

استخدام السلاسل الزمنية في التنبؤ - دراسة تطبيقية على طلاب التعليم الثانوي في سورية

الدكتور طالب أحمد *

كمال نجار **

الملخص

هدف البحث إلى استخدام أسلوب السلاسل الزمنية متمثلاً بنماذج ARIMA من أجل بناء نموذج للتنبؤ بأعداد طلاب المرحلة الثانوية في سورية المتوقع انتسابهم في كل عام دراسي. بهدف الاستفادة منه في السياسات التعليمية لوزارة التربية.

وكانت أهم نتائج الدراسة: تم بناء نموذج قياسي هو $ARIMA(2,1,2)$ للتنبؤ بعدد طلاب المرحلة الثانوية في سورية حتى عام 2030، وهو أفضل النماذج المقترحة للحصول على تنبؤات دقيقة. وتبين من خلال هذا النموذج أن عدد طلاب المرحلة الثانوية المقدر سيزداد خلال الأعوام القادمة.

كلمات مفتاحية

نماذج ARIMA، طلاب المرحلة الثانوية، التنبؤ .

* أستاذ مساعد - قسم الإحصاء والبرمجة - كلية الاقتصاد - جامعة تشرين - اللاذقية - سورية.

** طالب دراسات عليا (دكتوراه) - قسم الإحصاء والبرمجة - اختصاص إحصاء وبرمجة - كلية الاقتصاد - جامعة تشرين - اللاذقية - سورية.

Using time series in forecasting – applicative study on the students secondary education

Abstract

The aim of this study was to use time series method through ARIMA models to build a model to forecast the number of secondary students in Syria every year, It also aims to benefit in education politics.

The main results of the study: It was build ARIMA (2,1,2) model, which can be used to forecast the number of secondary students in Syria until 2030, it was the best model to get on accurate predictions. Through of this model, the estimated number of secondary students will increase through next years.

إن مجانية التعليم والزاميته التي يضمنها القانون في الجمهورية العربية السورية، ومعدلات النمو السكاني المرتفعة نسبياً وفتوة المجتمع، كل هذه العوامل مجتمعة فضلاً عن عوامل أخرى تؤدي إلى ظهور مشكلة حقيقية في أعداد المنتسبين سنوياً إلى المرحلة الثانوية في سورية، وما يرتبط بهذه الأعداد من إلتزامات يجب التخطيط لها ومواجهتها من خلال التنبؤ بأعداد الطلاب. يمكن أن تتم عملية التنبؤ بأعداد الطلاب بأساليب مختلفة: التحليل التقليدي للسلسلة الزمنية أو النماذج السببية، أو التحليل الحديث للسلسلة الزمنية متمثلاً بنماذج ARIMA، وهذا الأسلوب لا يفترض وجود أي نمط معين للبيانات التاريخية للسلسلة التي نتنبأ بها، حيث اختيار النموذج المناسب يتم بمقارنة توزيعات معاملات الارتباط الذاتي للسلسلة الزمنية بالتوزيعات النظرية للنماذج المختلفة، ويكون النموذج المختار جيداً إذا كانت الفروق (البواقي) بين القيم المقدرة والبيانات التاريخية صغيرة، وتتوزع بشكل طبيعي، ومستقلة عن بعضها. حيث سيتم بناء نموذج قياسي باستخدام نماذج ARIMA للتنبؤ بعدد طلاب المرحلة الثانوية في سورية حتى عام 2030.

الدراسات السابقة

1-دراسة بعنوان: دراسة تحليلية كمية لأساليب التنبؤ بأعداد الطلبة في المدارس الفلسطينية. بحث منشور، إعداد: البول، رجا، مجلة جامعة النجاح للأبحاث العلوم الانسانية، المجلد 32، العدد الأول، 2018.

هدفت هذه الدراسة إلى:

هدفت هذه الدراسة إلى مناقشة عدة أساليب تنبؤ رياضية معاصرة تستخدم للتنبؤ بأعداد السكان، بغية معرفة أي أسلوب من هذه الأساليب أكثر قدرة على التنبؤ بأعداد الطلبة الذين سيلتحقون بالتعليم في مراحلهم المختلفة مستقبلاً في فلسطين، وذلك باستخدام بيانات تاريخية للفترة 1994-2009 واستخدامها للتنبؤ بأعداد الطلبة للسنوات 2010 - 2016

وكانت أهم نتائج هذه الدراسة:

- عدم استخدام نفس الأسلوب للتنبؤ بأعداد الطلبة لفترات زمنية طويلة تزيد عن عقد.
- إجراء دراسة مقارنة بين عدة أساليب لكل عقد وذلك لأنه لا يمكن ضمان عدم تغير اتجاهات النمو في المجتمع.

- عمل دراسة مشابهة لمعرفة أكثر الأساليب ملائمة للتنبؤ بأعداد الطلبة لكل فوج على حدى إذ أن الأسلوب الذي يلائم التنبؤ بالعدد الكلي قد لا يناسب التنبؤ بالأفواج. [1]

2- دراسة بعنوان: استخدام منهجية بوكس- جينكنز لبناء نموذج قياسي للتنبؤ بعدد المواطنين السوريين. بحث منشور، إعداد: أحمد، طالب، مجلة جامعة تشرين للعلوم الاقتصادية والقانونية، المجلد 40 ، العدد السادس، 2018

هدفت هذه الدراسة إلى:

استعراض طرائق التنبؤ بالسلاسل الزمنية من خلال نماذج بوكس-جينكنز، واختبار امكانية تطبيق أسلوب ARIMA في التنبؤ بعدد المواطنين السوريين، بالاضافة لتحديد النموذج الأمثل بين نماذج ARIMA للتنبؤ بعدد المواطنين السوريين، كما هدفت الدراسة إلى التنبؤ بعدد المواطنين السوريين حتى عام 2035.

وكانت أهم نتائج هذه الدراسة:

يشكل عدد السكان في سورية خلال الفترة المدروسة سلسلة زمنية غير مستقرة، لأن اختبارات معاملات دالة الارتباط الذاتي أظهرت ذلك نتيجة التزايد الكبير في عدد السكان، كما توصل الباحث إلى أن النموذج الأفضل من بين النماذج المقترحة للتنبؤ بعدد المواطنين السوريين هو نموذج $ARIMA(2,1,1)$ ، كما توصل الباحث إلى توقع ازدياد عدد سكان سورية خلال العدين القادمين، حيث سيصل عدد المواطنين السوريين في عام 2035 إلى 28394938 نسمة، وذلك من خلال النموذج المقترح. [2]

3-دراسة بعنوان: منهجية في box-Jenkins في تحليل السلاسل الزمنية والتنبؤ، دراسة تطبيقية على أعداد تلاميذ الصف الأول من التعليم الأساسي في سورية. إعداد: نقار، عثمان و العواد ، منذر، مجلة جامعة دمشق للعلوم الاقتصادية والقانونية، المجلد 27 ، العدد الثالث، 2011 هدفت هذه الدراسة إلى:

- اختبار إمكانية تطبيق الأسلوب الحديث في تحليل السلاسل الزمنية في التنبؤ بأعداد التلاميذ.

- وضع نموذج قياسي للتنبؤ بأعداد تلاميذ الصف الأول من التعليم الأساسي

- التنبؤ بأعداد التلاميذ المتوقع انتسابهم إلى الصف الأول من التعليم الأساسي.

وكانت أهم نتائج هذه الدراسة:

- تشكيل سلسلة أعداد المنتسبين إلى الصف الأول من التعليم الاساسي سياقاً عشوائياً غير مستقر، وأظهر اختبار Dickey Fuller وجود جذر الوحدة، وقد أخذ مرشح الفروق الاولى لجعلها مستقرة.

- أن النموذج الأفضل بين النماذج التي وضعت في هذا البحث للتنبؤ بأعداد المنتسبين إلى الصف الاول الاساسي هو النموذج $ARIMA(0,1,1)$ [3]

4- دراسة بعنوان: استخدام أحد نماذج بوكس - جينكيز للتنبؤ بأعداد الطالبات في المرحلة الأساسية في محافظة أربيل. بحث منشور، إعداد: حنان عطروش، مجلة العلوم الإدارية - المجلد 11 - العدد 5 - 2015.

هدفت هذه الدراسة إلى:

- تطبيق الأسلوب الإحصائي الحديث للسلاسل الزمنية باستخدام أحد نماذج بوكس -جينكيز للتنبؤ بأعداد الطالبات في التعليم الأساسي في محافظة أربيل ، لوضع الخطط والبرامج التي تستهدف تطوير التعليم الأساسي للفتاة بشكل عام والفتاة الريفية بشكل خاص وتشجيع التعليم للفتاة.

وكانت أهم نتائج هذه الدراسة:

- تم التوصل إلى نموذج $ARIMA(1,1,1)$ وهو النموذج المناسب لتقدير بيانات السلسلة الزمنية، ومن

خلال هذه النموذج تم التنبؤ بأعداد الطالبات في المرحلة الأساسية في محافظة أربيل إلى عام 2027 [4].

مشكلة البحث

تكمن مشكلة البحث في إيجاد أسلوب إحصائي دقيق يمكن من خلاله معرفة سلوك ظاهرة تزايد أو تناقص عدد طلاب المرحلة الثانوية في سورية، من أجل الوصول إلى نتائج يمكن الاسترشاد بها في وضع السياسة التعليمية الصحيحة من قبل وزارة التربية. كما أن هناك صعوبة في أن تكون السلسلة

الزمنية لعدد طلبة المرحلة الثانوية في سورية مستقرة خلال الفترة المدروسة، ويمكن تمثيل مشكلة البحث بالتساؤل التالي:

ما هو النموذج الاحصائي الأمثل من بين نماذج ARIMA يمكنه التنبؤ بعدد طلاب المرحلة الثانوية في سورية ؟

أهمية البحث وأهدافه:

تكمن أهمية البحث في استخدام أسلوب السلاسل الزمنية من أجل تقدير نموذج يمثل ظاهرة تسجيل عدد الطلاب للمرحلة الثانوية بهدف الاستفادة منه في السياسات التعليمية لوزارة التربية. كما تكمن الأهمية باستنتاج نموذج قياسي يستخدم للتنبؤ بأعداد طلاب المرحلة الثانوية في سورية المتوقع انتسابهم في كل عام دراسي.

ويهدف البحث إلى:

- استعراض طرائق التنبؤ بالسلاسل الزمنية من خلال نماذج ARIMA.
- وضع نموذج قياسي للتنبؤ بأعداد طلاب المرحلة الثانوية في سورية.
- التنبؤ بأعداد الطلاب المتوقع انتسابهم إلى صفوف المرحلة الثانوية في سورية حتى عام 2030.

فرضيات البحث

الفرضية الرئيسية الاولى: لا يمكن التوصل إلى النموذج الذي يجعل السلسلة الزمنية مستقرة باستخدام نماذج ARIMA

الفرضية الرئيسية الثانية: لا يمكن بناء نموذج قياسي للتنبؤ بعدد طلاب المرحلة الثانوية في سورية باستخدام نماذج ARIMA

منهجية البحث

سيتم اتباع المنهج الوصفي التحليلي الذي يعتمد على جمع البيانات وتحليلها واستخلاص النتائج، مترافقاً مع الاعتماد على أسلوب ARIMA في تحليل السلاسل الزمنية، كما سنقوم بتحليل البيانات باستخدام برنامج SPSS 25، وتم الحصول على بيانات السلسلة الزمنية من المكتب المركزي للإحصاء في سورية .

مكان وزمان البحث

الجمهورية العربية السورية، (1986-2018).

مجتمع البحث:

الإطار النظري للدراسة:

أولاً: النماذج الإحصائية للسلاسل الزمنية

يعتبر أسلوب تحليل السلاسل الزمنية من الأساليب الإحصائية الهامة في التنبؤ، وقد تم استخدام هذا الأسلوب على نطاق واسع في الكثير من التطبيقات الإحصائية والاقتصادية، حيث يتم التنبؤ بالتغيرات المستقبلية للمتغير بالاعتماد فقط على سلوك هذا المتغير في الماضي. أو بعبارة أخرى فإن نموذج السلاسل الزمنية يأخذ في الاعتبار أنماط التغيرات في الماضي لمتغير معين و يستخدم هذه المعلومات للتنبؤ بالتغيرات المستقبلية لذلك المتغير، مما يجعل نموذج السلاسل الزمنية طريقة متطورة ووسيلة فعالة في التنبؤ. تركز هذه النماذج الإحصائية على الجانب العشوائي في السلسلة الزمنية، وتنقسم الى:

- نماذج الانحدار الذاتي (AR) : حيث تكتب القيمة الحالية كدالة خطية في القيم السابقة لنفس المتغير، حيث نجد نماذج الانحدار الذاتي للسلاسل الزمنية من الدرجة الأولى، وهي أبسط نموذج يرمز له بالرمز $AR(1)$ ويكتب من الشكل التالي:

$$Y_t = \theta Y_{t-1} + u_t$$

والافتراض خلف نموذج الانحدار الذاتي من الدرجة الأولى أن سلوك السلسلة الزمنية Y_t يحدد غالباً من قبل قيمها للفترة الزمنية السابقة. أي ما يحدث في الفترة T يعتمد على ما يحدث في الفترة $t-1$. وكذلك ما سوف يحدث في الفترة $T+1$ سوف يتحدد بسلوك السلسلة الزمنية في الفترة الحالية. ولتعميم نموذج الانحدار من الدرجة الأولى $AR(1)$ نستخدم $AR(p)$ حيث يمثل الرقم داخل القوس درجة عملية الانحدار الذاتي [5]، فعلى سبيل المثال $AR(2)$ سيكون من الدرجة الثانية:

$$Y_t = \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + u_t$$

وكذلك $AR(p)$ سيكون انحدار ذاتي من الدرجة P كما يلي:

$$Y_t = \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} + u_t$$

أو باستخدام رمز الجمع:

$$Y_t = \sum_{i=1}^p \theta_i Y_{t-i} + u_t$$

- نماذج المتوسطات المتحركة (MA) : حيث تكتب القيمة للمتغير كدالة خطية في القيم الحالية لعنصر الخطأ العشوائي وعدد من قيمه السابقة، ونموذج المتوسط المتحرك في أبسط أشكاله هو من الدرجة الأولى وهو يأخذ الشكل التالي:

$$Y_t = u_t + \theta u_{t-1}$$

أما نموذج المتوسطات المتحركة من الدرجة (q) يكتب من الشكل التالي :

$$Y_t = u_t + \vartheta_1 u_{t-1} + \vartheta_2 u_{t-2} + \dots \dots + \vartheta_q u_{t-q}$$

$$Y_t = u_t + \sum_{j=1}^q \vartheta_j u_{t-j}$$

- نماذج الانحدار الذاتي والمتوسطات المتحركة (ARMA): حيث تتميز هذه النماذج بجمع نماذج الانحدار الذاتي ونماذج المتوسط المتحرك لتتحصل على سلسلة زمنية جديدة تسمى ARMA(p,q) وتكتب من الشكل التالي :

$$Y_t = \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots \dots \theta_p Y_{t-p} + u_t + \vartheta_1 u_{t-1} + \vartheta_2 u_{t-2} + \dots \dots + \vartheta_q u_{t-q}$$

وتكتب باستخدام صيغة الجمع على النحو التالي :

$$Y_t = \sum_{i=1}^p \theta_i Y_{t-i} + u_t + \sum_{j=1}^q \vartheta_j u_{t-j}$$

- نماذج الانحدار الذاتي والمتوسطات المتحركة التكاملية (ARIMA)

تعد نماذج ARIMA أكثر نماذج السلاسل الزمنية استخداماً إذ أنه بالإمكان اشتقاق جميع النماذج منها سواء الانحدار الذاتي أو المتوسطات المتحركة أو المختلطة، وتتكون هذه النماذج من ثلاثة أجزاء، يمثل الجزء الأول منها نموذج الانحدار الذاتي $AR(p)$ الذي يستخدم عادة في عمليات التنبؤ للسلسلة الزمنية، أما الجزء الثاني فيمثل نموذج الأوساط المتحركة $MA(q)$ ، ويمثل الجزء الثالث (d) الفروق التي تتطلبها السلسلة لتكون مستقرة. وعندما تكون السلسلة الزمنية غير مستقرة يجب أولاً تحويلها إلى سلسلة زمنية مستقرة قبل بناء النموذج الرياضي وذلك بأخذ الفروق (d)، واستخدام أحد التحويلات وعدد الفروق المطلوب لتحويل السلسلة إلى سلسلة مستقرة تسمى بدرجة التكامل Integrated، حيث يتحول نموذج الانحدار الذاتي والمتوسطات المتحركة $ARMA(p,q)$ إلى نموذج الانحدار الذاتي المتكامل $ARIMA(p,d,q)$ [6]

ثانياً - مراحل تطبيق نماذج ARIMA : من أجل تطبيق هذه النماذج يجب اتباع المراحل التالية :

- دراسة استقرارية السلسلة الزمنية .

- تحديد النموذج (AR) ، (MA) ، (ARMA) . أو ARIMA

- تقدير معالم النموذج المحدد .

- دراسة صلاحية النموذج المقدر .

- التنبؤ .

1- استقرارية السلسلة الزمنية: تكون السلسلة العشوائية مستقرة، إذا تذبذبت حول وسط حسابي ثابت، مع تباين ليس له علاقة مع الزمن، وتتمثل أسباب عدم الاستقرار في مركبة الاتجاه العام والفصلية و للتخلص من مشكلة عدم الاستقرارية يجب أولاً معرفة مسبباته، ثم محاولة إزالتها بإحدى الطرق السالفة الذكر. ومن أهم المراحل الخاصة بدراسة الإستقرارية للسلسلة الزمنية نجد :

- التحليل البياني : من خلال الرسم البياني تكون لدينا فكرة حول استقرارية السلسلة الزمنية من عدمها أو احتوائها على احدى المركبات التي يمكن أن تظهر في الرسم البياني كالمركبة الفصلية.

- تحليل دالة الارتباط الذاتي (ACF) والارتباط الذاتي الجزئي (PACF): حيث تكون دالة الارتباط الذاتي (ACF) مؤشراً مهماً لكشف عدم استقرارية سلسلة زمنية، وهذا عندما لا تتعدم هذه الدالة بعد

فترة معينة تعادل $\frac{T}{4}$ (ربع عدد المشاهدات) وتتاقصها يكون في شكل أسّي نظرياً، بينما تطبيقياً يجب أن تقع معاملات هذه الدالة داخل مجال ثقة مناسب حتى تكون مستقرة ، كما أنها تعتبر كاشف مهم للفصلية من خلال القيم و التنبؤات التي تظهر في شكل منتظم على هذه الدالة.

- اختبارات ديكي فولر: وتعتبر المعيار الأكثر مصداقية، فإذا كانت السلسلة الزمنية غير مستقرة يجب معالجتها عن طريق الفروقات حسب درجة التكامل من أجل تحويلها الى سلسلة مستقرة ومن ثم القيام بالتقدير اللازم.

2- تحديد النموذج : بعد دراسة استقرارية السلسلة الزمنية تأتي مرحلة تحديد النموذج (AR) ،

(MA)، (ARMA)... إلخ وذلك بالاعتماد على دالة الارتباط الذاتي ودالة الارتباط الذاتي الجزئي .

- بالنسبة لنماذج المتوسطات المتحركة (MA) من الدرجة q تتعدم دالة الارتباط الذاتي (ACF) مباشرة بعد الدرجة q ، بينما دالة الارتباط الذاتي الجزئية (PACF) متناقصة ولكنها لا تتعدم لحظياً.
- بالنسبة لنماذج الانحدار الذاتي (AR) من الدرجة p فإن دالة الارتباط الذاتي الجزئية (PACF) تتعدم مباشرة بعد الدرجة p ، بينما دالة الارتباط (ACF) تبقى متناقصة ولكنها لا تتعدم بنفس السرعة.
- أما النماذج المختلطة فإن الدالتين تبقيان مستمرتي التدهور ولكنهما لا تتعدمان عند الدرجتين المذكورتين سابقاً [7].

3- تقدير معالم النموذج المحدد : بعد الانتهاء من مرحلة التعرف على نموذج السلسلة الزمنية و ذلك بتحديد كل من (p,q) ، يمكننا الانتقال إلى مرحلة التقدير لمعالم النموذج باستخدام طريقة المربعات الصغرى العادية أو طريقة المربعات الصغرى المعممة.

4- صلاحية النموذج المقدر: في هذه المرحلة نركز على المؤشرات والاختبارات الاحصائية التي على أساسها يتم قبول النموذج المقدر ومن أهمها :
أولاً- اختبار معنوية المعلمات المقدر.

ثانياً- اختبار طبيعية البواقي: حيث يعتبر هذا الاختبار مهماً جداً، ففي حالة عدم تتبع أخطاء النموذج للقانون الاحتمالي الطبيعي فلا يمكن استخدام طريقة بوكس جينكنز في التنبؤ وبالتالي وجود ارتباط ذاتي بين الأخطاء مما يجب تشخيص نماذج أخرى مثلاً: ARCH, GARCH.... إلخ [8].
إن تحليل البواقي يتم من خلال:

- رسم دالة الارتباط الذاتي الكلية والجزئية للبواقي:

يتم رسم دالة الارتباط الذاتي الكلية والجزئية للتأكد أن معالمها تقع داخل مجال المعنوية المعبر عنه بالصيغة التالية:

$$|\Gamma_k| \leq \frac{2}{\sqrt{T}}$$

- اختبار Box-Pierce

يسمح هذا الاختبار بإظهار وجود سياق الضجة البيضاء الذي يتميز بقيم معدومة لـ P_k وتتمثل فرضياته فيما يلي:

$$\begin{cases} H_0 : P_1 = P_2 = \dots = P_k \\ H_1 : \exists! P_k \neq 0 \end{cases}$$

ومن أجل تحقيق هذا الاختبار نستعمل Box-Pierce التي تتبع تقريبا توزيع χ^2 والمعرفة كما يلي:

$$\varphi = T \sum_{k=1}^h \Gamma_k^2(e_t)$$

حيث :

k : عدد التباطؤات.

Γ_k : الارتباط التجريبي من الرتبة k .

T : عدد المشاهدات.

ويتم اتخاذ القرار على الشكل التالي: إذا كانت $\varphi > \chi^2_{(1-\alpha)}$ نرفض فرضية الضجة البيضاء [9].

- اختبار Ljung-Box

يسمح هذا الاختبار بإظهار وجود سياق الضجة البيضاء، وتعطى صيغته في الشكل الرياضي التالي:

$$\varphi' = T(T+2) \sum_{k=1}^h \frac{\Gamma_k^2(e_t)}{T-k}$$

ويتم اتخاذ القرار على الشكل التالي:

تقارن قيمة إحصاءة Ljung-Box مع قيمة χ^2 الجدولية وإذا كانت $\varphi' > \chi^2_{(1-\alpha)}$ نرفض فرضية الضجة البيضاء.

ثالثاً: مقارنة النماذج:

إن النماذج غير المرفوضة بواسطة الأدوات الإحصائية تتم المفاضلة بينها بواسطة المعايير التالية:

1- معيار *AKAIKE*:

تعطى علاقة هذا المعيار على الشكل التالي:

$$AIC = \hat{\sigma}^2 \exp\left(2\left(\frac{p+q}{T}\right)\right)$$

$$\hat{\sigma}^2 = \frac{\sum e_t^2}{T}$$

كما يمكن كتابة هذا المعيار في شكله اللوغاريتمي كما يلي:

$$AIC = \ln(\hat{\sigma}^2) + \left(2\left(\frac{p+q}{T}\right)\right)$$

وبسبب إعطائه وزن أكبر للنماذج المستعملة لأكثر عدد من المشاهدات عدلت صيغته الرياضية

على النحو التالي:

$$NAIC = \frac{AIC}{T}$$

ويتم اختيار النموذج الذي يحقق أصغر قيمة للمعيار *AIC* [10].

2- معيار *Shwarz*:

رغبة في تحقيق خصائص تقاربية قام *Shwarz* بتعديل المعيار السابق كما يلي:

$$BIC = \ln \hat{\sigma}^2 + \frac{(p+q)}{T} \ln(T)$$

ويتم اختيار النموذج الذي يحقق أصغر قيمة لهذا المعيار.

ويستخدم أيضاً عدد من اختبارات الدقة التنبؤية التي تبين مدى صلاحية النموذج لإجراء التنبؤ ومن

هذه الاختبارات:

1- متوسط مربع الخطأ *Mean- square error (MSE)*: بحسب هذا الاختبار وفق

الصيغة التالية:

$$MSE = \sum_{t=1}^T (e_t)^2 / T$$

T عدد المشاهدات ، e الخطأ العشوائي.

2-متوسط الخطأ النسبي المطلق: Mean absolute percentage error (MAPE) :

يحسب هذا الاختبار وفق الصيغة التالية:

$$MAPE = \sum_{t=1}^T |e_t / y_t| / T * 100$$

ويتم اختيار النموذج الأفضل للتنبؤ عندما تكون قيم MSE و MAPE أقل ما يمكن.

5- التنبؤ: إن مرحلة التنبؤ تعد آخر و أهم مرحلة باعتبار أن التنبؤ هو عملية عرض حالي لمعلومات مستقبلية باستخدام معلومات مشاهدة تاريخية، وذلك باستعمال نماذج السلاسل الزمنية (AR) ، (MA) ، (ARMA).... إلخ والتي تم تحديدها مسبقاً ودراسة صلاحيتها حيث هدفها الأساسي هو تحقيق التنبؤ.

النتائج والمناقشة:

تم استخدام بيانات عدد طلاب المرحلة الثانوية في سورية خلال الفترة (1986-2018) لغرض تقدير نموذج ARIMA الملائم والتنبؤ من خلاله بعدد طلاب المرحلة الثانوية في المستقبل من خلال الجدول رقم (1):

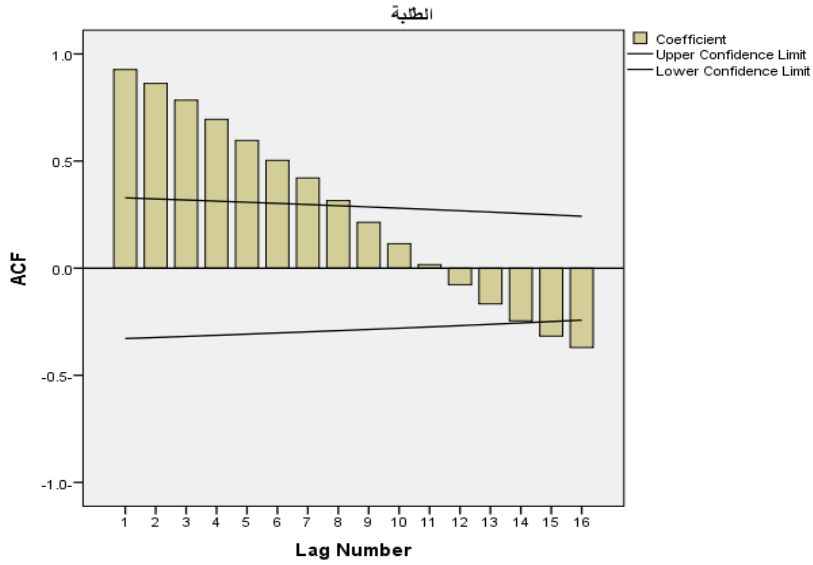
الجدول رقم (1): عدد طلاب المرحلة الثانوية في سورية خلال الفترة (1986-2018)

عدد الطلاب	العام	عدد الطلاب	العام	عدد الطلاب	العام
367572	2008	151681	1997	120675	1986
381702	2009	151545	1998	121341	1987
392960	2010	160147	1999	123896	1988
430702	2011	171725	2000	126254	1989
439463	2012	201874	2001	124198	1990
322726	2013	233838	2002	190654	1991
382923	2014	266196	2003	181568	1992
371375	2015	279633	2004	171679	1993
378255	2016	312132	2005	169566	1994
360896	2017	332659	2006	166142	1995
330810	2018	361745	2007	156099	1996

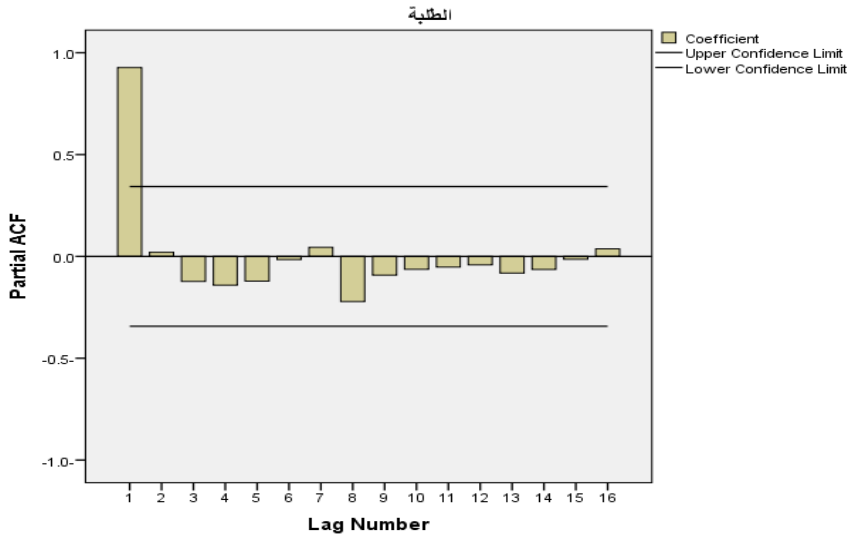
المصدر: المكتب المركزي

للإحصاء.

لتطبيق النموذج يجب أن تكون السلسلة الزمنية مستقرة، وللتأكد من ذلك نقوم برسم الشكل البياني لدالة الارتباط الذاتي Autocorrelation Function (ACF) ودالة الارتباط الذاتي الجزئي Partial Autocorrelation Function (PACF) لمتغير عدد طلاب المرحلة الثانوية خلال الفترة المدروسة.

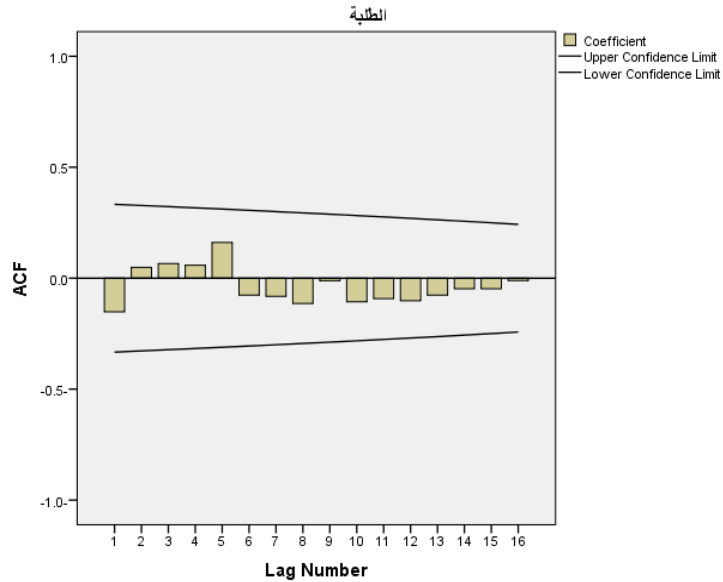


الشكل رقم (1) : دالة الارتباط الذاتي (ACF) لبيانات طلاب المرحلة الثانوية في سورية المصدر: من إعداد الباحث بالاعتماد على بيانات الجدول رقم (1) باستخدام برنامج SPSS.



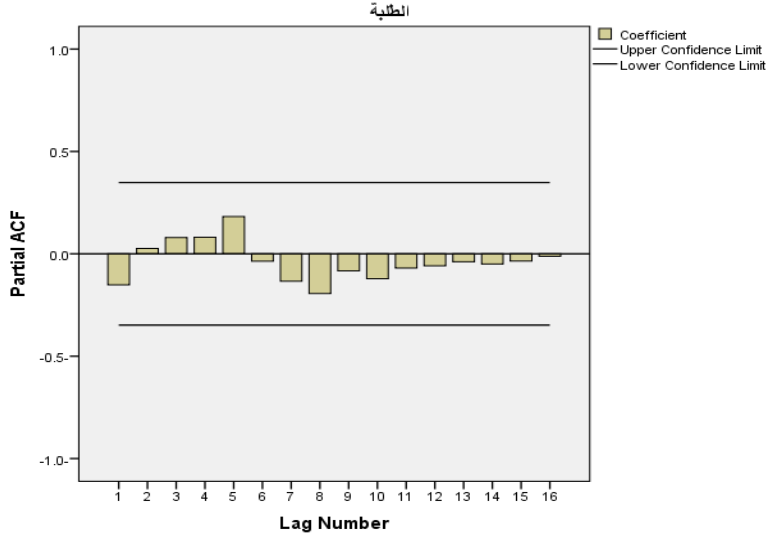
الشكل رقم (2): دالة الارتباط الذاتي الجزئي (PACF) لبيانات طلاب المرحلة الثانوية في سورية المصدر: من إعداد الباحث بالاعتماد على بيانات الجدول رقم (1) باستخدام برنامج SPSS.

نلاحظ من الشكل (1) أن أغلب قيم معاملات الارتباط الذاتي ACF تقع خارج حدود مجال الثقة، لذلك نعتبر أن السلسلة الزمنية لعدد طلاب المرحلة الثانوية في سورية غير مستقرة، لذلك نقوم بأخذ الفروقات من الدرجة الأولى، ونقوم برسم الشكل البياني لدالة الارتباط الذاتي (ACF) ودالة الارتباط الذاتي الجزئي (PACF) لمتغير الفروقات من الدرجة الأولى لأعداد طلاب المرحلة الثانوية خلال الفترة المدروسة.



الشكل رقم (3) : دالة الارتباط الذاتي (ACF) للفروقات من الدرجة الأولى لطلاب المرحلة الثانوية في سورية.

المصدر: من إعداد الباحث بالاعتماد على بيانات الجدول رقم (1) باستخدام برنامج SPSS.



الشكل رقم (4) : دالة الارتباط الذاتي الجزئي (PACF) للفروقات من الدرجة الأولى لطلاب المرحلة الثانوية في سورية.

المصدر: من إعداد الباحث بالاعتماد على بيانات الجدول رقم (1) باستخدام برنامج SPSS.

نلاحظ من الشكلين (3) و (4) أن جميع قيم معاملات الارتباط الذاتي ACF ومعاملات الارتباط الذاتي الجزئي PACF تقع ضمن حدود مجال الثقة، لذلك نعتبر أن السلسلة الزمنية لعدد طلاب المرحلة الثانوية في سورية مستقرة عند الفروقات من الدرجة الأولى. وتجريب جميع نماذج $ARIMA(p,d,q)$ المناسبة لتحديد رتبة الانحدار الذاتي P ورتبة المتوسطات المتحركة q ، نجد أن النموذج $ARIMA(2,1,2)$ هو المناسب لتمثيل سلسلة البيانات لأنه يحقق أقل قيمة لمعيار أكاي للمعلومات (AIC). ومعيار شوارترز (SBC) من النماذج الأخرى المقترحة، كذلك يحقق أقل قيمة من حيث الدقة التنبؤية للاختبار MSE و MAPE من النماذج الأخرى الموجودة في الجدول رقم(2)، حيث أن قيمة معيار $AIC= 777.21$ ، $SBC= 780.11$ لهذا النموذج. وقيمة اختبار $MSE = 30395.62$ ، $MAPE = 7.60$

لذلك نرفض الفرضية الأولى من فرضيات بحثنا، ونقبل البديلة لها التي تنص بأنه يمكن التوصل إلى النموذج الذي يجعل السلسلة الزمنية مستقرة باستخدام نماذج ARIMA.

جدول رقم (2): مقارنة المعايير الإحصائية للنماذج المقترحة لتمثيل سلسلة طلاب المرحلة الثانوية

النموذج	معيار AIC	معيار SBC	اختبار MSE	اختبار MAPE
ARIMA (1,1,0)	777.28	780.27	30560.64	7.719
ARIMA (1,1,1)	779.35	783.84	31060.98	7.688
ARIMA (2,1,0)	779.34	783.83	31056.05	7.649
ARIMA (2,1,1)	780.63	786.62	31169.42	8.103
ARIMA(2,1,2)	777.21	780.11	30395.62	7.60
ARIMA(3,1,0)	781.25	787.23	31478.99	7.652

المصدر: من إعداد الباحث بالاعتماد على مخرجات برنامج SPSS.

بعد تقدير النموذج من خلال برنامج SPSS حصلنا على الثوابت التالية:

جدول رقم (3): الثوابت المقدرة لنموذج ARIMA (2,1,2)

Variables	B	SEB	T-RATIO	APPROX. PROB.(Sig.)
AR1	0.967	0.346	2.8	0.009
AR2	-0.671	0.384	-1.746	0.042
MA1	1.267	2.958	0.428	0.672
MA1	-0.993	4.657	-0.213	0.833
CONSTANT	6239.71	5212.94	1.195	0.242

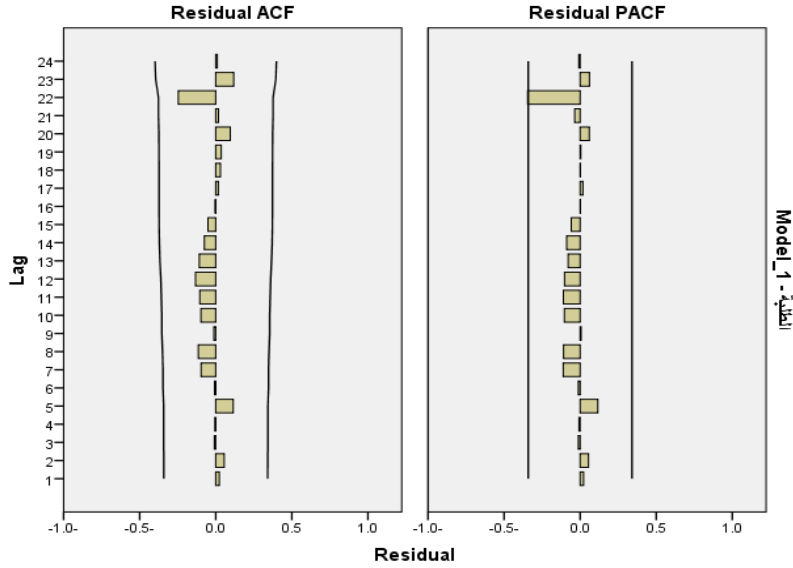
المصدر: من إعداد الباحث بالاعتماد على مخرجات برنامج SPSS.

من الجدول (3) نحصل على بيانات النموذج المقدر لعدد طلبة المرحلة الثانوية:

$$y_t = 6239.71 + 0.967 y_{t-1} - 0.671 y_{t-2} - 1.267 e_{t-1} + 0.993 e_{t-2}$$

حيث أن بعض معاملات النموذج والخاصة بالانحدار الذاتي معنوية، لأن Sig لكل منها أصغر من 0.05. وجميع معاملات النماذج الأخرى غير معنوية، لذلك تم اختيار هذا النموذج.

ومن أجل التأكد من أفضلية النموذج المقدر، يتم فحص معاملات دالتي الارتباط الذاتي والارتباط الجزئي لبواقي النموذج، حيث يبين الشكل (5) الرسم البياني لدالتي الارتباط الذاتي والارتباط الجزئي لبواقي النموذج $ARIMA(2,1,2)$.



شكل رقم (5): دالتي الارتباط الذاتي والارتباط الجزئي لبواقي النموذج $ARIMA(2,1,2)$

المصدر: من إعداد الباحث بالاعتماد على مخرجات برنامج SPSS.

نلاحظ من الشكل (5) أن قيم دالة الارتباط الذاتي ACF وقيم دالة الارتباط الذاتي الجزئي PACF لبواقي النموذج $ARIMA(2,1,2)$ غير معنوية باحتمال ثقة 0.95، لأن جميعها تقع ضمن حدي الثقة (Confidence Limits)، ويمكن التأكد من أفضلية النموذج المقدر من خلال استخدام اختبار Ljung-Box للتأكد من عدم وجود ارتباط ذاتي لبواقي.

جدول رقم (4) : اختبار Ljung-Box للتحقق من صلاحية النموذج

Ljung-Box Q(18)			Number of Outliers
Statistics	DF	Sig.	
5.002	14	.986	0

المصدر: من إعداد الباحث بالاعتماد على مخرجات برنامج SPSS.

من الجدول (4) نجد أن قيمة Sig. لهذا الاختبار هي 0.986 وهي أكبر من 0.05 ، لهذا يتم قبول فرضية عدم وجود ارتباط ذاتي بين البواقي.

كذلك نقوم بالتحقق فيما إذا البواقي تتبع التوزيع الطبيعي من خلال تطبيق الاختبار اللامعلمي (كلوموجوروف- سميرونوف). حيث يبين الجدول رقم (5) نتائج الاختبار.

جدول (5) : نتائج اختبار Kolmogorov-Smirnov لبواقي

النموذج

		Noise residual from الطلبة- Model_1
N		33
Normal Parameters ^{a,b}	Mean	140.67
	Std. Deviation	28432.141
	Most Extreme Differences	
	Absolute	.143
	Positive	.137
	Negative	-.143-
Test Statistic		.143
Asymp. Sig. (2-tailed)		.086 ^c

المصدر: من إعداد الباحث بالاعتماد على مخرجات برنامج SPSS.

نجد من الجدول (5) أن قيمة Sig. لهذا الاختبار هي 0.086 وهي أكبر من 0.05 ، لهذا يتم قبول فرضية عدم أي البواقي تتبع التوزيع الطبيعي. لذلك يعتبر النموذج المقترح هو النموذج الأفضل في

التنبؤ بعدد طلاب المرحلة الثانوية للأعوام القادمة. وبالتالي سيتم رفض الفرضية الثانية من البحث وقبول الفرضية البديلة لها التي تنص على أنه يمكن بناء نموذج قياسي للتنبؤ بعدد طلاب المرحلة الثانوية باستخدام نماذج ARIMA. ويظهر الجدول (6) عدد طلاب المرحلة الثانوية المقدر وفق النموذج ARIMA (2,1,2) حتى عام 2030.

جدول رقم (6) : عدد طلاب المرحلة الثانوية المقدر حتى عام 2030

العام	عدد طلاب المرحلة الثانوية المقدر وفق نموذج ARIMA(2,1,2)	العام	عدد طلاب المرحلة الثانوية المقدر وفق نموذج ARIMA(2,1,2)
2019	351891	2025	375262
2020	352488	2026	384872
2021	359393	2027	390034
2022	354332	2028	392969
2023	349192	2029	396736
2024	352000	2030	402798

المصدر: من إعداد الباحث بالاعتماد على مخرجات برنامج SPSS.

نلاحظ من بيانات الجدول (6) أن عدد طلاب المرحلة الثانوية المقدر سيرتفع حتى عام 2021 حيث يبلغ عندها عدد الطلاب (359393)، وسيخفض ابتداء من عام 2022 ويستمر هذا الانخفاض حتى عام 2024، حيث سيبلغ عدد الطلاب عندها (352000). ثم يبدأ بالارتفاع اعتباراً من عام 2025 حيث سيبلغ 402798 طالب عام 2030.

الاستنتاجات والتوصيات:

الاستنتاجات:

كانت أهم النتائج التي توصل إليها الباحث:

- 1 - تشكل سلسلة أعداد المنتسبين إلى مرحلة التعليم الثانوي في سورية سلسلة زمنية غير مستقرة، وتم أخذ الفروق من الدرجة الأولى لجعلها مستقرة.
- 2- إن النموذج الأفضل من بين النماذج المقترحة للتنبؤ بعدد طلاب المرحلة الثانوية في سورية هو نموذج $ARIMA(2,1,2)$.
- 3- يتوقع ارتفاع عدد طلاب المرحلة الثانوية في سورية للأعوام القادمة، حيث سيبلغ 402798 طالب لعام 2030 وفق النموذج المقترح.

التوصيات:

- 1 - يجب الأخذ بوسائل التنبؤ الحديثة لتحليل السلاسل الزمنية ولا سيما نماذج $ARIMA$ لما تمتاز به من دقة في النتائج.
- 2- استخدام النموذج الذي تم التوصل إليه للتنبؤ بعدد طلاب المرحلة الثانوية من قبل وزارة التربية، واعتماد التنبؤات التي يعطيها لوضع الخطط لمواجهة كل ما يخص سير العملية التعليمية للطلبة قبل بلوغهم مرحلة التعليم الجامعي.
- 3 - البحث عن طرائق إحصائية أخرى في تحليل السلاسل الزمنية مثل نماذج $GARCH$ ، وأسلوب الشبكات العصبية الاصطناعية، ومقارنتها مع نماذج $ARIMA$ في الدراسات المستقبلية.

المراجع

- [1] - البول، رجاء. دراسة تحليلية كمية لأساليب التنبؤ بأعداد الطلبة في المدارس الفلسطينية. مجلة جامعة النجاح للأبحاث العلوم الانسانية، المجلد 32 ، العدد 1، 2018 ص 174-190.
- [2] أحمد، طالب. استخدام بوكس- جينكنز لبناء نموذج قياسي للتنبؤ بعدد المواطنين السوريين، مجلة جامعة تشرين للعلوم الاقتصادية والقانونية، المجلد 40 ، العدد 6، 2018 ص 13-25.
- [3] نقار، عثمان و العواد، منذر. منهجية *Box- Jenkins* في تحليل السلاسل الزمنية والتنبؤ- دراسة تطبيقية على أعداد تلاميذ الصف الاول من التعليم الاساسي في سورية، مجلة جامعة دمشق للعلوم الاقتصادية والقانونية، المجلد 27، العدد 3، 2011 ص 125-152.
- [4] عطروش، حنان، استخدام أحد نماذج بوكس - جينكينز للتنبؤ بأعداد الطالبات في المرحلة الأساسية في محافظة أربين، مجلة العلوم الإدارية ، المجلد 11- العدد 5 ، 2015، ص 145-169 .
- [5] ADHIKARI, R., & AGRAWAL, R. K.; *An introductory study on time series modeling and forecasting. arXiv preprint arXiv*., 2013, 1302.6613.
- [6] SHUMWAY; RH ; *Applied Statistical Time Series Analysis, First Edition, prentice Hall New Jersey, USA, 1998, P.537.*
- [7] أبو ذر يوسف، أحمد، وآخرون، استخدام السلاسل الزمنية للتنبؤ بإنتاجية الصمغ العربي في سوق محاصيل الأبييض للفترة 1960-2012 ، مجلة البحث العلمي للعلوم والأداب، العدد 15 ، 2012 ، ص 211-238 ، السودان.
- [8] ياسين، فايق، التنبؤ الاقتصادي بالمساحات المزروعة بمحصول الحنطة في العراق باستخدام نماذج ARIMA للمدة 2008-2015 ، مجلة الأنبار للعلوم الزراعية، المجلد 9 ، العدد 2 ، 2011 ، ص 27-41.
- [9] SHUMWAY; RH ; *Applied Statistical Time Series Analysis, First Edition, prentice Hall New Jersey, USA, 1998, P.537.*
- [10] Kirchgassner, G. , Wolters J. *Introduction to Modern Time Series Analysis*, SPRINGER-Verlag, Berlin Heidelberg, 2007.

