Iraqi Journal Of Statistical Sciences

المجلة العراقية للعلوم الإحصائية

www.stats.mosuljournals.com

معالجة عدم استقرارية السلسلة الزمنية – مراجعة مقال–

الاء عبد الستار حمودات ، .ندى نزار محمد و زينب توفيق الدباغ

قسم الاحصاء والمعلوماتية، كلية علوم الحاسوب والرياضيات، جامعة الموصل، الموصل، العراق

معلومات النشر

تاريخ المقالة:

تم استلامه في 27 شباط 2022

تم القبول في 9 نيسان 2022

متاح على الإنترنت في 1 حزيران 2022

الكلمات الدالة:

السلاسل الزمنية،منهجية بوكس جنكينز ،الاستقرارية.

المراسلة:

الاء عبد الستار حمودات

alla-hamoodat@uomosul.edu.iq

الخلاصة

تعتبر السلاسل الزمنية غير المستقرة مشكلة في تحليل الاقتصاد القياسي حيث ان الخواص الاحصائية لتحليل السلاسل تفقد عند استخدام سلاسل زمنية غير مستقرة. يهدف البحث الى تقديم عدة طرق لمعالجة الاستقرارية ومنها (Box Jenkins ، Exponential Smoothing ، Double Exponential Smoothing ، Exponential Smoothing ، Moving Averages ، Fuzzy ، Neural Network) والمقارنة بين الطرق المقدمة من خلال تشخيص نماذج ARIMA بعد تحقق الاستقرارية واختيار الطريقة الافضل التي تقابل القيم الاقل للمعايير الاحصائية (MSE , AIC , BIC). ولقد تم تطبيق الطرق المذكورة اعلاه على بيانات يومية لعام 2020 لتوليد الكهرباء الناتج من الماء الوارد من نهر دجلة وتم التوصل الى ان طريقة (Fuzzy) هي الافضل لمعالجة الاستقرارية مقارنة بالطرق الاخرى لامتلاك النموذج ARIMA (0,1,3) المقابل لها اقل القيم للمعايير الاحصائية (MSE=0.572 , AIC=-196.4536 , BIC=-0.6931).

DOI: https://doi.org/10.33899/iqjoss.2022.174335 , ©Authors, 2022, , College of Computer and Mathematical Science, University of Mosul.
This is an open access article under the CC BY 4.0 license (http://creativecommons.org/licenses/by/4.0/).

مقدمة Introduction

من أهم المشاكل التي تواجه الباحثين عند القيام بتحليل سلسلة زمنية هو أستقرارية السلسلة من عدم أستقراريتها وانطلاقا من هذه الأهمية القصوى فقد ارتأينا معالجة عدم الاستقرارية في بيانات السلاسل الزمنية أولا وهو الأسلوب المتعارف عليه في معظم دراسات نماذج الانحدار الذاتي والمتوسطات المتحركة المتكاملة .(الصفاوي والطائي،2003) يمكن أن يؤثر على دقة النموذج الرياضي الذي يروم كل باحث الوصول أليه بأقل خطأ ممكن . تعد النماذج المستقرة Stationary Models من أهم الأصناف التي لقيت اهتماما من الباحثين، حيث تفترض هذه النماذج ثبات الوسط الحسابي للسلسلة الزمنية، وفي حالة عدم تحقق هذا الشرط تتحول هذه النماذج إلى نماذج غير مستقرة وهي السمة الغالبة لكثير من السلاسل الزمنية المستخدمة في قطاعي الصناعة والأعمال.

1- السلسلة الزمنية Time Series

تعرف السلسلة الزمنية بانها مجموعة من المشاهدات لظاهرة معينة خلال فترة زمنية غالباً ما تكون متساوية ومتتالية، وتعرف رياضياً بأنها متتابعة من المتغيرات العشوائية معرفة ضمن فضاء الاحتمالية متعدد المتغيرات ومؤشرة بالدليل (t) الذي يعود إلى مجموعة دليليه (T) ويرمز للسلسلة الزمنية عادة بالرمز $\{Z_t \setminus t \in T\}$ وان هذه المتتابعة تتوزع تبعاً لدالة التوزيع المشترك لمجاميع المتغيرات العشوائية $\{Z_t \setminus t \in T\}$ وهي $(Z_{t1}, Z_{t2}, \ldots, Z_{tn})$

ويمكن تمثيل السلسلة الزمنية بوضع الزمن على المحور الافقي وقيم الظاهرة على المحور الرأسي فأذا أظهرت السلسلة الزمنية إتجاهاً معيناً او طابعاً خاصاً خلال فترة ممتدة من الزمن فأننا نتوقع أن يستمر حدوث هذا الطابع أو الأنتظام في المستقبل أيضاً وذلك يعتبر أساساً معقولاً للتنبؤ. (Miljanovic , 2012).

2- استقرارية السلسلة الزمنية: Stationarity of Time Series

يقال للسلسلة الزمنية بأنها مستقرة (Stationary) إذا لم يكن هناك نمو أو انحراف في البيانات اي بمعنى عدم ظهور اتجاه عام أي تبعثر البيانات أفقيا حول متوسط ثابت ، وكذلك التباين يكون ثابتا عبر الزمن (Makridakis et al. , 1998) ، ويمكن القول بان السلسلة المستقرة التي تكون مبنية على أساس إفتراض أن السلسلة في حالة خاصة من الموازنة الإحصائية أي امتلاكها وسطاً حسابياً وتبايناً ثابتين مع استمرار الزمن، عندها يقال أن السلسلة مستقرة في الوسط والتباين. وتكون مستقرة إذا لم يكن هناك إتجاه إلى الأعلى أو الأسفل في المعدل عبر الزمن أو عدم ظهور اختلاف حول الوسط. المقصود بالإستقرارية إنّ مشاهداتها تتذبذب عشوائياً حول المتوسط وهناك نوعين من الاستقرارية وهما: [ثامر والشرابي ،2020]

4 – عدم استقرارية السلسلة الزمنية Non-Stationarity of Time Series

إن معظم الظواهر التطبيقية والعملية في السلاسل الزمنية تتصف بخاصية عدم الاستقرارية ومن احد اسباب عدم استقرارية السلسلة هو وجود اتجاه عام ووجود تقلبات موسمية وعدم استقرار التباين أو الوسط الحسابي عبر الزمن . لذلك يطلق عليها سلسلة غير مستقرة ، حيث تعد حالة عدم الاستقرارية في السلسلة مشكلة رئيسية، ألا أن هنالك حلولاً يمكن من خلالها تحويل السلاسل غير المستقرة إلى سلاسل زمنية مستقرة عن طريق الدراسة المتأنية للعوامل التي تؤثر فيها وتجعلها غير مستقرة. (المحمدي و طعمة ، 2011)

5 – معالجة عدم الاستقرارية في السلاسل الزمنية Treatment of Non – Stationarity in Time Series

إن الخطوة الأولى في تحليل أية سلسلة زمنية هي الرسم البياني لمشاهدات السلسلة مع الزمن ، حيث يُظهر الرسم الملامح الوصفية للبيانات مثل وجود اتجاه عام والبيانات الشاذة أو زيادة تشتت بيانات هذه السلسلة مع مرور الزمن أي أن تباين المشاهدات يزداد مع زيادة مستوى السلسلة ، نستنتج من ذلك بأن السلسلة الزمنية غير مستقرة Non–Stationary لذلك يستحسن تحسين أو تعديل السلسلة أي التخلص من مشكلة عدم الاستقرارية وهذا يتطلب أخذ بعض الإجراءات لجعل السلسلة مستقرة وهذه الإجراءات هي(فاندل ، 1983) :-

5-1 بوكس جنكيز Box-Jenkins :

تعد هذه الطريقة من أهم طرائق المعالجة وتسمى طريقة الفروق Differencing ، إن طريقة الفروق المتتالية Consecutive Differencing هي طريقة عظيمة الفائدة لوجود أنموذج لاتجاه عام عشوائي في عدد كبير من السلاسل الزمنية التجارية والاقتصادية . وعلى الرغم من وجود فهم عام لمعنى الاتجاه العام ، فأنه من الصعب إعطاء تعريف أكثر دقة من التعريف القائل بأن الاتجاه العام هو(تغير منتظم في مستوي السلسلة الزمنية) . ويعد هذا أفضل تعريف موجود للاتجاه العام على الرغم من عدم توافر الدقة الرياضية له . وترجع صعوبة تعريف الاتجاه العام إلى أن مشاهدات سلسلة زمنية قصيرة قد تظهر ما يعتقد انه تغير في مستوى السلسلة ، ولكن عند الحصول على بيانات سلسلة أطول يتبين لنا ان ما شاهدناه كان جزءاً من تغيرات دورية ولم يكن اتجاهاً عاماً .

5-2 نموذج ممداني المضبب Mamdani Fuzzy Model:

تم اقترح هذا النموذج من قبل العالم E.H.Mamdani في عام 1974 . فلنفرض أن لدينا عدداً محدداً من القواعد ولتكن $R1, R2, \dots, Rr$ وكانت $Ai1, Ai2, \dots, Ain, Bi1, \dots, Bim$ هي مجموعات مضببة متعلقة بالقاعدة R_i حيث إن $i=1,2,\dots,r$. إن نموذج ممداني في نظام الاستدلال يكون جزء الإخراج من القواعد عبارة عن متغيرات لغوية مضببة تعرف من خلال دالة العضوية الخاصة بكل منها ويكون له الشكل العام الآتي:

$$\begin{aligned}&\text{Ri: If } y(t) \text{ is Ai1 and } y(t-1) \text{ is Ai2} \dots \text{ and } y(t-n+1) \text{ is Ain}\\&\text{and } u(t) \text{ is Bi1 and } \dots u(t-m+1) \text{ is Bim}\\&\text{then } y(t+1) \text{ is Ci}\ ;\ i=1,2,\dots,r\end{aligned}$$

5-3 الشبكة العصبية ذات الانتشار العكسي: Backpropagation Neural Network

تعتمد الشبكة العصبية ذات الانتشار العكسي للخطأ على طريقة الانحدار التدريجي (Gradient Descent) وذلك لإيجاد القيمة الصغرى لمربع الخطأ الكلي لقيمة المخرجات المحسوبة من قبل الشبكة حيث تحدث الأوزان بين الطبقات لحين الوصول إلى الأوزان المثلى التي تعطي اقل خطأ بين مخرجات الشبكة العصبية وبيانات النموذج (قيم المعيار(Criterion Section)) حيث تستخدم هذه الأوزان لحساب تكهنات جديدة لم يسبق للشبكة العصبية إن تدربت عليها. إن الخطوات الأساسية وفق هذه المنهجية تتم بحساب أخطاء مستوى المخرجات لتحديث أوزان طبقة المستوى(المخفي–المخرجات) ثم حساب أخطاء المستوى المخفي لتحديث أوزان طبقة مستوى (المدخلات–المخفي) وبعد ذلك نحسب مخرجات الشبكة بالأوزان الجديدة لتستمر العملية في حساب الأخطاء لتحديث الأوزان لغرض الوصول إلى اقل خطأ في الشبكة العصبية. تتضمن طريقة التدريب للشبكة العصبية باستخدام الانتشار العكسي ثلاث مراحل:

- مرحلة الانتشار الأمامي للخطأ.
- مرحلة الانتشار الخلفي للخطأ.
- مرحلة توليف الأوزان.

4–5 التمهيد الاسي:

يعد موضوع التمهيد الاسي للتنبؤ بالسلاسل الزمنية من الاجراءات الاحصائية والاستدلالية المهمة والتي تعالج التشويش او الاخطاء العشوائية ويعرف التمهيد بنه عملية صقل او تنعيم البيانات وهو عبارة عن تقنية احصائية للكشف عن تقلبات معنوية لغرض جعل البيانات مستقرة ويمكن تحليلها ، واساس عمل هذه التقنية هو اعطاء البيانات الاقدم اقل الاوزارن والبيانات الاحدث اعلى الاوزان وبشكل تدريجي وبعد المعالجة يمكن اجراء التنبؤ للبيانات الممهدة اسياً، ويعتبر العالم (1958) holt c.c اول من وضع هذه التقنية وقد تطورت هذه الطرائق وتعددت واصبحت بأشكال عدة لذا وجب اختيار الطريقة الاكثر ملائمة (داؤد وصالح ،2018).ومن طرائق التمهيد الاسي:

4–5 –1 التمهيد الاسي البسيط simple Exponential Smoothing

اقترحت هذه الطريقة من قبل (1958) Holt c.c وكان في البداية يستخدم فقط للسلسة الزمنية غير الموسمية ثم بعد ذلك اكد الباحث (Browns , 1963) على امكانية استخدامه لاكثر انواع السلاسل الزمنية ، ثم اكمل طريقه الباحث (Harrison , 1965) وان اجراءات هذه الطريقة :(داؤود وصالح ،

$$Y_t = \mu_t + \beta_t t + e_t \qquad (1)$$

5–4–2 **نموذج التمهيد الاسي المزدوج Double Exponential Smoothing**

في التمهيد الاسي البسيط نفترض ان السلسلة ثابتة في بعض الحالات ، وان بعض السلاسل الزمنية تمتلك اتجاه عام (Overall trend) وان طريقة التمهيد الاسي الثنائي تكون مفيدة في هذه الحالة حيث تعمل هذه الطريقة على تنعيم بيانات السلسلة الزمنية مرتين ، ويستخدم التمهيد الاحصائي من الرتبة الاولى والثانية لحساب التنبؤ ويستخدم رمز (') للاشارة الى سلسلة التمهيد من الرتبة الاولى كما في المعادلة الاتية: (الجبوري، 2010)

$$s_t = \alpha z_t + (1-\alpha)(s_{t-1} + b_{t-1}), \quad t = 1,2,...,n$$
$$b_t = \gamma(s_t - s_{t-1}) + (1-\gamma) b_{t-1}, \quad t = 1,2,...,n$$

علما ان

$$\hat{z}_t = s_t + b_t t, \quad t = 1,2,...,n \qquad (2)$$

والتنبؤات للقيم المستقبلية من

$$z_n(\ell) = s_n + b_n \ell, \quad \ell > 0$$

نحسب القيم الأولية s_0 و b_0 من $s_0 = z_1$

$$b_0 = z_2 - z_1 \quad \text{or}$$
$$b_0 = \frac{(z_2 - z_1) + (z_3 - z_2)}{2} = \frac{(z_3 - z_1)}{2} \quad \text{or} \tag{3}$$
$$b_0 = \frac{(z_2 - z_1) + (z_3 - z_2) + (z_4 - z_3)}{3} = \frac{(z_4 - z_1)}{3}$$

3-4-5 التمهيد الآسي للأوساط المتحركة Exponential Smoothing Moving Averages

يستخدم المتوسط المتحرك لتمهيد الشاهدات وذلك بتقليل تباين الاخطاء فمثلا لو كان لدينا مشاهدات من متسلسلة زمنية $Y_1, Y_2, Y_3, \dots, Y_{n-1}, Y_n$ فالمتوسط المتحرك من الدرجة m للمشاهدات يعطى بالعلاقة التالية:

$$\hat{Y_t} = \frac{1}{m}\left(Y_t + Y_{t-1} + Y_{t-2} + \cdots + Y_{t-m+1}\right), t = m, m+1, \dots, n \tag{4}$$

لاحظ ان عدد المشاهدات أصبح بعد التمهيد $n - m + 1$.

ولكي نرى كيف يعمل التمهيد لتقليل تباين الأخطاء لنفترض ان المشاهدات تتبع النموذج

$$y_t = \mu + a_t, a_t \sim WN\left(0, \sigma^2\right), t = 1, 2, \dots, n$$

فيكون $V\left(y_t\right) = \sigma^2, \forall t$

وبالتالي

$$V\left(\hat{y}_t\right) = \frac{\sigma^2}{m}, t = m, m+1, \dots, n \tag{5}$$

إي ان المشاهدات الممهدة أصبح تباينها اصغر ب (m) ضعف من المشاهدات الأصلية وهذا التمهيد للأخطاء يظهر إي نمط في المتسلسلة كان مدفونا او مغطى من تاثير الاخطاء.

6- نماذج بوكس وجنكيز: Box-Jenkins

من النماذج المستخدمة في السلاسل الزمنية هو أنموذج الانحدار الذاتي والمتوسطات المتحركة الذي يعرف أحيانا بأنموذج بوكس جنكيز حيث يجمع منهجيتين في معادلة واحدة وسوف نقوم باستعراض النماذج للتعرف عليها وهي:

1-6 إنموذج الانحدار الذاتي Autoregressive Model (AR)

يمكن تعريف إنموذج الانحدار الذاتي بأنه يصف العلاقة بين المشاهدات السابقة والحالية، وأن دالة الارتباط الذاتي له تتناقص تدريجياً بشكل أسي متخذة شكلاً منحنياً تنازلياً في حين أن دالة الارتباط الذاتي الجزئي له تتقطع بعد الفترة p. وأن الصيغة العامة لهذا النموذج من الرتبة (p) . [Box et al.,2016]

$$z_t = \varphi_1 z_{t-1} + \varphi_2 z_{t-2} + \cdots + \varphi_p z_{t-p} + e_t \tag{6}$$

حيث أن:

e_t : تمثل الخطأ العشوائي أو ما يسمى بالتشويش Noise يتوزع توزيعاً طبيعياً بوسط حسابي صفر وتباين ثابت مقداره σ_e^2.

$\varphi_1, \varphi_2, \dots, \varphi_p$: تمثل معلمات إنموذج الأنحدار الذاتي.

وأن المتغيرات المفسرة تمثل القيم السابقة لمتغير الاستجابة z_t أي القيم المزحفة للمتغير z_t .

2-6 إنموذج الأوساط المتحركة Moving Average Model(MA)

أن فلسفة هذه النماذج تعتبر أن المشاهدة الحالية للسلسلة الزمنية z_t هي دالة خطية في الأخطاء السابقة $e_{t-1}, e_{t-2}, \dots, e_{t-q}$ بالإضافة إلى الخطأ العشوائية a_t وهنا نجد أن رتبة هذه العملية تساوي عدد معاملات الإنموذج التي يجب تقديرها. وان الصيغة العامة لهذا الإنموذج من الرتبة (q) واختصاراً MA(q) تاخذ الشكل الاتي: [Box et al., 2016]

$$z_t = e_t - \vartheta e_{t-1} - \vartheta_2 e_{t-2} - \cdots - \vartheta_q e_{t-q} \tag{7}$$

حيث أن:

$\vartheta_1, \vartheta_2, \vartheta_3, \dots, \vartheta_q$: تمثل معلمات المتوسطات المتحركة.

ويتم تحديد رتبة الأنموذج من دالة الارتباط الذاتي له حيث تنقطع بعد فترة q في حين دالة الارتباط الذاتي الجزئي تتناقص تدريجياً بشكل منحني تنازلي.

6-3 النماذج المختلطة (الانحدار الذاتي-المتوسطات المتحركة):

Mixed Models (Autoregressive-Moving Average) (ARMA)

وأن فلسفة هذه النماذج تعتبر أن المشاهدة الحالية للسلسلة الزمنية z_t هي دالة خطية في كل المتغيرات السابقة $z_{t-1}, z_{t-2}, \dots, z_{t-p}$ والأخطاء السابقة $a_{t-1}, a_{t-2}, \dots, a_{t-q}$ بالإضافة إلى الخطأ العشوائية e_t ويرمز لهذا الإنموذج باختصار ARMA(p,q). ويستخدم في حالة كون البيانات مستقرة، وأن دالتا الأرتباط الذاتي و الأرتباط الذاتي الجزئي له يتناقصان تدريجياً. الصيغة العامة لهذا الإنموذج من الرتبة (p,q) تأخذ الشكل الاتي: (ثامر والشرابي،2020)

$$z_t = \varphi_1 z_{t-1} + \varphi_2 z_{t-2} + \dots + \varphi_p z_{t-p} + e_t - \vartheta e_{t-1} - \vartheta_2 e_{t-2} - \dots - \vartheta_q e_{t-q} \quad (8)$$

6-4 النماذج المختلطة المتكاملة:

Autoregressive Integrated Moving Average Models(ARIMA)

ان هذا الإنموذج يعوض عن دراسة سائر النماذج ويرمز له بالرمز ARIMA(p,d,q). الصيغة العامة لهذا الإنموذج من الرتبة (p,d,q) يكتب بالشكل الآتي: [Box et al., 2016]

$$\varphi(B)(1-B)^d z_t = \vartheta(B) e_t \quad (9)$$

حيث أن :

$$\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots . - \varphi_p B^p \quad (10)$$

7- معايير اختيار افضل نماذج السلاسل الزمنية:

Criteria for Choosing the Best Time Series Models

ان طريقة Box-Jenkins من الطرق الشائعة الاستخدام ذات الكفاءة العالية في توفيق السلاسل الزمنية التي تعكس سلوك السلسلة الزمنية ان كانت موسمية او غير موسمية. ولتوفيق أفضل نموذج من نماذج السلاسل الزمنية للرطوبة النسبية تم استخدام معيارين:

أ-معيار اكاكي Akaike's Information Criterion(AIC)

اقترح معياره المستخدم والمسمى (AIC) الذي يعني (Akaike Information Criterion) والمعرف كالتالي:

$$AIC(k) = n ln(\sigma_e^2) + 2k \quad (11)$$

حيث ان (k) :تمثل عدد المعالم في النموذج،و σ_e^2 مقدار التباين و n عدد المشاهدات، فلمعرفة أفضل نموذج نختار اقل قيمة للمعيار AIC(k) ،وقد أُعتمد هذا المعيار لكونه اكثر واوسع المعايير استخداماً وملائمة للبيانات (الصفاوي والطائي ،2003).

ب- معيار معلومات بيز Bayesian Information Criterion

ويحسب بالصيغة الآتية

$$BIC(m) = n\log(\hat{\sigma}_a^2) + m\log(n)$$

ج- معيار متوسط مربعات الخطأ(MSE):

$$\mathrm{MSE} = \frac{\sum_{i=1}^{n}(Y_t - \hat{Y}_t)}{n-(k+1)} \quad (12)$$

حيث ان : n : تمثل عدد المشاهدات.

k :تمثل عدد المعلمات.

Y_t :تمثل قيم المشاهدة.

$\hat{Y}_t$:تمثل القيم بعد إجراء التمهيد للبيانات (Al-Safawi & Al-Taai,2003).

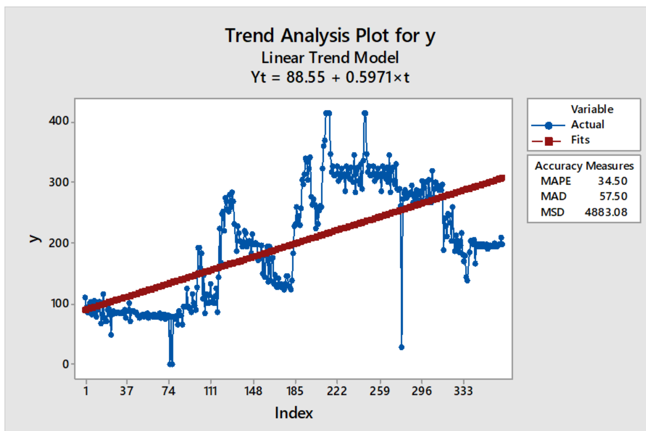
وكلما كانت قيمة (MSE) صغيرة كان النموذج المستخدم يمثل البيانات احسن تمثيل.

جمع البيانات

جمعت البيانات من مديرية محطات كهرباء سد الموصل والتي تتألف من (366) مشاهدة ، وتعود الى بيانات يومية لعام 2020 ، وهذه المشاهدات تمثل توليد الكهرباء الناتج من الماء الوارد من نهر دجلة ، اذ يعتمد التوليد على تدفق الماء الوارد من تركيا والامطار وذوبان الثلوج في فصل الربيع .

مرحلة تهيئة البيانات

ان الافتراض الاساسي لتحليل السلاسل الزمنية ومطابقة النمذجة Modeling Fitting هو ان البيانات تكون مستقرة Stationary وابسط تعريف للسلسلة المستقرة انها لا تحتوي على اتجاه عام Trend ، اذ يتم في هذه المرحلة تحضير البيانات من خلال رسم شكل الانتشار واستخراج معاملات الارتباط الذاتي والجزئي وكذلك رسم حدود الثقة لدالة الارتباط الذاتي للبيانات الاصلية (y) لمعرفة سلوك هذه البيانات وذلك باستخدام البرنامج الاحصائي Minitab ، فمن خلال الشكل (1) يلاحظ ان التباين يميل الى الثبات لكن يلاحظ وجود اتجاه عام متزايد مع الزمن ، اذ يوجد تذبذب للسلسلة يبدأ بالصعود التدريجي ثم الانخفاض التدريجي مما يدل على عدم استقرارية السلسلة الزمنية في المتوسط .



الشكل (1) : يمثل رسم الاتجاه العام للسلسلة الزمنية (y)

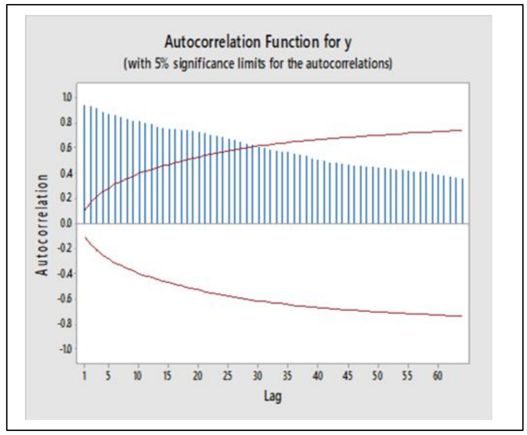
وقد اكدت عدم استقراريه السلسلة الزمنية ايضاً قيم معاملات الارتباط الذاتي والجزئي كما في الشكلين (2) و (3) على التوالي :

ولكي تكون السلسلة مستقرة لابد من دخول جميع معاملات الارتباط الذاتي للعينة ضمن حدود الثقة ما عدا عند الازاحة الاولى والثانية فممكن ان تقع خارج حدود الثقة .

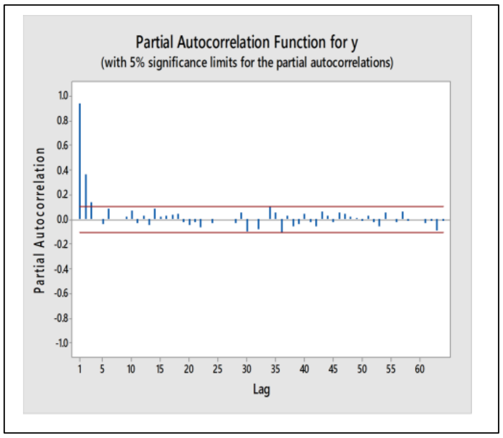
ولغرض تحديد افضل طريقة لتنعيم (تمهيد) البيانات وتحويل السلسلة الزمنية الى سلسلة مستقرة والتوصل لأفضل نموذج ، تم استخدام الطرائق التي التطرق اليها في الجانب النظري : (Box Jenkins ، Exponential Smoothing ، DoubleExponential Smoothing ، Exponential Smoothing Moving Averages ، Fuzzy ، Neural Network)

وبعد تطبيق الطرائق على سلسلة بيانات توليد الكهرباء (y) ورسم السلسلة اتضح انها غير مستقرة لذلك تم اخذ الفرق الاول على بيانات السلسلة الممهدة للحصول على سلسلة اكثر استقرارية :

$$\nabla_{z_t} = (z_t - z_{t-1})$$



الشكل (2) : يمثل معامل الارتباط الذاتي لـ (y)



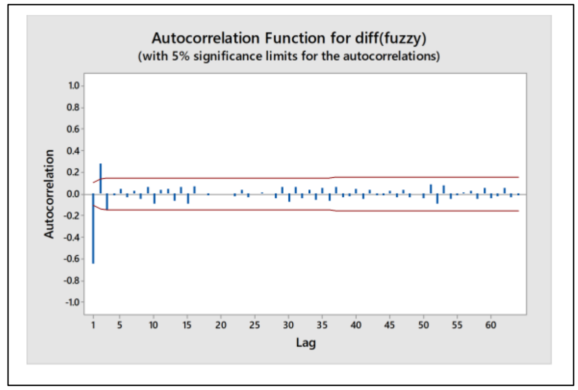
الشكل (3) : يمثل معامل الارتباط الذاتي لـ (y)

بعد استقرار السلسلة الزمنية يتم تشخيصها بنماذج ARIMA وذلك من خلال مشاهدة قيم معلمات دالة الارتباط الذاتي والارتباط الذاتي الجزئي لتحديد رتبة النموذج حيث يتم للحصول على سلسلة البواقي وذلك من خلال دالتي الارتباط الذاتي ACF و والارتباط الذاتي الجزئي PACF لكل طريقة ، ومن ثم يتم المفاضلة بين هذه الطرائق من خلال استخدام معايير المقارنة الاحصائية (MSE , AIC , BIC) لاختيار افضل نموذج وبالتالي افضل طريقة . والجدول (1) يوضح الانموذج الملائم لكل طريقة ومعايير المقارنة الثلاثة الانفة الذكر :

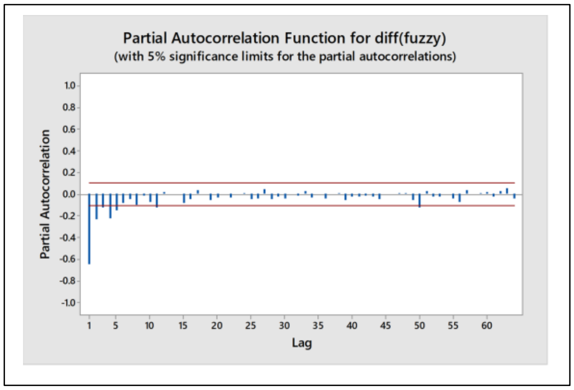
الجدول (1): يوضح نتائج المقارنة بين الطرائق المستخدمة لمعالجة استقرارية السلسلة الزمنية

الطرق المستخدمة لاستقراريه السلسلة الزمنية	النماذج	معايير المقارنة		
		MSE	AIC معيار	BIC معيار
Box Jenkins	ARIMA (0,1,1)	817	2466.2639	2477.874
Exponential Smoothing	ARIMA (4,1,1)	32.5	1282.1339	1297.744
Double Exponential Smoothing	ARIMA (0,1,3)	133.8	1800.0627	1815.6732
Exponential Smoothing Moving Averages	ARIMA (0,1,3)	204.2	1954.7906	1970.4011
Fuzzy	ARIMA (0,1,3)	0.572	-196.4536	-0.6931
Neural Network	ARIMA (1,1,4)	207	1959.7751	1975.385

من ملاحظة النتائج الجدول في (1) سجلت طريقة المضبب (Fuzzy) تفوقاً على طرائق(Box Jenkins ، Exponential Smoothing ، Double Exponential Smoothing ، Exponential Smoothing Moving Averages ، Neural Network) وذلك من خلال الانخفاض الواضح في قيم معايير المقارنة (MSE , AIC , BIC) . ومن الجدول (1) لوحظ ايضاً ان النموذج الملائم لطريقة الـ Fuzzy هو ARIMA (0,1,3) وذلك من خلال اقل قيمة لـ (MSE = 0.572) ، كذلك من خلال الاختبارات الاحصائية (معنوية المعالم المقدرة وتحليل دالة الارتباط الذاتي للبواقي) . والشكلين (3) و (4) يبينا دالتي الارتباط الذاتي والذاتي الجزئي لسلسلة البيانات المستقرة بطريقة Fuzzy وعلى التوالي :

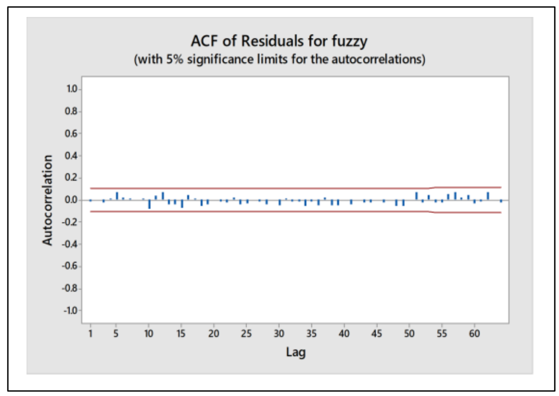


الشكل (4) : يمثل معامل الارتباط الذاتي لـ Fuzzy

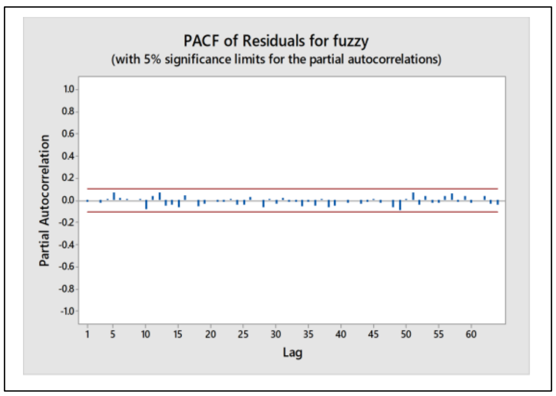


الشكل (5) : يمثل معامل الارتباط الذاتي لـ Fuzzy

والشكلان (6) و(7) يوضحان دالتي الارتباط الذاتي والذاتي الجزئي للبواقي :



الشكل (6) : يمثل معامل الارتباط الذاتي لـ Fuzzy



الشكل (7) : يمثل معامل الارتباط الذاتي لـ Fuzzy

ومن خلال الشكلين (6) و (7) لوحظ دخول جميع معاملات الارتباط الذاتي للعينة ضمن حدود الثقة .

<u>الاستنتاجات</u>

على ضوء ما تم التوصل اليه في الجانبين التجريبي والتطبيقي يمكن استخلاص اهم الاستنتاجات الخاصة بالبحث وكالاتي :-

1- عند استخدام بيانات توليد الكهرباء اظهر اسلوب Fuzzy تقدماً على طرائق (Box Jenkins ، Exponential Smoothing ، Double Exponential Smoothing ، Exponential Smoothing Moving Averages ، Neural Network) من خلال الانخفاض الواضح في قيم معايير المقارنة (MSE , AIC , BIC) .

2- ولجعل السلسلة الممهدة اكثر استقراراً تم اخذ الفرق الاول للبيانات ، ومطابقة معاملات الارتباط الذاتي والذاتي الجزئي للسلسلة الزمنية .

3- تم اختيار النموذج ARIMA (0,1,3) من بين عدة نماذج والتي تعود الى بيانات السلسلة الممهدة باستخدام اسلوب Fuzzy ، اذ ان (MSE = 0.572) .

Refference

1. AL-Jubury, Abear Hasan Ali,(2010)," Forecasting Iraqi oil prices for the year 2010 using time series",Babelon University Journal, Humanities, Issue 18 ,No.1.
2. AL-Safawi , Safaa Younis & AL-Taea ,Fadel Abbas,(2003)," Methods of dealing with the instability of some time series data by application to maize production in Iraq for the period (1949-1989),",Journal of AL-Rafeden Development, No.73.

3. Thamer , Noor AL-Huda Mahmood & AL-Sharabe, Najlaa Saad,(2020)," Prediction of transfer function models using the genetic algorithm with application to temperatures in Nineveh Governorate", MSc. Thesis , University of Mosul, Mosul ,Iraq.
4. Dwood ,Heba & Saleh ,Esraa Abd AL-jawad,(2018)," Use smoothing methods to choose the best model of the transfer function with the application", Iraqi Journal of Statistical Science ,University of Mosul ,Mosul, Iraq.
5. Fandel, Walter ,(1992)," Time series from the applied point of view and the models of Box and Jenkimz ", Mars Publishing House.
6. Box, G., Jenkins, G., Reinsel ,G. and Ljung G., (2016)," Time Series Analysis Forecasting and control", John wiley & Sons , Inc . Hoboken, New Jersey
7. Makridakis, S., Wheel, Wright , S. C. and Hyndman , R. J. (1998) : "Forecasting : Methods and Applications " ,$3rd$ ed., John Wiley And Sons ,New York , U.S.A.

Treatment of time series instability - review article-

Alaa Abd Al-Satar Hamodat, Nada Nazar Muhamad & Zainab Tawfeek Al-Dabag
Department of Informatics & Statistic, College of Computer & Mathematical Science, University of Mosul, Mosul, Iraq

Abstract:
The time series is a problem in econometric analysis as the statistical properties of series analysis are lost when using unstable time series. The research aims to present several methods for dealing with stability, including (Box Jenkins, Exponential Smoothing, Double Exponential Smoothing, Exponential Smoothing Moving Averages, Fuzzy, Neural Network) and to compare the methods presented through diagnosing ARIMA models after achieving stability and choosing the best method that corresponds to the lowest values of the criteria Statistics (MSE, AIC, BIC). The above-mentioned methods have been applied to daily data for the year 2020 to generate electricity from water coming from the Tigris River, and it was concluded that the (Fuzzy) method is the best for treating stability compared to other methods for having the ARIMA model (0,1,3) corresponding to the lowest values of the criteria Statistics (MSE=0.572, AIC=-196.4536, BIC=-0.6931).

Keyword: Time Series, Box-Jenkins Models, Stationary.