

إعادة تحويل التكهّنات عكسيا في السلاسل الزمنية

د. ظافر رمضان مطر  
استاذ مساعد  
مرّوة نبيل  
طالبة ماجستير  
جامعة الموصل / كلية علوم الحاسبات والرياضيات / قسم الإحصاء والمعلوماتية

المخلص

إن مما لا شك فيه أن قوة وفعالية النماذج الإحصائية تتحسن غالبا بعد إجراء التحويل المناسب على البيانات في حالة الحاجة لهذا التحويل، وان هذه الفعالية لا شك تنعكس أيضا على التحليل والتكهن باستخدام نماذج السلاسل الزمنية. من جانب آخر فان التعديل الرياضي للبيانات ينتج عنه إشكالات معينة في تفسير نتائج النماذج وعدم إمكانية الاستفادة من التكهّنات التي يتم الحصول عليها اعتمادا على البيانات المحولة لذا فانه سيتم في هذا البحث تناول طرق إعادة تحويل التكهّنات عكسيا باتجاه القياس الأصلي في السلاسل الزمنية وكذلك سيتم التطبيق العملي لهذا البحث على بيانات المعدل اليومي لاستهلاك الطاقة الكهربائية في محافظة نينوى.

**Retransformed Forecasts Back In Time Series**

Abstract

The power and activity of statistical models is often improved after doing a suitable transformation the efficacy is reflects also on analysis and forecasting using time series models in other side the mathematical modification of the data raises problems for the interpretation results of models also usefulness of forecasts based on transformed data, so in this paper we deal with methods for retransformed forecasts back into original scale. the application side of this paper will be in daily average data of consume energy electric in Nineveh conservatism.

## 1. المقدمة:

إن عملية تقدير السلوك المستقبلي لأي ظاهرة أو متغير تعد من الأهداف الأساسية للعلوم الإحصائية لأهميتها الكبيرة في عملية التخطيط للظاهرة أو المتغير قيد الدراسة وان التخطيط هو المفتاح لعملية اتخاذ القرار السليم، حيث بإمكاننا التعرف على الظروف والحالة المحيطة بالمشكلة قيد الدراسة قبل اتخاذ القرار، لذلك فان موضوع التكهّن لاقى ويلاقي اهتماما كبيرا ومتزايدا من قبل الباحثين وأصحاب القرارات (أمين بك، 2005). فالتكهّن أسلوب فني للتحليل الفعلي للبيانات الذي يقدر القيم المستقبلية لتأكيد أهمية المتغير. فبعد أن يتم التعرف على النموذج وتقدير معلماته يصبح جاهزا لاستخدامه في التطبيق ويعد التكهّن من ابرز الأهداف التي يصبو إليها معظم الباحثين ومتخذي القرار.

إن فعالية النماذج الإحصائية تتعزز غالبا من خلال استخدام البيانات المحولة ولا يستثنى من ذلك التحليل والتكهّن باستخدام نماذج السلاسل الزمنية. فمن المهم اختيار التحويل الملائم الذي يفى بالغرض. نفترض معظم الطرق الإحصائية أن المتغيرات موزعة توزيعا طبيعيا، وتحويل البيانات أداة مفيدة لجعل المتغيرات تحت الدراسة طبيعية. وان التعديل الرياضي للبيانات بهذا الأسلوب ينشئ نقاط خلاف أو نقاش ليست فقط في تفسير نتائج النماذج ولكن أيضا عدم فائدة التكهّن اعتمادا على البيانات المحولة (Liu، 2006) فحالما يتم بناء النموذج اعتمادا على المقياس المحول فان التكهّنات التي يتم الحصول عليها في النظام المترى تصبح بحاجة إلى إعادة تحويلها Retransformation وتظهر عند ذاك مشكلة أن خطوات إعادة التحويل تسبب أو تنتج تحيزا في التكهّنات ولتصحيح هذا التحيز اشتق (Guerrero، 1993) عوامل لحذف التحيز واقترح ضرب حدود التكهّن بعوامل حذف التحيز الخاص بكل نوع من تحويلات القوة Power Transformation.

إن اهتمام الباحثين ينصب في هذا البحث على تطبيق تحويلات القوى لتحسين دقة التكهّن عندما يعاد تحويل التكهّنات باتجاه القياس الأصلي، فتحويل القوة يحسن جوهريا من دقة التكهّن عندما يكون هناك حاجة للتحويل وان الضعف أو النقص في تحسين التكهّن يمكن أن يكون سببه الشواذ أو اختيار خاطئ لقيمة  $I$  (التي تمثل قيمة القوة في التحويل) أو بسبب اختيار النموذج الخاطئ أو كلاهما وسيتم التطبيق على بيانات المعدل اليومي لاستهلاك الطاقة الكهربائية في محافظة نينوى .

2. تحويلات القوة في السلاسل الزمنية:

في تحليل السلسلة الزمنية أحادية المتغير Univariate ناقش كل من ( Box and Jenkins 1976) عملية الحصول على نموذج ARIMA مع تجانس تباين البواقي اعتمادا على السلسلة الزمنية المشتقة من خلال سلسلة تحويلات القوة بشكل مشابه لمقترح (Box and Cox 1964) وهذا النوع من التحويلات يعبر عنه بالشكل التالي (Pankratz،1991) (Liu and، 2002) (Hudak، 2006) (Liu، 2006) :

$$z_t = \begin{cases} \frac{(y_t^I - 1)}{I} & ; \text{if } I \neq 0 \\ \ln(y_t) & ; \text{if } I = 0 \end{cases} \quad (1)$$

حيث أن :

$z_t$ : تمثل السلسلة المحولة .

$y_t$ : السلسلة الأصلية .

$I$ : قيمة القوة في التحويل  $|I| \leq 1$  . وان  $t = 1, 2, \dots, n$  .

إن الصيغة أعلاه لتحويل القوة مستمرة في المعلمة  $I$  ، وعلى الرغم من أن التحويل في (1) له خصائص مرغوبة مثل استمرارية  $I$  عند الصفر ، إلا إن تباينات السلسلة المحولة غير قابلة للمقارنة عند اختلاف قيم  $I$  ، ولمعالجة هذه المسألة اقترح كل من (Box and Jenkins 1976) (Ansley, Spivey and Wroblewski، 1977) تحويل قوة آخر ويعطى بالشكل التالي :

$$z_t = \begin{cases} \frac{(y_t^I - 1)}{(I \&^{I-1})} & ; \text{if } I \neq 0 \\ \& \ln(y_t) & ; \text{if } I = 0 \end{cases} \quad (2)$$

حيث أن :

$\& = (y_1 y_2 \dots y_n)^{\frac{1}{n}}$ : الوسط الهندسي للسلسلة .

ويطلق على الصيغتين (1) و(2) بالنوع الأول First Type من تحويلات القوة .

أما النوع الثاني من تحويلات القوة فيعبر عنه بالشكل التالي :

$$z_t = \begin{cases} y_t^I & ; \text{if } I \neq 0 \\ \ln y_t & ; \text{if } I = 0 \end{cases} \quad (3)$$

إن الصيغة أعلاه لتحويل القوة لاتستمر عند  $I = 0$ ، وقد تمتلك بعض الصفات الأخرى غير المرغوب بها، تعتبر هذه الصيغة لتحويل القوة بسيطة ولذا تستعمل كثيرا (Liu,2006).

إن طريقة Box-Cox تعظم دالة الترجيح الأعظم Maximum Likelihood Function والتي تعتمد على  $I$  والمعالم المجهولة للنموذج، فهي تحافظ على شرعية النموذج ككل، وبشكل عملي تساعد في تثبيت Stabilize تباين السلسلة (Guerrero,1993).

وهناك قضيتان أساسيتان في تطبيق تحويل القوة أعلاه، الأولى: اختيار قيمة ملائمة ل  $I$  التي ستحسن إما كفاءة النموذج Improve The Efficacy Of The Model أو دقة التكهّنات Accuracy Of The Forecasts، والثانية: تصحيح التحيزات التي تحدث خلال إعادة تحويل التكهّنات (إلى القياس الأصلي قبل التحويل) سنفترض أولاً بأن قيمة  $I$  مختارة (مثلاً  $I = 0$ ) للتحويل اللوغارتمي (وبالتالي يتطلب الأمر اختيار تحويل القوة المناسب (Liu,2006)).

### 3. طرق إعادة تحويل التكهّنات:

#### أ- طريقة إعادة التحويل المباشرة:

عندما تطبق تحويلات القوة على سلسلة زمنية، فإن التكهّنات تكون مستتدة على سلسلة محولة ولذا يجب إعادة تحويلها إلى القياس الأصلي (Liu,2006). ولكن هذه التكهّنات لاتحتفظ بخصائصها المثالية عندما تعاد إلى القياس الأصلي (Gurrero,1993)، ذلك لأن إعادة التحويل المباشر للتكهّنات يكون متحيزاً وهذا يحدث لأن المتوسط المقدر للتوزيع المتماثل في البيانات المحولة يصبح الوسيط المقدر بعد إعادة التحويل للمقدر (Liu,2006).

إن إعادة التحويل هذه للوسيط تنتج مايسمى "بالتكهّن البسيط" "Naive Forecasting" الذي قد يعتبر مثالي إذا كان اقتران دالة الكلفة Cost Function مع أخطاء التكهّن Forecast Errors خطياً ومتماثلاً حول الصفر (Granger and Newbold,1986). وعلى الرغم من هذا فإن تصحيح التحيز مطلوب إذا أردنا حقاً تقدير القيمة المتوقعة في (القياس) المتري الأصلي. مما يجب ذكره أنه في بعض الأحيان يكون تأثير التحيز عديم الفائدة عندما تصبح البيانات غير حساسة نسبياً للتحويل.

إن طريقة إعادة التحويل المباشرة للتكهنات يعبر عنها بمعادلات رياضية معينة، فإذا كانت  $\hat{Z}_n^{(l)}$  تمثل تكهن أقل متوسط مربع خطأ للتكهن بخطوة واحدة إلى الأمام ل  $Z_{n+1}$  عند نقطة أصل التكهن  $t=n$ ، فإن إعادة التحويل المباشر للتكهنات ل  $\hat{Z}_n^{(l)}$  اعتمادا على النوع الأول من تحويلات القوة Box-Cox هو :

$$\hat{y}_n^{(l)} = T^{-1}(\hat{z}_n^{(l)}) = \begin{cases} (I \hat{z}_n^{(l)} + 1)^{\frac{1}{I}} & ; \text{if } I \neq 0 \\ e^{\hat{z}_n^{(l)}} & ; \text{if } I = 0 \end{cases} \quad (4)$$

بالإضافة إلى إعادة التحويل المباشر للتكهنات ل  $\hat{Z}_n^{(l)}$  اعتمادا على النوع الأول (Non Box-Cox) من تحويلات القوة يعبر عنه بالشكل:

$$\hat{y}_n^{(l)} = T^{-1}(\hat{z}_n^{(l)}) = \begin{cases} (I (\hat{z}_n^{(l)})^{I-1} + 1)^{\frac{1}{I}} & ; \text{if } I \neq 0 \\ e^{\hat{z}_n^{(l)}} & ; \text{if } I = 0 \end{cases} \quad (5)$$

أما إعادة التحويل المباشر للتكهنات ل  $\hat{Z}_n^{(l)}$  اعتمادا على النوع الثاني من تحويلات القوة فيعطى بالشكل أدناه:

$$\hat{y}_n^{(l)} = T^{-1}(\hat{z}_n^{(l)}) = \begin{cases} \hat{z}_n^{(l)\frac{1}{I}} & ; \text{if } I \neq 0 \\ e^{\hat{z}_n^{(l)}} & ; \text{if } I = 0 \end{cases} \quad (6)$$

ب- طريقة إعادة التحويل الغير متحيزة :

إن تحويلات القوة Power Transformations يعبر عنها عموما بالشكل التالي:

$$z_t = T(y_t) \quad (7)$$

حيث أن:

$Z_t$ : تمثل السلسلة المحولة .

$y_t$  : السلسلة الأصلية .

$T(\cdot)$  : عامل التحويل .

أي بالإمكان الحصول على:

$$y_t = T^{-1}(z_t) \quad (8)$$

حيث أن:

$T^{-1}(\cdot)$  : معكوس عامل التحويل .

إن المقدر  $\hat{y}_n^{(l)}$  اعتمادا على المعادلات (3-6) متحيز لان المتوسط المقدر للتوزيع المتماثل في البيانات المحولة يصبح الوسيط المقدر بعد إعادة التحويل للمقدر. ولتصحيح التحيز اشتق Guerrero (1993) المقدر غير المتحيز  $\bar{y}_n^{(l)}$  التالي انظر (Liu,2006):

$$\bar{y}_n^{(l)} = T^{-1}(\hat{z}_n^{(l)}) \cdot C_l^{(l)} \quad (9)$$

حيث أن:

$$C_l^{(l)} = \begin{cases} \left\{ 0.5 + 0.5 \left[ \frac{1+2(l^{-1}-1)s_j^2}{m_j^2} \right]^{\frac{1}{2}} \right\}^{\frac{1}{l}} & ; \text{if } l \neq 0 \\ \exp \left( \frac{s_j^2}{2} \right) & ; \text{if } l = 0 \end{cases} \quad (10)$$

لتكن  $m_j = E_t [T(z_{t+j})]$  تمثل التكهّن المثالي للسلسلة المحولة. حيث أن الرمز السفلي t في مشغل التوقع يستخدم للإشارة شرطيا إلى المعلومات المعروفة حتى الزمن t. كذلك لتكن  $e_{t+j} = T(z_{t+j}) - M_j$  تمثل خطأ التكهّن Forecasting error ويعرف بأنه

$s_j^2 = Var_t(e_{t+j})$  وإذا كان  $E_t(z_{t+j})$ ;  $j = 1, 2, L$  تمثل التكهّنات التي نبحث عنها، فإن:

$$Var [T(x)] = \{T'[E(x)]\}^2 Var(x) \quad (11)$$

حيث أن:

$T'$ : تشير إلى مشتقة  $T$ .

فالتعبير المكافئ للمعادلة (11):

$$s_j^2 = \{T'[E_t(z_{t+j})]\}^2 Var_t[z_{t+j} - E(z_{t+j})] \quad (12)$$

وباستخدام متسلسلة تايلور نحصل على المعادلة التالية:

$$T(z_{t+j}) = T[E_t(z_{t+j})] + T'[E_t(z_{t+j})][z_{t+j} - E_t(z_{t+j})] + \frac{T''[E_t(z_{t+j})][z_{t+j} - E_t(z_{t+j})]^2}{2} \quad (13)$$

وبأخذ التوقع لطرفي المعادلة نحصل على:

$$m_j = T[E_t(z_{t+j})] + \frac{T''[E_t(z_{t+j})] \text{Var}[z_{t+j} - E_t(z_{t+j})]}{2} \quad (14)$$

وبتعويض المعادلة (11) في المعادلة (14)، نحصل على:

$$m_j = T[E_t(z_{t+j})] + \frac{T''[E_t(z_{t+j})] \{T'[E_t(z_{t+j})]\}^{-2} s_j^2}{2} \quad (15)$$

وبالتعويض عن كل  $T(z_t) = \log z_t$  نحصل على:

$$m_j = \log[E_t(z_{t+j})] + \frac{\log''[E_t(z_{t+j})] \{\log' E_t(z_{t+j})\}^{-2} s_j^2}{2} \quad (16)$$

من هذا نحصل على:

$$\log[E_t(z_{t+j})] = m_j + \frac{s_j^2}{2} \quad (17)$$

وبأخذ دالة exp للطرفين نحصل على:

$$E_t(z_{t+j}) = \exp\left(m_j + \frac{s_j^2}{2}\right) \quad (18)$$

أي إن النتيجة تكون بالضبط حيث  $\log(z_{t+j})$  يتبع التوزيع الطبيعي. بنفس الطريقة عندما  $T(z_t) = z_t^I$  مع  $I \neq 0$  نحصل عليهم من المعادلة (15):

$$m_j = E_t^I(z_{t+j}) + \frac{E_t^{I''}[E_t(z_{t+j})] \{E_t^{I'}[E_t(z_{t+j})]\}^{-2} s_j^2}{2} \quad (19)$$

الذي ينتج معادلة من الدرجة الثانية في  $E_t(z_{t+j})$ ، وبالحل نحصل على:

$$E_t(z_{t+j}) = m_j^{\frac{1}{I}} \left\{ 0.5 \pm 0.5 \left[ 1 + \frac{2(I^{-1} - 1)s_j^2}{m_j^2} \right]^{\frac{1}{2}} \right\}^{\frac{1}{I}} \quad (20)$$

وبما أن:

$$E_t(z_{t+j}) = T^{-1}(m_j) \cdot C_I^{(I)} \quad (21)$$

فان عامل حذف التحيز التقريبي An Approximate Debiasing Factor لتحويل القوة  
 الاعتيادي (النوع الثاني) (3) يؤخذ كما يلي :

$$C_I^{(j)} = \begin{cases} \left\{ 0.5 + 0.5 \left[ \frac{1 + 2(I^{-1} - 1)s_j^2}{m_j^2} \right]^{\frac{1}{2}} \right\}^{\frac{1}{I}} & ; \text{if } I \neq 0 \\ \exp \left( \frac{s_j^2}{2} \right) & ; \text{if } I = 0 \end{cases} \quad (22)$$

بنفس الطريقة يصبح عامل حذف التحيز التقريبي لتحويل Box-Cox (1) كما يلي:

$$C_I^{(j)} = \begin{cases} \left\{ 0.5 + 0.5 \left[ 1 + \frac{2(I^{-1} - 1)s_j^2}{(I^{-1} + m_j)^2} \right]^{\frac{1}{2}} \right\}^{\frac{1}{I}} & ; \text{if } I \neq 0 \\ \exp \left( \frac{s_j^2}{2} \right) & ; \text{if } I = 0 \end{cases} \quad (23)$$

كما أن عامل حذف التحيز التقريبي لتحويل Non Box-Cox (2) يصبح كما يلي:

$$C_i^{(j)} = \begin{cases} \left\{ 0.5 + 0.5 \left[ 1 + \frac{2(I^{-1} - 1)(s_j \&^{-1})^2}{(I^{-1} + m_j \&^{-1})^2} \right]^{\frac{1}{2}} \right\}^{\frac{1}{i}} & ; \text{if } I \neq 0 \\ \exp \left( \frac{s_j}{\bar{y}} \right)^2 & ; \text{if } I = 0 \end{cases} \quad (24)$$

ويجب أيضا تصحيح التحيز لفترة الثقة التالية والتي تقابل التكهن البسيط :

$$T^{-1} \left\{ \hat{m}_j \pm z_{\alpha/2} \left[ \hat{v}ar_t(e_{t+j}) \right]^{\frac{1}{2}} \right\} \quad (25)$$

يتم حذف التحيز بضرب نهاية النقاط من المعادلة (25) بعامل حذف التحيز (24) (23) (22)، بافتراض أن  $I$  قيمة ثابتة معلومة ، وعندما تقدر  $I$  من البيانات المشاهدة فإن تباين  $\hat{I}$  قد يتأثر بالنتائج السابقة ، بالرغم من أن هذا التأثير صغير عموما . (Taylor,1986).

#### 4. الشواذ في السلاسل الزمنية :

إن توفر بيانات خالية من أي شواذ قد يكون غير متوقع عبر الزمن ،وان هذه الملوثات تؤثر على عملية التحليل الإحصائي الدقيق مالم تدرس وتحدد ،فالملوثات هي تلك القيم التي تكون خارجة عن النسق العام للبيانات مما يجعلها لا تتوافق مع بعض الفرضيات الأساسية الواجب توافرها للنموذج الإحصائي (محمد،2008). حيث تتضمن السلسلة الزمنية في بعض الحالات مشاهدات غير عادية كبيرة سببها حوادث مجهولة أو غير متوقعة في السلسلة الزمنية هذه المشاهدات الغير عادية تدعى بالشواذ (Liu,2006).

إن تحليل السلسلة الزمنية ودقة المعلومات فيها يعتمد بشكل عالي على دقة البيانات وعدم احتوائها على الشواذ ،لذلك من المهم تحديد إن كان للسلسلة شواذ أم لا وماهي طبيعتها قبل القيام بتحليل السلسلة ،ومعرفة طرق معالجتها أو إزالة تأثيرها بعد معرفة موقعها (محمد، 2008)، فقد يكون للشواذ تأثير معتدل إلى ضخم على تحليل السلسلة الزمنية (Liu,2006)، كما إن وجود الشواذ في البيانات يجعل المقدرات حساسة وتقلل من صفاتها الجيدة وتعتمد دقة وكفاءة المقدرات على دقة

اختيار طريقة التقدير الملائمة. فهذه المقدرات قد تكون كفوءة ومناسبة في حالة توافر جميع شروط نماذج السلاسل الزمنية. كما أن للشواذ تأثير على أداء التكهّن للعينة المتقدمة Post Sample بالإضافة إلى تأثيرها على طرق إعادة تحويل التكهّنات فبسبب وجود الشواذ قد تعطي طريقة إعادة التحويل المباشرة Straight Retrtransformed تكهّنات أفضل من طريقة إعادة التحويل غير المتحيزة Unbiased Retrtransformed.

إن جميع التعاريف حول الشواذ تنصب في مفهوم واحد بكونها المشاهدة غير المتسقة مع بقية المشاهدات وان ظهور القيم الشاذة ناشئ عن أحد الأسباب التالية:

1. أخطاء القياس الناتجة من أخطاء التسجيل.
2. أخطاء المعاينة الناتجة من سوء اختيار العينة وعدم تمثيلها للمجتمع تمثيلاً جيداً أو بسبب التحيز في اختيار العينة.
3. أو تحدث لسبب طبيعي تفرضه حالة معينة . (الصفراوي، 2005).

أنواع الشواذ في السلاسل الزمنية:

إن عملية الكشف عن الشواذ وتحديد طبيعتها هي المرحلة الأكثر أهمية في عملية معالجة تأثيراتها السلبية على مقدرات معالم النماذج المختلفة. وبصورة عامة هناك أربعة أنواع من الشواذ (محمد، 2008، Liu، 2006):

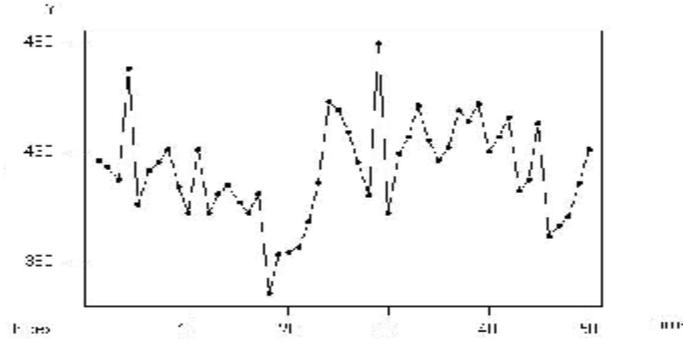
1. الشواذ المضافة Additive Outlier (AO)
2. الشواذ النمطية Innovational Outlier (IO)
3. تغيير المستوى Level Shift (LS)
4. تغيير وقتي (زائل) Temporary change (TC)

5. الجانب التطبيقي:

سيتم في هذا البحث تحليل السلسلة الزمنية للمعدل اليومي لاستهلاك الطاقة الكهربائية في محافظة نينوى للفترة من 2008/6/10 ولغاية 2008/8/7 (انظر الملحق 1) حيث سيتم استخدام أول خمسين مشاهدة في تقدير النموذج المناسب في حين تبقى المشاهدات التسعة الأخيرة كعينة متقدمة Post Sample لفحص دقة التكهّنات التي سيتم الحصول عليها من النموذج المختار وسيتم الاستفادة من البرنامج الإحصائي SCA (انظر الملحق 2).

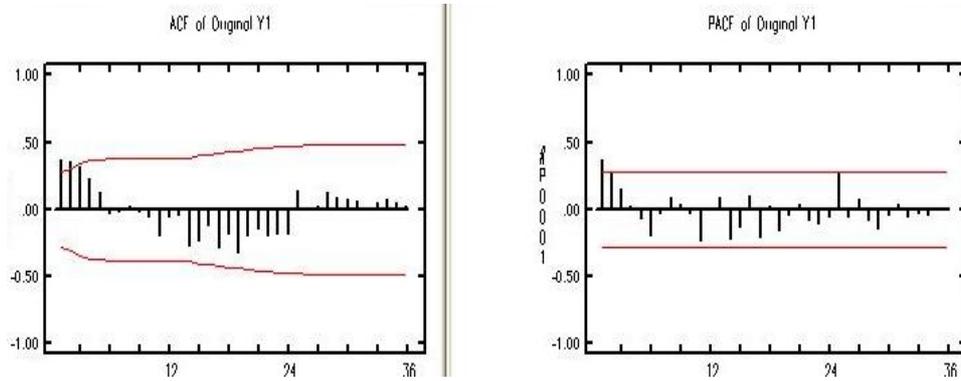
أ- وصف السلسلة الزمنية:

إن الخطوة الأولى في تحليل السلسلة الزمنية هي رسم السلسلة الزمنية Time Series Plot لمعرفة ما إذا كانت السلسلة الزمنية مستقرة أم لا، والشكل (1) أدناه يوضح رسم السلسلة الزمنية الخاصة بالمعدل اليومي لاستهلاك الطاقة الكهربائية في محافظة نينوى:



الشكل (1): رسم السلسلة الزمنية الخاصة بالمعدل اليومي لاستهلاك الطاقة الكهربائية في محافظة نينوى.

ومن خلال الرسم يتبين أن السلسلة الزمنية غير مستقرة في المتوسط، ويمكن التوثق من هذا الاستنتاج باختبار استقرار السلسلة من خلال رسم دالتي الارتباط الذاتي والارتباط الجزئي للشلسلة وكما مبين في الشكل (2) أدناه:



الشكل (2): دالتي الارتباط الذاتي (ACF) والارتباط الذاتي الجزئي (PACF) للسلسلة الزمنية الخاصة بالمعدل اليومي لاستهلاك الطاقة الكهربائية في محافظة نينوى.

ويمكن التأكد من عدم استقرار السلسلة الزمنية بإجراء اختبار ديكي-فولر لملاحظة ما إذا كانت السلسلة الزمنية تحوي جذر وحدة أم لا، وأدناه نتائج اختبار ديكي-فولر للسلسلة مدار البحث.

UROOT Y1. METHOD DF. ORDERS 0. SPAN 1,50.  
SIMPLE DICKEY-FULLER TEST  
SERIES TO BE TESTED FOR UNIT ROOT(S) IS: Y1  
TEST BASED ON SAMPLE SIZE: 50  
PROBABILITY OF A SMALLER VALUE IS SET TO: 0.050

(<---- INTERPOLATED D-F TABLE ---->)

CRITICAL VALUE BASED ON OLS t-STATISTIC

ORDER	TEST STATISTIC	0.010 LEVEL	0.050 LEVEL	0.100 LEVEL	0.050 LEVEL	UROOT
0	-0.18	-2.62	-1.95	-1.61	-1.95	YES

(<---- INTERPOLATED D-F TABLE ---->)

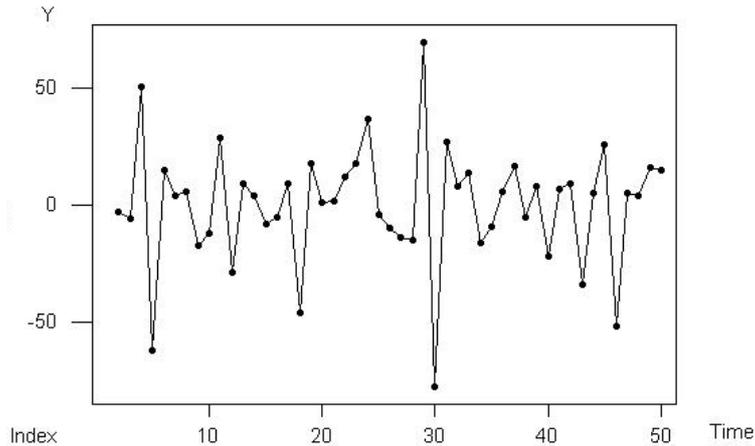
CRITICAL VALUE BASED ON OLS AR COEFF.

ORDER	TEST STATISTIC	0.010 LEVEL	0.050 LEVEL	0.100 LEVEL	0.050 LEVEL	UROOT
0	-0.08	-12.90	-7.70	-5.50	-7.70	YES

--

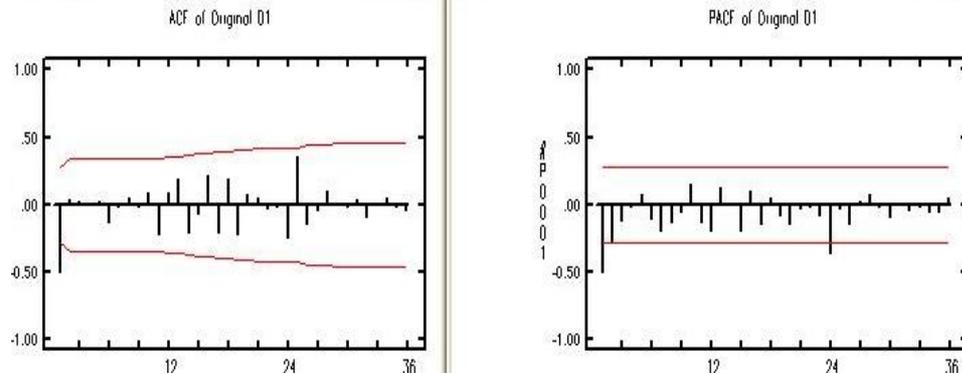
DIFFERENCE Y1. NEW D1. DFORDERS 1.  
DIFFERENCE ORDERS ARE (1-B)  
SERIES Y1 IS DIFFERENCED, THE RESULT IS STORED IN VARIABLE D1  
SERIES D1 HAS 50 ENTRIES

ومن خلال النتائج أعلاه يتبين أن للسلسلة الزمنية جذر وحدة حيث إن قيمة احصاء الاختبار أكبر من القيمة الحرجة عند جميع المستويات في كلا النتيجتين عليه تقبل فرضية عدم القائلة (للسلسلة الزمنية جذر وحدة) أي أن السلسلة غير مستقرة، ولتحقيق استقرار السلسلة الزمنية تم اخذ الفرق الأول لبيانات السلسلة الزمنية، والشكل (3) أدناه يوضح رسم السلسلة الزمنية بعد اخذ الفرق الأول.



الشكل (3): رسم السلسلة الزمنية بعد اخذ الفرق الأول .

وللتحقق من استقرارية السلسلة الزمنية تم رسم دالتي الارتباط الذاتي والارتباط الذاتي الجزئي بعد اخذ الفرق الأول للسلسلة الزمنية وكما مبين أدناه :



الشكل (4): دالتي الارتباط الذاتي (ACF) والارتباط الذاتي الجزئي (PACF) بعد اخذ الفرق

الأول للسلسلة الزمنية.

ومن خلال الفحص البصري يلاحظ أن دالة الارتباط الذاتي الجزئي تتناقص بشكل تدريجي في حين أن دالة الارتباط الذاتي تهبط إلى الصفر بعد الفجوة الزمنية الأولى مما يدل على أن السلسلة مستقرة. وبإجراء اختبار ديكي-فولر أدناه يتبين أن السلسلة الزمنية ليس لها جذر وحدة حيث أن قيمة احصاء الاختبار اقل من القيمة الحرجة عند كافة المستويات عليه ترفض فرضية العدم وتقبل الفرضية البديلة القائلة ( لاتحوي السلسلة الزمنية جذر وحدة ) أي أن السلسلة أصبحت مستقرة .

```
UROOT D1. METHOD DF. ORDERS 0. SPAN 1,50.
SIMPLE DICKEY-FULLER TEST
SERIES TO BE TESTED FOR UNIT ROOT(S) IS:          D1
TEST BASED ON SAMPLE SIZE:                        49
PROBABILITY OF A SMALLER VALUE IS SET TO:         0.050
```

(<---- INTERPOLATED D-F TABLE ---->)

CRITICAL VALUE BASED ON OLS t-STATISTIC

ORDER	TEST STATISTIC	0.010 LEVEL	0.050 LEVEL	0.100 LEVEL	0.050 LEVEL	UROOT
0	-11.81	-2.62	-1.95	-1.61	-1.95	NO

(<---- INTERPOLATED D-F TABLE ---->)

CRITICAL VALUE BASED ON OLS AR COEFF.

ORDER	TEST STATISTIC	0.010 LEVEL	0.050 LEVEL	0.100 LEVEL	0.050 LEVEL	UROOT
0	-71.97	-12.86	-7.68	-5.49	-7.68	NO

--

ب- تحليل السلسلة الزمنية ضمن مدى من قيم  $I$  :

في هذا الجزء سيتم البحث عن قيمة  $I$  ضمن مدى معين من (0.95 إلى -0.95) اعتماداً على تصغير الجذر التربيعي لمتوسط مربعات الخطأ (Root Mean Square Error (RMSE) وعندما تكون بيانات العينة المتقدمة متوفرة فإنه يمكن الحصول على RMSE لكل  $I$  وذلك من خلال التكهن ب9 خطوات للأمام ، إن مقياس RMSE المستخدم كإشارة للتحقق من كفاءة تحويل قوة معين هو .

$$RMSE = \sqrt{\frac{1}{m} \sum_{t=1}^m (Y_t - \hat{Y}_t)^2} \quad (26)$$

حيث أن:

$Y_t$  : المشاهدة الحقيقية عند الزمن  $t$  ،

$\hat{Y}_t$  : القيمة المتوقعة ل  $Y_t$  المعتمدة على نموذج معين أو طريقة معينة .

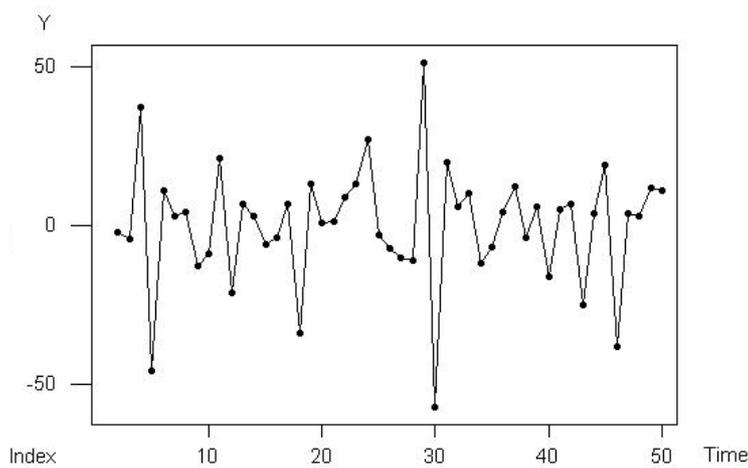
$m$  : عدد المشاهدات لفترة العينة المتقدمة.

جدول رقم (1) : يبين قيم RMSE ضمن مدى محدد من قيم  $I$  .

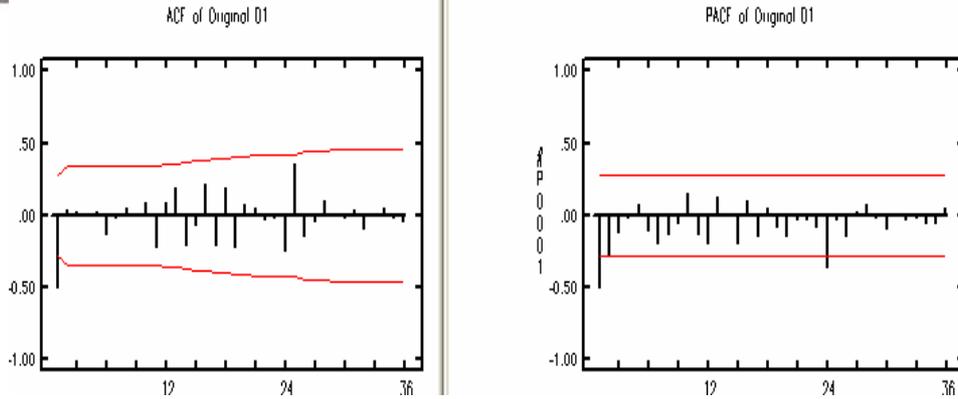
$I$	RMSE	$I$	RMSE
0.95	19.88885*	-0.95	21.08143
0.90	19.91616	-0.90	21.05056
0.85	19.94360	-0.85	21.01344
0.80	19.97134	-0.80	20.97798
0.75	19.99992	-0.75	20.94076
0.70	20.02722	-0.70	20.90531
0.65	20.05553	-0.65	20.87170
0.60	20.08414	-0.60	20.83640
0.55	20.11265	-0.55	20.80244
0.50	20.14167	-0.50	20.76827
0.45	20.17075	-0.45	20.73477
0.40	20.20014	-0.40	20.70022
0.35	20.22951	-0.35	20.66789

0.30	20.25929	-0.30	20.63505
0.25	20.28943	-0.25	20.60218
0.20	20.31937	-0.20	20.56990
0.15	20.34976	-0.15	20.52742
0.10	20.38058	-0.10	20.50582
0.5	20.14167	-0.5	20.76827
0.0	20.59126		

وكما هو واضح من الجدول أعلاه فإن اصغر قيمة للجذر التربيعي لمتوسط مربعات الخطأ هي عند  $I = 0.95$ ، عليه سيتم تحويل السلسلة الزمنية اعتمادا على هذه القيمة، حيث تم تحويل المشاهدات الخمسين الأولى باستخدام تحويل BOX-COX عند  $I = 0.95$  بعد ذلك نقوم بتحليل السلسلة الزمنية المحولة من خلال منهجية بوكس-جنكنز، حيث سيتم أولا رسم السلسلة الزمنية بعد اخذ تحويل BOX-COX عند  $I = 0.95$  بعد أن تم اخذ الفرق الأول للسلسلة الزمنية وكما موضح بالشكل (5) أدناه :



الشكل (5): رسم السلسلة الزمنية بعد اخذ التحويل المناسب للسلسلة والفرق الأول .



الشكل (6): دالتي الارتباط الذاتي (ACF) والارتباط الذاتي الجزئي (PACF) بعد اخذ التحويل المناسب للسلسلة والفرق الأول.

وبناء على ما ظهر من نتائج في أعلاه يمكن القول أن نموذج السلسلة الزمنية الملائم لبيانات السلسلة هو نموذج المتوسطات المتحركة من الرتبة الأولى MA(1)، ويمكن استخدام معيار معلومات أكاي المصحح للتأكد من اختيار النموذج الملائم للسلسلة، وتحسب قيمة ( $AIC_c$ ) كما يلي (Chatfield and Faraway, 1998):

$$AIC_c = n \ln \left( \frac{S^2}{n} \right) + 2p + \frac{2(p+1)(p+2)}{n-p-2} \quad (27)$$

حيث أن :

n: حجم العينة .

$S^2$ : مجموع مربعات البواقي Sum Of Square Residual، إذ أن البواقي تمثل خطأ التكهّن لخطوة واحدة للأمام داخل العينة (Within Sample) .  
P: عدد معلمات النموذج المقدر .

وفيما يلي حساب قيمة ( $AIC_c$ ) لبعض النماذج المرشحة وكما موضح بالجدول (2) أدناه:

الجدول (2): ملائمة نماذج ARMA(p,d,q) للبيانات المحولة وحساب قيمة  $AIC_c$  لها.

p	d	q	$AIC_c$
1	1	0	88.8911
2	1	0	87.6834
0	1	1	84.0690*
0	1	2	86.9240
1	1	1	87.1849
2	1	2	94.1811

\*أقل قيمة لمعيار  $AIC_c$ .

وكما هو واضح من الجدول فان قيمة  $(AIC_c)$  الصغرى عند النموذج  $ARIMA(0,1,1)$  مما يعزز النتيجة السابقة بان النموذج الملائم للسلسلة الزمنية هو  $MA(1)$ . وبعد تحديد نموذج السلسلة الزمنية فان النموذج المقدر هو كما موضح أدناه:

```
TSMODEL UTSMODEL. MODEL IS @
Y1(1) = (1)NOISE.
ESTIM UTSMODEL. HOLD RESID(_RES).
THE FOLLOWING ANALYSIS IS BASED ON TIME SPAN 1 THRU 50
```

NONLINEAR ESTIMATION TERMINATED DUE TO:  
RELATIVE CHANGE IN (OBJECTIVE FUNCTION)\*\*0.5 LESS THAN 0.1000D-02

SUMMARY FOR UNIVARIATE TIME SERIES MODEL - UTSMODEL

VARIABLE	TYPE OF VARIABLE	ORIGINAL OR CENTERED	DIFFERENCING
Y1	RANDOM	ORIGINAL	1 (1-B )

PARAMETER LABEL	VARIABLE NAME	NUM. / DENOM.	FACTOR	ORDER	CONS- TRRAINT	VALUE	STD ERROR	T VALUE
1	Y1	MA	1	1	NONE	.6516	.1075	6.06

EFFECTIVE NUMBER OF OBSERVATIONS . . . . . 49  
R-SQUARE . . . . . 0.117  
RESIDUAL STANDARD ERROR. . . . . 0.156999E+02  
--

وقبل إجراء التكهّن يجب توضيح أن السلسلة الزمنية قيد الدراسة تحوي قيم شاذة وكما موضح أدناه:

OUTLIER UTSMODEL. TYPES AO, IO, LS. SPAN 1,50.

TIME	ESTIMATE	T-VALUE	TYPE
29	41.67	3.14	AO
24	30.67	3.06	LS

وكما هو واضح فان المشاهدة 29 مشاهدة شاذة من نوع (AO) (شواذ مضافة) ، أما المشاهدة 24 فهي مشاهدة شاذة من نوع (LS) (تغيير المستوى) .

إن الخطوة الأخيرة من منهجية بوكس-جنكنز هي التكهّن بالسلسلة الزمنية حيث تم التكهّن ب9 خطوات للأمام ثم القيام بإعادة تحويل التكهّنات بطريقة إعادة التحويل الغير متحيز Unbiased Retransformed وبطريقة إعادة التحويل باتجاه القياس الأصلي المباشرة Straight Retransformed والنتائج مبينة في الجدول (3) أدناه مع مقارنتها بقيم العينة المتقدمة Post : Sample

جدول رقم (3) : يبين قيم التكهّن من خلال نموذج  $MA(1)$  والخطأ المعياري وإعادة تحويل التكهّنات بطريقة إعادة التحويل الغير متحيزة وطريقة إعادة التحويل باتجاه القياس الأصلي المباشرة بالإضافة إلى قيم العينة المتقدمة.

1	2	3	4	5	6
TIME	FORECAST	STD.ERROR	RETRANSFORMED UMBIASED	RETRANSFORMED STRAIGHT	POST SAMPLE
1	323.7577	15.6999	417.233	417.334	426
2	323.7577	16.6256	417.250	417.381	411
3	323.7577	17.5025	417.255	417.367	383
4	323.7577	18.3374	417.257	417.389	402
5	323.7577	19.1360	417.262	417.391	391
6	323.7577	19.9026	417.270	417.399	391
7	323.7577	20.6407	417.273	417.415	405
8	323.7577	21.3533	417.277	417.318	395
9	323.7577	22.0428	417.280	417.512	414

إن تكهّنات إعادة التحويل الغير متحيز أعطت اصغر جذر تربيعي لمتوسط مربعات الخطأ  $RMSE = 19.88873$  مقارنة مع تكهّنات إعادة التحويل المباشرة التي أعطت جذر تربيعي لمتوسط مربعات الخطأ  $RMSE = 19.97687$ ، عليه فان طريقة إعادة التحويل الغير متحيز أكثر كفاءة من طريقة إعادة التحويل المباشرة.

## 6. الاستنتاجات :

1. إن الضعف الواضح في التكهّنات بناء على القيم المحولة في العمود (2) سببه أن التكهّنات تمت بناء على البيانات المحولة والتي أصبحت رياضياً نتيجة التحويل تختلف عن البيانات الأصلية .
2. إن تكهّنات إعادة التحويل الغير متحيز أعطت اصغر جذر تربيعي لمتوسط مربعات الخطأ  $RMSE = 19.88873$  مقارنة مع تكهّنات إعادة التحويل المباشرة التي أعطت جذر تربيعي لمتوسط مربعات الخطأ  $RMSE = 19.97687$ ، عليه فان طريقة إعادة التحويل الغير متحيز أكثر كفاءة من طريقة إعادة التحويل المباشرة.
3. إن وجود الشواذ في السلسلة الزمنية يؤدي عادة إلى أن تكون تكهّنات الطريقة المباشرة أفضل من تكهّنات الطريقة غير المتحيزة ، ولكن تبين من الحالة الدراسية انه ليس بالضرورة دائماً وجود الشواذ يقود إلى النتيجة أعلاه وإنما قد تتأثر النتيجة بنوع وحجم وعدد الشواذ الموجودة في السلسلة الزمنية .
4. عدم إمكانية الاستفادة من نموذج  $MA(1)$  المستحصل عليه لأنه تم الحصول عليه بناء على البيانات المحولة.

المصادر:

1. أمين بك، عزة حازم زكي (2004)، "استخدام الشبكات العصبية في التكهّن للسلاسل الزمنية بتطبيق على استهلاك الطاقة الكهربائية في محافظة نينوى"، رسالة ماجستير غير منشورة، كلية علوم الحاسبات والرياضيات، جامعة الموصل.
2. محمد، سالم بدر، (2008)، "أساليب الكشف عن الشوارد في السلاسل الزمنية ذات الدرجات الدنيا (دراسة تجريبية)"، مجلة التقني، المجلد 21، العدد 6، العراق.
3. الصفاوي، صفاء يونس طليح، (2005). "مقارنة بين المقدرات الاعتيادية والحصينة لنماذج السلاسل الزمنية المختلطة الثنائية من الرتب الدنيا"، أطروحة دكتوراه غير منشورة، كلية علوم الحاسبات والرياضيات، جامعة الموصل.
4. Guerrero, V.M. (1993). "Time-series analysis supported by power transformations". Journal of forecasting 12:37-48.
5. Liu, L.-M., 2006, "Time Series Analysis and Forecasting ", 2<sup>nd</sup> ed., Scientific Computing Associates Crop., Illinois ,U.S.A..
6. Taylor, J, M, G., "The retransformed mean after a fitted power transformation " , Journal of the American Statistical Association , 81, (1986) , 114-18 .
7. Chatfield C., and Faraway, J., 1998, "time series Forecasting With Neural Networks :a Comparative Study Using the Airline Data ". Applied Statistics. Vol.47, part 2, 231-250.
8. Granger, C. W. J. and Newbold, P. (1986). Forecasting Economic Time Series, 2nd ed. San Diego: Academic Press, p. 4.
9. Ansley ,C.F., Spivey, W.A. and Wroblewski, W.J. (1977) . "A class of transformations for Box-Jenkins seasonal models". Applied Statistics 26: 173-178.
10. Liu, L.-M. and Hudak, G.B. , 2002, " THE SCA STATISTICAL SYSTEM REFERENCE MANUAL FOR GENERAL STATISTICAL ANALYSIS" , Illinois: Scientific Computing Associates Corp. .
11. Pankratz ,A., 1991, "Forecasting With Dynamic Regression Models ", John Wiley and Sons , U.S.A..

الملحق (1)

المشاهدات الحقيقية لاستهلاك الطاقة الكهربائية في مدينة الموصل للفترة 2008/6/10 ولغاية 2008/7/29:

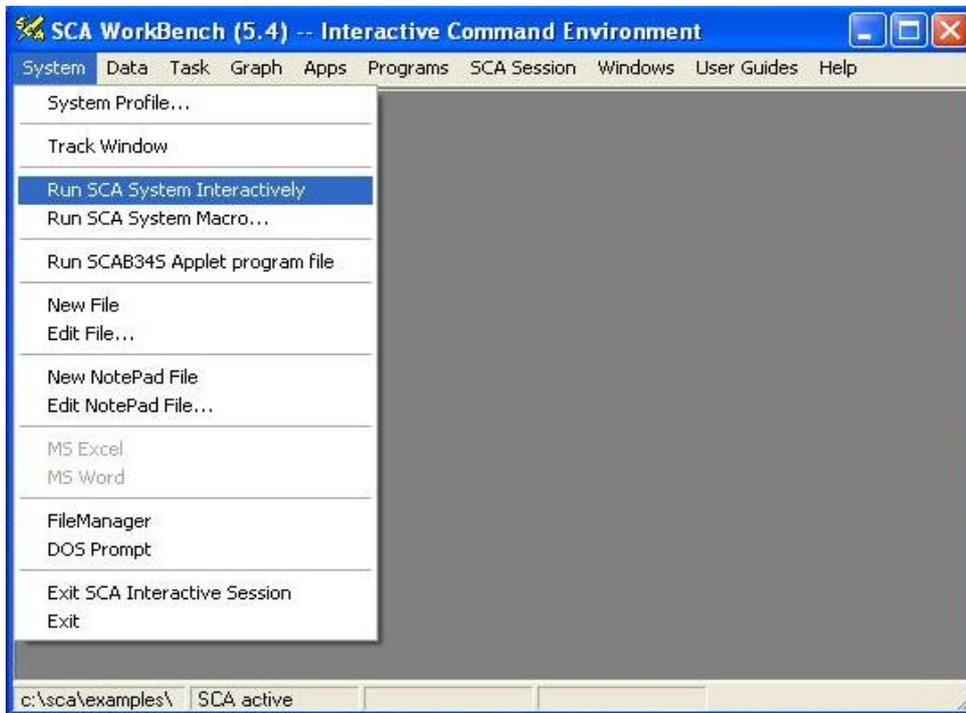
التسلسل	التاريخ	المشاهدة	التسلسل	التاريخ	المشاهدة
1	2008/6/10	426	26	2008/7/5	439
2	2008/6/11	423	27	2008/7/6	425
3	2008/6/12	417	28	2008/7/7	410
4	2008/6/13	468	29	2008/7/8	480
5	2008/6/14	406	30	2008/7/9	402
6	2008/6/15	421	31	2008/7/10	429
7	2008/6/16	425	32	2008/7/11	437
8	2008/6/17	431	33	2008/7/12	451
9	2008/6/18	414	34	2008/7/13	435
10	2008/6/19	402	35	2008/7/14	426
11	2008/6/20	431	36	2008/7/15	432
12	2008/6/21	402	37	2008/7/16	449
13	2008/6/22	411	38	2008/7/17	444
14	2008/6/23	415	39	2008/7/18	452
15	2008/6/24	407	40	2008/7/19	430
16	2008/6/25	402	41	2008/7/20	437
17	2008/6/26	411	42	2008/7/21	446
18	2008/6/27	365	43	2008/7/22	412
19	2008/6/28	383	44	2008/7/23	417
20	2008/6/29	384	45	2008/7/24	443
21	2008/6/30	386	46	2008/7/25	391
22	2008/7/1	398	47	2008/7/26	396
23	2008/7/2	416	48	2008/7/27	400
24	2008/7/3	453	49	2008/7/28	416
25	2008/7/4	449	50	2008/7/29	431

## الملحق (2): نبذة عن نظام SCA Statistical System

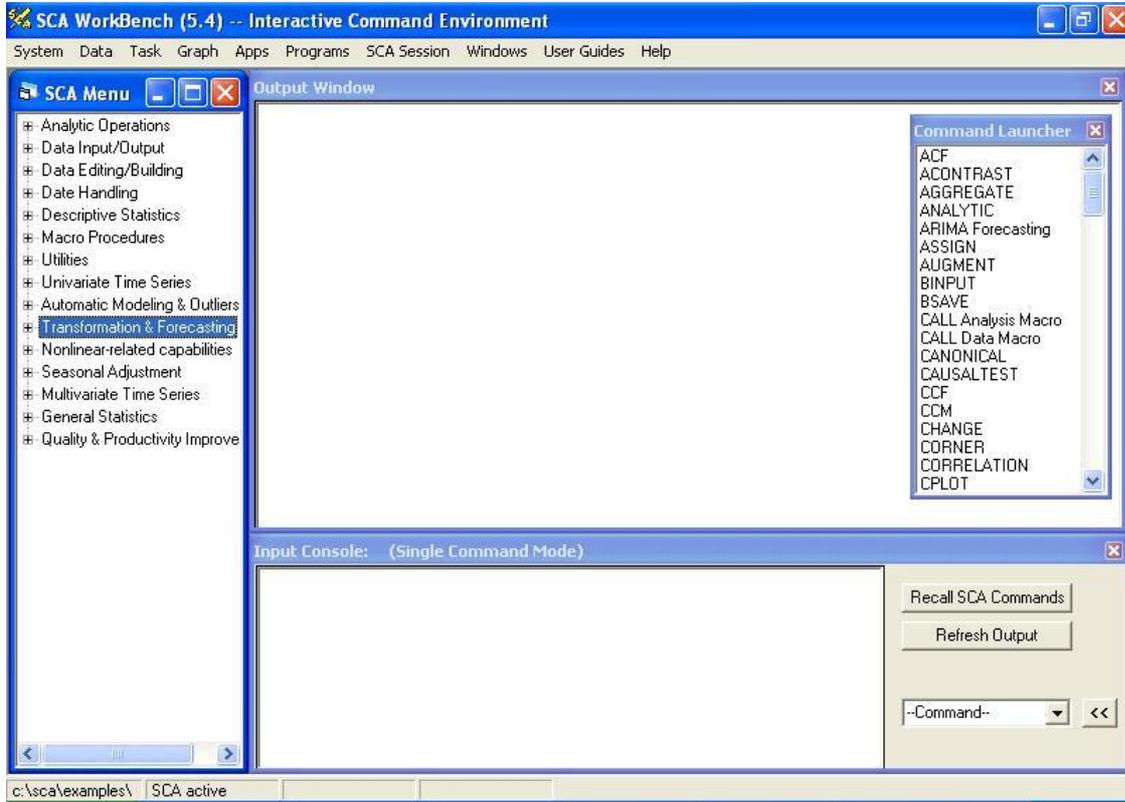
إن نظام SCA Statistical System نظام إحصائي يهتم بمجال تحليل السلاسل الزمنية خاصة بالإضافة إلى اهتمامه بالمجالات الإحصائية الأخرى. وبتشغيل الأيقونة الخاصة بالنظام ستظهر واجهة النظام بالشكل التالي:



وبتنفيذ الإيعاز Run SCA System Interactively من قائمة System كما في الشكل التالي:



لتصبح واجهة النظام بالشكل التالي:



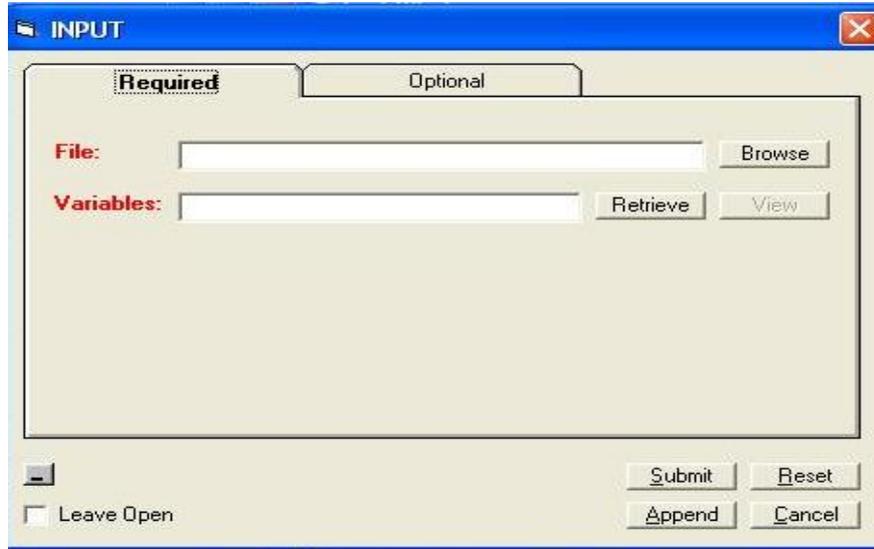
ويمكن تنفيذ الأوامر الخاصة بالنظام بعدة أساليب منها:

1. عن طريق اختيار الإيعاز من قائمة Command Luncher والتي تحوي كافة إيعازات النظام.
2. عبر اختيار الإيعاز من قائمة SCA Menu التي تحوي كافة الإيعازات.
3. عن طريق طباعة الإيعاز بشكل صحيح بكافة مدخلاته في مربع النص الخاص بإيعازات الإدخال والمسمى Input Console.

وبعد كتابة الإيعاز بشكل صحيح يظهر تنفيذه في شاشة الإخراجات Output Windows. كما يحوي النظام على قوائم الأوامر التي من الممكن اختيارها من شريط القوائم في الجهة العليا من الواجهة.

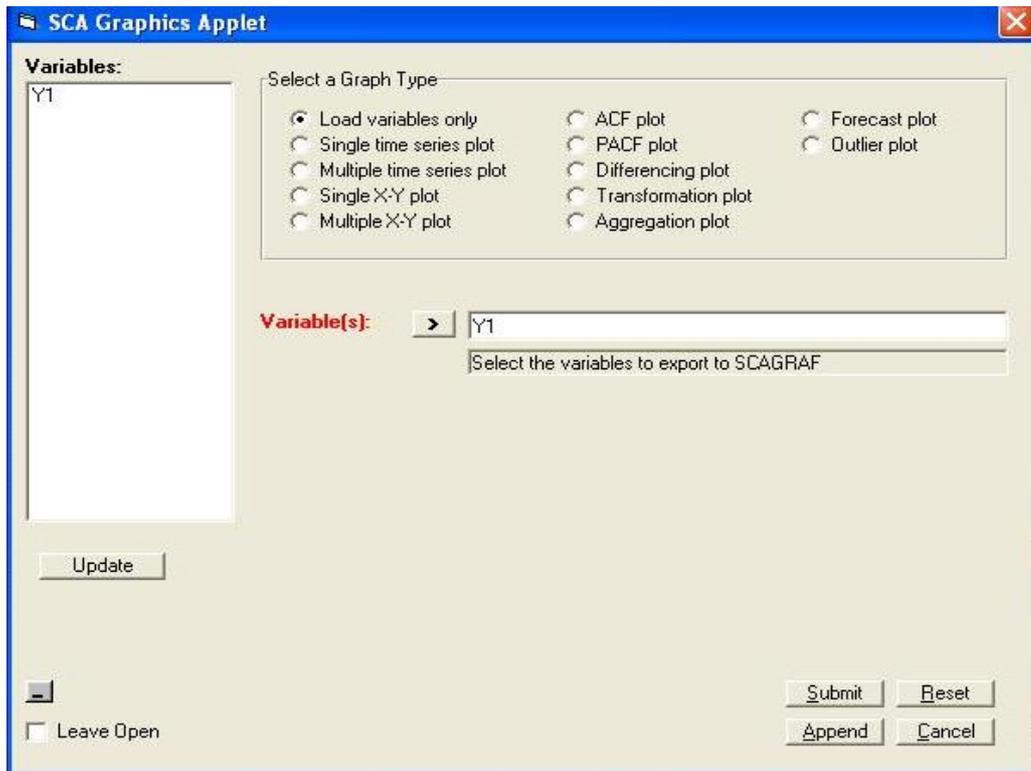
أما بالنسبة لإدخال البيانات إلى النظام فهناك عدة طرق منها عن طريق استيرادها من أوراق عمل خاصة بأنظمة أخرى مثل Microsoft Excel وذلك بعد تحويلها إلى أوراق عمل خاصة بنظام SCA.

أو عن طريق الإدخال المباشر في شاشة Input Console. كما يمكن إدخال البيانات إلى النظام عن طريق اختيار الإيعاز Input حيث يتم استيراد البيانات من ملفات مخزونة في الحاسوب كما في الشكل التالي:



INPUT dialog box with tabs for Required and Optional. It contains fields for File and Variables, with buttons for Browse, Retrieve, and View. At the bottom, there are buttons for Submit, Reset, Append, and Cancel, and a checkbox for Leave Open.

كما يحوي النظام على إمكانية إجراء الرسوم البيانية المختلفة وذلك عن طريق الإيعاز Graph من قائمة Descriptive Statistics وبتنفيذ هذا الأمر يظهر مربع الحوار التالي:



SCA Graphics Applet dialog box. It features a Variables list containing 'Y1'. A 'Select a Graph Type' section includes radio buttons for: Load variables only (selected), Single time series plot, Multiple time series plot, Single X-Y plot, Multiple X-Y plot, ACF plot, PACF plot, Differencing plot, Transformation plot, Aggregation plot, Forecast plot, and Outlier plot. Below this is a 'Variable(s):' field with a right arrow and 'Y1' entered, and a sub-field 'Select the variables to export to SCAGRAF'. At the bottom, there are buttons for Update, Submit, Reset, Append, and Cancel, and a checkbox for Leave Open.

كما يحوي النظام على ملحقات خاصة تفيد في عمليات تحليل البيانات ومعالجتها بالإضافة إلى شرح وافٍ يستفاد منه في التطبيق العملي.

وللنظام عدة إصدارات منها الإصدار التعليمي Educational Edition والإصدار الاحترافي بنوعيه Professional Edition (A) and Professional Edition (B) والإصدار المتقدم Advanced Edition وإصدارات أخرى وكل إصدار له مواصفاته الخاصة وسعره الخاص ومدة انتهائه المحددة.

ولشراء نسخة من النظام يتم مراسلة شركة Scientific Computing Associates Corp عن طريق موقع الشركة عبر الانترنت [www.scausa.com](http://www.scausa.com) أو عبر البريد الإلكتروني [sca@scausa.com](mailto:sca@scausa.com) تلفون 834-4567- (630) فاكس 834-1450- (630).

