمنية

م.م سمانة عزيز عبد الحسن

م.اسيل عبد الرزاق رشيد

الملخص

ركز هذا البحث على استخدام اسلوب السلاسل الزمنية (منهجية بوكس - جينكنز) للتنبؤ بأعداد الزائرين في مدينة كربلاء وبمقارنة نماذج مختلفة تبين أن النموذج الملائم هو نموذج $ARIMA(0,1,1)$ و تم استخدامه في التنبؤ بالقيم المستقبلية لخمسة اشهر .

Abstract

This research has focused on the use of method time series (Box-Jenkins methodology) to predict the numbers of visitors in the city of Karbala By comparing several different models, it shows that the best model is $ARIMA(0,1,1)$. Therefore, the research used $ARIMA(0,1,1)$ model for forecasting the future values during five months.

١-١ المقدمة

إن دراسة السلسلة الزمنية لا تقتصر على الامور الاقتصادية فقط بل أخذت حيزاً واسعاً في العلوم الطبيعية والصناعية والتجارة والتربة ونمو السكان والرعاية الصحية وغيرها حيث يعد تحليل السلاسل الزمنية من أهم الاساليب الاحصائية للاستدلال حول مستقبل الظاهرة بناءً على البيانات الاحصائية لأحداث الماضي والحاضر والتنبؤ لبيانات السلسلة الزمنية هو أحد الادوات الاحصائية القوية التي تساعد في بناء الاستراتيجيات ورسم الخطط المستقبلية وصنع القرارات. وقد ركز هذا البحث على استخدام احد الوسائل الاحصائية (السلاسل الزمنية) للتنبؤ بأعداد الزائرين في مدينة كربلاء للسنوات القادمة.

٢-١ هدف البحث

يهدف هذا البحث الى استخدام اسلوب تحليل السلاسل الزمنية لبناء نموذج تنبؤي بأعداد الزائرين غير العراقيين في محافظة كربلاء المقدسة باستخدام نموذج بوكس-جينكنز للسلاسل الزمنية وتطبيقها على بيانات شهرية تم جمعها من مديرية الامن السياحي في محافظة كربلاء المقدسة تمثل اعداد الزائرين غير العراقيين للمدة من عام ٢٠١٠ الى عام ٢٠١٣ .

٣-١ فرضية البحث

١- ان السلسلة الزمنية لأعداد الزائرين لمدينة كربلاء المقدسة للمدة من ٢٠١٠-٢٠١٣ تكون مستقرة.

٢- امكانية استخدام طرق تنبؤ حديثة مثل منهجية بوكس جينكنز للتنبؤ بأعداد الزائرين لمدينة كربلاء المقدسة للمدة من ٢٠١٠-٢٠١٣.

١-٤ الدراسات السابقة

هنالك كثير من الدراسات التي تناولت أسلوب السلاسل الزمنية واختباراتها واختيار النموذج الامثل للتنبؤ ببيانات تلك الدراسات وركزت دراسات اخرى على منهجية بوكس جينكنز في التنبؤ بالسلاسل الزمنية ذات الطبيعة الاقتصادية والاجتماعية المختلفة واكثر المصادر الموجودة في نهاية البحث تحوي امثلة تطبيقية على ذلك. أما فيما يخص الى استخدام منهجية بوكس- جينكنز للتنبؤ بأعداد الزائرين وخصوصا" في مدينة كربلاء المقدسة لم يتمكن الباحثان من الوصول الى دراسة سابقة قد تناولت هذا الموضوع.

الجانب النظري

٢-١ نماذج السلاسل الزمنية

تعرف السلسلة الزمنية على انها مجموعة من المشاهدات لقيم ظاهرة ما خلال مدة زمنية معينة في الغالب تكون متساوية، وتعرف السلسلة الزمنية رياضياً بانها متتابعة من المتغيرات العشوائية معرفة ضمن فضاء الاحتمالية متعدد المتغيرات. (٥) وتكون السلسلة الزمنية مستقرة إذا كانت الخصائص الاحتمالية لانتأثر بالزمن أو تكون غير مستقرة.

وأن الدالة التي تربط قيم السلسلة الزمنية بالقيم السابقة لها مع إضافة الاخطاء العشوائية، تسمى نماذج السلسلة الزمنية والتي تكون أما نماذج تصادفية أو نماذج احتمالية. هنالك أنواع متعددة من نماذج السلاسل الزمنية وكالاتي:

٢-١-١-١ نموذج الانحدار الذاتي (AR) Autoregressive Model

يرمز لنموذج الانحدار الذاتي من الرتبة p بالرمز AR(p) بالصيغة العامة هي:

$$\phi(B)X_t = \epsilon_t$$

حيث ان X_t تمثل مشاهدات السلسلة الزمنية عند الزمن t

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad \text{وان}$$

حيث ان ϕ_i تمثل معالم نموذج الانحدار الذاتي عندما $i = 1, 2, \dots, p$

وان ϵ_t متغيرات عشوائية غير مرتبطة بوسط حسابي صفر وتباين σ_ϵ^2

٢-١-٢-٢ نموذج الاوساط المتحركة (MA) Moving Average model

يرمز لنموذج الاوساط المتحركة من الرتبة q بالرمز MA(q) بالصيغة العامة هي:

$$X_t = \theta(B) e_t$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad \text{وان:}$$

حيث ان θ_i تمثل معالم نموذج الاوساط المتحركة عندما $i=1,2,\dots,q$ (٢)

٢-١-٣- نموذج الانحدار الذاتي والأوساط المتحركة (النموذج المختلط).

Autoregressive Moving Average Methods (ARMA)

يعرف هذا النموذج أحيانا "بنموذج بوكس-جينكنز ويرمز له بالرمز $ARMA(p,q)$ إن الصيغة العامة هي:

$$\phi(B)X_t = \theta(B)e_t$$

ان $AR(p)$ هي حالة خاصة من النموذج $ARMA$ عندما $q=0$ ، إن $MA(q)$ هي ايضا حالة خاصة من النموذج $ARMA$ عندما $p=0$.

ويتطلب تقدير النموذج أن تكون السلسلة الاصلية مستقرة، أما اذا كانت السلسلة غير مستقرة فيجب تحويلها لسلسلة مستقرة عن طريق أخذ الفروق ويعتبر عدد مرات الفروق المطلوبة لتحويل السلسلة الى سلسلة مستقرة درجة تكامل السلسلة فيقال السلسلة متكاملة من الدرجة d اذا تطلب أخذ الفروق d مرة لكي تصبح السلسلة مستقرة. وعندها يتحول الى نموذج الانحدار الذاتي والأوساط المتحركة المتكامل $Autoregressive Integrated Moving Average$

$$ARIMA(p, d, q) \quad \text{ويرمز له:}$$

حيث ان :

p : تمثل رتبة الانحدار الذاتي

q : تمثل رتبة الاوساط المتحركة

d : تمثل رتبة التكامل (٤)

٢-١-٤- النموذج المختلط المتكامل.

Autoregressive Integrated Moving Average Methods (ARIMA)

يرمز له بالرمز $ARIMA(p,d,q)$ إن الصيغة العامة هي:

$$\phi(B)(1-B)^d X_t = \theta(B)e_t$$

ويمثل هذا النموذج السلسلة الزمنية المستقرة بعد ان تم إجراء عدد من التحويلات أو الفروق على النموذج الاصيلي للسلسلة (سلسلة غير مستقرة). (٢)

٢-٢ منهجية بوكس-جينكنز

تعد هذه الطريقة التي اقترحها بوكس-جينكنز عام ١٩٧٦ من أوسع الطرق استخداماً في تحليل السلسلة الزمنية المستقرة وغير المستقرة والموسمية وغير الموسمية وتتضمن هذه الطريقة أربع مراحل:

١-٢-٢ التحديد أو التشخيص

تعد مرحلة التشخيص المرحلة الأهم في تحليل السلاسل الزمنية. وتشمل معرفة نوع النموذج وتحديد الرتبة للنموذج المحدد، وذلك من خلال استخدام دالتي الارتباط الذاتي (ACF) والارتباط الذاتي الجزئي (PACF) حيث يتم الرسم البياني لهما ثم يتم مطابقة معاملات الارتباط الذاتي والجزئي مع السلوك النظري لدالتي الارتباط الذاتي (ACF) والارتباط الذاتي الجزئي (PACF). والجدول التالي يوضح سلوك كل من الدالتين:

النموذج	دالة الارتباط الذاتي (ACF)	الارتباط الذاتي الجزئي (PACF)
AR (P)	تناقص أسّي	يقطع بعد p من الازاحات
MA(q)	يقطع بعد q من الازاحات	تناقص أسّي
ARMA(p,q)	تناقص أسّي	تناقص أسّي

٢-٢-٢ تقدير النموذج

يتم في هذه المرحلة تقدير معالم النموذج باستخدام إحدى طرق التقدير التامة أو التقريبية والتي تختلف باختلاف النموذج المستخدم وهناك عدة طرق للتقدير منها:

أ-طريقة Yule-Walker والتي تعتمد على الارتباط الذاتي للنموذج وتدعى بطريقة العزوم .

ب-طريقة الامكان الاعظم وتعتمد على تعظيم الدالة لجعل مجموع مربعات الاخطاء أقل ما يمكن.

ج-طريقة المربعات الصغرى غير الخطية والتي تستخدم عادة في نماذج ARMA(p,q) وتعتمد على فكرة البحث عن قيم المعالم التي تجعل مجموع مربعات الأخطاء أقل ما يمكن. (١)

٣-٢-٢ اختبار ملائمة النموذج

يتم في هذه المرحلة فحص النموذج من خلال معاملات الارتباط الذاتي و معاملات الارتباط الذاتي الجزئي للبوافي في النموذج المقدر وليس للسلسلة الاصلية فإذا كانت واقعة ضمن حدي الثقة وبمستوى ثقة معين فهذا يعني بان بوافي النموذج عشوائية وعليه فان النموذج المحدد يكون ملائماً. ويمكن اختبار عشوائية البوافي وذلك باستخدام إحدى المعايير الآتية:

أ-معيار (Box-Pierce 1970)

ب-معيار (Ljung-Box ١٩٧٨)

ج-اختبار (MA nti 1994)

وجميع الاختبارات السابقة تتبع بصورة تقريبية توزيع (χ^2) بدرجة حرية $(k-p-q)$

حيث ان :

p:تمثل رتبة الانحدار الذاتي، q:تمثل رتبة الاوساط المتحركة، k:تمثل عدد الازاحات المدروسة.

٢-٢-٤ التنبؤ بالقيم المستقبلية

في هذه المرحلة يتم إيجاد القيم المستقبلية للسلسلة الزمنية باستخدام النموذج الملائم الذي تم الحصول عليه من المراحل السابقة، إن التنبؤ الأمثل هو ذلك التقدير الذي يكون الخطأ الناتج عنه صغيراً جداً وتباينه أقل ما يمكن، وهناك طريقتان أساسيتان للتنبؤ بالسلاسل الزمنية حسب طريقة بوكس-جينكنز وهما:

أ- طريقة الاسقاط الذاتي

ب- طريقة السبب والتقدير

ان الصيغة العامة للتنبؤ هي:

$$E(X_{t+h}) = \phi_1 E(X_{t+h-1}) + \phi_2 E(X_{t+h-2}) + \dots + \phi_p + q E(X_{t+h-p} = d) - \theta_1 E(e_{t+h-1}) - \theta_2 E(e_{t+h-2}) - \dots - \theta_q E(e_{t+h-q}) + E(e_{t+h})$$

حيث ان: h تمثل الفترة الزمنية المستقبلية.

٢-٣ مقاييس دقة التنبؤ

يمكن قياس دقة التنبؤ والمعتمد على خطأ التنبؤ من خلال عدة مقاييس والتي تأخذ الصيغ التالية (٣):

أ- متوسط الأخطاء والذي يأخذ الصيغة التالية:

$$ME = (1/n) \sum_{i=1}^n e_i$$

ب- متوسط القيمة المطلقة للخطأ والذي يأخذ الصيغة التالية:

$$MAE = (1/n) \sum_{i=1}^n |e_i|$$

ج- متوسط نسبة الأخطاء المطلقة والذي يأخذ الصيغة التالية:

$$MAPE = (1/n) \sum_{i=1}^n \left| \frac{e_i}{y_t} \right| 100$$

الجانب العملي

1-3 عينة البحث

تم الحصول على البيانات من مديرية الامن السياحي في محافظة كربلاء المقدسة وهي عبارة عن اعداد الزائرين غير العراقيين (العرب والاجانب) الساكنين في الفنادق فقط وهي بيانات شهرية من شهر كانون الثاني من عام ٢٠١٠ وحتى شهر كانون الاول عام ٢٠١٣ وكما موضح في الجدول رقم (١).

جدول رقم (١)

بيانات اعداد الزائرين الغيرعراقيين الساكنين في الفنادق

الاعداد	السنوات	الاعداد	السنوات	الاعداد	السنوات	الاعداد	السنوات
461801	1-2013	375981	1-2012	465448	1-2011	205552	1-2010
109284	2-2013	294321	2-2012	317759	2-2011	607547	2-2010
211699	3-2013	536461	3-2012	361254	3-2011	199802	3-2010
183514	4-2013	272723	4-2012	302719	4-2011	307766	4-2010
57616	5-2013	235672	5-2012	307081	5-2011	356288	5-2010
169109	6-2013	135604	6-2012	261408	6-2011	211977	6-2010
45159	7-2013	217341	7-2012	248239	7-2011	195057	7-2010
128223	8-2013	142984	8-2012	305285	8-2011	290052	8-2010
170148	9-2013	244731	9-2012	310309	9-2011	195930	9-2010
314581	10-2013	228417	10-2012	334166	10-2011	224079	10-2010
292436	11-2013	464647	11-2012	317638	11-2011	85919	11-2010
258856	12-2013	310021	12-2012	313941	12-2011	901731	12-2010

المصدر: مديرية الامن السياحي في محافظة كربلاء المقدسة

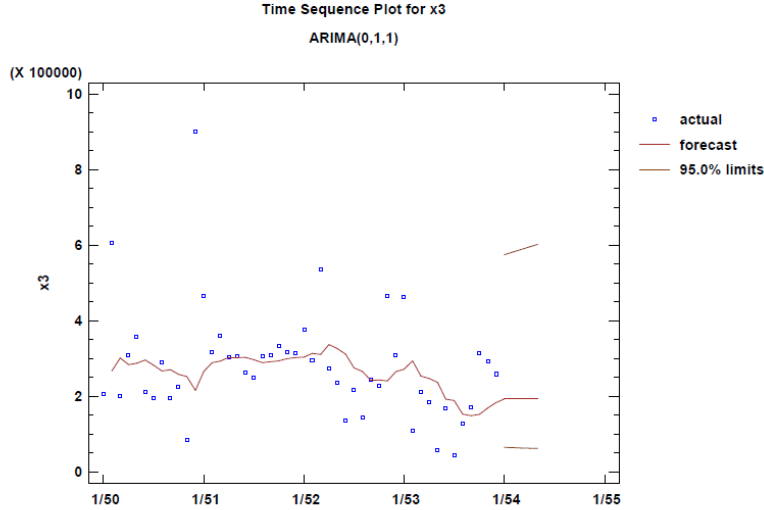
٢-٣: التحليل الاحصائي

١-٢-٣: استقرارية السلسلة

من خلال رسم سلسلة البيانات وجد ان السلسلة غير مستقرة من حيث المتوسط والتباين وذلك لوجود تذبذب في بعض قيم السلسلة وبناءً عليه تم اخذ اللوغارتم الطبيعي وكذلك الفرق الاول لبيانات السلسلة الزمنية لاجل تحقق شرط الاستقرارية في المتوسط والتباين وكما موضح في الشكل رقم (١).

شكل رقم (١)

استقرارية السلسلة بالاعتماد على الانموذج الافضل



٣-٢-٢: اختيار الانموذج الملائم:- تم مقارنة نتائج عدة نماذج مختلفة موضحة بالجدول رقم (٢)

جدول رقم (٢) مقارنة النماذج

1	ARIMA	(0,1,1)
2	ARIMA	(2,1,1)
3	ARIMA	(0,2,2)
4	ARIMA	(2,1,2)
5	ARIMA	(1,1,1)

اذ تم حساب مقاييس المقارنة والمتمثلة بـ (RMSE) والذي يمثل جذر متوسط مربعات الخطأ واختبار (MAE) والذي يمثل متوسط مطلق الاخطاء واختبار (MAPE) الذي يمثل متوسط مطلق الاخطاء النسبي وكذلك (ME) و (MPE) واللذان يمثلان متوسط الاخطاء ومتوسط الاخطاء النسبي و (AIC) حيث يمثل معيار أكيكي (AKAIAE) ومعيار (SBIC) متمثل بمقياس شوارز وقد تم اعتماد الانموذج الذي يقابل اقل قيمة لمعيار (AIC) لاجل التنبؤ بعدد الزائرين في محافظة كربلاء المقدسة الاوقات لاحقة وكما موضح في الجدول رقم (٣) وقد تبين ان الانموذج الملائم هو ARIMA(0,1,1)

جدول رقم (٣) مقاييس المقارنة

Model	RMSE	MAE	MAPE	ME	MPE	AIC	SBIC
(A)	151259	96482.3	46.6491	18110.5	-22.196	23.8952	23.9341
(B)	153531	92087.3	40.9273	20023.1	-16.6649	24.0083	24.1253
(C)	158707	95197.7	44.0536	18330.2	-14.119	24.033	24.1109
(D)	152728	93336.3	41.2635	26876.4	-15.2846	24.0395	24.1954
(E)	160172	102216	43.2381	17334.3	-14.8099	24.0513	24.1293

٣-٢-٣: تقدير معلومات الانموذج

بعد ان تم تحديد الانموذج الملائم للتنبؤ بأعداد الزائرين تم استخراج تقديرات لمعاملات هذا الانموذج وكما يأتي:

$$Z_t = at + 0.852543at - 1$$

٣-٢-٤: كفاءة الانموذج

تم احتساب معاملات الارتباط الذاتي والارتباط الذاتي الجزئي والمتمثلة بالجدولين رقم (٤) ورقم (٥) والتي من الممكن اعتماد قيمها لاجل اتخاذ القرار للتنبؤ بأعداد الزائرين من خلال الانموذج ARIMA(0,1,1) حيث بينت انها غير معنوية مما يشير ان سلسلة البواقي هي سلسلة اخطاء عشوائية وكما موضح بالشكلين رقم (٢) و(٣).

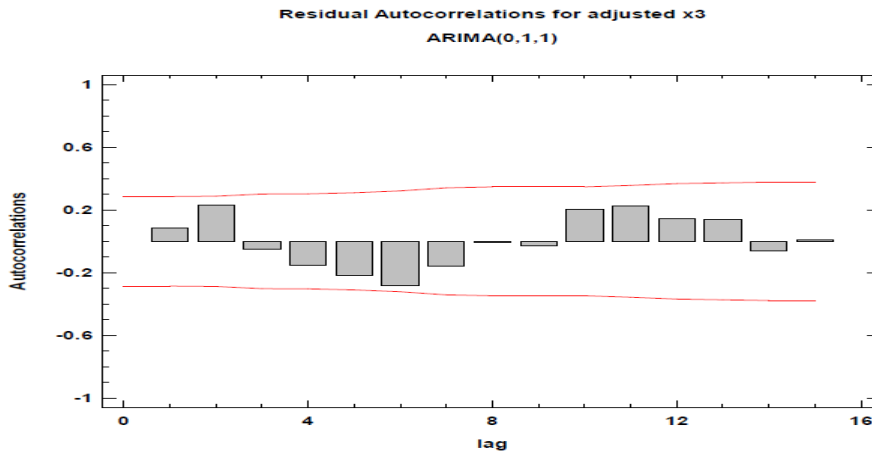
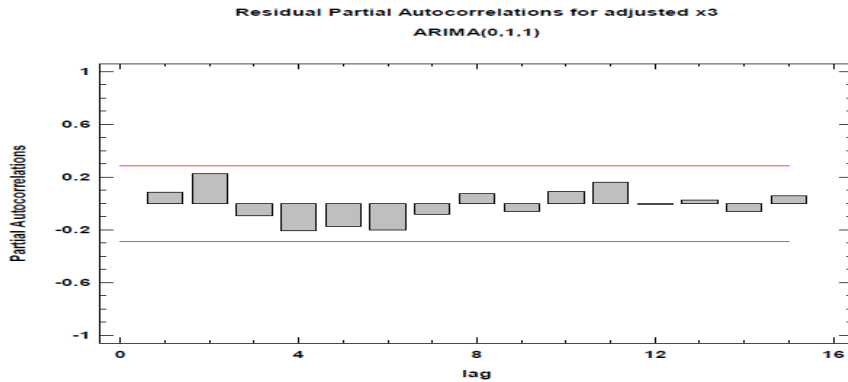
جدول رقم (٤) معاملات الارتباط الذاتي

Lag	Autocorrelation	Stnd. Error	Lower 95.0% Prop. Limit	Upper 95.0% Prop. Limit
1	0.0830517	0.145865	-0.285891	0.285891
2	0.229756	0.146868	-0.287856	0.287856
3	-0.0525082	0.154326	-0.302473	0.302473
4	-0.154883	0.154705	-0.303217	0.303217
5	-0.215248	0.15797	-0.309616	0.309616
6	-0.280153	0.164092	-0.321614	0.321614
7	-0.159805	0.173971	-0.340978	0.340978
8	-0.00890756	0.177067	-0.347045	0.347045
9	-0.0288432	0.177076	-0.347064	0.347064
10	0.202896	0.177176	-0.34726	0.34726
11	0.225585	0.182053	-0.356817	0.356817
12	0.144714	0.187906	-0.36829	0.36829
13	0.14039	0.190262	-0.372908	0.372908
14	-0.062572	0.192454	-0.377203	0.377203
15	0.00996316	0.192886	-0.378051	0.378051

جدول رقم (٥) معاملات الارتباط الذاتي الجزئي

			Lower 95.0%	Upper 95.0%
Lag	Autocorrelation	Stnd. Error	Prop. Limit	Prop. Limit
1	0.0830517	0.145865	-0.285891	0.285891
2	0.229756	0.145865	-0.285891	0.285891
3	-0.0911312	0.145865	-0.285891	0.285891
4	-0.209317	0.145865	-0.285891	0.285891
5	-0.17466	0.145865	-0.285891	0.285891
6	-0.200756	0.145865	-0.285891	0.285891
7	-0.0835714	0.145865	-0.285891	0.285891
8	0.0730763	0.145865	-0.285891	0.285891
9	-0.0634148	0.145865	-0.285891	0.285891
10	0.909679	0.145865	-0.285891	0.285891
11	0.158216	0.145865	-0.285891	0.285891
12	-0.00888763	0.145865	-0.285891	0.285891
13	0.0281238	0.145865	-0.285891	0.285891
14	-0.0631587	0.145865	-0.285891	0.285891
15	0.0594813	0.145865	-0.285891	0.285891

شكل رقم (٢) معاملات الارتباط الذاتي



شكل رقم (٣) معاملات الارتباط الذاتي الجزئي

٣-٢-٥: التنبؤ وحدود الثقة

بالاعتماد على الانموذج المقدر $ARIMA(0,1,1)$ تم ايجاد القيم التنبؤية بأعداد الزائرين لخمسة اشهر بالاضافة الى احتساب الحد الاعلى والحد الادنى لكل قيمة تنبؤية وكما موضح بالجدول رقم (٦) .

جدول رقم (٦) القيم التنبؤية وحدود الثقة لأعداد الزائرين

		Lower 95.0%	Upper 95.0%
Period	Forecast	Limit	Limit
1/54	192786.	64655.6	574839.
2/54	192786.	63896.3	581670.
3/54	192786.	63153.8	588509.
4/54	192786.	62427.5	595356.
5/54	192786.	61716.8	602212.

الاستنتاجات والتوصيات

٤-١ الاستنتاجات

- 1- تم رفض الفرضية الاولى حيث ظهر ان السلسلة الزمنية لأعداد الزائرين غير العراقيين لمدينة كربلاء المقدسة للمدة من ٢٠١٠-٢٠١٣ كانت غير مستقرة.
- 2- تم قبول الفرضية الثانية باستخدام طرق تنبؤ حديثة مثل منهجية بوكس جينكنز للتنبؤ بأعداد الزائرين غير العراقيين لمدينة كربلاء المقدسة للمدة من ٢٠١٠-٢٠١٣.
- 3- ان سلسلة اعداد الزائرين كانت غير مستقرة في التباين والمتوسط لذلك تم اعتماد التحويل اللوغاريتمي الطبيعي والفرق الاول لأجل تحقق شرط الاستقرار .
- 4- ان الانموذج الملائم للسلسلة هو $ARIMA(0,1,1)$.
- 5- فيما يخص اختبارات البواقي وجد انها عشوائية الحدوث مما يدل على كفاءة الانموذج المعتمد لغرض التنبؤ بأعداد الزائرين .
- 6- من خلال التنبؤ بأعداد الزائرين لفترة خمسة اشهر ظهرت القيم الفعلية للتنبؤ متساوية وذلك يدل على ان البيانات المعتمدة في هذا البحث تم الحصول عليها من سجلات فنادق محافظة كربلاء إن هذا العدد من الفنادق محدود .
- 7- تم اعتماد الحدود العليا للتنبؤ حيث نلاحظ زيادة في اعداد الزائرين

٤-٢ التوصيات

- 1:- اعتماد آلية علمية لتوثيق اعداد الزائرين الوافدين الى محافظة كربلاء من خلال قاعدة بيانات يتم تحديثها بعد كل زيارة. وذلك لأجل استخدامها في التنبؤ لتوفير الخدمات المناسبة لهم، ويقع هذا على عاتق الجهات المسؤولة في مجلس محافظة كربلاء.
- 2:- من خلال نتائج هذا البحث نوصي الجهات الاستثمارية في محافظة كربلاء بزيادة اعداد الفنادق وانشاء المدن السياحية في تلك المحافظة لاستيعاب اعداد الزائرين الوافدين وخاصة في اوقات الزيارة.

المصادر

- ١- البياتي، عصام حسين و المخلافي فؤاد عبدة إسماعيل(٢٠٠٧)" استخدام أسلوب بوكس- جينكنز للتنبؤ بإنتاجية العمل في مصنع اسمنت عمران في القطاع الصناعي اليمني " العراق، مجلة الادارة والاقتصاد، العدد ٣.
- ٢- الشاروط، محمد حبيب وصاحب، دعاء عبدالكريم(٢٠١٤)"التنبؤ باستهلاك الطاقة الكهربائية في محافظة القادسية باستخدام نماذج السلاسل الزمنية ونماذج الشبكات العصبية الاصطناعية". العراق، مجلة القادسية لعلوم الحاسبات والرياضيات، المجلد ٦، العدد ١.
- ٣- الشمراي، محمد موسى(٢٠١٣)"مقارنة بين بعض الاساليب الاحصائية التقليدية ونماذج بوكس- جينكنز في تحليل بيانات السلاسل الزمنية ".السعودية، مجلة جامعة ام القرى للعلوم التربوية والنفسية، المجلد ٥، العدد ١.
- ٤- الغنام، حمد عبدالله(٢٠٠٣)" تحليل السلسلة الزمنية لمؤشر أسعار الاسهم في المملكة العربية السعودية باستخدام منهجية بوكس- جينكنز (Box-Jenkins Method)". السعودية، مجلة الملك عبد العزيز: الاقتصاد والادارة، المجلد ١٧، العدد ٢.
- ٥- الطائي، فاضل عباس(٢٠٠٩)"التنبؤ والتمهيد للسلاسل الزمنية باستخدام التحويلات مع التطبيق ". العراق، المجلة العراقية للعلوم الاحصائية، المؤتمر الرابع، المجلد ١٠، العدد ١٧.