

تحقيق التوازن بين الاشتراكات والتعويضات لصندوق الضمان الاجتماعي في العراق¹

زيد خليل إسماعيل

zaid.khalil.ismail@gmail.com

أ.م.د. فراس احمد محمد

drfirasmohana@gmail.com

جامعة بغداد - كلية الادارة والاقتصاد - قسم الاحصاء

المستخلص

ان للسلاسل الزمنية اهمية كبيرة وهي عبارة عن مجموعة من المشاهدات لفترات زمنية متعاقبة تستخدم في التحليل عن طريق دراسة التغيرات التي تطرأ على الظاهرة ومن اهم نماذج السلاسل الزمنية هي نماذج الصندوق الأسود نموذج "ARMAX" وهو نموذج مختلط مكون من الانحدار الذاتي مع متوسطات متحركة مع مدخلات خارجية وهذا يعطي للنموذج دقة عالية. وان عملية بناء النموذج تتالف من عدة مراحل وهي تحديد رتبة النموذج وعملية تقدير معاملات النموذج ثم عملية التنبؤ بحالة الصندوق لمعرفة حجم التعويضات الممنوحة للعمال مستقبلا وامكانية تحقيق التوازن بين الاشتراكات والتعويضات من اجل ايفاء الصندوق بالتزاماته المستقبلية. وتم تحديد رتبة النموذج وفق المعايير (AIC, MDL, FPE)، و تقدير النموذج بطريقة المربعات الصغرى الاعتيادية وطريقة المربعات الصغرى التكرارية (RLS) بواسطة عامل التغاضي وايضا اقتراح طريقة تحسن نتائج التنبؤ من خلال تقدير الخطأ وادخاله بنموذج انحدار كعامل مؤثر على التنبؤ للمخرجات للحصول على افضل طريقة يمكن التنبؤ من خلالها للمستقبل باستخدام بيانات اخذت من دائرة الضمان الاجتماعي للعمال ومن الفترة 2013/1 الى 2019/6 حيث تفوق النموذج المقترح على طريقة المربعات الصغرى الاعتيادية OLS لانه اعطى نتائج دقيقة وقريبة من القيم الحقيقية.

الكلمات المفتاحية: بناء النموذج، معايير، تقدير النموذج، التنبؤ.

Striking A Balance Between Contributions and Compensation for the Social Security Fund in Iraq

Assist. Prof. Dr. Firas Ahmed Mohammed

drfirasmohana@gmail.com

University of Baghdad - College of Administration and Economics

Received 13/7/2020

Zaid Khalil Ismail

zaid.khalil.ismail@gmail.com

Accepted 18/8/2020

Abstract: Time series have gained great importance and have been widely applied in various fields such as economic, financial, health and social fields, which is a set of observations for successive time periods used in the analysis by studying the changes in the phenomenon and predicting the future of the phenomenon. The most important time series models are the black box models "ARIMAX" which is a mixed model consisting of self-regression with moving averages with external inputs and this gives the model a high accuracy. The process of building the model consists of several stages which are determining the model's rank and the process of estimating the model's parameters and then the process of predicting the condition of the fund to know the size of compensation granted to workers in the future and the possibility of balancing contributions and compensation in order to fulfill the fund's future obligations. The rank of the model was determined according to the criteria, AIC, MDL, FPE and estimating the model using ordinary least square method (OLS) and recursive least square (RLS) and also proposing a way to improve the prediction results by estimating the error and inserting it in a regression model as a factor influencing the prediction of the outputs to obtain the best predictable way for the future is by using

¹ مستل من رسالة ماجستير

data taken from the Social Security System for workers and from the period 1/2013 to 6/2019, where the proposed model outperformed the regular least squares method because it gave accurate results close to the real values.

Keywords: Model building, Standards, Model estimation, Prediction

1. المقدمة Introduction

ان التنوع الهائل في البيانات ادى الى ظهور انواع عديدة من البيانات منها التجريبية، و بيانات الحصر الشامل والبيانات المتعلقة بالزمن، حيث ان لكل منها آلية في التعامل و اكتشاف اسلوب تطور الظاهرة واستخدامه فيما بعد للتنبؤ بتلك الظاهرة في المستقبل ويسمى هذا النوع من البيانات بالسلاسل الزمنية، ومن هنا اكتسبت السلاسل الزمنية اهمية كبيرة وتطبيقاً واسعاً في شتى المجالات ومن ضمنها الاقتصادية والمالية والصحية والاجتماعية. و تستخدم في عملية التحليل من خلال دراسة التغيرات التي تطرأ على الظاهرة والتنبؤ بمستقبل تلك الظواهر ولفترات زمنية مختلفة من خلال الاعتماد على التغيرات التي حدثت في الفترات السابقة او الحالية للظاهرة والتي تتعامل مع بيانات لها فترات زمنية ثابتة (ساعة، يوم، اسبوع، شهر، سنة). ومن اهم نماذج السلاسل الزمنية نماذج الصندوق الاسود ونموذج ARMAX وهو من النماذج الكفوة في التنبؤ لذلك تم توظيف هذا النموذج للتنبؤ بسلسلة المخرجات والتي تمثل التعويضات بتأثير سلسلة المدخلات التي تمثل الاشتراكات لمعرفة القيم المستقبلية حسب النموذج ARMAX من اجل تحقيق التوازن بين الاشتراكات والتعويضات لصندوق الضمان الاجتماعي وتلافي اخطار العجز واستثمار الفائض.

يهدف هذا البحث في تحديد نموذج تنبؤي لتقدير علاقة التوازن بين الاشتراكات والتعويضات لنظام التأمينات الاجتماعية من جراء تحقق الاخطار وصرف الرواتب التقاعدية والمكافآت من اجل الايفاء بالتزامات الصندوق المالية التي قد تؤدي الى ازمت تدخل بها البلاد جراء عدم الايفاء لتلافي خطر العجز في الاموال لوضع الخطط للاستناد عليها من قبل اصحاب القرار والاستراتيجيات لضمان الوفرة المالية من اجل حياة كريمة مضمونة للعمال في القطاع الخاص.

ومن الدراسات السابقة دراسة Richard حيث قدم سنة 1979 بحثاً ناقش فيه اسلوب تنبؤ جديد تمثل بعملية دمج ما بين النماذج حيث ARMA مع متغير نموذج الانحدار لتكوين نموذج هجين اطلق عليه اسم ARIMAX حيث بين الباحث ان النموذج الجديد يعطي اقل متوسط مربعات خطأ باستخدام طريقة المكان الاعظم [10]. وفي سنة 2013 قدم الباحثون Sulman .N وآخرون دراسة باستخدام نماذج ARMAX في التنبؤ للعملة واستخدام اختبار Ljung – Box واختبار اثر نموذج ARCH ووجدوا ان النموذجين كفوئين مع فرق ليس كبير بينهما وتمت التوصية باستخدام نموذج ARMAX في التنبؤ [12].

وفي سنة 2015 قدم الباحث Hamed Shakoure بحثاً قام من خلاله بالتنبؤ بكمية الطلب على الطاقة من خلال تطوير نموذج مناسب للتنبؤ، وقد تم اختيار النموذج الافضل من ARMAX بناء على متغيرات خارجية متعددة مع مدخلات مختلفة للنموذج الانسب [4].

وفي سنة 2017 قدم الباحثون Oluwadare & Bolanle دراسة تضمنت المقارنة ما بين قدرة التنبؤ لنماذج السلسلة الزمنية ARMAX , ARMA لكل من المواليد الاحياء في نيجيريا بدءاً نموذج ARIMA احادي المتغير المتعدد [1].

2. الطرائق والاساليب

• نموذج الانحدار الذاتي والمتوسطات المتحركة مع مدخلات خارجية ARMAX Model

Autoregressive Moving Average With Exogenous Input Model

يقوم هذا النموذج بوصف سلسلة المخرجات للمتغير التابع $y(t)$ للظاهرة المراد التنبؤ بها من خلال دمج نموذج ARIMA مع سلسلة المدخلات الخارجية $x(t)$ وتم تطوير هذا النموذج وهو مطبق في مجالات الرقابة والاقتصاد القياسي والنموذج يكون بالصيغة الاتية [9][7][3][2]:

$$y(t) + a_1y(t-1) + \dots + a_ny(t-n_a) = b_1u(t- n_k) + \dots + b_{n_b}u(t - n_k - n_b + 1) + c_1e(t-1) + \dots + c_{n_c}e(t-n_c) + e(t)$$

حيث ان:

na = رتبة AR للانحدار الذاتي

nb = عدد حدود المدخلات

nc = رتبة MA للمتوسطات المتحركة

nk = زمن التأخير

وبهذا يمكن كتابة النموذج بالشكل الاتي :

$$A(L) y(t) = B(L) u(t - n_k) + C(L) e(t)$$

حيث ان

$$A(L) = 1 + a_1 L^{-1} + a_2 L^{-2} + \dots + a_{na} L^{-na}$$

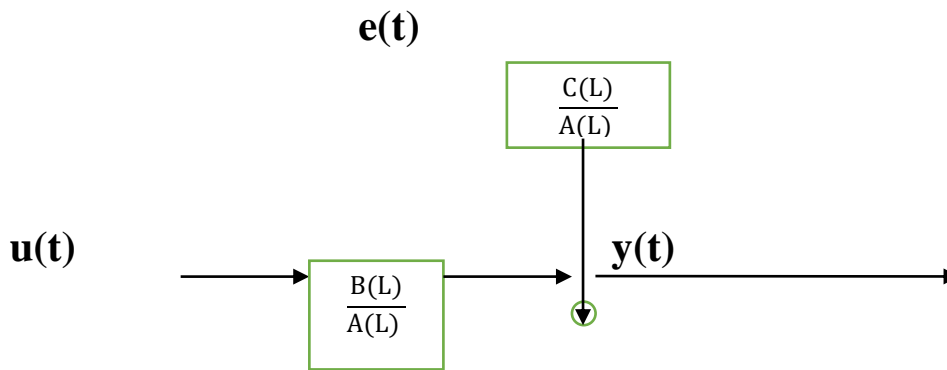
$$B(L) = b_1 L^{-1} + b_2 L^{-2} + \dots + b_{nb} L^{-nb+1}$$

$$C(L) = 1 + c_1 L^{-1} + c_2 L^{-2} + \dots + c_{nc} L^{-nc}$$

ومعلمات النموذج :

$$\theta = [a_1 \ a_2 \dots a_{na} \ b_1 \ b_2 \dots b_{nb} \ c_1 \ c_2 \dots c_{nc}]^T$$

$L =$ الازاحة الى الخلف



شكل (1): يوضح هيكلية انموذج ARMAX

معايير تحديد رتبة النموذج

يجب على الباحث اجراء عملية اختيار النموذج الملائم والذي يكون مناسباً للتنبؤ وذلك عن طريق عدة معايير يقوم باستخدامها لكي يصل الى النموذج المطلوب وهذه العملية ليست سهلة وتتطلب بعض الجهد لان الباحث سوف يجد نفسه امام عدة نماذج وجميعها تكون مرشحة وغير مرفوضة ولهذا السبب وضعت هذه المعايير للمساعدة في تحديد النموذج الافضل [11].

• معيار أكايكي (Akaike Information Criteria AIC)

اقترح هذا المعيار من قبل العالم Akaike عام 1973-1974 ومن خلاله يتم اختيار رتبة النموذج الملائم من بين النماذج المرشحة بحيث يقابل النموذج الافضل اقل قيمة من معيار AIC ويمثل الرتبة الاكثر ملائمة ويتم التعبير عنه كالآتي [9]:

$$AIC(B) = \ln(\sigma^2_{(B)}) + \frac{2B}{n}$$

حيث ان:

$B =$ عدد معلمات النموذج

$n =$ عدد المشاهدات

$\sigma^2_{(B)} =$ مقدار تباين الخطأ

• معيار (FPE) معيار خطأ التنبؤ

ووصف من قبل العالم اكاكي في عام 1969 وهو خطأ التباين ويحسب بالصيغة الآتية [7]:

$$FPE(B) = (\sigma^2_{(B)}) + 2 \lambda \frac{B}{n}$$

وان صيغة λ كما يلي :

$$\lambda = \sigma^2_{(B)} \frac{1}{1 - \frac{B}{n}}$$

وعند تعويض قيمة λ في المعادلة السابقة

$$FPE(B) = (\sigma^2_{(B)}) \frac{n+B}{n-B}$$

حيث ان:

عدد معلمات النموذج = B

عدد المشاهدات = n

مقدار تباين الخطأ = $\sigma^2_{(B)}$

• معيار اكاكي البيزي (BIC)

يسمى هذا المعيار بمعيار الوصف الاصغر ويكتب (MDL) اختصارا الذي اقترحه العالم Akaike عام 1979 وايضا العالم Rissanen عام 1978 ويكتب بالصيغة الاتية [7]:

$$MDL = \ln(\sigma^2_{(B)}) + \frac{B}{n} \ln(n)$$

حيث ان:

عدد معلمات النموذج = B

عدد المشاهدات = n

مقدار تباين الخطأ = $\sigma^2_{(B)}$

$$\text{تباين الخطأ} = \frac{\sum (Y - \hat{Y})^2}{N - 1}$$

Y = قيمة الحقيقية

\hat{Y} = قيمة التقديرية

N = عدد المشاهدات

تقدير معلمات النموذج

يتم تقدير معلمات الى النموذج ARMAX عن طريق طريقة المربعات الصغرى OLS وكالاتي:

• طريقة المربعات الصغرى الاعتيادية (Ordinary least square method (OLS)

تم تطوير طريقة المربعات الصغرى بواسطة Gauss في عام 1795 وهي من الحلول المثلى والاكثر تطبيقا في التحسين الخطي والهدف هو ايجاد ناتج النموذج \hat{Y} الذي يقدر مخرجات النموذج y ويمكن استخدام هذه الطريقة لتقدير معلمات نموذج ARMAX لجعل مجموع مربعات الخطأ اقل ما يمكن. وعلى افتراض ان النموذج يكون بالصورة التالية [13] [9] [5]:

$$y(t) + a_1y(t-1) + \dots + a_nay(t-n_a) = b_1u(t-n_k) + \dots + b_nbu(t-n_k-n_b+1) + \dots + c_1e(t-1) + \dots + c_nce(t-n_c) + e_t$$

$$y(t) + a_1y(t-1) + \dots + a_nay(t-n_a) \quad Y = \text{المخرجات}$$

$$b_1u(t-n_k) + \dots + b_nbu(t-n_k-n_b+1) \quad U = \text{المدخلات}$$

$$c_1e(t-1) + \dots + c_nce(t-n_c) = \quad e = \text{الايخطاء}$$

وان

$$\phi^T(t) = [-y(t-1) - \dots - y(t-n_a) \quad u(t-n_k) + \dots + u(t-n_k-n_b+1) \quad e(t-1) + \dots + e(t-n_c)]$$

والمعلمات الخاصة بالنموذج هي:

$$\theta = [a_1 \quad a_2 \quad \dots \quad a_{na} \quad b_1 \quad b_2 \quad \dots \quad b_{nb} \quad c_1 \quad c_2 \quad \dots \quad c_{nc}]^T$$

فان

$$y(t) = -a_1y(t-1) - \dots - a_nay(t-n_a) + b_1u(t-n_k) + \dots + b_nbu(t-n_k-n_b+1) + c_1e(t-1) + \dots + c_nce(t-n_c)$$

$$y(t) = \phi^T(t)\theta + e(t)$$

وبما ان الاخطاء غير متوفرة كمشاهدات لذلك نقوم بتقدير المعلمات كالاتي:

$$\theta = [a_1 \ a_2 \ \dots \ a_{na} \ b_1 \ b_2 \ \dots \ b_{nb}]$$

$$\varphi^T(t) = [-y(t-1) - \dots - y(t-n_a) \ u(t-1) + \dots + u(t-n_b)]$$

ولكي نحصل على مقدر خطأ للتنبؤ:

$$\hat{\theta}_{PE} = \arg \min_{\theta} V_n$$

ولكي نجعل مجموع المربعات للخطأ اقل مقدار:

$$V_n = \frac{1}{n} \sum_{t=1}^n (y(t) - \varphi^T(t)\hat{\theta})^2$$

$$\hat{y}(t) = \varphi^T(t)\hat{\theta}$$

ومن خلال اشتقاق المعادلة ومساواتها للصفر

$$\frac{d}{d\theta} V_n = -\frac{2}{n} \sum_{t=1}^n \varphi(t) (y(t) - \varphi^T(t)\hat{\theta}) = 0$$

$$\sum_{t=1}^n \varphi(t)y(t) = \sum_{t=1}^n \varphi(t)\varphi^T(t)\hat{\theta}$$

$$\hat{\theta} = [\sum_{t=1}^n \varphi(t)\varphi^T(t)]^{-1} [\sum_{t=1}^n \varphi(t)y(t)]$$

ويمكننا اعادة الصياغة بالشكل التالي

$$y(t) = \varphi^T(t)\theta + (c_1 e(t-1) + \dots + c_{nc} e(t-n_c)) + e(t)$$

ونحصل على المقدرات بالصيغة الاتية:

$$\hat{\theta} = [\sum_{t=1}^n \varphi(t)\varphi^T(t)]^{-1} [\sum_{t=1}^n \varphi(t)y(t)]$$

ويتم التقدير عن بطريقة المربعات الصغرى التكرارية وكالاتي:

• طريقة المربعات الصغرى التكرارية (Recursive Least Square (RLS)

تقوم هذه الطريقة بحساب المعلمة $\hat{\theta}(t)$ في الوقت t وحساب المعلمة $\hat{\theta}(t-1)$ عند الفترة $t-1$ عن طريق بعض التعديل البسيط وهذه الصيغة هي تعاقب من طريقة المربعات الصغرى وخطوات هذه الطريقة تبدأ من تعريف تقدير المعلمة حسب طريقة المربعات الصغرى الاعتيادية وكالاتي: ان تقدير المعلمة في الزمن t هو [8] [13] [14]

$$\hat{\theta}(t) = [\varphi(t)\varphi^T(t)]^{-1} [\varphi(t)y(t)]$$

اما في الزمن $t+1$ هو :

$$\hat{\theta}(t+1) = [\varphi(t+1)\varphi^T(t+1)]^{-1} [\varphi(t+1)y(t+1)]$$

وان حدود $y(t+1)$ و $\varphi(t+1)$ تكتب بالشكل الاتي :

$$\varphi^T(t+1) = \begin{bmatrix} \varphi^T(t) \\ \varphi(t+1) \end{bmatrix}$$

$$y(t+1) = \begin{bmatrix} y(t) \\ y(t+1) \end{bmatrix}$$

$$\varphi(t+1)\varphi^T(t+1) = [\varphi(t) \ \varphi^T(t+1)] \begin{bmatrix} \varphi^T(t) \\ \varphi(t+1) \end{bmatrix}$$

$$= \varphi(t)\varphi^T(t) + \varphi^T(t+1)\varphi(t+1)$$

$$\varphi(t+1) y(t+1) = [\varphi(t) \ \varphi^T(t+1)] \begin{bmatrix} y(t) \\ y(t+1) \end{bmatrix}$$

$$= \varphi(t)y(t) + \varphi^T(t+1)y(t+1). \quad (1)$$

ويتم تعريف مصفوفة التغيرات في الزمن t :

$$p(t) = [\varphi(t)\varphi^T(t)]^{-1} = [\sum_{t=1}^n (\varphi(t)\varphi^T(t))]^{-1}$$

اما في حالة الزمن $t+1$ فان مصفوفة التغيرات تصبح:

$$p(t+1)^{-1} = p(t)^{-1} + \varphi^T(t+1)\varphi(t+1) \quad (2)$$

ويمكن اعادة كتابة مصفوفة التغيرات

$$\hat{\theta}(t) = p(t)\varphi(t)y(t)$$

$$\hat{\theta}(t+1) = p(t+1)\varphi(t+1)y(t+1) \quad (3)$$

ومن خلال توظيف المعادلة (1) و (2) وبعد الترتيب نجد ان :

$$\begin{aligned} \hat{\theta}(t+1) &= p(t+1)[\varphi(t)y(t) + \varphi^T(t+1)y(t+1)] \\ &= p(t+1)\varphi(t)y(t) + p(t+1)\varphi^T(t+1)y(t+1) \end{aligned} \quad (4)$$

وبضرب طرفي المعادلة بمصفوفة التغيرات $p(t)^{-1}$ نجد :

$$\varphi(t)y(t) = p(t)^{-1}\hat{\theta}(t)$$

وعند دمج معادلة (3) و (2) نحصل على :

$$\varphi(t)y(t) = p(t+1)^{-1}\hat{\theta}(t) - \varphi^T(t+1)\varphi(t+1)\hat{\theta}(t)$$

ومن خلال تعويض الحد $\varphi(t)y(t)$ في العلاقة رقم (4)

$$\begin{aligned} \hat{\theta}(t+1) &= p(t+1)[p(t+1)^{-1}\hat{\theta}(t) - \varphi^T(t+1)\varphi(t+1)\hat{\theta}(t)] \\ &\quad + p(t+1)\varphi^T(t+1)y(t+1) \end{aligned}$$

وبعد تبسيط الصيغة:

$$\hat{\theta}(t+1) = \hat{\theta}(t) + p(t+1)\varphi^T(t+1)[y(t+1) - \varphi(t+1)\hat{\theta}(t)] \quad (5)$$

الحد $[y(t+1) - \varphi(t+1)\hat{\theta}(t)]$ يمثل الفرق بين سلسلة المخرجات وقيمة التنبؤ بخطوة واحدة , والذي يمثل تقدير خطأ التنبؤ في الزمن $t+1$ اي ان :

$$e(t+1) = y(t+1) - \varphi(t+1)\hat{\theta}(t)$$

وان متجه المعلمات $\hat{\theta}(t+1)$ في المعادلة (5) سيكون :

$$\hat{\theta}(t+1) = \hat{\theta}(t) + p(t+1)\varphi^T(t+1)e(t+1)$$

وان مصفوفة $p(t+1)$ يتم ايجادها عن طريق قانون معكوس المصفوفة

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$$

وحسب المعادلة رقم (2) فان الحدود تكون

$$A = p(t)^{-1} \quad C = 1 \quad B = \varphi^T(t+1) \quad D = \varphi(t+1)$$

والتي تمثل مصفوفة مربعة يكون المحدد لها غير مساوي للصفر اي ان المصفوفة تمتلك معكوس وبالتالي فان القيمة تكون :

$$p(t+1) = p(t)[Im - \varphi^T(t+1)(1 + \varphi^T(t+1)p(t)\varphi(t+1))^{-1}\varphi(t+1)p(t)]$$

وبعد التبسيط:

$$p(t+1) = p(t) - \frac{p(t)\varphi^T(t+1)\varphi(t+1)p(t)}{1 + \varphi^T(t+1)p(t)\varphi(t+1)}$$

• نهج عامل التغاضي Forgetting factor لطريقة المربعات الصغرى التكرارية

يتلخص هذا النهج بتقليل دالة الخسارة الى اقل حد و يعرف بعامل التغاضي من خلال اضافة (λ) الى دالة الخسارة وبالشكل الاتي [13] [9] [7]:

$$V_t(\theta) = \sum_{i=1}^t \lambda^{t-i} e^2(i)$$

$$= \sum_{i=1}^t \lambda^{t-i} (y(i) - \varphi^T(i)\theta)^2$$

سابقا كانت دالة الخسارة تستخدم $(\lambda = 1)$ لكن مع نهج عامل التغاضي تكون فيه قيمة λ عدداً موجباً اقل من واحد $(\lambda < 1)$ ويقع بين $(0.99 - 0.95)$ الذي يعطي اهمية نسبية للبيانات عن البيانات السابقة من خلال زيادة (t) وتسقيط البيانات التي حصلنا عليها مسبقاً

$$\hat{\theta}(t+1) = \hat{\theta}(t) + K(t+1)[y(t+1) - \varphi(t+1)\hat{\theta}(t)]$$

او نكتب الصيغة بدلالة خطأ التنبؤ

$$\hat{\theta}(t+1) = \hat{\theta}(t) + K(t+1)e(t+1)$$

حيث ان $K(t+1)$ يمثل حد الكسب او معامل المنفعة او الربحية الذي يفسر مقدار قيمة خطأ التنبؤ وتكون صيغته كالاتي:

$$K(t+1) = p(t+1)\varphi^T(t+1) = \frac{p(t)\varphi^T(t+1)}{[\lambda + \varphi(t+1)p(t)\varphi^T(t+1)]}$$

لذا فان مصفوفة التغيرات لعامل التغاضي تكون بالشكل الاتي :

$$p(t+1) = \frac{1}{\lambda} [1 - K(t+1)\varphi(t+1)]p(t)$$

• النموذج المقترح

دمج المقدرات \hat{y} لطريقة الافضل من طرائق التقدير التي تقابل اقل قيمة للمعايير RMSE, MAPE لنموذج ARMAX

مع الخطأ المقدر \hat{e} بواسطة التمهيد الاسي لمعلمتي هولت بنموذج انحدار واحد وبالصيغة الاتية :

$$\hat{y} = B_0 + B_1 \hat{y}(\text{ARMAX}) + B_2 \hat{e}$$

• خطوات التنبؤ بالنموذج المقترح

ان التنبؤ هو المرحلة الاخيرة والتي نريد الوصول اليها عن طريق الطرق المذكورة لنموذج ARMAX ولغرض زيادة الدقة في التنبؤ فيجب ملاحظة ان كل نموذج يحتوي على نسبة من الخطأ والتي في العادة لا تعطي اهمية كبيرة ولذلك سوف نولي اهتماماً بهذا الخطأ ونرى تأثيره على النموذج من خلال النموذج المقترح اذ يتم التنبؤ لطرق التقدير المربعات الصغرى الاعتيادية والطريقة التكرارية بواسطة عامل التغاضي واختيار الطريق الافضل منها عن طريق المفاضلة بمعيار المقارنة RMSE لتقدير قيم المخرجات \hat{y} ليصبح النموذج كالاتي :

$$\hat{y} = Y_{\text{ARMAX}} + e$$

وهنا نستخرج الخطأ عن طريق الفرق بين القيمة الحقيقية والقيمة المقدره وكالاتي :

$$e_{(t)} = y - \hat{y}_{\text{ARMAX}}$$

ومن ثم تقدير الخطأ ايضا وتمهيده بواسطة طريقة التمهيد الاسي لمعلمتي هولت الذي اقترحه العالم Holt سنة 1957 ويعتمد على معلمتين هما α و γ وان قيم المعلمتين محصورة بين الصفر والواحد $0 \leq \alpha \leq 1$ $0 \leq \gamma \leq 1$ وهذا التمهيد هو لتقدير الخطأ من اجل اعطاء طريقة قوية للتنبؤ ونقوم باحتساب تمهيد هولت كالاتي:

$$a(t) = \alpha e_t + (1 - \alpha) (a_{t-1} + b_{t-1})$$

$$b(t) = \gamma (a_t - a_{t-1}) + (1 - \gamma) b_{t-1}$$

$$a_1 = e_1$$

$$b_1 = \frac{(e_2 - e_1) + (e_4 - e_3)}{2}$$

$$\hat{e} = a(t) + b(t)w$$

حيث ان :

$a(t)$ = معامل التمهيد $b(t)$ = معامل تمهيد الانحدار w = عدد خطوات التنبؤ
وبعد تقدير المعلمات للنموذج المقترح

$$\hat{y} = B_0 + B_1 \hat{y}(ARMAX) + B_2 \hat{e}$$

نجد قيمة \hat{y} ثم نجد قيمة MSE ونقارن مع قيمة MSE للطريقة الافضل لنموذج ARMAX من الطرق السابقة الثلاثة ونحدد الاقل ويكون هو الافضل .

• فحص ملائمة النموذج

بعد تحديد النموذج وتقدير معلماته تاتي مرحلة فحص ملائمة النموذج للتنبؤ عن طريق تحسينه او تطويره او الابقاء عليه وتعد هذه المرحلة من اهم المراحل لانها تحدد القبول بالنموذج او القيام باستبداله , ويتم اختبار النموذج عن طريق الاختبار الاتي :

• اختبار (Ljung – Box)

في هذا الاختبار قام الباحثان Ljung & Box في عام 1978 بتعديل اختبار Box & Prece ويستخدم الاختبار للتأكد من عشوائية البواقي بحساب معاملات الارتباط للبواقي لعدد من الازاحات حسب الفرضية [6] :

$$H_0 : P_1 = P_2 = \dots = P_k = P_h \quad \text{for } k= 1, 2, \dots, h$$

$$H_1 : P_h \neq 0$$

بان الاختبار ياخذ بنظر الاعتبار اقتراب او ابتعاد الارتباط الذاتي عن اعلى قيمة h التي دخلت في احتساب إحصاءه الاختبار كما يلي :

$$Q^* = n(n+2) \sum_{k=1}^h (n-k)^{-1} r^2 k$$

ويتم المقارنة مع توزيع مربع كاي من الدرجة $(h-m)$ اي $X^2(h-m)$
حيث ان :

$r^2 k$ = معامل الارتباط الذاتي n = حجم العينة h = عدد الازاحات m = عدد المعالم
فاذا كانت القيمة المحسوبة $Q^* < X^2(h-m)$ نقبل فرضية العدم اي ان البواقي عشوائية غير مرتبطة اما اذا كانت اكبر نرفض فرضية العدم اي ان البواقي غير عشوائية

• التنبؤ

بعد الحصول على القيم (n_a, n_b, n_c) لمعلمات النموذج من خلال طرق التقدير يصبح النموذج كالآتي [2] :

$$y(t) + a_1 y(t-1) + \dots + a_{n_a} y(t-n_a) = b_1 u(t-1) + \dots + b_{n_b} u(t-n_b) + \dots + c_1 e(t-1) + \dots + c_{n_c} e(t-n_c) + e_t$$

اما التنبؤ لخطوة واحدة فقط يكون بالصيغة الاتية :

$$\hat{y}_t(I) = \hat{y}_t(t+1) = -a_1 y(t) - \dots - a_{n_a} y(t-n_a+1) + b_1 u(t) + \dots + b_{n_b} u(t-n_b+1) + c_1 e(t) + \dots + c_{n_c} e(t-n_c+1)$$

وهذه تعتبر الخطوة الاخيرة والمراد تحقيقها من اجل معرفة القيم المستقبلية لحالة صندوق الضمان ودراسة الحالة لتجاوز وتلافي الاخطار التي من الممكن مواجهتها وتعيق الصندوق في الابقاء بالتزاماته

3. تحليل البيانات

• متغيرات الدراسة والعلاقة بينهما

ان الصندوق يتأثر في متغيرين اساسيين يعتمد عليها الصندوق في ايراداته وفي تعويضاته وهذان المتغيران:

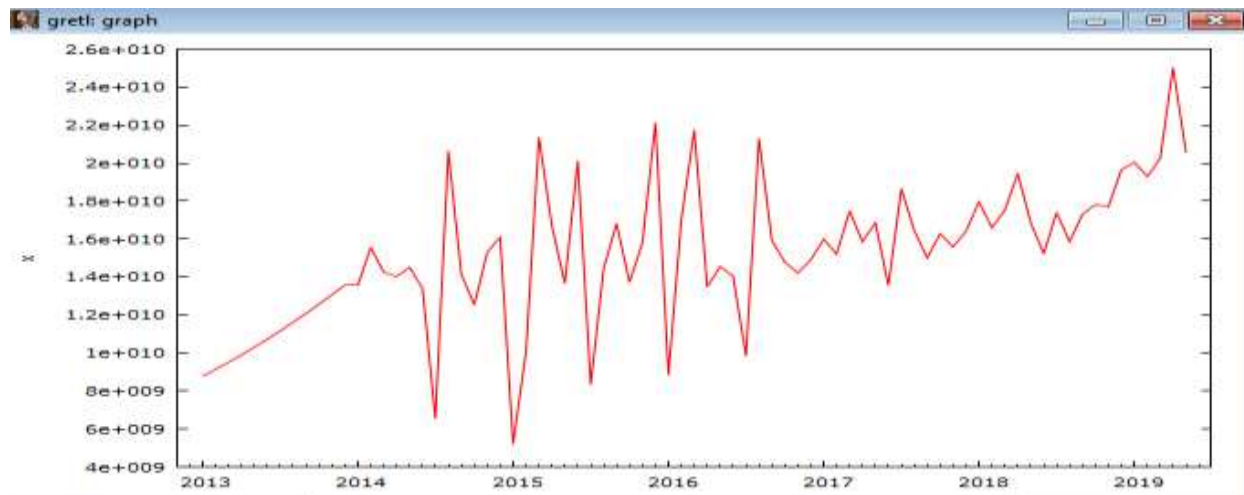
1. الاشتراكات: ان الاشتراكات هي المورد الرئيسي للصندوق وهي تجمع من العمال عبر الية تسديد شهرية ومقدارها 17% من راتب العامل.

2. التعويضات: ان التعويضات تمنح من قبل دائرة الضمان الاجتماعي للعامل عند تحقق احد الاخطار العجز او المرض او الاصابة او الوفاة اثناء العمل. ان العلاقة بين الاشتراكات والتعويضات علاقة تأثير قوية وذلك لان الاشتراكات تؤثر في التعويضات فعند ازياج العمال المشتركين تحت مظلة الضمان الاجتماعي فذلك يعني ازياج منح التعويضات من جراء زيادة عدد المشتركين المشمولين بالضمان الاجتماعي. ولكن زيادة التعويضات لاتؤثر في زيادة الاشتراكات لان منح التعويضات هو من جراء تحقق احد الاخطار التي يغطيها الضمان الاجتماعي وذلك لايحني ان ايرادات الاشتراكات قد زادت وهنا نقول ان الاشتراكات تؤثر في التعويضات ولكن التعويضات لاتؤثر في الاشتراكات .

• مراحل بناء نموذج ARMAX (na,nc,nb,nk)

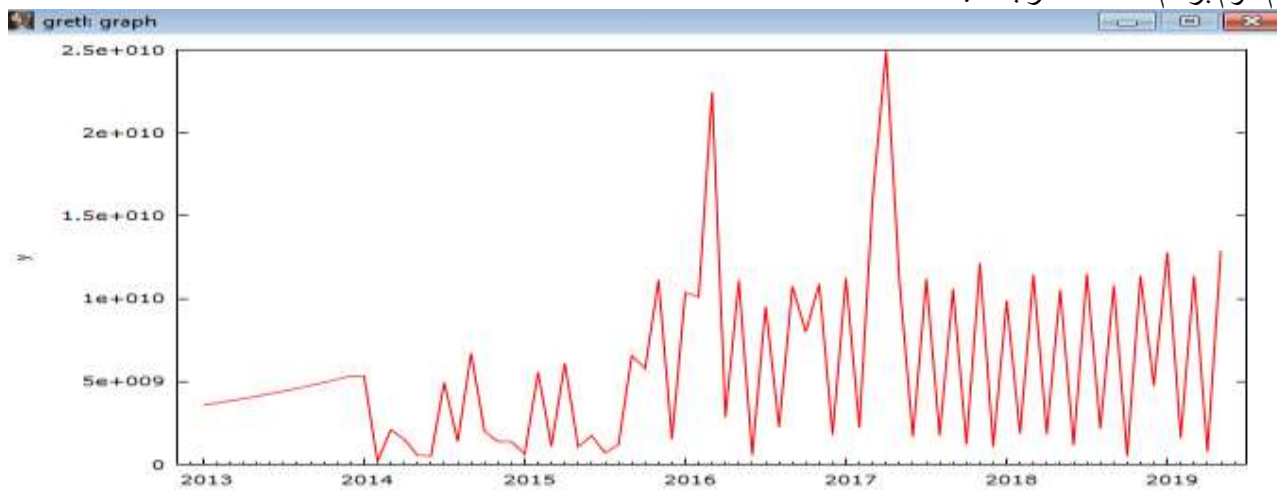
✓ مرحلة التعرف على السلسلة

تمثلت بيانات سلسلة المدخلات من سلسلة الاشتراكات من شهر كانون الثاني 2013 الى شهر حزيران 2019 وتم الحصول على البيانات من وزارة العمل والشؤون الاجتماعية نبدا بالتعرف على السلسلة من خلال الرسم لسلسلة المدخلات:



شكل (2): سلسلة المدخلات من سلسلة الاشتراكات من شهر كانون الثاني 2013 الى شهر حزيران 2019

ثم نقوم برسم سلسلة المخرجات :



شكل (3): سلسلة المخرجات من سلسلة الاشتراكات من شهر كانون الثاني 2013 الى شهر حزيران 2019

✓ تحديد رتب النموذج $ARMAX(na,nc,nb,nk)$

يتم التعرف على النموذج من خلال تحديد رتب na, nc, nb وتم تحديد الرتب وفق برنامج بلغة ماتلاب MATLAB من قبل الباحث وبعد اجراء عملية تبديل للنماذج وايجاد معيار اكايكي لاختيار الافضل من بين النماذج المرشحة المبينة في الجدول الاتي :

جدول (1): مقارنة رتب النماذج المرشحة

مقارنة نماذج $ARMAX(na, nc, nb, nk)$ المرشحة			
MODELS	AIC	FPE	MDL
ARMAX(1,1,1,0)	44.72131121	2.643e+19	44.81195386
ARMAX(1,1,1,1)	44.7475171	2.714e+19	44.83815975
ARMAX(1,1,1,2)	44.73308799	2.675e+19	44.82373064
ARMAX(1,2,1,0)	44.59457014	2.329e+19	44.715427
ARMAX(1,2,1,1)	44.62502459	2.402e+19	44.74588145
ARMAX(1,2,1,2)	44.6893612	2.56e+19	44.81021806
ARMAX(2,1,2,0)	44.24057515	1.635e+19	44.39164623
ARMAX(2,1,2,1)	44.2598595	1.667e+19	44.41093058
ARMAX(2,1,2,2)	44.36098185	1.844e+19	44.51205293
ARMAX(2,2,2,0)	44.24014919	1.634e+19	44.42143448
ARMAX(2,2,2,1)	44.24867802	1.649e+19	44.42996332
ARMAX(2,2,2,2)	44.27107224	1.686e+19	44.45235754
ARMAX(3,1,3,0)	44.29947761	1.735e+19	44.51097712
ARMAX(3,1,3,1)	44.28487881	1.71e+19	44.49637832
ARMAX(3,1,3,2)	44.31590508	1.763e+19	44.52740459
ARMAX(3,2,3,0)	44.27994511	1.702e+19	44.52165884
ARMAX(3,2,3,1)	44.29000267	1.718e+19	44.53171639
ARMAX(3,2,3,2)	44.31121987	1.756e+19	44.5529336

نلاحظ من الجدول ان رتبة النموذج الافضل الذي يمثل العلاقة بين المدخلات والمخرجات هو $ARIMAX(2,2,2,0)$ وذلك لان اقل المعايير هي في هذا النموذج وهي :

$FPE = 1.634e+19, AIC = 44.24014919$ اي ان $na = 2, nc = 2, nb = 2, nk = 0$ من بين عدة نماذج تم تشخيصها للوصول للنموذج الافضل .

✓ مرحلة تقدير النموذج $ARIMAX(2,2,2,0)$

بعد الانتهاء من تحديد رتبة النموذج وتحديد المعايير الاحصائية المذكورة مسبقا يتم تقدير المعلمات باستخدام المربعات الصغرى الاعتيادية OLS وطريقة التكرارية بعامل التغاضي وكانت النتائج وفق متعددات الحدود

➤ طريقة المربعات الصغرى الاعتيادية OLS

$$A(L) = 1 + 0.3236 L^{-1} - 0.713 L^{-2}$$

$$B(L) = 0.06223 + 0.1755 L^{-1}$$

$$C(L) = 1 + 0.5463 L^{-1} - 0.4537 L^{-2}$$

وبهذا يمكننا كتابة النموذج بالشكل الاتي :

$$\hat{Y}(t) = -0.3236 Y_{(t-1)} + 0.713 Y_{(t-2)} + 0.06223 u(t) + 0.1755 u_{(t-1)} + e_{(t)} + 0.5463 e_{(t-1)} - 0.4537 e_{(t-2)}$$

➤ طريقة المربعات الصغرى التكرارية بواسطة عامل التغاضي

كانت نتائج التقدير وفق متعددات الحدود :

$$A(L) = 1 + 0.1294 L^{-1} - 0.2187 L^{-2}$$

$$B(L) = 0.07029 + 0.335 L^{-1}$$

$$C(L) = 1 - 0.1992 L^{-1} + 0.1673 L^{-2}$$

وبهذا يمكننا كتابة النموذج بالشكل الاتي :

$$\hat{Y}(t) = -0.1294 Y_{(t-1)} + 0.2187 Y_{(t-2)} + 0.07029 u(t) + 0.335 u_{(t-1)} + e_{(t)} - 0.1992 e_{(t-1)} + 0.1673 e_{(t-2)}$$

جدول (2): نتائج التقديرات لمعاملات نموذج ARIMAX

طريقة التقدير	ϕ_1	ϕ_2	θ_1	θ_2	ϕ_1	ϕ_2
OLS	0.3236	- 0.713	0.5463	- 0.4537	0.06223	0.1755
RLS-EF	0.1294	- 0.2187	- 0.1992	0.1673	0.07029	0.335

✓ مقارنة طرق التقدير

تمت مقارنة طرق التقدير لنموذج ARMAX(2,2,2,0) باستعمال مقياس الجذر التربيعي لمتوسط مربعات الخطأ RMSE ومقياس الخطأ المطلق المئوي MAPE بعد ايجاد الاخطاء عبر برنامج ماتلاب من اعداد الباحث للطرق الثلاثة وكانت النتائج كما ياتي :

جدول (3): يوضح مقارنة طرق التقدير

طريقة التقدير	RMSE	MAPE
OLS	3766961640	1.376553632
RLS-EF	4817340001	2.225197022

نلاحظ ان الطريقة الافضل حسب الجدول اعلاه هي طريقة المربعات الصغرى الاعتيادية OLS وفق مقاييس RMSE , MAPE

✓ النموذج المقترح

يتم التنبؤ داخل السلسلة بواسطة طريقة OLS لنموذج ARIMAX(2,2,2,0) الى 76 مشاهدة وايجاد الاخطاء ثم نقوم بتمهيد البيانات وفق التمهيد الاسي لهولت لايجاد الخطأ المقدر بمعلمة:

$$\alpha = 0.127197942938255 \quad \gamma = 0.034859518448255 \quad \text{وكالاتي:}$$

$$a(t) = \alpha e t + (1 - \alpha) (a_{t-1} + b_{t-1})$$

$$b(t) = \gamma (a_t - a_{t-1}) + (1 - \gamma) b_{t-1}$$

$$a_1 = e_1$$

$$b_1 = \frac{(e_2 - e_1) + (e_4 - e_3)}{2}$$

$$\hat{e}(t) = a(t) + b(t)w$$

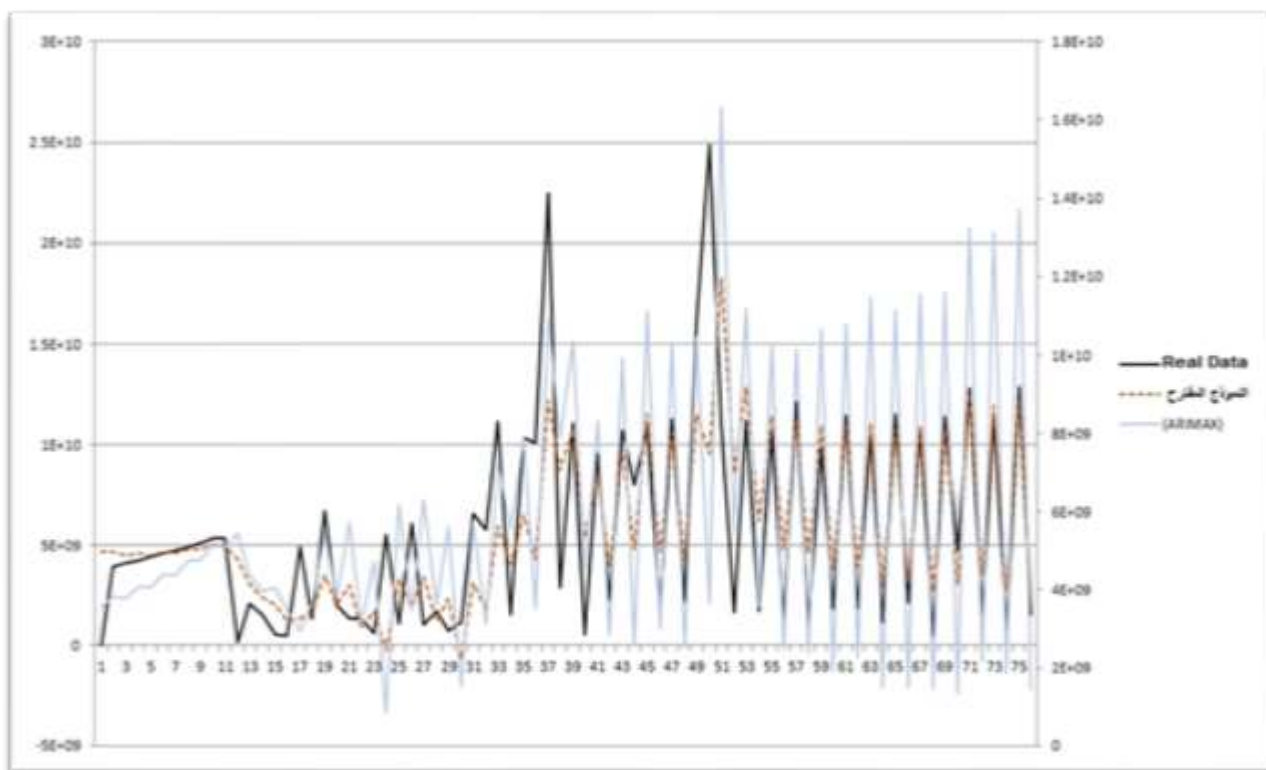
نقوم بدمج التقديرات بطريق OLS لنموذج ARMAX مع الخطأ بنموذج واحد لمعرفة تأثير الخطأ على النموذج وكالاتي :

$$\hat{y} = B_0 + B_1 \hat{y}(\text{ARMAX}) + B_2 \hat{e}(\text{holt})$$

وبعد تقدير معاملات النموذج يصبح النموذج :

$$\hat{y} = 1655281862 + 0.80142957 \hat{y}(\text{ARMAX}) + 1.291183164 \hat{e}(\text{holt})$$

ثم نقوم برسم البيانات الاصلية مع البيانات التي تم التنبؤ بها بواسطة نموذج ARMAX OLS والنموذج المقترح لمقارنة النتائج من خلال الرسم :



شكل (4): علاقة البيانات الاصلية مع البيانات التي تم التنبؤ بها

ومن ثم نقارن بالجزر التربيعي لمتوسط مربعات الخطأ للنموذج المقترح وطريقة نموذج المربعات الصغرى ARIMAX الاعتيادية :

$$RMSE\ ARMAX = 3796063751 \quad RMSE\ Holt\ Reg = 3385941523$$

في نموذج التنبؤ وكما نلاحظ في الرسم اعلاه ان قيم التنبؤ بالنموذج المقترح هو الاقرب للقيم الاصلية. $RMSE$ نجد ان النموذج المقترح هو الافضل لأنه يملك قيم اقل.

✓ اختبار Ljung –Box

لغرض اختبار عشوائية النموذج المقترح والجدول الاتي يوضح نتائج اختبار القيم التي تم تقديرها :

جدول (4): يوضح اختبار Ljung –Box

TEST	Q* (Lag)	Q* - Statistic	$X^2(h - m)$	p-value
Ljung –Box	Q*(8)	3.551121522	3.84146	0.059505
Ljung –Box	Q*(10)	5.619058725	7.81473	0.131688
Ljung –Box	Q*(12)	6.77931899	11.0705	0.237577

ويتضح من الجدول ان القيمة المحسوبة اصغر من القيمة الجدولية $X^2(h - m)$ وهذا يعني قبول فرضية العدم بان البواقي للنموذج عشوائية وهذا يدل على ملائمة النموذج للتنبؤ .

✓ التنبؤ بخطوة الى الامام

ثم نقوم بالتنبؤ لخطوة الى الامام بواسطة النموذج المقترح ونموذج ARIMAX وكالاتي :

جدول (5): مقارنة القيمة التنبؤية

القيمة الحقيقية	ARIMAX(OLS)	النموذج المقترح
10,784,790,130	14,732,238,590	12,962,186,249

الخطأ المطلق e	النموذج المقترح	
	ARIMAX(OLS)	
	3,947,448,460	2,177,396,119

4. مناقشة النتائج

بعد تحديد رتبة النموذج وتقدير المعلمات بطريقة المربعات الصغرى العادية والتكرارية وتفوق طريقة OLS واختباره للملاءمة بطريقة OLS تم تطبيق طريقة المربعات الصغرى الاعتيادية OLS للتنبؤ بنموذج ARMAX وتحسين النموذج من خلال اقتراح نموذج يحسن التنبؤ عن طريق تقدير الخطأ المقدر بالتمهيد الاسي لمعلمتي هولت ويعدها من خلال ادخال الخطأ المقدر كعامل مؤثر بنموذج انحدار على مخرجات النموذج نلاحظ من مقارنة القيمة الحقيقية مع التنبؤية بطريقة OLS والنموذج المقترح ان النموذج المقترح اقل فرقا وهذا يدل على انه الاقرب الى القيمة الحقيقية وانه الطريقة الافضل للتنبؤ وبهذا يتم اعتماد النموذج المقترح للتنبؤ بالتعويضات لصندوق الضمان الاجتماعي لتحقيق التوازن بين الاشتراكات المستحقة والتعويضات الممنوحة للعمال المضمونين.

5. الاستنتاجات Conclusions

بعد بناء نموذج ARMAX واجراء عملية التنبؤ توصل الباحث من خلال النتائج التي تم الحصول عليها في الجانب التطبيقي الى عدة استنتاجات وكالاتي :
تم تحديد رتبة النموذج باستخدام ثلاثة معايير وهما (AIC , FPE,MDL) والتي اتفق اثنان منها على نفس الرتبة وتفوق طريقة المربعات الصغرى OLS على طريقة المربعات الصغرى التكرارية RLS بعامل التغاضي، وتم تحسين نموذج ARIMAX(2,2,2,0) وفق النموذج المقترح ثم اختبار جودة نموذج ARIMAX (2,2,2,0) بواسطة اختبار Ljung-Box , وكانت النتائج ملائمة و تمت مقارنة القيم الاصلية مع القيم التي تم التنبؤ بها لنموذج ARIMAX (2,2,2,0) والنموذج المقترح وتم اعتماد النموذج المقترح لأنه اعطى افضل واقترب الى القيم الاصلية. وهذا يدل على دقة التنبؤ وملاحظة وجود تزايد في منح التعويضات من اموال الصندوق وذلك من جراء تحقق الاخطار والوضع الاقتصادي المتذبذب وهذا يسبب انخفاض راس المال للصندوق وذلك قد يندر بخطر في المستقبل على امكانية ايفاء الصندوق بالتزاماته المالية وحدث عجز، وهذا يؤدي الى عدم تحقيق التوازن بين الاشتراكات والتعويضات مستقبلا .

6. الاعمال المستقبلية

استثمار راس المال لصندوق التامينات الاجتماعية وفق دراسات لاحقة لتعظيم راس المال ودراسة حالة الصندوق بصورة دورية كل ثلاثة او ستة اشهر واعتماد نموذج ARMAX للتنبؤ للجهات الرقابية واصحاب القرار وتوسيع مظلة الضمان الاجتماعي من خلال زيادة الشمول للعمال من ضمنهم العقود والاجور اليومية في جميع دوائر الدولة من اجل ضمان تحقيق احتياطي لمجابهة الاخطار وتحقيق التوازن نظرا لازيد اعداد المتقاعدين وتأمين رواتبهم على المدى البعيد وايضا هناك دراسة لقانون الضمان الاجتماعي من اجل تعديله ليشمل العمال في القطاع غير المنتظم والضمان الاختياري الذي سيعظم الايرادات بسبب زيادة اعداد المشتركين من العمال .

المصادر References

- [1] Adekanmbi Damilola Bolanle & Akinyemi Oluwadare, (2017), "Arima And Arimax Stochastic Models For Fertility In Nigeria", International Journal of Mathematics and Computer Applications Research, Vol. 7, Issue 5.
- [2] Chi-man & Jacob (2008), "Statistical modeling and Forecasting Schemes for Air-Conditioning System", Pao Yue-Kong library, Hong kong Polytechnic University, Hong Hom, Kowloon, Heng Kong .
- [3] Gautam, Bijaya (2010), "Spectral Estimation of Electroencephalogram Signal Using ARMAX Model and Particle Swarm Optimization", The Faculty of College of Graduate Studies Lamar University .
- [4] Hamed Shakouri G. & Aliyeh Kazemi, (2015), "Selection of the best ARMAX model for forecasting energy demand: case study of the residential and commercial sectors in Iran", Energy Efficiency , University of Tehran
- [5] Kumar P.R. & Varaiya Pravin, (1986), "Stochastic Systems : Estimation, Identification, and Adaptive Control " , Prentice-Hall, Inc., Englewood Gifts, New Jersey .
- [6] Ljung Greta, Box E.P, (1978), "On a measure of lack of fit in time series models " .Biometrika, Vol. (65), No(2) , pp. (297-303)
- [7] Ljung Lennart, (1999), "System Identification : Theory for the User", 2nd edition, Linkoping University, Sweden, Prentice Hall International

- [8] Morata, A. & Loira, L., (2017), Yeast-Industrial Applications, Intech, Croatia, 1st edition.
- [9] Nelles Oliver, (2001), Nonlinear System Identification, Springer – Verlag, 1st edition.
- [10] Richard T. Baillie, (1979), "Predictions From ARMAX Models", Wayne State University, Detroit, MI 48202, USA University of Aston, Birmingham 84 6TE, UK , Journal of Econometrics 12 (1980) 365-374. 0 North-Holland Publishing Company.
- [11] Saeed Matroushi, (2011), "Hybrid Computational Intelligence Systems based on Statistical and Neural Networks Methods for Time Series Forecasting: The Case of Gold Price , M. Sc. Thesis, Lincoln University.
- [12] Suleman, N. & Albert .L & Lea .A., (2013), "The Efficacy of ARIMAX and SARIMA Model in Predicting Monthly Currency in Circulation in Ghana", Mathematical Theory and Modeling Vol. 3, No. 5.
- [13] Soderstrom Torsten & Stoica Peter, (2001), System Identification, Prentice Hall International.
- [14] Wellstead, P.E & Zarrop, M.B (1991), Self – Tuning Systems : Control and Signal Processing, Chichester , UK:John Wiley & Sons