السلاسل الزمنية – مفهوم نماذج ARIMA

السلاسل الزمنية

يعد تحليل السلاسل الزمنية من الأساليب الإحصائية المهمة التي تستخدم في التنبؤ بقيم الظواهر العشوائية في المستقبل. وتتطلب فهماً عميقاً وإلماماً واعياً بالجوانب النظرية في أساليب تحليل السلاسل الزمنية . فالسلاسل الزمنية تمثل ظواهر معتمدة على الزمن، وقيمها المشاهدة تمثل قيم الظاهرة المعتمدة على الزمن، حيث يكون الزمن هو الظاهرة المستقلة. وقد جرت العادة أن تكون القيم المشاهدة للسلسلة الزمنية هي المردود للزمن في فترات زمنية متعاقبة ومتساوية بالرغم من أن ذلك لا يمنع أن تكون الفترات الزمنية غير متساوية (السعدي، ٢٠٠٤ ،ص٤٧٣.) إن دراسة السلاسل الزمنية وتحليلها يشكل أهمية كبيرة في الوقت الحاضر وذلك لأن هذه الدراسات تشير إلى التغيرات والعوامل التي تسببها وتمكن المخططين من وضع المعالجات المطلوبة

تعريف السلسلة الزمنية

كما ذكرنا سابقاً فإن السلسلة الزمنية هي عبارة عن مجموعة من المشاهدات التي تتولد على التوالي خلال الزمن. وتتميز أية سلسلة زمنية بأن بياناتها مرتبة بالنسبة للزمن، وأن المشاهدات المتتالية عادة ما تكون غير مستقلة، أي تعتمد على بعضها بعضا . ويمكن تعريفها أيضاً على أنها: "قيم أو مقادير هذه الظاهرة في سلسلة تواريخ متتابعة مثل أشهر أو أيام أو سنين، وفي العادة تكون الفترات بين التواريخ متتالية متساوية

ويمكن تعريفها أيضاً على أنها: "مجموعة من القياسات أو المشاهدات أو البيانات والمرتبة بحسب فترات زمنية متعددة. ويفضل لقراءة هذه السلاسل استخدام عدد مناسب وليس قليلاً من تلك الفترات، حيث أن التغيرات والتأثيرات يمكن أن تظهر وبشكل واضح لسلسلة زمنية بعدد من الفترات ولتكن ٥٣ فترة أفضل من العدد ١٥ فترة

والمقصود من تحليل السلسلة الزمنية هو معرفة التغيرات التي تطرأ على الظاهرة خلال مدة معينة، حيث يمكن مقارنة قيم الظاهرة بعضها ببعض لأنها مقيسة بنفس الوحدات وبنفس الطريقة في التواريخ المختلفة. ويتم رسم الخط البياني للسلسلة الزمنية حيث يوضح سير الظاهرة وتغيرها مع الزمن. والخط البياني للسلسلة هو عبارة عن نقطة تتحرك بمرور الزمن تماماً كجزء مادي يتحرك تحت تأثير قوى مادية . وبدلاً من القوى المادية فإن الحركة في السلسلة الزمنية تعزى إلى مجموعة من القوى الاقتصادية والنفسية والسياسية وغيرها. والخط البياني للسلسلة الزمنية يسمى المنحنى التاريخي للظاهرة.

أهداف تحليل السلاسل الزمنية

١- الحصول على وصف دقيق للملامح الخاصة للعملية التي تتولد منها السلسلة الزمنية .

٢- إنشاء نموذج لتفسير وشرح سلوك السلسلة بدلالة متغيرات أخرى يربط القيم المشاهدة ببعض قواعد سلوك السلسلة .

٣- استخدام النتائج التي نحصل عليها في رقم (١) أو رقم (٢) للتنبؤ بسلوك السلسلة في المستقبل وذلك اعتماداً على معلومات الماضي. من رقم (١) نفترض وجود قوة دافعة كافية في النظام تؤكد أن سلوك السلسلة في الماضي هو

نفس سلوكها في المستقبل . ومن رقم (٢) يكون لدينا تبصر أكبر بالقوى المؤثرة في عملية السلسلة الزمنية واستغلال ذلك في الحصول على تنبؤات أكثر دقة .

٤- التحكم في العملية التي تتولد منها السلسلة الزمنية بفحص ما يمكن حدوثه عند تغيير بعض معالم النموذج، أو بالتوصل إلى سياسات تستخدم فقط للتدخل عندما تنحرف عملية السلسلة عن الهدف المحدد بأكثر من مقدار معين.

٥- معرفة الماضي وتحديد نماذج التغيير الحالية للسلسلة الزمنية

٦- إعطاء فكرة عن النماذج المستقبلية، وهذه النماذج تستخدم من قبل الإدارة في التخطيط والضبط والتنبؤ .

بناءعلى ما سبق تم التفكير باتباع منهجية مترابطة متكاملة، لأنها تمثل حلقة الربط بين ما هو متحقق من تراكم معرفي نظري، وتطبيقي، وبين إمكانية تجسيد ذلك التراكم، واقعاً وحاضراً ومستقبلاً. ويعتمد تحديد مسارات المنهجية على ما يتيسر من ذلك التراكم الذي ينبغي أن يخضع للانتقاء والاختبار، بهدف التحقق من إمكانية استخدامها ضمن رؤى حالية ومستقبلية

الفروض التي يقوم عليها التنبؤ

يقوم التنبؤ على مجموعة من الفروض كما يلي

١- أن المستقبل لا يمكن التأكد منه تماماً ويبقى عدم التأكد هذا قائماً بغض النظر عن الطريقة التي استخدمت فيه إلى أن يمر الزمن ويمكن حينذاك رؤية الواقع الحقيقي .

٢- أن هناك نقاط اً غير واضحة في التنبؤ، فنحن على سبيل المثال لا نستطيع التنبؤ بمستجدات التكنولوجيا التي لا تتوفر لدينا معلومات تشير إليها الآن .

٣- أن التنبؤ يستخدم لوضع السياسات سواء كانت اجتماعية أو اقتصادية وأن هذه السياسات نفسها إذا ما نفذت ستؤثر على المستقبل وتجري عليه تغيرات لم يتكلم عنها التنبؤ نفسه، مما يحدث الافتراق بين ما جاء في التنبؤ وما سيتحقق على أرض الواقع.

لقد أظهرت التجارب عدم وجود طريقة تنبؤية معينة ملائمة لجميع الحالات، بل إن لكل حالة طريقة تنبؤ خاصة بها يتعين البحث عنها واستخدامها. إلا أن الأخذ بأكثر من طريقة قد يؤدي إلى رفع درجة دقة التقديرات المستقبلية، كما أن الشواهد تشير إلى أن إضافة الطرق الوصفية إلى الكمية قد يؤدي أيضاً إلى رفع درجة الدقة إذا أحسن اختيار الشروط التي يمكن بموجبها الجمع بين الطريقتين للحصول على أحسن توفيق fitting أمثل بينهما

مفهوم نماذج ARIMA

يقصد بنماذج ARIMA تلك المنهجية التي طبقها كل من George Box و Gwilyn Jenkins على السلاسل الزمنية عام 1970 ، وهذه المنهجية تعتمد في صياغتها على ثلاثة اجزاء هي كالاتي:

أ-نموذج الانحدار الذاتيAR

ب-نموذج المتوسط المتحركMA

ج- نموذج الانحدار الذاتي والمتوسطات المتحركة المتكاملة ARIMA

ويمكن صياغة النموذج بالشكل الاتي

تقدير معادلة الاتجاه العام للتنبؤ بالهجرة غير النظامية في المملكة العربية السعودية باستخدام نماذج ARIMA نماذج الانحدار الذاتي المتكاملة مع المتوسطات المتحركة.

$$\check{Y}_t = \mu + \boldsymbol{\varphi_1 Y_{t-1}} + \cdots + \cdots \boldsymbol{\varphi_p Y_{t-p}} - \boldsymbol{\theta_1 e_{t-1}} - \cdots \boldsymbol{\theta_q e_{t-q}}$$

حيث ان

$\check{Y}_t$ تمثل قيم المتغير Y المتنبأ بها

μ ثابت الذي يمثل المتوسط للمهاجرين غير منتظم خلال الفترة

φ_p معاملات معادلة الانحدار (نموذج الانحدار الذاتيAR)

θ_q تمثل معامل حد الخطأ أو معامل المتوسط التحرك (نموذج المتوسط المتحركMA)

e_{t-1} حد الخطأ لفترة ماقبل الفترة الحالية

استقرارية السلسلة الزمنية

لكى يمكن وصف السلسلة الزمنية محل الدراسة بالسكون لابد وأن يتسم كل من المتوسط والتباين بالثبات .ويقصد بثبات المتوسط :ألا تعبر السلسلة الزمنية عن اتجاه عام مع الزمن. وتعد طريقة الفروق هي أشهر الطرق المستخدمة في التخلص من أثر الاتجاه العام .أما ثبات التباين فيقصد به ألا يكون التباين متزايدًا أو متناقصًا مع الزمن، وتعتبر التحويلة اللوغاريتمية وتحويله الجذر التربيعي هي أكثر التحويلات استخدامًا لتثبيت التباين. وتستخدم دالة الارتباط الذاتي للكشف عن استقرارية او عدم استقرارية السلسة الزمنية.

ولتحويل السلاسل الزمنية غير مستقرة الي مستقرة نقوم بعملية التفاضل للسلسلة الأصلية من خلال طرح قيم YTمن YT- 1 و Y_{T-1} من Y_{T-2} وهكذا حتى نحصل على سلسلة زمنية جديدة ونختبر السلسلة الزمنية الجديدة فإذا أصبحت مستقرة تكون درجة التفاضل d=1 وإذا كانت السلسلة لازالت تعاني من عدم الاستقرارية تقوم بالتفاضل مرة ثانية وهكذا حتى نصل إلى سلسلة زمنية مستقرة.

- **تشخيص رتبةAR و MA**

يتم تحديد رتبة كل منAR و MA من خلال مشاهدة دالة الارتباط الذاتي ودالة الارتباط الذاتي الجزئي فاذا كانت دالة الارتباط الذاتي لا تتنازل بسرعة مع زيادة درجات الابطاء فهذا يعني ان السلسة الزمنية غير مستقرة وتحتاج الى التفاضل.

اختبارات دقة النتائج التنبؤية

هنالك اختبارات عديدة تعتمد في تقيم النماذج المتنبأ بها ومنها الأتي:

متوسط القيم المطلقة للأخطاءmean absolute error

$$MAE = \Sigma|e_T| / n$$
$$e_T = Y_T - F_T$$

حيث أن:

e_T :تمثل الخطأ أو البواقي

Y_T : تمثل القيم الحقيقة للمتغير $\sum e_t^2 / n$

F_T : تمثل القيم المتنبأ بها

مربع الأخطاء المطلقة mean squared error

وتحسب بالشكل الأتي

$$MSE= \sum e_t^2 / n$$

ويستخدم الاختبار الأول والثاني لمعرفة القوة التنبؤية للنموذج المستخدم

ج-النسبة المطلقة لمتوسط الأخطاء mean absolute percentage error

وتحسب بالشكل الأتي

$$MAPE = \Sigma(|e_T|/Y_T)/ n$$

وتستخدم هذه الصيغة للمقارنة بين عدة نماذج تنبؤية

تطبيق ما هو الاتجـاه العام للهجرة غير النظامية في المملكة العربية السعودية خلال الفترة 2002-2013؟

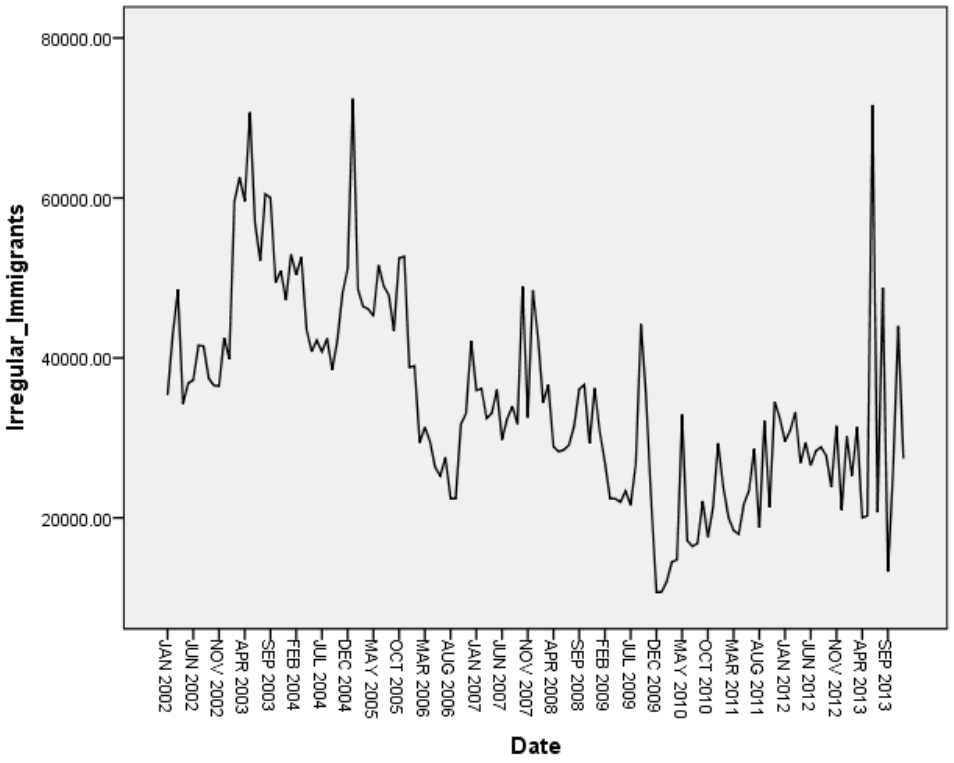
تم استخدام البيانات الشهرية (144 شهرا) في هذا التحليل وقد تم توفيره من قبل وزارة حرس الحدود.

Date	Observed irregular immigrants	Date	Observed irregular immigrants	Date	Observed irregular immigrants
Jan-02	35,369	Feb-06	29,398	Apr-10	14,759
Feb-02	42,781	Mar-06	31,311	May-10	32,920
Mar-02	48,542	Apr-06	29,568	Jun-10	17,137
Apr-02	34,254	May-06	26,321	Jul-10	16,443
May-02	**36,842**	Jun-06	25,259	Aug-10	16,817
Jun-02	37,233	Jul-06	27,574	Sep-10	22,094
Jul-02	41.57	Aug-06	22,407	Oct-10	17,561
Aug-02	41.46	Sep-06	22,396	Nov-10	21,320

Sep-02	37,418	Oct-06	31,702	Dec-10	29,313
Oct-02	36,560	Nov-06	33,065	Jan-11	23,834
Nov-02	36,480	Dec-06	42,120	Feb-11	20,076
Dec-02	42,494	Jan-07	35,949	Mar-11	18,414
Jan-03	39,863	Feb-07	36,155	Apr-11	17,960
Feb-03	59,600	Mar-07	32,453	May-11	21,726
Mar-03	62,561	Apr-07	33,075	Jun-11	23,425
Apr-03	59,585	May-07	36,014	Jul-11	28,614
May-03	70,706	Jun-07	29,774	Aug-11	18,754
Jun-03	56,877	Jul-07	32,408	Sep-11	32,131
Jul-03	52,178	Aug-07	33,944	Oct-11	21,356
Aug-03	60,490	Sep-07	31,757	Nov-11	34,515
Sep-03	60,034	Oct-07	48,937	Dec-11	32,413
Oct-03	49,433	Nov-07	32,523	Jan-12	29,563
Nov-03	50,912	Dec-07	48,449	Feb-12	30,931
Dec-03	47,217	Jan-08	42,425	Mar-12	33,212
Jan-04	52,946	Feb-08	34,415	Apr-12	26,871
Feb-04	50,392	Mar-08	36,643	May-12	29,379
Mar-04	52,611	Apr-08	28,890	Jun-12	26,546
Apr-04	43,596	May-08	28,280	Jul-12	28,375
May-04	40,806	Jun-08	28,494	Aug-12	28,860
Jun-04	42,198	Jul-08	29,097	Sep-12	27,760
Jul-04	40,761	Aug-08	31,464	Oct-12	23,871
Aug-04	42,424	Sep-08	36,104	Nov-12	31,497
Sep-04	38,527	Oct-08	36,646	Dec-12	20,982
Oct-04	42,264	Nov-08	29,310	Jan-13	30,178
Nov-04	48,101	Dec-08	36,200	Feb-13	25,285
Dec-04	51,135	Jan-09	30,761	Mar-13	31,341
Jan-05	72,425	Feb-09	26,915	Apr-13	20,019
Feb-05	48,586	Mar-09	22,407	May-13	20,265
Mar-05	46,434	Apr-09	22,407	Jun-13	71,623
Apr-05	46,121	May-09	21,962	Jul-13	20,694
May-05	45,228	Jun-09	23,368	Aug-13	48,763
Jun-05	51,588	Jul-09	21,607	Sep-13	13,277
Jul-05	48,987	Aug-09	26,812	Oct-13	25,900
Aug-05	47,824	Sep-09	44,204	Nov-13	43,971
Sep-05	43,390	Oct-09	35,016	Dec-13	27,420
Oct-05	52,491	Nov-09	22,099	Jan-14	
Nov-05	52,689	Dec-09	10,694	Feb-14	
Dec-05	38,817	Jan-10	10,720	Mar-14	
Jan-06	39,002	Feb-10	12,045	Apr-14	
		Mar-10	14,502	May-14	
				Jun-14	

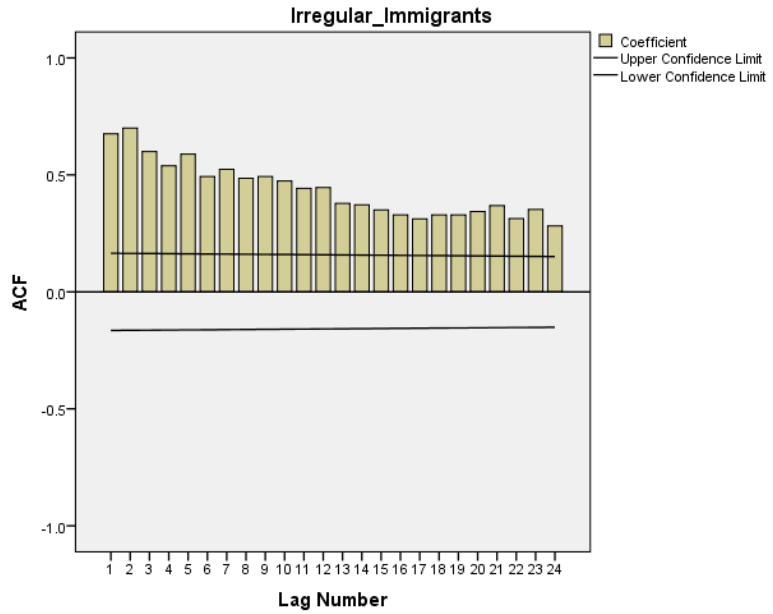
				Jul-14	
				Aug-14	
				Sep-14	
				Oct-14	
				Nov-14	
				Dec-14	

الشكل العام للبيانات الأصلية

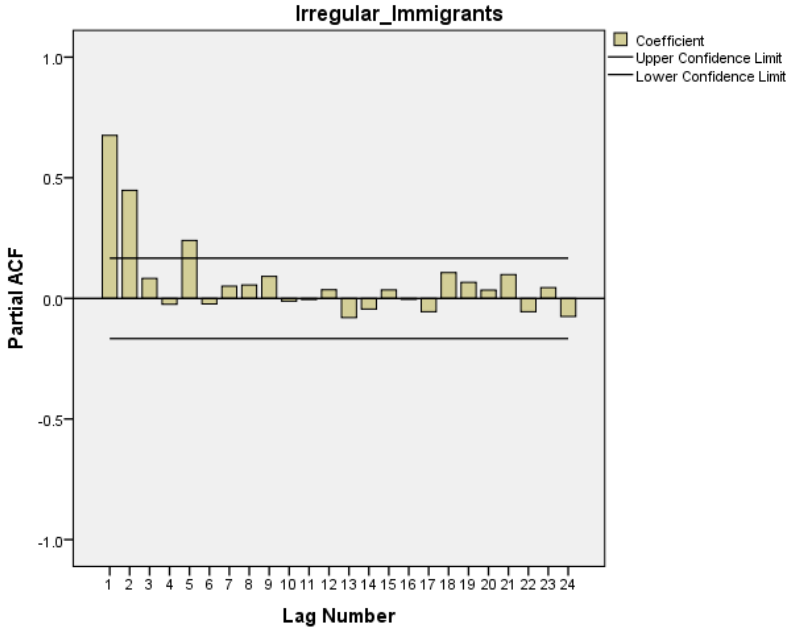


يلاحظ من الرسم وجود اتجاه عام نحو الانخفاض، كما يوجد نقاط عالية ونقاط منخفضة ويوضح ذلك الرسم التالي:

معاملات الارتباط الذاتي



معاملات الارتباط الذاتي الجزئي

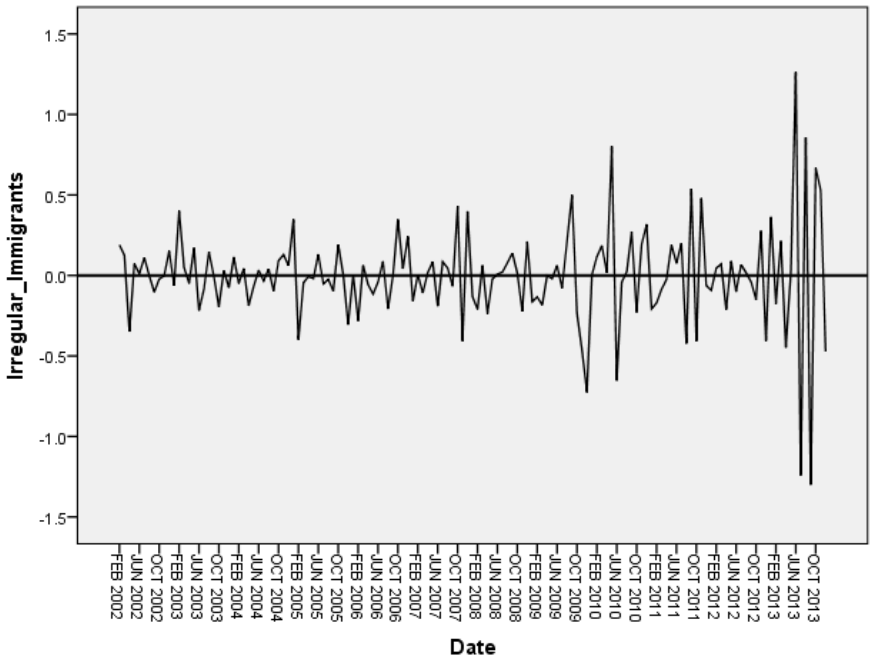


اختيار نموذج السلاسل الزمنية للتنبؤ الهجرة غير النظامية

كما ألمحنا في وقت سابق، تعتمد الدراسة الحالية على طريقة بوكس جنكنز (ARIMA) لتقدير نموذج أو التنبؤ بالهجرة غير النظامية في المملكة العربية السعودية. في محاولة لاختيار النموذج الأنسب، ومن أجل الحصول على نتائج قوية يجب اختبار خصائص بيانات السلاسل الزمنية وكما يتضح من الرسم البياني السابق أن

السلسلة الزمنية غير ثابتة، ولذلك تم استخدام تحويلة اللوغاريتم بجانب أخذ الفروق الأولى، وكان الناتج كما بالرسم

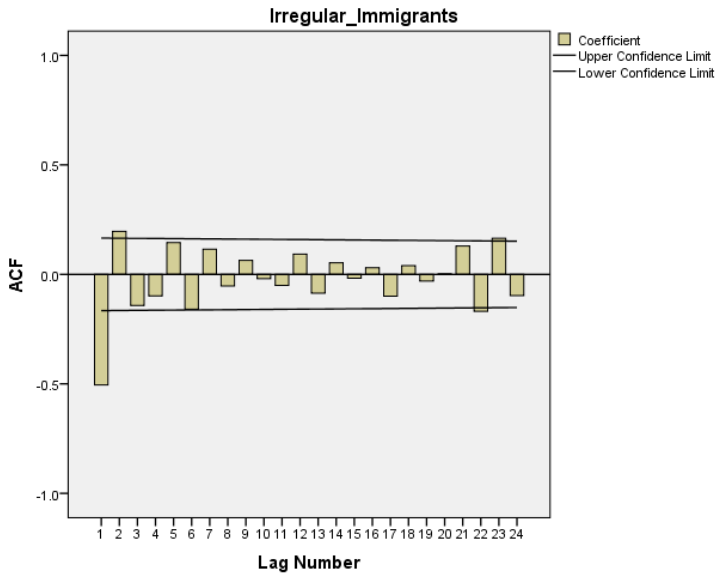
Taking the logarithm and first difference



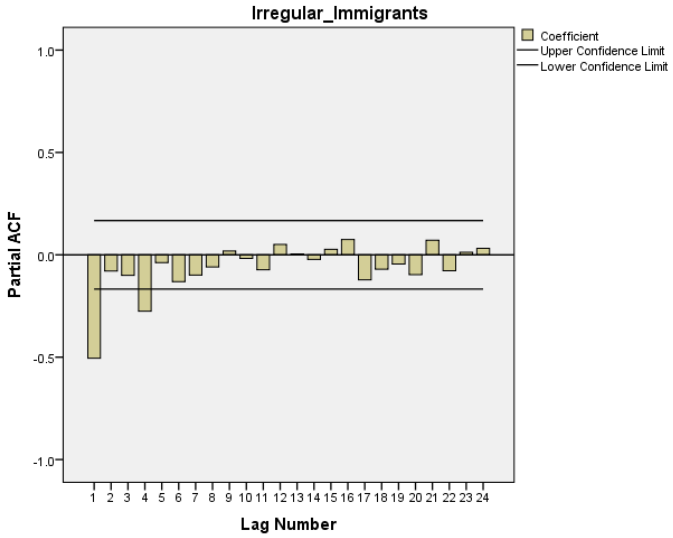
Transforms: natural logarithm, difference(1)

بعد تحويل البيانات نلاحظ هناك تحسن كبير في السكون. ومع ذلك، هناك زيادة التباين في نهاية هذه السلسلة والتي ظهرت في نهاية المنحنى. وهذا يشير إلى أنه الفروق والتحويل اللوغاريتمي ساعدت في جعل سلسلة أكثر ثابتا، ويظهر ذلك الارتباط الذاتي والارتباط الذاتي الجزئي في الشكلين التاليين.

Autocorrelation of the series with the lags of transformed data (log and differencing)



Partial autocorrelation of the series with the lags of transformed data (log and differencing)



بعد ذلك تم التحقيق موسمية السلسلة وكذلك اختبار طبيعية البيانات

. **– تشخيص رتبة** ARIMA

يتم تحديد رتبة كل من MA و AR من خلال مشاهدة دالة الارتباط الذاتي ودالة الارتباط الذاتي الجزئي السابقة فاذا كانت دالة الارتباط الذاتي لاتتنازل بسرعة مع زيادة درجات الابطاء فهذا يعني ان السلسة الزمنية غير مستقرة وتحتاج الى التفاضل. اما تحديد رتبةMA و AR فيمكن من خلال

دالة الارتباط الذاتي ودالة الارتباط الذاتي الجزئي السابقتين تحديد العدد اللازم من AR و MA ولذلك تقترح دالة الارتباط الذاتي الجزئي فترتين من AR في حين تشير دالة الارتباط الذاتي إضافة أربع قطع من MA فيكون النموذج المقترح ARIMA (2,1,4)

وتكون الدالة هي:

$$Ln\Delta\check{Y}_t = \mu + \varphi_1 \Delta Y_{t-1} + \varphi_2 \Delta Y_{t-2} - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \theta_3 e_{t-3} - \theta_4 e_{t-4}$$

حيث $Ln\Delta\check{Y}_t$ هي لوغاريتم الفروق الأولى لسلسلة المهاجرين

$\Delta\check{Y}_t$ الفروق الأولى للمتغير لسلسلة المهاجرين $Y_t - Y_{t-1}$

μ المتوسط لسلسلة المهاجرين

$\varphi_1 \; and \; \varphi_2$ are the coefficients of lags 1 and 2 (Autoregressive part of the model or the AR) respectively

$\Delta Y_{t-1} \; and \; \Delta Y_{t-2}$ are the first the first and second lags of the differenced outcome variable respectively

$\theta_1, \theta_2, \theta_3, and \; \theta_4$ are the coefficients of the first, second, third, and fourth lag of the error terms (Moving averages part of the model) respectively

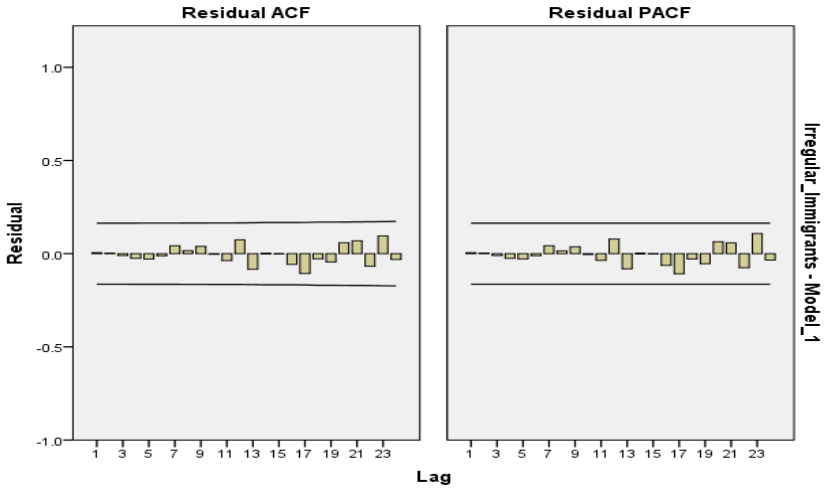
$e_{t-1}, e_{t-2}, e_{t-3}, and \; e_{t-4}$ are the first, second, third, and fourth lags of the error terms (moving averages or MA).

القوة التنبؤية للنموذج

Model Statistics

Model				Ljung-Box Q(18)			Number of Outliers
	R-squared	MAPE	Normalized BIC	Statistics	DF	Sig.	
Irregular_Immigrants-Model_1	.592	18.574	18.272	5.546	12	.937	0

- تشير الإحصاءات تناسب النموذج المقدر حيث أن النموذج الحالي لديه قوة التنبؤ 59.2٪ كما يلاحظ من R^2. وهذا يعني أن 59.2٪ من الاختلافات في سلسلة الزمنية الحالية يمكن التنبؤ من خلال النموذج الحالي.
- كما يلاحظ أن قيمة MAPE=18.5 يمكن اعتبارها صغيرة.
- كما تشير إحصائية ليونغ بوكس أنه لا يوجد فروق ذات دلالة إحصائية بين النموذج الأصلي والمقدر عند مستوى معنوية 5%



Autocorrelation of the final model

من الرسم يتضح أن السلسلة ساكنة

معاملات النموذج

ARIMA Model Parameters							
				Estimate	SE	t	Sig.
Irregular_Immigrants-Model_1	Natural Logarithm	AR	Lag 1	-1.192	.239	-4.994	.000
			Lag 2	-.450	.221	-2.039	.043
		Difference		1			
		MA	Lag 1	-.582	.232	-2.502	.014
			Lag 2	.118	.171	.689	.492
			Lag 3	.386	.144	2.680	.008
			Lag 4	.326	.096	3.409	.001

من الجدول السابق نلاحظ أن أن النموذج النهائي ARIMA (2,1,4) بدون الثابت حيث انه كان غير معنوي. كما يلاحظ أن جميع معلمات AR وMA معنوية حيث أن

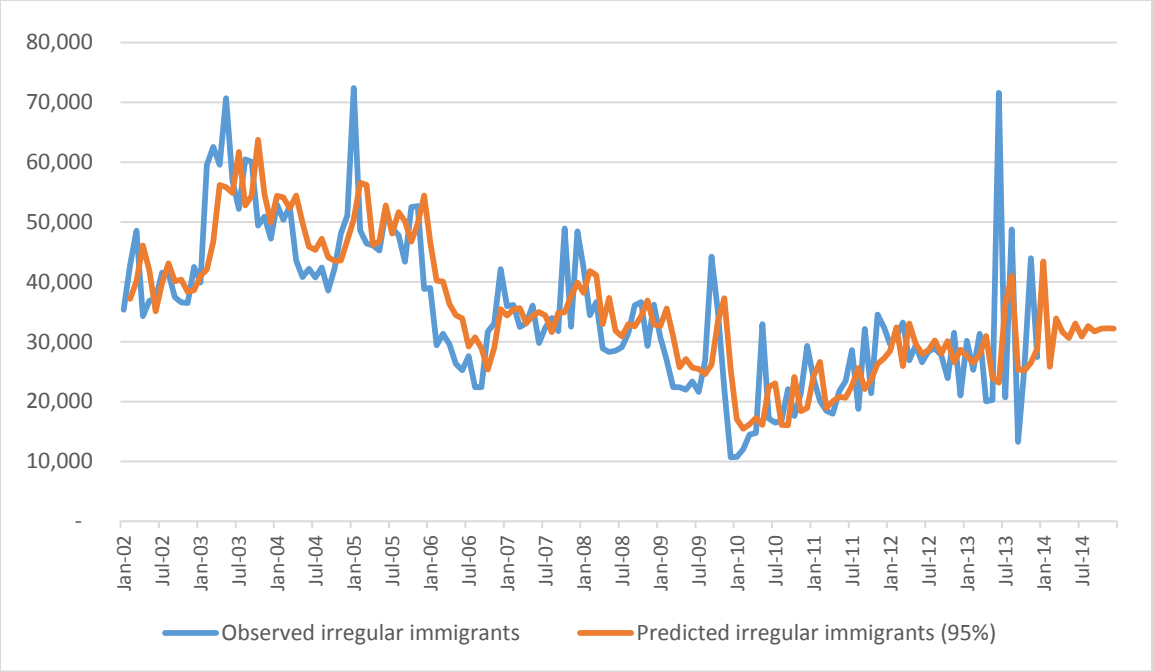
القيمة الاحتمالية أقل من 5% ماعدا Lag2 فهي غير معنوية حيث أن القيمة الاحتمالية أكبر من 5% وبالتالي يمكن التنبؤ بالهجرة غير النظامية عن طريق المعادلة التالية:

$$\mathbf{Ln\Delta\breve{Y}_t = -1.92\Delta Y_{t-1} - 0.45\Delta Y_{t-2} + 0.582e_{t-1} - 0.86e_{t-3} - 0.326e_{t-4}}$$

جزء القيم المشاهدة والمتنبأ بها باستخدام المعادلة السابقة

Part of the Observed and predicted values

Date	Observed irregular immigrants	Lower confidence interval (95%)	Predicted irregular immigrants	Upper confidence interval (95%)
Jan-13	30,178	16,398	27,597	43,678
Feb-13	25,285	15,763	26,528	41,987
Mar-13	31,341	16,549	27,851	44,080
Apr-13	20,019	18,404	30,973	49,022
May-13	20,265	14,277	24,026	38,027
Jun-13	71,623	13,749	23,138	36,621
Jul-13	20,694	21,552	36,269	57,405
Aug-13	48,763	24,373	41,017	64,920
Sep-13	13,277	15,058	25,342	40,110
Oct-13	25,900	14,962	25,179	39,852
Nov-13	43,971	15,712	26,442	41,851
Dec-13	27,420	17,116	28,804	45,590
Jan-14		25,799	43,417	68,717
Feb-14		14,736	25,823	42,165
Mar-14		17,965	33,908	58,537
Apr-14		16,519	31,622	55,161
May-14		15,845	30,637	53,835
Jun-14		16,531	33,065	59,545
Jul-14		15,301	30,849	55,871
Aug-14		15,765	32,611	60,151
Sep-14		15,127	31,706	59,026
Oct-14		15,086	32,191	60,687
Nov-14		14,857	32,227	61,453
Dec-14		14,608	32,175	62,003



الرسم البياني يمثل السلسلة الأصلية والسلسلة المتنبأ بها