المجلة العراقية للعلوم الاحصائية (17) 2010
عدد خاص بوقائع المؤتمر العلمي الثاني للرياضيات-الاحصاء والمعلوماتية
ص ص [171– 194]

إعادة تحويل التكهنات عكسيا في السلاسل الزمنية

ظافر رمضان مطر*　　　　مروة نبيل**

الملخص

إن مما لاشك فيه أن قوة وفعالية النماذج الإحصائية تتحسن غالبا بعد إجراء التحويل المناسب على البيانات في حالة الحاجة الى هذا التحويل ،وان هذه الفعالية لاشك تنعكس أيضا على التحليل والتكهن باستخدام نماذج السلاسل الزمنية. من جانب أخر فان التعديل الرياضي للبيانات ينتج عنه إشكالات معينة في تفسير نتائج النماذج وعدم إمكانية الاستفادة من التكهنات التي يتم الحصول عليها اعتمادا على البيانات المحولة لذا فانه سيتم في هذا البحث تناول طرائق إعادة تحويل التكهنات 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 reflected also on analysis and forecasting using time series models. The mathematical modification of the data raises problems for the interpretation of models results and 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 electric power consumption in Nineveh governorate.

---

* استاذ مساعد/قسم الاحصاء والمعلوماتية/كلية علوم الحاسبات والرياضيات
**طالبة ماجستير/ قسم الاحصاء والمعلوماتية/كلية علوم الحاسبات والرياضيات
تاريخ التسلم :2009/7/1 ———— تاريخ القبول : 6 /12 /2009

1. المقدمة:

إن عملية تقدير السلوك المستقبلي لأية ظاهرة أو متغير تعد من الأهداف الأساسية للعلوم الإحصائية لأهميتها الكبيرة في عملية التخطيط للظاهرة أو المتغير قيد الدراسة وان التخطيط هـو المفتاح لعملية اتخاذ القرار السليم، اذ بإمكاننا التعرف على الظروف والحالة المحيطة بالمشكلة قيد الدراسة قبل اتخاذ القرار،لذلك فان موضوع التكهن لاقى ويلاقي اهتماما كبيرا ومتزايدا من قبـل الباحثين وأصحاب القرارات (أمين بك،2005) .فالتكهن 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 التكهن من ابرز الأهداف التي يصبو إليها معظـم البـاحثين ومتخذي القرار.

إن فعالية النماذج الإحصائية تتعزز غالبا من خلال استخدام البيانات المحولة ولا يستثنى من ذلك التحليل والتكهن باستخدام نماذج السلاسل الزمنية.فمن المهم اختيار التحويل الملائم الـذي يفـي بالغرض .تفترض معظم الطرائق الإحصائية أن المتغيرات موزعة توزيعـا طبيعيـا ،وتحويـل البيانات أداة مفيدة لجعل المتغيرات تحت الدراسة طبيعية.وان التعديل الرياضـي للبيانـات بهـذا الأسلوب ينشئ نقاط خلاف أو نقاشا ليست فقط في تفسير نتائج النماذج ولكن أيضا عـدم فائـدة التكهن اعتمادا على البيانات المحولة (Liu،2006) فحالما يتم بناء النموذج اعتمادا على المقيـاس المحول فان التكهنات التي يتم الحصول عليها في النظام المتري تصبح بحاجة إلى إعادة تحويلهـا Retransformation وتظهر عند ذاك مشكلة أن خطوات إعادة التحويل تسبب أو تنتج تحيزا في التكهنات ولتصحيح هذا التحيز اشتق (Guerrero،1993) عوامل لحذف التحيز واقترح ضـرب حدود التكهن بعوامل حـذف التحيـز الخـاص بكـل نـوع مـن تحـويلات القـوة Power Transformation.

إن اهتمام الباحثين ينصب في هذا البحث على تطبيق تحويلات القوى لتحسين دقة التكهن عنـدما يعاد تحويل التكهنات باتجاه القياس الأصلي ,فتحويل القوة يحسن جوهريا من دقة التكهن عنـدما تكون هناك حاجة للتحويل وان الضعف أو النقص في تحسين التكهن يمكن أن يكون سببه الـشواذ أو اختيار خطأ لقيمة λ (التي تمثل قيمة القوة في التحويل) أو بسبب اختيار النمـوذج الخطـأ أو كليهما 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 الاتي (Pankratz،1991) (2002، Liu and Hudak) (Liu،2006) :

$$z_t = \begin{cases} \dfrac{\left(y_t^{\lambda} - 1\right)}{\lambda} & ; if \quad \lambda \neq 0 \\ \ln\left(y_t\right) & ; if \quad \lambda = 0 \end{cases} \qquad (1)$$

حيث أن :

z_t: تمثل السلسلة المحولة .

y_t: السلسلة الأصلية .

λ: قيمة القوة في التحويل $|\lambda| \leq 1$. وان $t = 1,2,\cdots,n$.

إن الصيغة أعلاه لتحويل القوة مستمرة في المعلمة λ ، وعلى الرغم من أن التحويل في (1) له خصائص مرغوب فيها مثل استمرارية λ عند الصفر ، فان تباينات السلسلة المحولة غير قابلة للمقارنة عند اختلاف قيم λ ،ولمعالجة هذه المسالة اقترح كل من Box and Jenkins (1976) Ansley, Spivey and Wrobleski(1977) تحويل قوة آخر ويعطى بالشكل الاتي :

$$z_t = \begin{cases} \dfrac{\left(y_t^{\lambda} - 1\right)}{\left(\lambda \ \dot{y}^{\lambda - 1}\right)} & ; if \quad \lambda \neq 0 \\ \dot{y} \ln\left(y_t\right) & ; if \quad \lambda = 0 \end{cases} \qquad (2)$$

حيث أن :

$\dot{y} = \left(y_1 \ y_2 \cdots y_n\right)^{\frac{1}{n}}$: الوسط الهندسي للسلسلة .

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

أما النوع الثاني من تحويلات القوة فيعبر عنه بالشكل الاتي :

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

إن الصيغة أعلاه لتحويل القوة لاتستمر عند $\lambda = 0$، وقد تمتلك بعض الـصفات الأخـرى غيـر المرغوب فيها ، 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 λ 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 λ التي ستحسن إما كفاءة النموذج Improve The Efficacy Of The Model أو دقة التكهنات Accuracy Of The Forecasts ، والثانية : تصحيح التحيزات التي تحدث خلال إعادة تحويل التكهنات (إلى القياس الأصلي قبل التحويل) سنفترض أولا بأن قيمة λ مختارة (مثلا $\lambda = 0$ للتحويل اللوغارتمي) ومن ثم يتطلب الأمر اختيار تحويل القوة المناسب (Liu،2006).

3. <u>طرائق إعادة تحويل التكهنات:</u>

أ- طريقة إعادة التحويل المباشرة:

عندما تطبق تحويلات القوة على سلسلة زمنية، فان التكهنات تكون مستندة الى سلسلة محولة ولذا يجب إعادة تحويلها إلى القياس الأصـلي (Liu،2006) . ولكـن هـذه التكهنـات لاتحـتفظ بخصائصها المثالية عندما تعاد إلى القياس الأصلي (Gurrero،1993)،ذلك لان إعـادة التحويـل المباشر للتكهنات يكون متحيزا وهذا يحدث لان المتوسط المقدر للتوزيع المتماثـل فـي البيانـات المحولة يصبح الوسيط المقدر بعد إعادة التحويل للمقدر(Liu،2006) .

إن إعادة التحويل هذه للوسيط تنتج مايسمى "بالتكهن البسيط " "Naive Forecasting" الـذي قـد يعتبر مثاليا إذا كان اقتران دالة الكلفة Cost Function مع أخطـاء الـتكهن Forecast Errors خطيا ومتماثلا حول الصفر (Granger and Newbold،1986).وعلى الرغم من هذا فان تصحيح

التحيز مطلوب إذا أردنا حقا تقدير القيمة المتوقعة في (القياس) المتري الأصلي .مما يجب ذكره انه في بعض الأحيان يكون تأثير التحيز عديم الفائدة عندما تصبح البيانات غير حساسة نسبيا للتحويل.

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

$$\hat{y}_n^{(l)} = T^{-1}\left(\hat{z}_n^{(l)}\right) = \begin{cases} \left(\lambda\, z_n^{(l)} + 1\right)^{\frac{1}{\lambda}} & ;\ if \quad \lambda \neq 0 \\ e^{\hat{z}_n^{(l)}} & ;\ if \quad \lambda = 0 \end{cases} \qquad (4)$$

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

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

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

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

ب- طريقة إعادة التحويل غير المتحيزة :

إن تحويلات القوة Power Transformations يعبر عنها عموما بالشكل الاتي:

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

حيث أن:

Z_t : تمثل السلسلة المحولة .

y_t : السلسلة الأصلية .

$T(\cdot)$: عامل التحويل.

أي بالإمكان الحصول على:

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

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 $\bar{y}_n^{(l)}$ الاتي انظر (Liu،2006):

$$\bar{y}_n^{(l)} = T^{-1}\left(\hat{z}_n^{(l)}\right).C_\lambda^{(l)} \tag{9}$$

حيث أن:

$$C_\lambda^{(l)} = \begin{cases} \left\{0.5 + 0.5\left[\dfrac{1 + 2\left(\lambda^{-1} - 1\right)\sigma_j^2}{\mu_j^2}\right]^{\frac{1}{2}}\right\}^{\frac{1}{\lambda}} & ;\ if\quad \lambda \neq 0 \\ \exp\left(\dfrac{\sigma_j^2}{2}\right) & ;\ if\quad \lambda = 0 \end{cases} \tag{10}$$

لتكن $\mu_j = E_t\left[T\left(z_{t+j}\right)\right]$ تمثل التكهن 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 التكهن Forecasting error ويعرف بأنه $\sigma_j^2 = Var_t(e_{t+j})$ وإذا كان $E_t(Z_{t+j});\ j = 1,2,\cdots,$ تمثل التكهنات التي نبحث عنها ,فإن:

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

حيث أن :

T' : تشير إلى مشتقة T .

فالتعبير المكافئ للمعادلة (11) :

$$\sigma_j^2 = \left\{T'\left[E_t\left(z_{t+j}\right)\right]\right\}^2 Var_t\left[z_{t+j} - E\left(z_{t+j}\right)\right] \tag{12}$$

وباستخدام متسلسلة تايلور نحصل على المعادلة الاتية:

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

وبأخذ التوقع لطرفي المعادلة نحصل على:

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

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

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

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

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

من هذا نحصل على :

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

وبأخذ دالة exp للطرفين نحصل على:

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

أي إن النتيجة تكون بالضبط حيث $\log(z_{t+j})$ يتبع التوزيع الطبيعـي. بالطريقـة نفـسها عنـدما $T(z_t)=z_t^\lambda$ مع $\lambda \neq 0$ نحصل عليها من المعادلة (15):

$$\mu_j = E_t^\lambda(z_{t+j})+\frac{E_t^{\lambda''}[E_t(z_{t+j})]\,\{E_t^{\lambda'}[E_t(z_{t+j})]\}^{-2}\sigma_j^2}{2} \quad (19)$$

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

$$E_t\left(z_{t+j}\right)=\mu_j^{\frac{1}{\lambda}}\left\{0.5\pm 0.5\left[1+\frac{2\left(\lambda^{-1}-1\right)\sigma_j^2}{\mu_j^2}\right]^{\frac{1}{2}}\right\}^{\frac{1}{\lambda}} \qquad (20)$$

وبما أن :

$$E_t\left(z_{t+j}\right)=T^{-1}\left(\mu_j\right).\,C_\lambda^{(l)} \qquad (21)$$

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

$$C_\lambda^{(j)}=\begin{cases}\left\{0.5+0.5\left[\dfrac{1+2\left(\lambda^{-1}-1\right)\sigma_j^2}{\mu_j^2}\right]^{\frac{1}{2}}\right\}^{\frac{1}{\lambda}} & ;\ if\quad \lambda\neq 0\\[2ex] \exp\left(\dfrac{\sigma_j^2}{2}\right) & ;\ if\quad \lambda=0\end{cases} \qquad (22)$$

بالطريقة نفسها يصبح عامل حذف التحيز التقريبي لتحويل Box-Cox (1) كما ياتي:

$$C_\lambda^{(j)}=\begin{cases}\left\{0.5+0.5\left[1+\dfrac{2\left(\lambda^{-1}-1\right)\sigma_j^2}{\left(\lambda^{-1}+\mu_j\right)^2}\right]^{\frac{1}{2}}\right\}^{\frac{1}{\lambda}} & ;\ if\quad \lambda\neq 0\\[2ex] \exp\left(\dfrac{\sigma_j^2}{2}\right) & ;\ if\quad \lambda=0\end{cases} \qquad (23)$$

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

$$C_{\lambda}^{(j)} = \begin{cases} \left\{ 0.5 + 0.5\left[1 + \dfrac{2\left(\lambda^{-1} - 1\right)\left(\sigma_j\, \dot{y}^{\lambda-1}\right)^2}{\left(\lambda^{-1} + \mu_j\, \dot{y}^{\lambda-1}\right)^2} \right]^{\frac{1}{2}} \right\}^{\frac{1}{\lambda}} & ;\ if\ \ \lambda \neq 0 \\ \exp \dfrac{\left(\dfrac{\sigma_j}{\bar{y}}\right)^2}{2} & ;\ if\ \ \lambda = 0 \end{cases} \qquad (24)$$

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

$$T^{-1}\left\{ \hat{\mu}_j \pm z_{\alpha/2}\left[V\hat{a}r_t\left(e_{t+j}\right)\right]^{\frac{1}{2}} \right\} \qquad (25)$$

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

4. الشواذ في السلاسل الزمنية :

إن توافر بيانات خالية من أي شواذ قد يكون غير متوقع عبر الزمن ،وان هذه الملوثات تؤثر في عملية التحليل الإحصائي الدقيق مالم تدرس وتحدد ،فالملوثات هي تلك القيم التي تكون خارجة عن النسق العام للبيانات مما يجعلها لاتتوافق 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 غير العادية تدعى بالشواذ(Liu،2006).

إن تحليل السلسلة الزمنية ودقة المعلومات فيها يعتمد بشكل عال على دقة البيانات وعدم احتوائها على الشواذ ،لذلك من المهم تحديد إن كان للسلسلة شواذ أم لا وماهي طبيعتها قبل القيام بتحليل السلسلة ،ومعرفة طرائق معالجتها أو إزالة تأثيرها بعد معرفة موقعها (محمد، 2008)،فقد يكون للشواذ تأثير معتدل إلى ضخم في 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 حالـة تـوافر جميـع شروط نماذج السلاسل الزمنية.كما أن للشواذ تأثيرا فـي أداء الـتكهن للعينـة المتقدمـة Post Sample فضلا عن تأثيرها على طرائق إعادة تحويل التكهنات فبسبب وجود الشواذ قد تعطـي طريقة إعادة التحويل المباشرة Straight Retransformed تكهنات أفضل من طريقـة إعـادة التحويل غير المتحيزة Unbiased Retransformed.

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

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

أنواع الشواذ في السلاسل الزمنية:

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

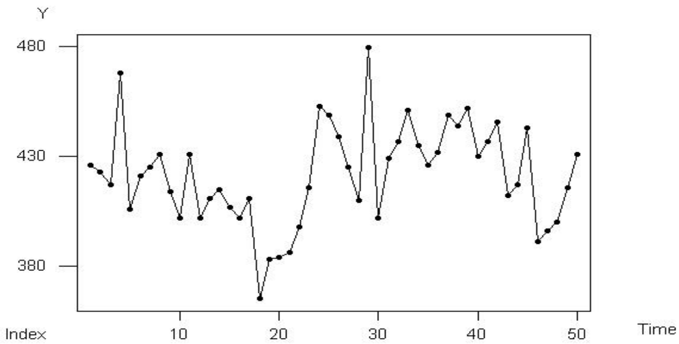
1. الشواذ المضافة Additive Outlier (AO)
2. الشواذ النمطية Innovational Outlier (IO)
3. تغيير المستوى Level Shift (LS)
4. تغيير وقتي (زائل) Temporary change (TC)

5. <u>الجانب التطبيقي:</u>

سيتم في هذا المبحث تحليل السلسلة الزمنية للمعدل اليومي لاسـتهلاك الطاقـة الكهربائيـة فـي محافظة نينوى للفترة من 2008/6/10 ولغاية 2008/8/7 (انظر الملحق 1) حيث سيتم اسـتخدام أول خمسين مشاهدة في تقدير النموذج المناسب في حين تبقى المشاهدات التسعة الأخيرة كعينـة متقدمة Post Sample لفحص دقة التكهنات التي سيتم الحصول عليها مـن النمـوذج المختـار وستتم الاستفادة من البرنامج الإحصائي**SCA** (انظر الملحق 2).

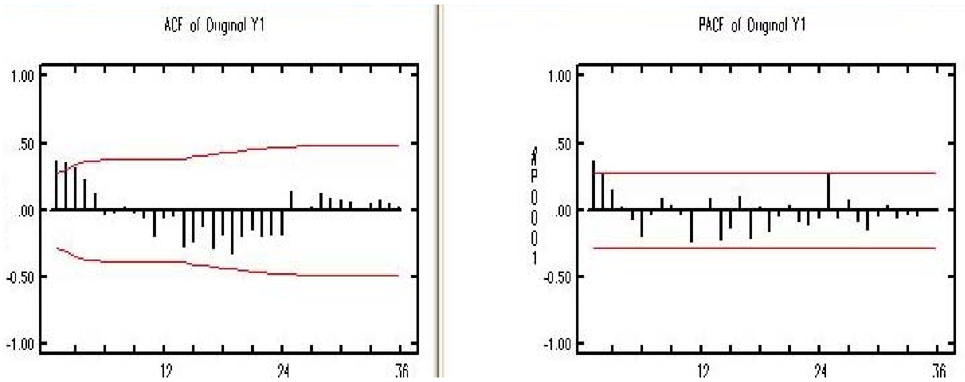
أ- وصف السلسلة الزمنية:

إن الخطوة الأولى في تحليل السلسلة الزمنية هي رسم السلسلة الزمنية Time Series Plot لمعرفة ما إذا كانت السلسلة الزمنية مستقرة أم لا،و الشكل (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 استقرارية السلسلة من خلال رسم دالتي الارتباط الذاتي والارتباط الذاتي الجزئي للسلسلة وكما مبين في الشكل (2) أدناه:

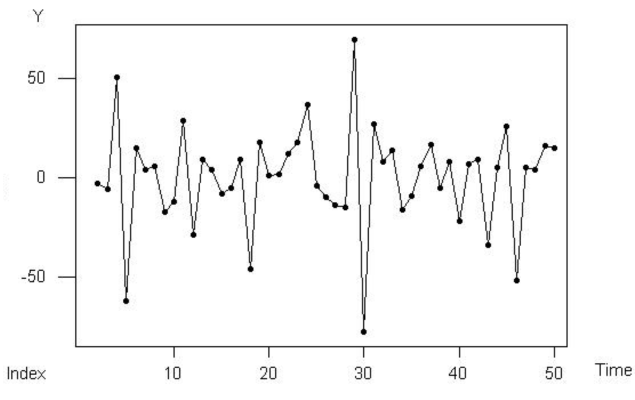


الشكل (2) : دالتا الارتباط الذاتي(ACF) والارتباط الذاتي الجزئي(PACF) للسلسلة الزمنية الخاصة بالمعدل اليومي لاستهلاك الطاقة الكهربائية في محافظة نينوى.

ويمكن التأكد من عدم استقرارية 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.

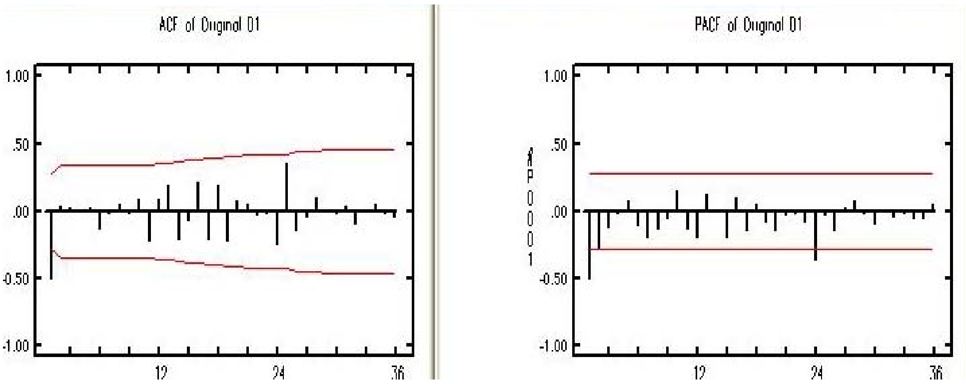
```
UROOT Y1. METHOD DF. ORDERS 0. SPAN 1,50.
SIMPLE DICKEY-FULLER TEST
SERIES TO BE TESTED FOR UNIT ROOT(S) IS:        Y1
TEST BASED ON SAMPLE SIZE:                      50
PROBABILITY OF A SMALLER VALUE IS SET TO:    0.050

                    (<---- INTERPOLATED D-F TABLE ---->)
                  CRITICAL VALUE BASED ON OLS t-STATISTIC
               TEST        0.010     0.050     0.100     0.050
       ORDER   STATISTIC     LEVEL     LEVEL     LEVEL     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.
               TEST        0.010     0.050     0.100     0.050
       ORDER   STATISTIC     LEVEL     LEVEL     LEVEL     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 احصاءة الاختبار اكبر من القيمة الحرجة عند جميع المستويات في كلا النتيجتين عليه تقبل فرضية العدم القائلة(للسلسلة الزمنية جذر وحدة) أي أن السلسلة غير مستقرة،ولتحقيق استقرارية السلسلة الزمنية تم اخذ الفرق الأول لبيانات السلسلة الزمنية،والشكل (3) أدناه يوضح رسم السلسلة الزمنية بعـد اخـذ الفـرق 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): دالتا 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 احصاءة الاختبار اقل من القيمة الحرجة عند المستويات كافة 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
              TEST        0.010     0.050     0.100     0.050
      ORDER   STATISTIC    LEVEL     LEVEL     LEVEL     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.
              TEST        0.010     0.050     0.100     0.050
      ORDER   STATISTIC    LEVEL     LEVEL     LEVEL     LEVEL    UROOT
       0      -71.97      -12.86    -7.68    -5.49      -7.68     NO
--
```

ب- تحليل السلسلة الزمنية ضمن مدى من قيم λ :

في هذا الجزء سيتم البحث عن قيمة λ ضمن مدى معين من (0.95 إلى -0.95) اعتمـادا علـى تصغير الجذر ألتربيعي لمتوسـط مربعـات الخطـأ Root Mean Square Error (RMSE) وعندما تكون بيانات العينة المتقدمة متوافرة فانه يمكن الحصول على RMSE لكل λ وذلك من

خلال التكهن بـ9 خطوات للأمام ، إن مقياس RMSE المستخدم كإشارة للتحقق من كفاءة تحويل قوة معين هو.

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

حيث أن:

Y_t : المشاهدة الحقيقية عند الزمن t،

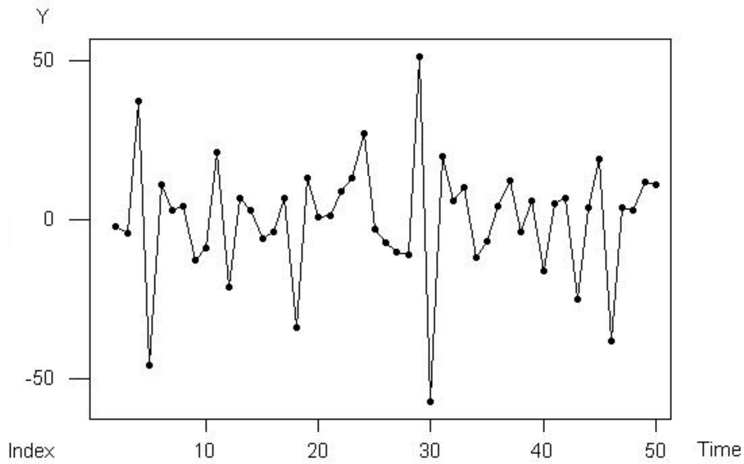
$\hat{Y}_t$: القيمة المتوقعة ل Y_tالمعتمدة على نموذج معين أو طريقة معينة .

m :عدد المشاهدات لفترة العينة المتقدمة.

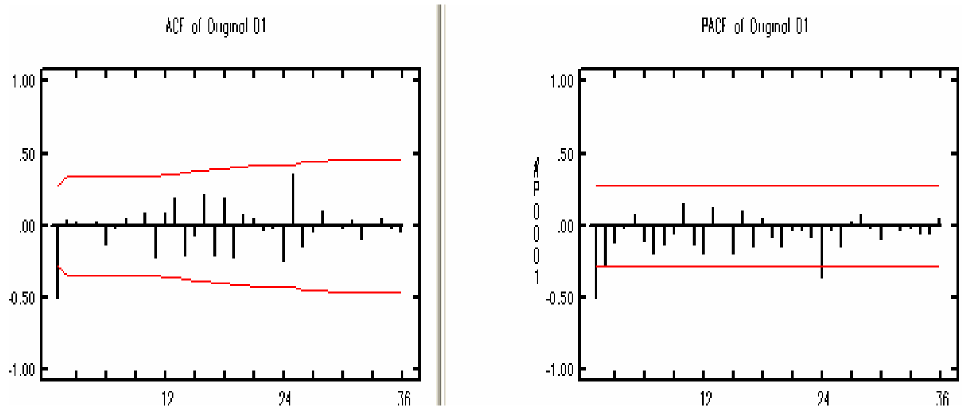
الجدول (1) : يبين قيم RMSE ضمن مدى محدد من قيم λ .

λ	RMSE	λ	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 فان اصغر قيمة للجذر ألتربيعي لمتوسط مربعات الخطأ هـي عند $\lambda = 0.95$ ، عليه سيتم تحويل السلسلة الزمنية اعتمادا على هـذه القيمـة،حيث تـم تحويـل المشاهدات الخمسين الأولى باستخدام تحويل BOX-COX عند $\lambda = 0.95$ بعد ذلك نقـوم بتحليـل السلسلة الزمنية المحولة من خلال منهجية بوكس-جنكنز،حيث سيتم أولا رسم السلسلة الزمنية بعد اخذ تحويل BOX-COX عند $\lambda = 0.95$ بعد أن تم اخذ الفرق الأول للسلـسلة الزمنيـة وكمـا موضح بالشكل (5) أدناه :



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



الشكل (6): دالتا 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 الملائـم لبيانـات السلسلة هو نموذج المتوسطات المتحركة من الرتبة الأولى MA(1)، و يمكن اسـتخدام معيـار معلومات أكاكي المصحح للتأكد من اختيار النموذج الملائم للسلسلة،وتحسب قيمـة (AIC_c) كمـا ياتي (Chatfield and Faraway،1998) :

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

حيث أن :

n: حجم العينة .

S^2:مجموع مربعات البواقي Sum Of Square Residual، إذ أن البواقي تمثل خطـأ الـتكهن لخطوة واحدة للأمام داخل العينة (Within Sample) .

P: عدد معلمات النموذج المقدرة .

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

الجدول (2): ملاءمة 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) ممـا يعزز النتيجة السابقة بان النموذج الملائم للسلسلة الزمنية هـو MA(1).وبعـد تحديـد نمـوذج 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    ORIGINAL     DIFFERENCING
          VARIABLE   OR CENTERED
                                         1
   Y1      RANDOM     ORIGINAL     (1-B  )
 -----------------------------------------------------------------------

PARAMETER   VARIABLE  NUM./  FACTOR  ORDER   CONS-     VALUE      STD     T
LABEL       NAME    DENOM.                 TRAINT               ERROR  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 التكهن يجب توضيح أن السلسلة الزمنية قيد الدراسة تحوي قيما 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 التكهن بالسلسلة الزمنية حيث تم التكهن ب9 خطوات للأمام ثم القيام بإعادة تحويل التكهنات بطريقة إعادة التحويل غير المتحيز Unbiased Retransformed وبطريقة إعادة التحويل باتجاه القياس الأصلي المباشرة Straight Retransformed والنتائج مبينة في الجدول (3) أدناه مع مقارنتها بقيم العينة المتقدمة Post Sample :

الجدول (3) : يبين قيم التكهن من خلال نموذج MA(1) والخطأ المعياري وإعادة تحويل التكهنات بطريقة إعادة التحويل غير المتحيزة وطريقة إعادة التحويل باتجاه القياس الأصلي المباشرة فضلا عن 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 تكهنات إعادة التحويل غير المتحيز أعطت اصغر جذر تربيعي لمتوسط مربعات الخطأ $RMSE = 19.88873$ مقارنة مع تكهنات إعادة التحويل المباشرة التي أعطت جذرا تربيعيا لمتوسط مربعات الخطأ $RMSE = 19.97687$، عليه فان طريقة إعادة التحويل غير المتحيز أكثر كفاءة من طريقة إعادة التحويل المباشرة.

6. الاستنتاجات :

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

المصادر:

1. أمين بك,عزة حازم زكي(2004),"استخدام الشبكات العصبية في التكهن للسلاسل الزمنية بتطبيق على استهلاك الطاقة الكهربائية في محافظة نينـوى", رسـالة ماجـستير غيـر منشورة,كلية علوم الحاسبات والرياضيات,جامعة الموصل.

2. محمد ،سالم بدر،(2008) ،"أساليب الكشف عن الـشوارد فـي الـسلاسل الزمنيـة ذات الدرجات الدنيا ((دراسة تجريبية))"، مجلة التقني ، المجلد 21، العدد 6،العراق.

3. الصفاوي، صفاء يونس طليع،(2005)." مقارنة بين المقـدرات الاعتياديـة والحـصينة لنماذج السلاسل الزمنية المختلطة الثنائية من الرتب الـدنيا"، أطروحـة دكتـوراه غيـر 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^{nd} 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. SanDiego: Academic Press, p. 4.
9. Ansley ,C.F., Spivey, W.A. and Wrobleski, 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..

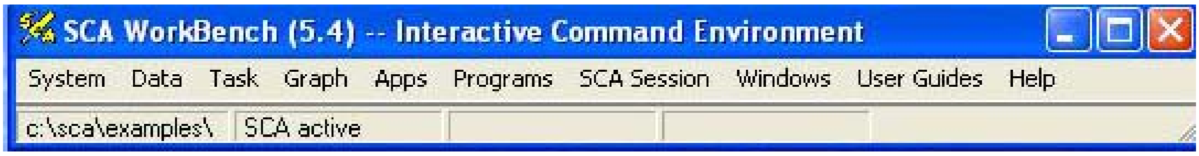
الملحق (1)

المشاهدات الحقيقية لاستهلاك الطاقة الكهربائية في مدينة الموصل للفترة2008/6/10 ولغاية 2008/7/29:

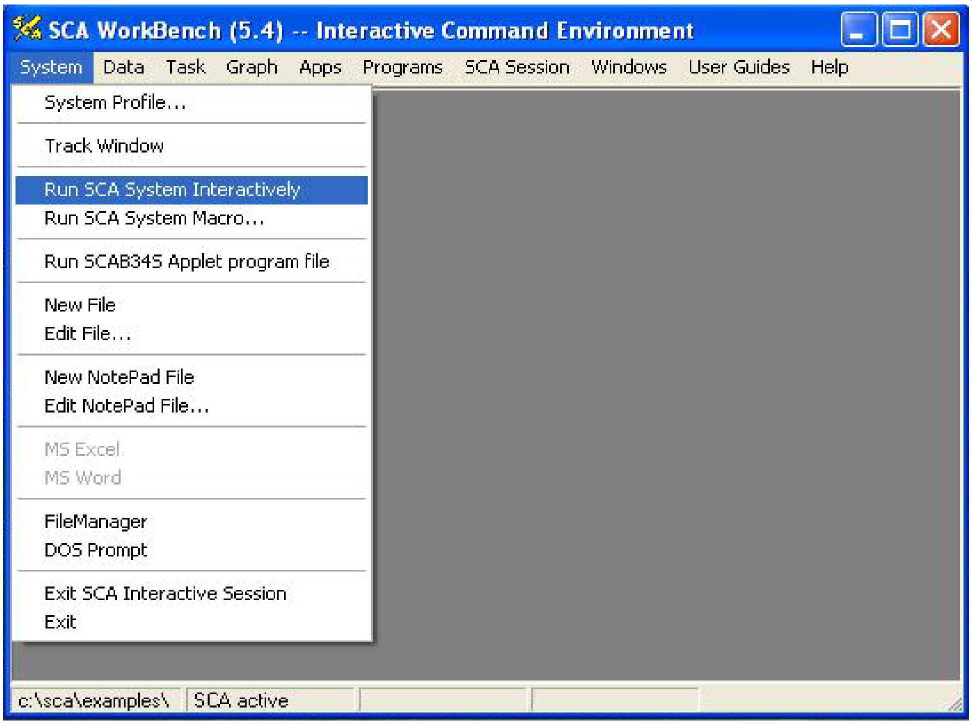
التسلسل	التاريخ	المشاهدة	التسلسل	التاريخ	المشاهدة
1	2008/6/10	426	26	2008/7/5	439
2	2008/6/11	423	27	2008/7/6	425
3	2008/6/12	417	28	2008/7/7	410
4	2008/6/13	468	29	2008/7/8	480
5	2008/6/14	406	30	2008/7/9	402
6	2008/6/15	421	31	2008/7/10	429
7	2008/6/16	425	32	2008/7/11	437
8	2008/6/17	431	33	2008/7/12	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

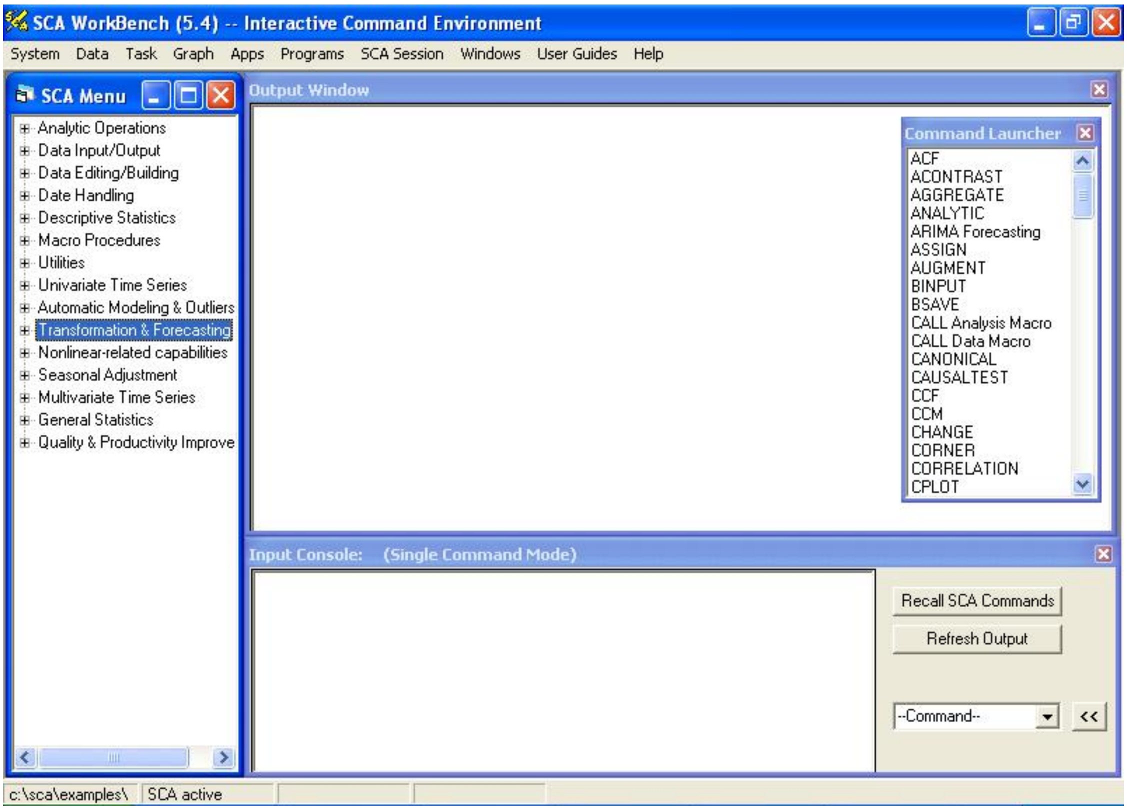
إن نظام SCA Statistical System نظام إحصائي يهتم بمجال تحليل السلاسل الزمنية خاصة فضلا عن اهتمامه بالمجالات الإحصائية الأخرى .وبتشغيل الآيكونة الخاصة بالنظام ستظهر واجهة النظام بالشكل الاتي:



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



لتصبح واجهة النظام بالشكل الاتي:



ويمكن تنفيذ الأوامر الخاصة بالنظام بعدة أساليب منها:

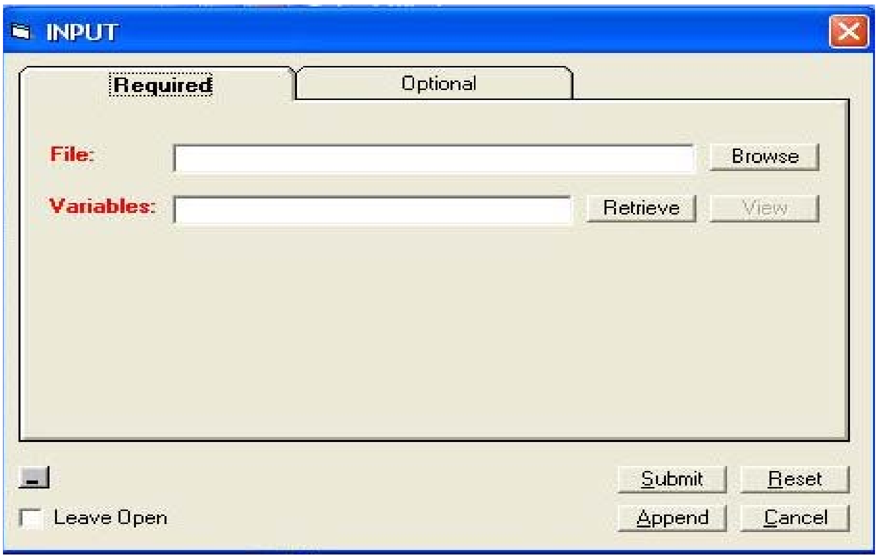
1.عن طريق اختيار الإيعاز من قائمة Command Luncher والتي تحوي ايعازات النظام كافة.
2.عبر اختيار الإيعاز من قائمة SCA Menuالتي تحوي كافة الايعازات.
3.عن طريق طباعة الإيعاز بشكل صحيح بمدخلاته كافة في مربع النص الخاص بايعازات الإدخال والمسمىInput Console.

وبعد كتابة الإيعاز بشكل صحيح يظهر تنفيذه في شاشة الاخراجات Output Windows.

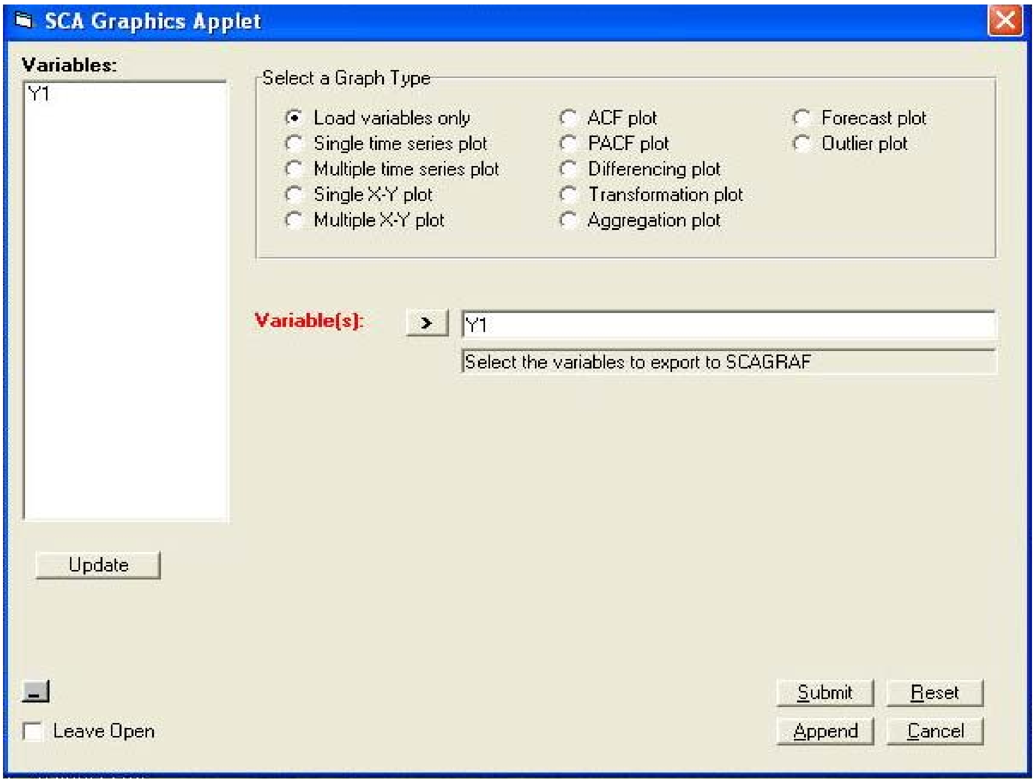
كما يحوي النظام على قوائم الأوامر التي من الممكن اختيارها من شريط القوائم في الجهة العليا من الواجهة.

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

أو عن طريق الإدخال المباشر في شاشةInput Console . كما يمكن إدخال البيانات إلى النظام عن طريق اختيار الإيعازInput حيث يتم استيراد البيانات من ملفات مخزونة في الحاسوب كما في الشكل الاتي:



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



كما يحوي النظام على ملحقات خاصة تفيد في عمليات تحليل البيانات ومعالجتها فضلا عن شرح واف يستفاد منه في التطبيق العملي.

وللنظـام عـدة إصـدارات منهـا الإصـدار التعليمـي Educational Edition والإصـدار الاحترافـي بنوعيـه Professional Edition (A) and Professional Edition (B) والإصـدار المتقدم Advanced Edition وإصدارات أخرى وكل إصدار له مواصفاته الخاصـة وسعره الخـاص ومـدة انتهائه المحددة.

ولشراء نسخة من النظـام تتم مراسلة شـركة Scientific Computing Associates Corp. عن طريق موقع الشركة عبر الانترنت www.scausa.com.

أو عبر البريد الالكتروني sca@scausa.com

تلفون4567 -834 (630). فاكس 1450- 834 (630).