Using models Box-Jenkins ARIMA forecasting produce electric power

استخدام نماذج بوكس-جينكنز ARIMA في التنبؤ بإنتاج الطاقة الكهربائية

Abstract:
This research includes a study of Box-Jenkins ARIMA models for the forecasting of electric power production for Baghdad city. These models are used to build a model of the time series and then choose the best model to predict values for future electricity production with practical application for electric power production for 6 years time series (i.e. 69 months) using statistical software STATGRAPHICS. The best model was ARIMA (1,0,2) than ARIMA (1,0,1) model and ARIMA (1,0,0) from performance predict methods.

Résumé:
Cette recherche comprend une étude des modèles Box-Jenkins ARIMA pour la prédiction de la production électrique pour la ville de Bagdad. Ces modèles sont utilisés pour construire un modèle de la série temporelle et ensuite choisir le meilleur modèle pour prédire les valeurs de production future de l’électricité avec une application pratique pour la production électrique de 6 ans (soit 69 mois) utilisant le logiciel de statistique STATGRAPHICS. Le meilleur modèle était ARIMA (1,0,2) par rapport à ARIMA (1,0,1) et ARIMA (1,0,0) en fonction des méthodes de prédiction.
\[ X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + \epsilon_t \quad \text{(1)} \]

حيث أن:
- مشاهدات السلسلة \( X \)
- معامل النموذج \( \phi \)
- رتبة الاتجاه الالتحاصي \( p \)

ويمكن كتابة النموذج بالشكل التالي:

\[ \phi_p (\beta) X_t = \epsilon_t \quad \text{(2)} \]

حيث أن
- \( \phi_p (\beta) \) هو شكل متعدد الحدود من الدرجة \( p \) للدالة \( \phi_p (\beta) = (1 - \phi_1 \beta - \phi_2 \beta^2 - \ldots - \phi_p \beta^p) \)
- \( \beta \) هو المتغير

### 2-3 نماذج الأوساط المتحركة

إن نماذج الأوساط المتحركة هي عبارة عن ارتباط مشاهدات السلسلة الزمنية الحالية مع خطأ السلاسل السابقة لمدة سابقة.

\[ X_t = \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \ldots - \theta_q \epsilon_{t-q} \quad \text{(3)} \]

حيث أن
- مشاهدات السلسلة \( X \)
- معامل النموذج \( \theta \)
- رتبة الأوساط المتحركة \( q \)

ويمكن كتابة النموذج بالشكل التالي:

\[ \theta_q (\beta) \epsilon_t = \epsilon_t \quad \text{(4)} \]

وهي دالة متعددة الحدود من الدرجة \( q \) للدالة \( \theta_q (\beta) = (1 - \theta_1 \beta - \theta_2 \beta^2 - \ldots - \theta_q \beta^q) \)

### 3-3 نماذج الانحدار الذاتي والأوساط المتحركة

نماذج الانحدار الذاتي والأوساط المتحركة هي عبارة عن ارتباط قيم السلسلة الزمنية الحالية مع القيم السابقة لسلسلة نفس السلاسل:

\[ X_t = \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + \ldots + \phi_p \epsilon_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \ldots - \theta_q \epsilon_{t-q} \quad \text{(5)} \]

ويرمز لنموذج الانحدار الذاتى والأوساط المتحركة (ARMA(p,q)) حيث أن النموذج يكتب بالشكل التالي:

\[ \phi(\beta) \epsilon_t = \theta(\beta) \epsilon_t \quad \text{(6)} \]

ويعتمد أخذ الفرق أو الفرق (Integrated) في أسلوب النموذج لتصبح النموذج من نوعين من النماذج، والذي يسمى النموذج المختلفة غير مستقر. ويرمز له بالعديد ARIMA(p,d,q).

\[ \phi(\beta) \nabla^d X_t = \phi(\beta)(1 - \beta)^d X_t = \theta(\beta) \epsilon_t \quad \text{(7)} \]

حيث أن
- مشاهدات السلسلة \( X \)
- معامل النموذج \( \phi \)
- رتبة الاتجاه الالتحاصي \( p \)
- رتبة الفرق الداخلي غير المستقر \( d \)
- عامل الفرق \( \nabla \)
واسطة عوضنا عن نماذج $\phi(\beta)W_t = \theta(\beta)a_t$ نماذج ARMA(p,q)
لسلسلة (w_t) $X_t$ (تحول إلى نموذج ARIMA(p,d,q) لسلسلة $X_t$)
أي أن نستطيع تطبيق نماذج ARMA(p,q)
لذلك فإن كافة المفاصل النظرية لنموذج ARIMA(p,d,q)

4- منهجية بوكس-جينكنز

عين سلوك السلسلة الزمنية، ويجب التعرف على الأقل (50) مشاهدة لتحليل وبناء نموذج للسلسلة الزمنية Forecasting Model، ومن ثم تشخيص السلسلة الزمنية.

4-1- استقرارية السلسلة الزمنية:

الاستقرارية معناها عدم وجود نماذج في البيانات، أي أن تكون البيانات متكررة حول وسط ثابت لا يعتمد على الزمن.

\[ \text{VAR}(X_t) = \sigma^2 \]
\[ \text{E}(X_t) = \mu \]
\[ \text{VAR}(X_t) = \sigma^2 \]
\[ \text{E}(X_t) = \mu \]

ويمكن ملاحظة الاستقرارية للسلسلة الزمنية من خلال معامل الارتباط الذاتي فإذا كانت معاملات الأرتباط الذاتي تتحدد بأنها الصفر بعد الأزاحة الثانية فзамен أن السلسلة الزمنية مستقرة، أما عدم الاستقرارية فهي عدم تحقق أي من شروط الأرتباط الذاتي.

وإن معامل الأرتباط يأخذ قيمة كبيرة لمعناؤه الزمن.

وألعب السلسلة الزمنية مستقرة حول وسط حسابي ثابت فيمكن أجزاء تعديل الفروق إذا الفرق من الدرجة (d) يمثل Differences بالمعادلة:

\[ \nabla^d X_t = (1-B)^d X_t, \ldots (8) \]

والمعادلة التالية تحل الفرق للجزء الموسمي للسلسلة الزمنية الموسيوية:

\[ \nabla^D S X_t = (1-B^S)^D X_t, \ldots (9) \]

أما إذا كانت السلسلة الزمنية غير مستقرة بالتبابين يمكن أخذ اللوغاريتم أو اخذ الجذر التربيعية للسلسلة الزمنية أو أي تحويلات ثلاث السلسلة الزمنية.

ويعود الأصل في العددان أن السلسلة الزمنية مستقرة إذا كانت الخصائص الأحتمالية لها لانترنت بالزمن، أو غير مستقرة إذا تأثرت خصائصها الأحتمالية بالزمن.

4-2- تشخيص النموذج

يعتبر مراحل تطوير النموذج من المراحل المهمة والصعبة للوصول للنموذج المألوم، إذ الخطة الأولى لهذه المرحلة هي تسمية السلسلة الزمنية الأصلية لتفريق على كيفية أستخبار السلسلة الزمنية لتصحيح مدى جاهزه إلى تحقيق (8) في استخدام نموذج (p,q)

أي أن النموذج الرئيسي لهذه المرحلة هو التشخيص واستخبار النموذج الأساسي، الذي يمثل السلسلة الزمنية في النماذج العامة. ونكون النموذج يتأثر عن طريق دراسة سلوك دلالي الأرتباط الذاتي والأرتباط ARIMA(p,d,q) أو ARMA(p,q)

الذاتي الجزئي على وفق الجدول التالي:

<table>
<thead>
<tr>
<th>النموذج</th>
<th>دالة الأرتباط الذاتي الجزئي</th>
<th>دالة الأرتباط الذاتي الجزئي</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(p)</td>
<td>بين (p من الأرتباطات)</td>
<td>بين (p من الأرتباطات)</td>
</tr>
<tr>
<td>MA(q)</td>
<td>بين (q من الأرتباطات)</td>
<td>بين (q من الأرتباطات)</td>
</tr>
<tr>
<td>ARMA(p,q)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

وذلك من خلال مقياس AIC، وهو مقياس AIC Bayesian Information Criterion (BIC)يعتبر مقياس عالي في معيار، وهو مقياس AIC بيز، وهو مقياس عالي في معيار، وهو مقياس عالي في معيار.

4-3- تقدير معلمات النموذج

تعد معلمات النموذج هو طرازات منها، تعتمد على معرفة التوزيع الأحتمالي للسلسلة الزمنية، ومن هذه الطرق: طريقة الأكتر من الأكتر، وبطريقة الأكتر من الأكتر، هناك طرازات تعتمد على معرفة التوزيع الأحتمالي للسلسلة الزمنية، وهي طريقة ملهمات التصوير، وتسخدم البرامجات الجاهازة الكفاءة لهذا الغرض.

4-4- اختيار ملامحة النموذج المشهور

ARIMA(p,d,q) في عام 1970 اقترح كل من (Box and Jenkins)
النحو الآتي:

\[ E_t(X_{t+L}) = \hat{X}_t(L) = \phi_0 \hat{X}_{t-L} + \ldots + \phi_{p+d} \hat{X}_{t-(p+d)L} - \theta_0 \hat{a}_{t-L} - \ldots - \theta_q \hat{a}_{t-q} + a_t \]

وبكتابة المعادلة (14) عند الفترة (t+L) وآخذ التوقع الشرطي عند زمن الأصل (t) نحصل على معادلة التنبؤ عند زمن الأصل (t+L).

\[ e_t = X_t - \hat{X}_t \]

\[ \text{MSE} = \frac{1}{n} \sum_{i=1}^{n} e_i^2 \]
Mean Absolute Error متوسط مطلق الخطأ

\[ MAE = \frac{1}{n} \sum_{i=1}^{n} |e_i| \quad \text{(17)} \]

Mean Absolute Percentage Error متوسط مطلق الخطأ المنوية

\[ MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{e_i}{X_i} \right| \times 100 \quad \text{(18)} \]

Mean Error متوسط الخطأ

\[ ME = \frac{1}{n} \sum_{i=1}^{n} e_i \quad \text{(19)} \]

Mean Percentage Error متوسط الخطأ المنوية

\[ MPE = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{e_i}{X_i} \right) \times 100 \quad \text{(20)} \]

ويمكن توضيح مراحل بناء النموذج حسب منهجية بوكس-جينز في التنبؤ كما في المخطط الآتي

---

6- التطبيق العملي

البيانات الأتية تخص السلسلة الزمنية لنتائج الطاقة الكهربائية خلال (69) شهرا لسنوات (2012-2007) ووحدة القياس المستخدمة (كيلو واط) لأحد محطات التوليد (محمية القدس) التابعة لمدينة بغداد.
### مجلة جامعة كربلاء العلمية – المجلد الحادي عشر - العدد الثاني / علوي / 2013

جدول رقم(2)البيانات تمثل أنتاج الطاقة الكهربائية لكل شهر للسنوات (2012-2007)

<table>
<thead>
<tr>
<th>السنة</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>17276800</td>
<td>128386400</td>
<td>196251900</td>
<td>185973500</td>
<td>129261400</td>
<td>175737800</td>
<td>196573100</td>
<td>192156850</td>
<td>197884150</td>
<td>135686200</td>
<td>135683600</td>
<td>199514700</td>
</tr>
<tr>
<td>2008</td>
<td>142412700</td>
<td>131711700</td>
<td>130372800</td>
<td>114898900</td>
<td>138463200</td>
<td>149929700</td>
<td>181456380</td>
<td>195238400</td>
<td>76206200</td>
<td>84351700</td>
<td>76206200</td>
<td>199514700</td>
</tr>
<tr>
<td>2009</td>
<td>176326600</td>
<td>201865040</td>
<td>206327130</td>
<td>184737000</td>
<td>184448560</td>
<td>173999000</td>
<td>183354640</td>
<td>195731300</td>
<td>177874900</td>
<td>193680240</td>
<td>160096240</td>
<td>172513340</td>
</tr>
<tr>
<td>2010</td>
<td>339668424</td>
<td>199449010</td>
<td>217379090</td>
<td>223557500</td>
<td>296196210</td>
<td>277586870</td>
<td>316534300</td>
<td>322943400</td>
<td>309039700</td>
<td>298132921</td>
<td>277196850</td>
<td>323345550</td>
</tr>
<tr>
<td>2011</td>
<td>291274422</td>
<td>275219370</td>
<td>261928220</td>
<td>288383230</td>
<td>333969030</td>
<td>345933109</td>
<td>343116100</td>
<td>341527200</td>
<td>318376900</td>
<td>322465500</td>
<td>309550290</td>
<td>327231100</td>
</tr>
<tr>
<td>2012</td>
<td>339668420</td>
<td>324381970</td>
<td>364062100</td>
<td>336161900</td>
<td>305208200</td>
<td>332962890</td>
<td>327283000</td>
<td>343708000</td>
<td>314490000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
- التنبؤ بنموذج ARIMA

تم افتراض عدة نماذج منها:
ARIMA(1,0,0) - 1
ARIMA(1,0,1) - 2
ARIMA(1,0,2) - 3

لقيام بعمليات التنبؤ لنماذج لابد من رسم السلسلة الزمنية الأصلية (رسم مشاهدات السلسلة الزمنية قبل الدراسة) للتعرف على خصائصها ثم اختيار السلسلة الزمنية لمعرفة مدى حاجتها إلى تحويل لعدم استقراريتها وتم رسم دالة الارتباط الذاتي ودالة الارتباط الذاتي الجزئي لمعرفة استقراريتها باستخدام الحاسب الألكتروني وبالاعتماد على البرنامج الإحصائي Statgraphics.

شكل رقم (1) يمثل السلسلة الزمنية لأنتاج الطاقة الكهربائية في محطة الفسد للفترة من 2007 إلى 2012.

![Time Series Plot for prod kwh](X1E8)

![Estimated Autocorrelations for prod kwh](X1E8)
يتضح أن دالة الارتباط الذاتي تتناقص آسيًا وتقترب من الصفر. وهو دليل على استقرار السلسلة.

ARIMA(1,0,0)

<table>
<thead>
<tr>
<th>Statistic</th>
<th>period</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>2.77226E7</td>
</tr>
<tr>
<td>MAPE</td>
<td>15.0721</td>
</tr>
<tr>
<td>ME</td>
<td>645464.0</td>
</tr>
<tr>
<td>MPE</td>
<td>-4.54233</td>
</tr>
</tbody>
</table>

ARIMA Model Summary

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std.Error</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>0.86283</td>
<td>0.0628256</td>
<td>13.7337</td>
<td>0.000000</td>
</tr>
<tr>
<td>Mean</td>
<td>2.39902E8</td>
<td>3.477E7</td>
<td>6.89969</td>
<td>0.000000</td>
</tr>
<tr>
<td>Constant</td>
<td>3.29075E7</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ARIMA(1,0,0) مع الثابت

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>1.50932E15</td>
</tr>
<tr>
<td>MAE</td>
<td>2.75924E7</td>
</tr>
<tr>
<td>MAPE</td>
<td>14.8987</td>
</tr>
<tr>
<td>ME</td>
<td>4.40734E</td>
</tr>
<tr>
<td>MPE</td>
<td>-1.7950</td>
</tr>
</tbody>
</table>

ARIMA Model Summary

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Stnd.Error</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>1.01224</td>
<td>0.0175735</td>
<td>57.600</td>
<td>0.000000</td>
</tr>
<tr>
<td>MA(1)</td>
<td>0.656561</td>
<td>0.09954</td>
<td>6.59595</td>
<td>0.000000</td>
</tr>
<tr>
<td>Mean</td>
<td>1.52311E8</td>
<td>4.65632E7</td>
<td>3.27106</td>
<td>0.001705</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.86451E6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

من نتائج التقدير نجد أن قيمة المقدرة تساوي (1.01224) وقيمة 0 المقدرة تساوي (0.656561) ونلاحظ أن معامل (1) لا يعبر عن متوسط معنوي (0.05) وبذلك يكون النموذج (ARIMA(1,0,0)) ملائماً للتنبؤ.
과학적 기술 연구 - العدد الثاني عشر - العدد الثاني / عام 2013

ARIMA(1,0,1)

شكّل رقم (5) القيم الحقيقية والممتلكة بالنموذج (1,0,1)

### ARIMA(1,0,1) - 7-3

#### Estimation

<table>
<thead>
<tr>
<th>Statistic</th>
<th>period</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>1.49347E15</td>
</tr>
<tr>
<td>MAE</td>
<td>2.71183E7</td>
</tr>
<tr>
<td>MAPE</td>
<td>14.6052</td>
</tr>
<tr>
<td>ME</td>
<td>3.93334E6</td>
</tr>
<tr>
<td>MPE</td>
<td>-1.9674</td>
</tr>
</tbody>
</table>

#### ARIMA Model Summary

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Stnd.Error</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>1.01264</td>
<td>0.0163027</td>
<td>62.1148</td>
<td>0.000000</td>
</tr>
<tr>
<td>MA(1)</td>
<td>0.561357</td>
<td>0.124665</td>
<td>4.50293</td>
<td>0.000029</td>
</tr>
<tr>
<td>MA(2)</td>
<td>0.126213</td>
<td>0.125575</td>
<td>1.00507</td>
<td>0.318588</td>
</tr>
<tr>
<td>Mean</td>
<td>1.29364E8</td>
<td>5.98567E7</td>
<td>2.16123</td>
<td>0.034367</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.63473E6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

من نتائج التقدير نجد أن قيمة المقدرة تساوي (1.01264) وقيمة مني المقدرة تساوي (0.561357) وقيمة مني المقدرة من 0.126213 P-value لمكون معنوية (0.000029) وأنتي أصغر من قيمة (0.05). ننصح أن نعمل معامل (1) P-value. ومنهما على التوالي تعمل معامل (0.318588) و направленة تكون من القيمة يمكن أن تكون瞄准 (8.858) وننصح أن يكون النموذج ملائمًا.
مجلة جامعة كبراء العلمية – المجلد الحادي عشر- العدد الثاني/ علمي/ 2013

ARIMA(1,0,2)

شك رقم (6) القيم الحقيقية والمتنبأ بها للنموذج (2,0,0) و (1,0,0)

8- المقياسات التي تقيم أداء التنبؤ

جدول رقم (3) مقاييس أداء التنبؤ اللاخير أفضل نموذج للتنبؤ

<table>
<thead>
<tr>
<th>المقياس</th>
<th>ARIMA(1,0,2)</th>
<th>ARIMA(1,0,1)</th>
<th>ARIMA(1,0,0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>1.4935</td>
<td>2.7118</td>
<td>14.6052</td>
</tr>
<tr>
<td>MAE</td>
<td>1.5093</td>
<td>2.7592</td>
<td>14.8987</td>
</tr>
<tr>
<td>MAPE</td>
<td>1.7815</td>
<td>2.7722</td>
<td>15.0721</td>
</tr>
<tr>
<td>MPE</td>
<td>-1.9674</td>
<td>-1.7950</td>
<td>-1.9674</td>
</tr>
</tbody>
</table>

نلاحظ أن النموذج (2,0,0) أفضل نموذج للتنبؤ بالقيم المستقبلية لأنه امتلك أقل متوسط مربع الخطأ ومتوسط مطلقة الخطأ. MAE من التموذجين (1,0,0) وARIMA(1,0,1)

9- استخدام النموذج (2,0,0) في التنبؤ

وعند تطبيق نماذج ARIMA على السلسلة الزمنية لإنتاج الطاقة الكهربائية تبين أن النموذج (2,0,0) كان أفضل نموذج للتنبؤ من النموذجين الآخرين من خلال مقاييس أداء التنبؤ وذلك يمكن استخدام هذا النموذج للتنبؤ بإنتاج الطاقة الكهربائية لـ (12) شهر قادماً من القيم المتنبأ بها لـ (12) شهر قادماً.

<table>
<thead>
<tr>
<th>Period</th>
<th>Forecast</th>
</tr>
</thead>
<tbody>
<tr>
<td>70</td>
<td>3.1449E8</td>
</tr>
<tr>
<td>71</td>
<td>3.1449E8</td>
</tr>
<tr>
<td>72</td>
<td>3.1449E8</td>
</tr>
<tr>
<td>73</td>
<td>3.1449E8</td>
</tr>
<tr>
<td>74</td>
<td>3.1449E8</td>
</tr>
<tr>
<td>75</td>
<td>3.1449E8</td>
</tr>
<tr>
<td>76</td>
<td>3.1449E8</td>
</tr>
<tr>
<td>77</td>
<td>3.1449E8</td>
</tr>
<tr>
<td>78</td>
<td>3.1449E8</td>
</tr>
<tr>
<td>79</td>
<td>3.1449E8</td>
</tr>
<tr>
<td>80</td>
<td>3.1449E8</td>
</tr>
<tr>
<td>81</td>
<td>3.1449E8</td>
</tr>
</tbody>
</table>

من نتائج القيم المتنبأ بها نلاحظ أن قيمة أنتاج الطاقة الكهربائية ثابتة لـ التنبؤ لمدة قريبة لـ عشرة أشهر والبيانات حقيقية لذلك تظهر النتائج متقاربة جداً ومتساوية.
الاستنتاجات:

1- إن نماذج بوكس – جينزكر تستخدم في حالات السلسلة الزمنية المستقرة. وعند السلسلة الزمنية غير مستقرة بعد أن يتم تحويلها إلى سلسلة مستقرة.

2- تتاسب طريقة بوكس – جينزكر للسلسلة الزمنية المعقدة وحالات التنبؤ التي توجد فيها أماكن مختلفة في أن واحد مما يؤدي إلى اكتشاف نموذج ملائماً للبيانات.

3- أن السلسلة الزمنية لإنتاج الطاقة الكهربائية من (2007) إلى (2012) لسلاسل من نوع (ARIMA(1,0,2)) هو أفضل نموذج تم اختياره على ضوء البيانات الحقيقية.

المصادر:


