مقارنة بين الطرائق الاعتيادية (LS,4SIV) و الطرائق الحصينة (2SWLS,LTS,RA) لتقدير معلمات أنموذج ARX(1,1,1) للأحمال الكهربائية

أ.م.د فراس احمد محمد / كلية الادارة والاقتصاد / جامعة بغداد
الباحث / زينب حامد حميد /

تاريخ التقديم:2018/5/9
تاريخ القبول:2018/6/24

المستخلص

تعاني نماذج السلاسل الزمنية في اغلب الاحيان من مشكلة وجود القيم الشاذة التي ترافق عملية جمع البيانات ولأسباب عديدة والتي قد يسبب وجودها تأثيراً كبيراً على تقدير معلمات الأنموذج المدروس، حيث ان الحصول على مقدرات تمتلك كفاءة عالية يعتبر من اهم مراحل التحليل الاحصائي، وعليه يجب الاهتمام باختيار الطرائق المناسبة للحصول على مقدرات جيدة. وان الهدف من هذا البحث هو اجراء المقارنة بين المقدرات الاعتيادية والمقدرات الحصينة لتقدير معلمات أنموذج الانحدار الذاتي بوجود متغير خارجي (ARX) ذات الرتبة (1,1,1) باستعمال بيانات حقيقية تحتوي على القيم الشاذة حيث تم استعمال الرتبة (1,1,1) اعتماداً على عدد من معايير تحديد الرتبة والتي تم توضيحها ضمن الرسالة قيد الانشاء، وقد تبين لنا من خلال الدراسة ان الطريقة التي تم توظيفها وهي طريقة المربعات الصغرى المشذبة (LTS) هي افضل طريقة تقدير، حيث تمت عملية المقارنة باستعمال الجذر التربيعي لمتوسط مربعات الخطأ (RMSE) و متوسط مطلق الخطأ النسبي (MAPE) ونسبة الخطأ المتوقعة (EEP)، كما تم اجراء اختبار للتأكد من دقة الأنموذج الذي تم التوصل اليه وبعدها تم استعماله للتنبؤ بالقيم المستقبلية.

المصطلحات الرئيسة للبحث/ أنموذج الانحدار الذاتي بوجود متغير خارجي (ARX)، الكشف عن القيم الشاذة، طريقة المربعات الصغرى الاعتيادية (OLS)، طريقة المتغيرات المساعدة ذات الخطوات الاربعة (4SIV)، طريقة المربعات الصغرى الموزونة ذات المرحلتين (2SWLS)، طريقة المربعات الصغرى المشذبة (LTS)، مقدرات التباينات الذاتية للبواقي(RA).

مجلة العلوم
الاقتصادية والإدارية
العدد 109 المجلد 24
الصفحات 496ـ 514

*البحث مستل من رسالة ماجستير

المبحث الاول/ المقدمة وهدف البحث:

يعد موضوع تحليل السلاسل الزمنية من المواضيع الاحصائية المهمة والتي تُعتمد لتفسير الكثير من الظواهر الحياتية المهمة المرتبطة بالفترات الزمنية ولاسيما تلك التي تكون فتراتها متساوية (يومية، شهرية، فصلية، سنوية)، وقد شملت مجالات واسعة اقتصادية وطبية وبيئية وغيرها، ويعتبر أنموذج السلسلة الزمنية بأنه الدالة التي تربط القيم الحالية للمتغير بالقيم السابقة له وبالأخطاء وبالإضافة الى ذلك بقيم سلسلة زمنية اخرى.

يتم بناء أنموذج السلاسل الزمنية من خلال عدد من المراحل تبدأ اولاً بمرحلة تهيئة البيانات ومن ثم تحديد الرتبة للأنموذج ومن ثم تقدير معلماتهِ بعدها يتم اختبار دقة الأنموذج واخيراً مرحلة التنبؤ بالقيم المستقبلية، وتعد مرحلة التقدير (Estimation) احد المراحل المهمة في بناء أنموذج السلسلة الزمنية ومن اكثر الطرائق المستعملة في تقدير معلمات نماذج السلاسل الزمنية هي الطرائق الاعتيادية ومنها طريقة المربعات الصغرى (LS) وطريقة الامكان الاعظم (ML) وغيرها، وان مقدرات هذهِ الطرائق قد تكون كفوءة ومناسبة في حالة توفر شروط السلاسل الزمنية ، ولكن عند اختلاف الشروط نتيجة عامل خارجي يطرأ على السلسلة الزمنية فأنه من الضروري البحث عن طرائق تقدير اخرى تستطيع التعامل مع السلسلة الزمنية التي لاتتوفر فيها الشروط المطلوبة وهذهِ الطرائق تعرف بالطرائق الحصينة، فمن المعروف ان طرائق التقدير الاعتيادية تكون حساسة لوجود القيم الشاذة (Outliers) في البيانات.

يهدف هذا البحث الى بناء أنموذج لأحمال الذروة الكهربائية بالاعتماد على متغير خارجي متمثل بدرجات الحرارة العظمى من خلال فلترة البيانات التي تحتوي على قيماً شاذة ومن ثم المقارنة بين مختلف طرائق التقدير الاعتيادية والحصينة للأنموذج بالاعتماد على مجموعة من مقاييس المقارنة واستعمال الطريقة الافضل التي يتم التوصل اليها للتنبؤ بالقيم المستقبلية للأحمال الكهربائية.

المبحث الثاني/ الجانب النظري

1- أنموذج الانحدار الذاتي بوجود متغير خارجي (ARX):

(Autoregressive With Exogenous Input Model)

عبارة عن أنموذج انحدار ذاتي ((Autoregressive (AR) (القيم السابقة للمُخرج) مع متغير خارجي (متغير اخر غير المُخرج) او هو عبارة عن انموذج يتضمن نظام من المدخلات والتي تتألف من المتغير المؤثر (u) مع مجموعة من ارتدادات هذا المتغير فضلا عن بعض ارتدادات المتغير التابع او متغير الاستجابة (y) ومخرج يمثل المتغير المعتمد (y_t)، ويمكن صياغة الصيغة العامة لهذا الانموذج من خلال معادلة الفرق الخطية التالية:[II:pp.81][III:pp.219-220]

$$y_t = -a_1 y_{t-1} - \cdots - a_{n_a} y_{t-n_a} + b_1 u_{t-n_k} + \cdots + b_{n_b} u_{t-n_k-n_b+1} + e_t \qquad \dots (1)$$

حيث ان :

y_t يمثل مُخرج الانموذج او المتغير المعتمد .

u_t يمثل مُدخل الانموذج او المتغير التوضيحي .

e_t يمثل حد الخطأ .

t يمثل مؤشر الزمن .

n_a تمثل رتبة الانحدار الذاتي (AR) من الانموذج (عدد معلمات متعدد الحدود (A(q) .

n_b تمثل رتبة المُدخل (u_t) وارتداداته في الانموذج (عدد معلمات متعدد الحدود (B(q) .

n_k يمثل تأخير المُخرج (y_t) بالنسبة الى المُدخل (u_t).

ويمكن كتابة أنموذج ARX بشكل عام (أنموذج انحدار) وكما يأتي :

$$y_t = \varphi_t^T \ \theta + e_t$$

حيث ان

$$\varphi_t = [-\mathrm{y}_{t-1} \cdots - \mathrm{y}_{t-n_a} \ \mathrm{u}_{t-n_k} \cdots \mathrm{u}_{t-n_k-n_b+1}]^{\mathrm{T}}$$

$$\theta = [a_1 \cdots a_{n_a} \ b_1 \cdots b_{n_b}]^T$$

اذ ان

φ_t يمثل متجه القياسات (عناصر الانحدار) بالبعد $(n_a + n_b) \times 1$.

θ يمثل متجه المعلمات بالبعد $(n_a + n_b) \times 1$.

2- القيم الشاذة للسلسلة الزمنية: (Time Series Outliers)

قد تتأثر مشاهدات السلسلة الزمنية في بعض الاحيان بأحداث غير عادية، او اضطرابات، او الاخطاء التي تُنتج تأثيرات زائفة في السلسلة وتُنتج انماط استثنائية في المشاهدات والتي لا تتفق مع معظم المشاهدات في السلسلة الزمنية. ويطلق على هذه المشاهدات الغير عادية اسم القيم الشاذة. وقد تكون نتيجة احداث خارجية غير عادية مثل الإضرابات، والحروب والازمات الاقتصادية المفاجئة وموجات الحرارة او البرد غير المتوقعة، او بسبب اخطاء الكتابة والتسجيل التي لا يمكن ملاحظتها. وان اول من درس الكشف عن القيم الشاذة في السلاسل الزمنية هو الباحث (Fox) في عام (1972) حيث قدم نوعين من القيم الشاذة هما: [IV:pp.536][V:pp.223-228]

القيم الشاذة المضافة (Additive Outliers): ويقصد بها القيم التي تؤثر في مشاهدة واحدة فقط وليس لها تأثير على بقية المشاهدات بحيث ان السلسلة بعدها تعود الى مسارها الطبيعي.

القيم الشاذة المُجددة (Innovation Outliers): وهي القيم التي تؤثر في المشاهدة التي تقع عندها وكذلك المشاهدات التي تليها او هي القيم الشاذة التي ينتشر تأثيرها في المشاهدات التي تليها.

2-1 الكشف عن القيم الشاذة: (The Outliers Detection)

يتم استعمال قاعدة (three sigma) للكشف عن القيم الشاذة والتي قدمت من قبل الباحث (Friedrick Pukelsheim) في عام 1992 ومن بعده من قبل الباحث (Patrick D.Spagon) في عام 1997. اذ ان قاعدة (3σ) او طريقة الدرجة المعيارية عبارة عن اداة تستعمل على نطاق واسع في الابحاث لغرض الكشف عن القيم الشاذة وان هذا المصطلح يعد مصطلح عام لبعض اختبارات الفرضيات الاحصائية التي تختبر البواقي. وان الشروط التي بموجبها تم اثبات هذهِ القاعدة احصائياً تم تحليلها والتحقق في مدى تطبيقها على المربعات الصغرى والتأكد من كفاءتها. وان قاعدة (3σ) تفترض تقدير المشاهدة او القيمة الشاذة اذا كان الاحتمال للقيمة المشاهدة خارج المدى [μ-3σ,μ+3σ] وتحسب من الصيغة الاتية:[VI][I:pp.52]

$$Z_i = \frac{x_i - \bar{x}}{s} \qquad \ldots (2)$$

اذ ان

x_i : قيمة المشاهدة.

$\bar{x}$: متوسط السلسلة او متوسط المشاهدات.

S : الانحراف المعياري.

فاذا كانت $|Z_i| > 3$ تعد المشاهدة x_i مشاهدة شاذة (outlier).

3- طرائق التقدير:

1-3 طريقة المربعات الصغرى الاعتيادية (ols):

(Ordinary Least Squares Method)

ان الغرض من المربعات الصغرى الاعتيادية (OLS) هو تقليل مربعات الخطأ كما في المعادلة الاتية: [III:pp.220]

$$e^{T}e = (y-\Phi\theta)^{T}(y-\Phi\theta) = y^{T}y - 2y^{T}\Phi\theta^{T} + \theta^{T}\Phi^{T}\Phi\theta$$

وعند اخذ المشتقة الاولى بالنسبة الى المعلمة θ ومن ثم مساواة المشتقة الى الصفر ينتج الاتي:

$$\frac{\partial e^{T}e}{\partial\theta} = -2\Phi^{T}y + 2\Phi^{T}\Phi\hat{\theta} = 0$$

$$2\Phi^{T}y = 2\Phi^{T}\Phi\hat{\theta}$$

وبأخذ المعكوس الى المصفوفة $(\Phi^{T}\Phi)$ يتم الحصول على مقدر المربعات الصغرى الاعتيادية (OLS) والذي يعطى بالمعادلة الاتية:

$$\hat{\theta} = (\Phi^{T}\Phi)^{-1}\Phi^{T}y \qquad \dots(3)$$

2-3 طريقة المتغيرات المساعدة ذات الخطوات الاربعة (4SIV):

(Four-Step Instrumental Variables Method)

ن طريقة (4SIV) هي الامتداد الى طريقة المتغيرات المساعدة (IV) والتطوير لها لتعطي مقدرات للمعلمات بتباين مقارب (Asymptotic Variance) اقل من التباين المقارب لطريقة المتغيرات المساعدة (IV) ومساوي الى حدود Cramer-Rao. ويمكن الوصول الى مقدرات طريقة (4SIV) من خلال تطبيق خوارزمية تتكون من الخطوات الاربعة الاتية: [II:pp.487]

Step 1: يتم تقدير θ بأستعمال طريقة المربعات الصغرى (LS) ويرمز للتقدير ب $\hat{\theta}^{(1)}$ ودالة التحويل المقابلة للتقدير ب $\hat{G}^{(1)}(q)$.

Step 2: توليد المساعدات كما يأتي:

$$x_t^{(1)} = \hat{G}^{(1)}(q)u_t \qquad \dots(4)$$

$$\zeta_t^{(1)} = [-x_{t-1}^{(1)} \cdots - x_{t-n_a}^{(1)}\; u_{t-n_k} \cdots u_{t-n_k-n_b+1}]^{T}$$

ثم يتم ايجاد تقدير θ بطريقة المتغيرات المساعدة (IV) واستعمال المساعدات ${\zeta_t}^{(1)}$ ويرمز للتقدير ب $\hat{\theta}^{(2)}$ ودالة التحويل المقابلة للتقدير ب $\hat{G}^{(2)}(q)$ وتكون بالصيغة الاتية:

$$\hat{G}^{(2)}(q) = \frac{\hat{B}^{(2)}(q)}{\hat{A}^{(2)}(q)}$$

Step 3: ليكن

$$\hat{W_t}^{(2)} = \hat{A}^{(2)}(q)y_t - \hat{B}^{(2)}(q)u_t$$

ويفترض انموذج انحدار ذاتي AR ذات رتبة $(n_a + n_b)$ ل $\hat{W_t}^{(2)}$ ويكون كما يأتي:

$$L(q)\hat{W_t}^{(2)} = e_t$$

ويتم تقدير L(q) باستعمال طريقة المربعات الصغرى (LS) ويرمز للناتج ب $\hat{L}(q)$.

Step 4: يتم تعريف ${x_t}^{(2)}$ بشكل مماثل الى ${x_t}^{(1)}$ وليكن

$${\zeta_t}^{(2)} = \hat{L}(q)[-{x_{t-1}}^{(2)} \cdots - {x_{t-n_a}}^{(2)}\ u_{t-n_k} \cdots u_{t-n_k-n_b+1}]^T$$

وباستعمال المساعدات ${\zeta_t}^{(2)}$ يتم الوصول الى المقدرات النهائية للمعلمات والتي تكون بالصيغة الاتية :

$$\hat{\theta} = \left[\sum_{t=1}^{n} {\zeta_t}^{(2)} \varphi_t^T(F)\right]^{-1} \sum_{t=1}^{n} {\zeta_t}^{(2)} y_t(F) \qquad \dots(5)$$

حيث ان

$$\varphi_t(F) = \hat{L}(q)\varphi_t \qquad و \qquad y_t(F) = \hat{L}(q)y_t$$

3-3 طريقة المربعات الصغرى الموزونة ذات المرحلتين (2SWLS):[VII:pp.49]

(Two-Stage Weighted Least Squares Method)

تستعمل طريقة المربعات الصغرى الموزونة ذات المرحلتين لتقدير معلمات الأنموذج حيث يتم في المرحلة الاولى انشاء مصفوفة الوزن الاولية وتمييز القيم الشاذة باستعمال المربعات الصغرى المعدلة الوزن تكراريأ **(Iteratively reweighted least squares) (IRWLS)** مع دالة ترجيح **(bisquare)** **(The bisquare weighting function)**.

اما في المرحلة الثانية سيتم اعادة وزن مصفوفة الوزن الاولية وذلك باستعمال قيم مختلفة من عامل التغاضي الاسي (The Exponential Forgetting Factor)، والذي يكون فيه قيمة λ عدد موجب واقل من الواحد $(0 < \lambda < 1)$ علماً ان القيم المثالية او النموذجية ل λ هي $(\lambda = 0.95, \cdots, 0.99)$ وان عامل التغاضي الاسي يعطي اهمية نسبية للبيانات الحالية عن البيانات السابقة وهذا يدل على انه في حالة زيادة n يتم التغاضي عن البيانات السابقة [XII:pp.52]. يتم حساب مجموع مربعات الاخطاء باستعمال معاملات الانحدار تلك مع القيم المختلفة من عامل التغاضي الاسي (The Exponential Forgetting Factor). والمعاملات المثلى مع مجموع مربعات الاخطاء الصغرى تستعمل للتنبؤ. وتضمن المرحلة الثانية هذهِ تقدير معامل له القدرة على التكيف.

وبشكل عام، أنموذج الانحدار لعدد من الحالات n يمكن ان يكون مكتوب بشكل مصفوفات كما يأتي

$$y = \Phi\theta + e$$

حيث ان

y يمثل متجه المُخرج او المتغير المعتمد $(n \times 1)$ و $y = [y_1, y_2, \cdots, y_n]^T$.

Φ يمثل مصفوفة المُدخل او المتغير المستقل $(n \times (n_a + n_b))$. ويمكن ان يكون Φ مكتوب ايضاً كما $[\varphi_1, \varphi_2, \cdots, \varphi_n]^T$ ، حيث $\varphi_1, \varphi_2, \cdots, \varphi_n$ تمثل متجهات بالبعد $((n_a + n_b) \times 1)$ لكل حالة.

θ يمثل متجه معاملات الانحدار $((n_a + n_b) \times 1)$.

e يمثل متجه الاخطاء $(n \times 1)$ مع الانحراف المعياري σ.

ان طريقة المربعات الصغرى التقليدية تقلل مجموع مربعات الاخطاء وتأخذ الصيغة الاتية:

$$\hat{\theta} = (\Phi^T\Phi)^{-1}\Phi^T y$$

واذا تم اضافة وزن لكل مربع خطأ، فأن انحدار المربعات الصغرى التقليدية يصبح انحدار المربعات الصغرى الموزونة وكما في المعادلتين الاتية :

$$min\ z = \sum_{t=1}^{n} W_{tt} e_t^2 = \sum_{t=1}^{n} W_{tt} (y_t - \varphi_t^T \theta)^2$$

$$\hat{\theta} = (\Phi^T W \Phi)^{-1}\Phi^T W y \qquad \dots (6)$$

حيث ان $W = diag(W_{11}, W_{22}, \cdots, W_{nn})$.

اذا اشرنا الى $\Phi' = w\Phi$ و $y' = wy$ مع $w = (\sqrt{W_{11}}, \sqrt{W_{22}}, \cdots, \sqrt{W_{nn}})^T$ ، سيكون لدينا شكل جديد للمربعات الصغرى مع مدخل موزون Φ' و مُخرج موزون y' .

$$y' = \Phi'\theta' + e' \qquad \dots (7)$$

$$\min z' = \sum_{t=1}^{n} e_t'^2 = \sum_{t=1}^{n} (y_t' - \varphi_t^{T\prime} \theta')^2 \qquad \dots(8)$$

$$\hat{\theta}' = (\Phi'^T \Phi')^{-1} \Phi'^T y' \qquad \dots(9)$$

في المرحلة الاولى من اجراء المربعات الصغرى الموزونة ذات المرحلتين، يتم حساب مصفوفة الوزن الاولية من خلال **IRWLS** حيث يركز نهج **IRWLS** للانحدار الحصين على تقليل مجموع مربعات اخطاء التقديرات عن طريق استعمال دالة الوزن التي تقلل من تأثير نقاط البيانات البعيدة (الشاذة). وسوف نستخدم هنا دالة **bisquare** كدالة وزن. حيث ان دالة الوزن **bisquare** تم اقتراحها من قبل **Tukey** عام **(1974)** كالاتي :

$$W_{tt}(r_t) = \begin{cases} (1 - l_t^2)^2 & |l_t| \le 1 \\ 0 & otherwise \end{cases} \qquad \dots(10)$$

حيث ان

$$l_t = \frac{(y_t - \hat{y}_t)}{cS\sqrt{1 - h_t}}$$

وان

h_t يمثل الرافعة **(Leverage)** والتي تكون عبارة عن العناصر القطرية للمصفوفة $\Phi(\Phi^T\Phi)^{-1}\Phi^T$ **[XIII:pp.225]**،
c=4.685 تمثل ثابت الضبط اختيرت على اساس كفاءة 95% من الانحدار الحصين الى طريقة المربعات الصغرى،

و $S = \frac{MAD}{0.6745}$ حيث ان MAD هو الوسيط لمطلق انحراف البواقي $MAD = median|e_i - median(e_i)|$.

يبدأ اجراء الحل مع التقديرات الاولية ل θ و σ. في كل تكرار يتم تطبيق اجراء المربعات الصغرى الموزونة وتقديرات θ و σ ويتم تحديث r_t. ويستمر الاجراء حتى التقارب. حيث يكون معيار التقارب كالاتي :

$$\frac{(\sqrt{W}\theta)^T(\sqrt{W}\sigma r)}{\|\sqrt{W}\theta\|_2 \|\sqrt{W}\theta\|_2} \qquad \dots(11)$$

حيث ان ‖.‖ هو مبدأ **Euclidean**

في المرحلة الثانية، يتم استعمال مصفوفة الوزن الجديدة $W' = diag[W_{tt}'] = diag[\lambda^{n-t}]$ في المربعات الصغرى الموزونة مرة اخرى للمشكلة الجديدة التي وصفت من خلال المعادلات (7) و(8) و(9). لذلك، تعرف مشكلة المربعات الصغرى الجديدة على النحو الاتي :

$$y'' = \Phi''\theta'' + e'' \qquad \dots(12)$$

$$\min z'' = \sum_{t=1}^{n} e_t''^2 = \sum_{t=1}^{n} (y_t'' - \varphi_t^{T''} \theta'')^2 \qquad \dots(13)$$

$$\hat{\theta}'' = (\Phi''^T \Phi'')^{-1} \Phi''^T y'' \qquad \dots(14)$$

حيث ان $\Phi'' = w'\Phi'$ و $y'' = w'y'$.

ومن ثم، يتم تقدير معاملات الانحدار النهائية من خلال المربعات الصغرى الموزونة ذات المرحلتين.

4-3 مقترح توظيف طريقة المربعات الصغرى المشذبة في أنموذج ARX (LTS):

(proposal of Using Least Trimmed Squares Method in ARX)

طور Rousseeuw عام (1984) طريقة تقدير المربعات المشذبة الصغرى (LTS) وهي احدى طرائق تقدير الانحدار الحصين، حيث ان المقدر الناتج من هذهِ الطريقة يدعى بمقدر المربعات الصغرى المشذبة والذي يأخذ الرمز (LTS)، وهذه الطريقة تعطى من خلال الصيغة الاتية: [VIII:pp.413]

$$\hat{\theta}_{LTS} = \arg\min Q_{LTS}(\theta) \qquad \dots(15)$$

حيث ان

$$Q_{LTS} = \sum_{i=1}^{m} e_i^2 \qquad \dots(16)$$

$e_1^2 \le e_2^2 \le \dots \le e_m^2$ تمثل مربعات البواقي المرتبة من الاصغر الى الاكبر. يتم حساب LTS عن طريق تقليل مربعات البواقي المرتبة m،

حيث ان

$$m = \left[\frac{n}{2}\right] + \left[\frac{(h+1)}{2}\right]$$

n تمثل حجم العينة.

h تمثل عدد المعلمات.

في هذه الطريقة يتم استبعاد مربعات البواقي الاكبر من عملية الجمع، والتي تسمح الى ان تكون نقاط البيانات المتطرفة تلك مستبعدة تماماً عن مقاييس المقارنة للتقدير. اعتمادا على قيمة m وترتيب البيانات المتطرفة (الشاذة) وفي الواقع، ان طريقة LTS يمكن ان تكون كفوءة جداً، اذا شذبت الارقام المضبوطة لنقاط البيانات المتطرفة، وهذه الطريقة مكافئة حسابياً الى طريقة OLS. كذلك اذا كانت نقاط البيانات الشاذة التي يتم تشذيبها كثيرة، فأن هذه الطريقة تكون غير كفوءة. على العكس من ذلك، اذا كان التشذيب اكبر من نقاط البيانات المتطرفة (الشاذة)، يتم استبعاد بعض البيانات الجيدة من الحساب.

5-3 مقدرات التباينات الذاتية للبواقي (RA):[IX:PP.52-61]

(Residual Auto-covariances Estimators)

تم اقتراح مقدرات **RA** من قبل العالمان **Bustos** و **Yohai** في عام **(1986)** لنماذج **ARMA**. واظهروا انه من الممكن ان تظهر مقدرات المربعات الصغرى **(LS)** بشكل يمكن ان يتضمن على عينة التباينات الذاتية الاعتيادية للبواقي. ثم اقترحوا مقدرات **RA** كحل للنظام حيث يتم استبدال التباينات الذاتية الاعتيادية هذهِ بأخرى حصينة.

ومن الممكن صياغة خوارزمية عامة لحل مقدرات **RA** فبالنسبة للدالة η من نوع **Mallows** يمكن وضع مخطط تكراري مبسط للمربعات الصغرى. ويمكن تعديل نفس الخوارزمية العامة لمقدرات **RA** ل **ARX** **(RA-ARX)** حيث يكون التركيز على تعديل الخوارزمية التكرارية للدالة η من نوع **Mallows** والتي تأخذ الصيغة التالية:

$$\eta_M(U,V) = \psi(U)\psi(V) \qquad Mallows$$

ومثال على الدالة ψ هي دالة **bisquare** والتي تكون بالشكل الاتي:

$$\psi_B(U;\vartheta) = \begin{cases} U\left(1-\frac{U^2}{\vartheta^2}\right)^2 & if \quad 0 \le |U| \le \vartheta \\ 0 & elsewhere \end{cases}$$

حيث ان $\vartheta = 4.685$.

ليكن

$$r_t^* = \psi\left(\frac{r_t(\bar{\theta})}{\hat{\sigma}_e}\right)\hat{\sigma}_e$$

$$y_t^* = \frac{\hat{B}(q)}{\hat{A}(q)}u_t + \frac{1}{\hat{A}(q)}r_t^*$$

كذلك وان

$$\hat{\sigma}_e = \frac{med\left(|r_{n_a+1}|, \dots, |r_n|\right)}{0.6745}$$

$$r_t(\bar{\theta}) = \hat{A}(q)y_t - \hat{B}(q)u_t$$

$$A = [a_1, \dots, a_{n_a}]^T$$

$$B = [b_1, \dots, b_{n_b}]^T$$

$$\theta = [A^T, B^T]^T$$

بافتراض ان السلسلة الزمنية المشاهدة للفترة هي y_t^* مع المتغير الخارجي u_t.

وان الخوارزمية الحسابية التكرارية التالية تستعمل لحساب مقدرات (RA) وهي كما يأتي:

Step 1: لتكن $\hat{\theta}^{(0)}$ و $\hat{\sigma}_{\varepsilon}$ مقدرات اولية، ويمكن ان تكون هذهِ المقدرات على سبيل المثال مقدرات LS المعتادة.

Step 2: يتم حساب $\hat{\theta}^{(n)}$ و $\hat{\sigma}_{\varepsilon}{}^{(n)}$ مقابلة للتكرار n كذلك حساب

$$r_t^{*(n)} = \psi\left(\frac{r_t\left(\hat{\theta}^{(n)}\right)}{\hat{\sigma}_{\varepsilon}{}^{(n)}}\right)\hat{\sigma}_{\varepsilon}{}^{(n)}$$

واستعمالها في حساب الصيغة الاتية:

$$y_t^{*(n)} = \frac{\hat{B}^{(n)}(q)}{\hat{A}^{(n)}(q)}u_t + \frac{1}{\hat{A}^{(n)}(q)}r_t^{*(n)}$$

Step 3: ثم تحسب مقدرات LS للسلسلة $y_t^{*(n)}$ مع متغيرات خارجية u_t.

Step 4: يتم تكرار الخطوتين 2 و 3 حتى يتم التوصل الى التقارب.

4- مقاييس المقارنة

تتم عملية المقارنة بين طرائق التقدير من خلال المقاييس الاتية:

4-1 الجذر التربيعي لمتوسط مربعات الخطأ (RMSE):

وهو عبارة عن الجذر التربيعي لمجموع مربعات الفرق بين القيمة الحقيقية والقيمة التنبؤية مقسوماً على حجم العينة والذي يتم حسابه وفق الصيغة الاتية:[X:pp.24]

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(\hat{y}_t - y_t)^2}{n}} \qquad \cdots (17)$$

4-2 متوسط مطلق الخطأ النسبي (MAPE):

هو عبارة عن متوسط مطلق نسبة الاخطاء الى القيم الحقيقية ويكتب بالصيغة الاتية:
[V:PP.182][XI:pp.403]

$$MAPE = \frac{\sum_{t=1}^{n}\left|\frac{\hat{y}_t - y_t}{y_t}\right|}{n} \qquad \dots (18)$$

4-3 نسبة الخطأ المتوقعة (EEP):

وهو المقياس الذي يعتمد على الجذر التربيعي لمتوسط مربعات الخطأ مقسوماً على اكبر قيمة حقيقية ويتم حسابه وفق الصيغة الاتية:[VII:pp.50]

$$EEP = \frac{RMSE}{Max_{t=1,2,\cdots,n}\{y_t\}} \qquad \dots (19)$$

حيث ان

y_t : تمثل القيم الحقيقية.

$\hat{y}_t$: تمثل القيم التقديرية.

n : تمثل عدد البيانات الحقيقية.

وان طريقة التقدير الافضل هي التي تقابل اقل قيمة للمقاييس الثلاثة.

5- اختبار دقة الأنموذج:

سيتم التحقق من مصداقية او دقة الأنموذج عن طريق اجراء اختبارين هما اختبار معنوية الارتباط الذاتي بين سلسلة الاخطاء e_t واختبار معنوية الارتباط الذاتي بين سلسلة الاخطاء e_t وسلسلة المُدخل u_t كما يأتي:

[II:pp.509-514]

1-5 اختبار Whiteness: (Whiteness Test)

يتضمن اختبار Whiteness اختبار فرضية العدم والفرضية البديلة الاتية:

$H_0: R_e = 0$

$H_0: R_e \neq 0$

اذ ان

H_0 **: يمثل فرضية العدم.**

H_1 **: يمثل الفرضية البديلة.**

R_e **: يمثل الارتباط الذاتي بين الاخطاء.**

ويمكن تقدير R_e عند الارتداد T كالاتي:

$$\hat{R}_e(T) = \frac{1}{n}\sum_{t=1}^{n} e_t e_{t-T} \qquad \cdots (20)$$

وبالتالي فان احصاءة الاختبار ستكون بالصيغة الاتية:

$$\zeta = \frac{n}{(\hat{R}_e(0))^2}\sum_{T=1}^{m} (\hat{R}_e(T))^2 \qquad \cdots (21)$$

اذ تتم مقارنة قيمة $\chi^2(m)$ بدرجة حرية m ومستوى معنوية α مع قيمة ζ فاذا كان

$\zeta \leq \chi^2(m)$

ففي هذهِ الحالة ينبغي قبول الفرضية العدم والتي تنص على عدم وجود ارتباط ذاتي بين عناصر سلسلة الخطأ او يكون الارتباط غير معنوي اما اذا كان

$\zeta > \chi^2(m)$

فعندها ينبغي قبول الفرضية البديلة والتي تعني وجود الارتباط الذاتي بين عناصر سلسلة الخطأ.

2-5 اختبار الاستقلالية بين الاخطاء والمدخلات السابقة:
(Independence Between Residuals And Past Inputs)

من اجل اختبار الاستقلالية بين الاخطاء والمُدخل يتم وضع واختبار كل من فرضية العدم والفرضية البديلة والتي تكون بالشكل الاتي:

$$H_0: R_{\varepsilon u} = 0$$

$$H_1: R_{\varepsilon u} \neq 0$$

اذ ان

H_0 **: يمثل فرضية العدم.**

H_1 **: يمثل الفرضية البديلة.**

$R_{\varepsilon u}$**: يمثل الارتباط بين حد الخطأ والمُدخل حيث يمكن تقديره عند الارتداد الزمني T كما يأتي:**

$$\hat{R}_{\varepsilon u}(T) = \frac{1}{n}\sum_{t=1}^{n} e_t u_{t-T} \qquad \cdots (22)$$

ففي حالة تحقق اختبار البياض اي عدم وجود ارتباط بين عناصر سلسلة الخطأ عندها يكون:

$$\sqrt{n}R_{\varepsilon u}^2(T) \sim N(0,P)$$

حيث تعرف

$$P = \sum_{k=-\infty}^{\infty} R_{\varepsilon}(k) R_u(k)$$

وان كل من $R_{\varepsilon}(k)$ **و** $R_u(k)$ **تكون بالشكل الاتي:**

$$R_{\varepsilon}(k) = E e_t e_{t-k}$$

$$R_u(k) = \bar{E} u_t u_{t-k}$$

وفي حالة $N_{\propto}$ **يدل على التوزيع** $N(0,1)$ **عند المستوى** $\propto$ **يمكننا بالتالي التحقق من**

$$|R_{\varepsilon u}(T)| \leq \sqrt{\frac{P}{n}} N_{\propto} \qquad \cdots (23)$$

فاذا تحققت المتراجحة اعلاه عندها ينبغي قبول الفرضية العدم والتي تنص على ان الارتباط بين الخطأ والمُدخل ارتباط غير معنوي اما في حالة عدم تحقق المتراجحة عندها ينبغي قبول الفرضية البديلة والتي تبين بأن الارتباط بين الخطأ والمُدخل ارتباطاً معنوياً.

ان الطريقة المناسبة لتنفيذ الاختبار هي رسم $\hat{R}_{\varepsilon u}(T)$ كدالة ل T وان حدود الثقة ستكون بشكل خطوط افقية، مثل هذا الرسم يعطي معلومات قيمة حول صحة هيكل الأنموذج.

6- التنبؤ: (Forecasting)

بعد ما تم تحديد معلمات أنموذج ARX عن طريق استعمال طرائق التقدير التي تم ذكرها سابقا يتم البدء بمناقشة كيف يمكن التنبؤ بالقيم المستقبلية لظاهرة معينة حيث تعد عملية التنبؤ من الاهداف المهمة في عملية بناء السلاسل الزمنية ودراسة السلوك المستقبلي للظاهرة. ويتم التنبؤ لأنموذج ARX بأخذ التوقع الشرطي للأنموذج ويكون كما يأتي:[II:pp.64-71]

لكي يتم التنبؤ بخطوة واحدة الى الامام ل y يكون باستعمال الصيغة الاتية:

$$H(q)\hat{y}_{t|t-1} = G(q)u_t + [H(q) - 1]y_t \qquad \cdots (24)$$

اذ ان

$$G(q) = \frac{B(q)}{A(q)} \quad , \quad \mathrm{H}(q) = \frac{1}{A(q)}$$

وان

$$A(q) = 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a}$$

$$B(q) = b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b+1}$$

المبحث الثالث/ الجانب التطبيقي

1- المقدمة

من اجل تحقيق الهدف من البحث سوف يتم في هذا المبحث تطبيق الخطوات التي تم توضيحها في المبحث الثاني باستعمال برامج تم كتابتها بلغة (Matlab) على بيانات حقيقية تتمثل بمتغيرين، المتغير الاول (y_t) ويعبر عن سلسلة المخرجات ويتمثل بأحمال الذروة الكهربائية اليومية لمحافظة بغداد والمتغير الثاني (u_t) ويعبر عن سلسلة المدخلات ويتمثل بدرجات الحرارة العظمى لمحافظة بغداد، حيث كانت البيانات اليومية لفصل الصيف والذي يضم الاشهر (6-7-8) من السنوات (2013-2014-2015-2016) اي بعدد (368) مشاهدة يومية، وقد تم الحصول على البيانات من وزارة الكهرباء/دائرة توزيع الطاقة.

2- تفسير النتائج

2-1 اختبار استقرارية البيانات والكشف عن القيم الشاذة:

تم اختبار استقرارية البيانات من خلال تطبيق اختبار Dickey-Fuller الموسع (Augmented Dickey-Fuller test (ADF)) باستعمال برنامج (gretl) والجداول الاتية تبين نتائج الاختبار:

جدول (3-1) يبين P-value لاختبار (ADF) لكل من سلسلة المخرجات y_t وسلسلة المدخلات u_t

مستوى المعنوية	سلسلة المتغير	P-value لاختبار Dickey-Fuller الموسع		
		بدون الحد الثابت	بوجود الحد الثابت	بوجود الحد الثابت والاتجاه الزمني
0.05	y_t	0.7329	0.2867	0.1559
	u_t	0.8181	0.0048	0.008294

الجدول (3-1) يبين نتائج اختبار (ADF) لكل من سلسلة المخرجات y_t وسلسلة المدخلات u_t، اذ تم استعمال الحلات الثلاثة للاختبار وهي حالة عدم وجود الحد الثابت وحالة وجود الحد الثابت وحالة وجود كل من الحد الثابت والاتجاه الزمني معاً، ومن خلال مقارنة القيم الاحتمالية (p) في الجدول اعلاه مع القيمة (0.05) نستنتج ان الاختبار غير معنوي اي ان السلسلتين غير مستقرتين لذلك سيتم معالجة بيانات السلسلتين من خلال اخذ اللوغاريتم واخذ الفروق الاولى، والجدول الاتي يوضح نتائج اختبار (ADF) بعد معالجة البيانات:

جدول (3-2) يبين P-value لاختبار (ADF) لكل من سلسلة المخرجات y_t وسلسلة المدخلات u_t بعد اخذ (log) والفروق الاولى

مستوى المعنوية	سلسلة المتغير	P-value لاختبار Dickey-Fuller الموسع		
		بدون الحد الثابت	بوجود الحد الثابت	بوجود الحد الثابت والاتجاه الزمني
0.05	d-l-y_t	4.043e-013	5.281e-012	2.509e-011
	d-l-u_t	4.937e-021	3.988e-021	3.51e-022

ومن خلال القيم في الجدول (3-2) اعلاه نستنتج ان السلسلتين اصبحت مستقرة بعدما تمت معالجتها.

ويتم تطبيق قاعدة (σ3) للكشف عن القيم الشاذة الموجودة في كل من سلسلة المدخلات وسلسلة المخرجات بعد اخذ اللوغاريتم والفروق الاولى للسلسلتين لتحقيق شروط الاستقرارية، حيث تم تحديد القيم ذات التسلسل (18, 19, 294, 295, 318, 319, 332, 333) في سلسلة (y_t) والمتمثلة بالقيم (1425, 2120, 1425, 3450, 1825, 3550, 1877, 3500) وذات التسلسل (215,276,325) في سلسلة (u_t) والمتمثلة بالقيم (46,35,50).

2-2 تقدير معلمات الأنموذج:

تم تقدير معلمات الأنموذج باستعمال الطرائق (OLS) و (4SIV) و (2SWLS) و (LTS) و (RA)، والجدول (3-3) ادناه يوضح نتائج المعلمات المقدرة للأنموذج حسب كل طريقة من طرائق التقدير وكما يأتي:

جدول (3-3) يبين قيم المعلمات المقدرة للأنموذج

طرائق التقدير	قيم المعلمات	
	a	B
OLS	-0.9933	0.1107
4SIV	-0.9993	0.2512
2SWLS	-0.9931	0.2898
LTS	-0.9999	-7.27e-06
RA	-0.9933	-0.1706

3-2 مقارنة طرائق التقدير:

يتم مقارنة طرائق التقدير عن طريق استعمال مقاييس المقارنة الموضحة في المبحث الثاني وهي (RMSE) و (MAPE) و (EEP) والنتائج موضحة في الجدول الاتي:

جدول (3-4) يبين نتائج مقاييس المقارنة بين طرائق التقدير للأنموذج

طرائق التقدير	مقاييس المقارنة		
	RMSE	MAPE	EEP
OLS	0.0134	0.0067	0.0049
4SIV	0.0116	0.0047	0.0042
2SWLS	0.0183	0.0081	0.0067
LTS	3.33e-07	1.32e-07	1.21e-07
RA	0.1841	0.0619	0.0671

ومن خلال مقاييس المقارنة اعلاه والمتمثلة بمقياس الجذر التربيعي لمتوسط مربعات الخطأ (RMSE) ومقياس مطلق الخطأ النسبي (MAPE) ومقياس نسبة الخطأ المتوقعة (EEP) تبين لنا ان طريقة المربعات الصغرى المشذبة (LTS) هي افضل طريقة تقدير حيث جاءت في المرتبة الاولى اعتماداً على مقاييس المقارنة الثلاثة ومن ثم جاءت طريقة (4SIV) بالمرتبة الثانية ومن ثم طريقة (OLS) وتليها طريقة (2SWLS) وجاءت طريقة (RA) في المرتبة الاخيرة.

وبذلك فأن الصيغة التقديرية للأنموذج وفق طريقة (LTS) تكون بالشكل الاتي:

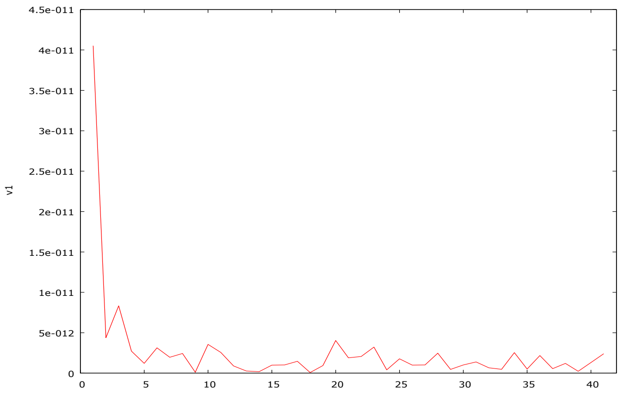
$$y_t = (0.9999)y_{t-1} - (7.27e^{-6})u_{t-1} + e_t$$

3-3 اختبار مصداقية الأنموذج:

يتم اختبار مصداقية الأنموذج عن طريق التحقق من الفرض الاساسي الذي تم افتراضه والذي ينص على ان حد الخطأ يتوزع توزيعاً طبيعياً بوسط صفر وتباين ثابت وهذا يعني انه تشويش ابيض، حيث يتم التحقق من هذا الفرض عن طريق اجراء الاختبارين التاليين:

*** يتم في الاختبار الاول حساب قيم البواقي (e_t) واختبار معنوية الارتباط بين قيم البواقي، حيث كانت القيمة المحسوبة هي (5.18) يتم مقارنتها مع القيمة الجدولية ل χ^2 بدرجة حرية (40) ومستوى معنوية (0.05) والتي هي (26.51)، وبما ان القيمة المحسوبة اقل من القيمة الجدولية فيتم قبول الفرضية العدم والتي تنص على عدم وجود ارتباط بين عناصر سلسلة الخطأ.**

* اما في الاختبار الثاني فيتم فيه اختبار الارتباط بين الخطأ والمُدخل وذلك من خلال حساب المعادلة (23) والشكل التالي يوضح نتيجة الاختبار حيث ان حدود الثقة هي (0.0000018) و (0.0000018-).

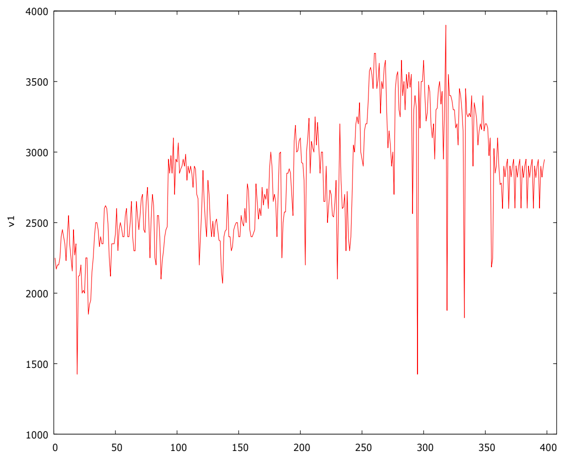


شكل (3-1) يبين قيم الارتباط بين حد الخطأ (e_t) والمُدخل (u_t)

وبما ان جميع قيم الارتباط بين الخطأ والمُدخل ضمن حدود الثقة فيتم قبول الفرضية العدم اي ان الارتباط يكون غير معنوي. وبما ان النتائج اعلاه التي تم التوصل اليها بالنسبة لطريقة (LTS) والتي تتضمن عدم وجود ارتباط معنوي بين قيم الاخطاء مع بعضها البعض من جهة وحد الخطأ والمُدخل من جهة اخرى، وهذا يؤكد صحة افتراض بياض الخطأ (white noise).

3-4 التنبؤ:

وبعد ان تم انهاء عملية بناء الأنموذج والتحقق من مصداقيته، حيث تبين انه يحقق جميع الشروط الواجب توفرها لذلك يمكن استعماله في عملية التنبؤ لثلاثين قيمة مستقبلية بالنسبة للمتغير y_t، حيث تم الحصول على ثلاثين قيمة مستقبلية بالنسبة للمتغير u_t واستعمالها لغرض التنبؤ وباستعمال طريقة التقدير (LTS) وكما في الشكل التالي:



شكل (3-2) يوضح البيانات الحقيقية والبيانات التي تم التنبؤ بها للمتغير (y_t)

المبحث الرابع/ الاستنتاجات والتوصيات

ومن خلال نتائج البحث التي تم الحصول عليها تم التوصل الى عدد من الاستنتاجات والتوصيات التي يعتقد انها جديرة بالاهتمام.

4-1 الاستنتاجات:

1- تم اجراء اختبار استقرارية السلاسل الزمنية للمتغيرات التي تناولها البحث باستعمال احد اختبارات جذر الوحدة وهو اختبار (ADF)، وتم التوصل الى استقرارية سلسلة المتغيرات بعدما تمت معالجة البيانات من خلال اخذ اللوغاريتم والفروق الاولى.

2- تفوقت طريقة (LTS) والتي جاءت في المرتبة الاولى على الطرائق (4SIV) و (OLS) الاعتيادية والتي جاءت في المرتبة الثانية والثالثة على التوالي وجاءت الطرائق (2SWLS) و (RA) الحصينة في المراتب الاخيرة من خلال استعمال مقاييس المقارنة والتي تتمثل ب (RMSE,MAPE,EEP) وفقاً للبيانات المدروسة والتي تحتوي قيماً شاذة.

3- تم اختبار جودة أنموذج (ARX) باستعمال طريقة (LTS) الحصينة لتقدير المعلمات حيث بينت النتائج عدم وجود ارتباط ذاتي بين عناصر الخطأ وكذلك عدم وجود ارتباط بين الخطأ والمُدخل.

4- ان اعلى حمل كهربائي يمكن توقعه باستعمال أنموذج ARX(1,1,1) خلال 30 يوم هو (2949) واقل قيمة حمل يمكن توقعها هي (2601).

4-2 التوصيات:

1- نوصي بالمقارنة بين طرائق التقدير الاعتيادية وطرائق التقدير الحصينة لأنموذج ARX باستعمال اسلوب المحاكاة.

2- نوصي بدراسة أنموذج ARX متعدد المُدخلات ومخرج واحد (MISO) وأنموذج ARX متعدد المُدخلات ومتعدد المُخرجات (MIMO) عند وجود مشكلة تلوث البيانات.

3- نوصي بمقارنة أنموذج ARX مع بقية نماذج الصندوق الاسود في حالة وجود مشكلة التلوث في البيانات.

4- نوصي وزارة الكهرباء بالأخذ بنظر الاعتبار أن ما سيتم توقعه بأعلى معدل للاستهلاك قدره (2949) ميكاواط واقل معدل استهلاك قدره (2601) ميكاواط ضمن المدة التي تم التنبؤ بها (شهر حزيران من عام 2017).

المصادر العربية والاجنبية:

I- عنبر، جنان عبد الله (2016) "مقارنة بعض المقدرات البيزية الحصينة مع مقدرات اخرى لانموذج GARCH(1.1) مع تطبيق عملي" اطروحة مقدمة إلى مجلس كلية الادارة والاقتصاد / جامعة بغداد للحصول على درجة "دكتوراه فلسفة في علوم الإحصاء".

II- Ljung, L. (1999) "System Identification : Theory FOR THE User" 2nd Edition, Prentice Hall PTR, Upper Saddle River, NJ.

III- Silva, A. J., Neto, J. V. F., & Nagem, N. F. (2009) "Parametric ARX Modeling of the Electrolytic Smelter Pot" IEEE, pp. 217-222.

IV- BOX, G. E. P., JENKINS, G. M. & REINSEL, G. C. (2008) "Time Series Analysis: Forecasting and Control" 4th Edition, John Wiley & Sons, Inc.

V- Wei, W. W. S. (2006) "Time Series Analysis: Univariate and Multivariate Methods" 2nd Edition, Pearson Education, Inc.

VI- Lehman, R. (2013) "The 3σ-Rule for Outlier Detection from the Viewpoint of Geodetic Adjustment" Journal of Surveying Engineering, Vol. 139, (4), pp.157-165.

VII- Guo, Y., Nazarian, E., Ko, J., & Rajurkar, K. (2014) "Hourly cooling load forecasting using time-indexed ARX models with two-stage weighted least squares regression" Energy Conversion and Management, Vol. 80, pp. 46-53.

VIII- Alma, Ö. G. (2011) "Comparison of Robust Regression Methods in Linear Regression" Int. J. Contemp. Math. Sciences, Vol. 6, No. 9, pp.409-421.

IX- Duchesne, P. (2001) "ROBUST AND POWERFUL SERIAL CORRELATION ESTS WITH NEW ROBUST ESTIMATES IN ARX MODELS" JOURNAL OF TIME SERIES ANALYSIS, Vol. 26, No. 1, PP.49-81.

X- Awaludin, I., Ibrahim, R., & Rao, K. S. R. (2009) "Conventional ARX and Artificial Neural Networks ARX Models for Prediction of Oil Consumption in Malaysia" IEEE, Symposium on Industrial Electronics and Applications, pp.23-28.

XI- Othman, M. F. B., & Yusoff, M. K. B. M. (2009) "System Identification to Forecast Electricity Loads" IEEE, SCOReD, PP.403-406.

XII- Mahmoud, M. S., & Xia, Y. (2012) "Applied Control Systems Design" New York, Springer.

XIII- Rousseeuw, P. J. & Leroy, A. M. (1987) "Robust Regression Outlier Detection" John Wiley & Sons, Inc.

Comparison Between Ordinary Methods (LS,IV) and Robust Methods (2SWLS,LTS,RA) to estimate the Parameters of ARX(1,1,1) Model for Electric Loads

Abstract:

The models of time series often suffer from the problem of the existence of outliers that accompany the data collection process for many reasons, their existence may have a significant impact on the estimation of the parameters of the studied model. Access to highly efficient estimators is one of the most important stages of statistical analysis, And it is therefore important to choose the appropriate methods to obtain good estimators. The aim of this research is to compare the ordinary estimators and the robust estimators of the estimation of the parameters of the Autoregressive with exogenous variable (ARX) model with the order of (1,1,1) using real data containing outliers, the order (1,1,1) has been used based on a number of criteria for determining the rank, which were explained in the thesis under construction. The study showed that the method employed The Least Trimmed Squares (LTS) method is the best method of estimation. The comparison was done using the Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and Expected Error Percentag (EEP), A test was also carried out to ascertain the accuracy of the model reached and then used to predict future values.

Key Words: **(ARX); The Outliers Detection; Ordinary Least Squares Method; Four-Step Instrumental Variables Method; Two-Stage Weighted Least Squares Method; Least Trimmed Squares Method; Residual Auto-covariances Estimators.**