

استعمال تقدير الشرائح التجميعية Additive Splines Estimation
لتشخيص أنموذج الانحدار الذاتي التجميعي اللاخطي بوجود متغير
خارجي NAARX مع تطبيق عملي

أ.م.د. فراس أحمد محمد / كلية الإدارة والاقتصاد / جامعة بغداد

م.م. علي سلمان حبيب / جامعة سومر

المستخلص :

يُعد موضوع تحليل السلاسل الزمنية اللاخطية من المسائل المعقدة وبشكل خاص أنموذج الانحدار الذاتي اللاخطي بوجود متغير خارجي **Nonlinear Autoregressive With Exogenous Variable (NARX)** ومن ثم فإن مسألة تشخيص الأنموذج وتحديد الرتبة الصحيحة من المسائل المهمة في تحليل السلاسل الزمنية ، وطبقاً لذلك أقترح طريقة (**Splines Estimation**) لتشخيص الأنموذج ومن ثم استعمال ثلاثة معايير لتحديد الرتبة الصحيحة . أن الطريقة المقترحة لتقدير الشرائح التجميعية لتشخيص النموذج ومن ثم تحديد الرتبة والتي تعتمد على خاصية التجميع (**Additive**) للتخلص من مشكلة زيادة الأبعاد **Curse (Dimensionally)** ومن ثم يسهل تطبيق تقدير دالة الانحدار باستخدام طرائق تقدير لامعلمية مختلفة عند تشخيص النموذج وتحديد الرتبة بشكل صحيح . وقد استنتجنا أن الطريقة المقترحة هي من الطرق اللامعلمية الجيدة في تقدير دالة الانحدار الذاتي اللاخطي ، أن المعيار (**BIC**) هو أفضل معيار لتحديد الرتبة الصحيحة للنماذج المفترضة وهو معيار دقيق وأثبت كفاءته .

المصطلحات الرئيسية للبحث / أنموذج الانحدار الذاتي اللاخطي مع متغير خارجي ، التجميع ، تقدير الشرائح ، العقد الداخلية ، التشخيص.



مجلة العلوم

الاقتصادية والإدارية

العدد 96 المجلد 23

الصفحات 320-340

*البحث مستل من أطروحة دكتوراه.



1-1 المقدمة :

أن الهدف الرئيسي من نمذجة السلاسل الزمنية هو تحليل بيانات السلسلة الزمنية للملاحظات السابقة ومحاولة دراستها بشكل دقيق لتطوير أنموذج مناسب ليصف الهيكلية المتأصلة داخل السلسلة الزمنية للظاهرة المعنية فمثلاً تحليل أنموذج الانحدار الذاتي يبحث عن هيكلية الدالة داخل بيانات السلسلة نفسها والارتباط الذاتي بين مشاهداتها بعيداً عن الارتباط الذاتي داخل سلسلة الخطأ العشوائي والمسمى بالأوساط المتحركة أو الجمع بينهما ، فإذا لم تكن لدينا معرفة كافية عن دالة التوزيع لتلك الظاهرة وليس هناك افتراضات محددة عليها والاعتماد فقط على البيانات عند ذلك تسمى طرائق تقدير لامعلمية .
أفرض أن :

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-p}) + \varepsilon_t \dots \dots \dots (1)$$

هو أنموذج انحدار ذاتي لاخطي من الرتبة (p) ويرمز له (NAR(p)) ، وعند إضافة متغيرات خارجية للنموذج (1) يصبح النموذج بالشكل الآتي :

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-p}, u_t, u_{t-1}, \dots, u_{t-q}) + \varepsilon_t \dots \dots \dots (2)$$

حيث أن : $y_{t-1}, y_{t-2}, \dots, y_{t-p}$: متغيرات التأخر الزمني (Lag's) للسلسلة y_t ، وأن $p \geq 1$.

$u_t, u_{t-1}, \dots, u_{t-q}$: متغيرات التأخر الزمني (Lag's) للسلسلة للمتغير الخارجية u_t وأن $q \geq 0$

ε_t : سلسلة من المتغيرات العشوائية المستقلة والمتماثلة التوزيع $\varepsilon_t \sim N(0, \sigma^2)$.

النموذج المذكور انفا يوصف بأنه نموذج انحدار ذاتي لاخطي مع متغير خارجي NARX(p,q,L) ولو فرضنا إن معلمة delay غير موجودة هذا يعني ($L = 0$) أي أن النموذج يكون NARX(p,q) ومن ثم يمكن استخدام الأساليب العددية للحصول على تقديرات جيدة للنموذج مثل استخدام أسلوب الشبكات العصبية الاصطناعية . لكن النموذج (2) يعاني من صعوبة في التطبيق باستعمال الأساليب الجبرية بسبب مشكلة تسمى زيادة الأبعاد " curse of dimensionality " كون الدالة تتضمن سلاسل زمنية بارتدادات مختلفة للمتغير الداخل والمتغير الخارج أي وجود لأن الدالة لاخطية وهي متعددة الإبعاد ومن ثم فإن تحليلها يتطلب تمهيد متعدد المتغيرات ، وإذا كان النموذج يحتوي على العديد من المتغيرات التوضيحية المعنوية فإن الصعوبة تكون أكثر تعقيداً ، ولأجل التغلب على هذه الصعوبة في عملية بناء النموذج تستخدم خاصية تسمى "التجميع Additive" . أن استعمال خاصية التجميع تبسط من تعقيد النموذج وتسهل عملية التعامل معاً لأنها تفترض عدم وجود حدود تفاعل بين المتغيرات أو التعامل مع حدود التفاعل على أنها متغيرات خارجية ، ومن ثم يأخذ النموذج الشكل الآتي^[4] :

$$y_t = f_1(y_{t-1}) + f_2(y_{t-2}) + \dots + f_p(y_{t-p}) + g_1(u_{t-1}) + \dots + g_q(u_{t-q}) + \varepsilon_t \dots (3)$$

$f_i(\cdot)$'s : دوال قيمة حقيقية قابلة للقياس. $g_j(\cdot)$'s : دوال قيمة حقيقية قابلة للقياس للمتغيرات الخارجية.

ε_t : تمثل سلسلة المتغيرات العشوائية مستقلة ومتماثلة التوزيع $\varepsilon_t \sim N(0, \sigma^2)$.

ويسمى هذا النموذج بـ نموذج الانحدار الذاتي التجميعي اللاخطي من الرتبة (p,q) NAARX(p,q)

وفي أبسط الحالات تأخذ كل دالة متغير واحد فقط وهنا يمكن التقدير باستعمال الطرائق اللامعلمية. ولأجل الوصول إلى مقدرات لامعلمية مثلى للدوال المجهولة في النموذج رقم (3) يمكن استعمال أساليب تمهيد المتغير الواحد لأنه يعتمد على الخواص الموضوعية للبيانات

2-1 مشكلة البحث : كيف التغلب على مشكلة زيادة الأبعاد باستعمال خاصية التجميع وتشخيص

الأنموذج باستعمال طريقة لامعلمية وهي تقدير الشرائح التجميعية لتشخيص الأنموذج وتحديد الرتبة الصحيحة.



3-1 **هدف البحث** : تشخيص أنموذج الانحدار الذاتي اللاخطي بوجود متغير خارجي بطريقة لاعلمية من ثم تحديد رتبة الأنموذج باستعمال ثلاثة معايير لتحديد الرتبة الصحيحة للأنموذج ومن ثم يمكن استعمال طرائق تقدير لاعلمية لتقدير دالة الانحدار للأنموذج (NAARX)

2- الجانب النظري :

1-2 **المتوسط الشرطي** : أفرض أن (X_t, Y_t) , $t=0, \pm 1, \dots$

تمثل سلسلة زمنية مستقرة وبدقة حيث أن $X_t = (X_{t1}, X_{t2}, \dots, X_{td})$ معرفة على $(R^d, d \geq 1)$ وأن Y_t معرفة على (R) وهي تمثل مجموعة الأعداد الحقيقية للسلسلة.

وأن X_t : تمثل المتغيرات المتأخرة زمنياً للسلسلة $(Y_{t-i}, i=1, 2, \dots, p)$ والمتغيرات الخارجية للسلسلة $(U_{t-j}, j=1, 2, \dots, q)$ وأن $(d = p + q)$ ، لذلك فإن النموذج اللامعلمي للسلسلة المذكور انفا سيكون بالشكل الآتي :

$$Y_t = \sum_{i=1}^d f_i(X_{ti}) + \varepsilon_t \dots \dots \dots (4)$$

حيث أن Y_t : يمثل متغير الاستجابة (السلسلة الزمنية). و $f_i(X_{ti})$: دوال مجهول قابلة للقياس .

ε_t : سلسلة الخطأ العشوائي . وبأخذ التوقع الشرطي للنموذج أعلاه ينتج

$$E(Y_t / X_t = x) = \sum_{i=1}^d f_i(X_{ti}) + E(\varepsilon_t / X_t) \dots \dots \dots (5)$$

حيث أن $E(\varepsilon_t / X_t) = 0$ الأخطاء العشوائية مستقلة عن المتغيرات التوضيحية في النموذج .

$$\mu(x) = \sum_{i=1}^d f_i(X_{ti}) \dots \dots \dots (6)$$

ومن ثم فإن النموذج اللامعلمي يسمى أنموذج انحدار ذاتي لاعلمي ويصبح ممثل بالشكل الآتي :

$$Y_t = \mu(x) + \varepsilon_t , t = 0, \pm 1, \dots \dots \dots (7)$$

2-2 التقارب 1 : Ergodicity

أن البيانات المتوفرة ما هي الا دالة عينة وفي الجانب النظري تضم توقعات المتغيرات العشوائية ، وفي حالة حجم العينة كبير جداً فإن المعدل الزمني يتساوى مع المعدل الاعتيادي للمتغيرات العشوائية فأذا أقترب المعدل الزمني من المعدل الاعتيادي عند زيادة مجال المعاينة ، فإن العملية التصادفية تسمى بـ "عملية تقارب Ergodicity" نسبة الى الكمية التي يحسب المعدل الزمني لها [3].

أفرض أن الصيغة العامة لأنموذج الانحدار الذاتي اللاخطي (NAR) المعرفة في المعادلة (1) هي

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-p}) + \varepsilon_t$$

¹ لمزيد من المعلومات والتفاصيل حول Ergodicity بالتقارب راجع المصدر [1].



العملية $\{y_t\}$ ولدت من أنموذج بالترتيب الزمني الطبيعي نعرف^[5] :

$$Y_t = (y_t, \dots, y_{t-p+1})^T, \quad \varepsilon_t = (\varepsilon_t, 0, \dots, 0)^T \quad \text{and for}$$

$$y = (y_1, \dots, y_{n-1})^T \in \mathbb{R}^p, \quad f(y) = f(y_1, y_2, \dots, y_{n-1})^T$$

فإن الصيغة الأتية للمعادلة (1) التي $\{Y_t\}$ تكون سلسلة ماركوف وتعرف كالآتي :

$$Y_t = f(Y_{t-1}) + \varepsilon_t \dots \dots \dots (8)$$

تعريف 1^[5] : إذا وجد توزيع مثل F وثابت مثل $\rho \in (0, 1]$ بحيث أن :

$$\rho^{-n} \|F_n(. / y) - F(.)\| \rightarrow 0 \quad \text{for any } y.$$

فإن أنموذج ماركوف (8) يسمى تقاربي (Ergodicity) عندما $\rho = 1$ وتقاربي هندسي (Geometrically Ergodicity) عندما $\rho < 1$. وأن F يسمى توزيع مستقر، وأن $\|\cdot\|$ يدل على (norm) وهو قياس بُعد المتجه وأن $F_n(. / y)$ تدل على التوزيع الشرطي .

3-2 الخلط α (α -Mixing) :

يقال للمتسلسلة أنها تمتلك شرط الخلط القوي α أي معامل (الخلط α) إذا كان :

أفرض أن $X = (X_t, t \in \mathbb{Z})$ تمثل سلسلة متغيرات عشوائية معرفة على الفضاء الاحتمالي (Ω, F, P) . وأفرض أن F_j^l (بحيث أن لكل $-\infty \leq j \leq l \leq \infty$) يمثل حقل جبري (σ -field) لحوادث ولدت من خلال متغيرات عشوائية X_t لكل $j \leq t \leq l$ ($t \in \mathbb{Z}$). فإن لأي حقلين من النوع (σ -field) مثل A, B نعرف "مقياس الاعتمادية" لأي سلسلة عشوائية معطاة مثل X_t كالآتي^[2]:

$$\alpha(A, B) = \sup_{A \in \mathcal{A}, B \in \mathcal{B}} |P(A \cap B) - P(A)P(B)| \dots \dots \dots (9)$$

فإن لأي عدد صحيح موجب مثل (n) نعرف معامل الاعتمادية كالآتي :

$$\alpha(n) = \alpha(X, n) = \sup_{j \in \mathbb{Z}} \alpha(F_{-\infty}^j, F_{j+n}^{\infty}) \dots \dots \dots (10)$$

مما سبق ، فإن سلسلة الإعداد $(\alpha(n), n \in \mathbb{N})$ تكون غير متزايدة ومن ثم فإن السلسلة العشوائية X_t

نقول بأن تكون تحت الخلط القوي " الخلط α " إذا كان $\alpha(n) \rightarrow 0$ as $n \rightarrow \infty$

نظرية (1) : لمتتابعين من الإعداد الموجبة a_n, b_n ، فإن^[6] :

$$b_n \lesssim a_n, \quad a_n \lesssim b_n \quad \text{إذا كانت} \quad a_n \cdot b_n \text{ محددة بانتظام. و} \quad \frac{a_n}{b_n}$$

نفرض الشروط الآتية :

1- الكثافة الحدية للـ U_t تكون محددة بعيداً عن الصفر واللانهاية المنتظمة على C^2 .

2- القيم الذاتية للـ $E(X_t X_t^T | U_t = u)$ تكون محددة بانتظام بعيداً عن الصفر واللانهاية لكل $u \in C$.

² تمثل مجموعة الأعداد المركبة Complex Numbers



$$-3 \quad K_n \cdot n^r \text{ for } 0 < r < 1$$

4-العمليات $\{Y_t, X_t, U_t\}_{-\infty}^{\infty}$ تمتلك استقرارية قوية مشتركة، ومعامل الخلط α

$$\alpha(t) \leq C t^{-\alpha} \text{ for } \alpha > \frac{5}{2} \left(\frac{r}{1-r} \right) \text{ وتحقق الآتي: } \{Y_t, X_t, U_t\}$$

$$-5 \text{ لبعض قيم } m \text{ الكبيرة بشكل كافي لـ } m > 0 \text{ فإن } j=1,2,\dots,d, E\left(\left|X_{t_j}\right|^m\right) < \infty$$

تعريف 2: العملية $\{X_t\}$ نقول بأن تكون تحت الخلط من النوع $(\alpha - \text{Mixing})$ إذا كان:

$$\alpha(n) \rightarrow 0 \text{ for each } n \rightarrow \infty$$

4-2 المركبات الممهدة : Smoothing compounds

أن طريقة تحديد الرتبة تتم من خلال تحديد أي المتغيرات المعنوية التوضيحية (التأخر أو المتغيرات الخارجية) التي تؤثر في النموذج، فلو نظرنا إلى النموذج رقم (3) نرى أنه متكون من مركبات تجميعية وكل مركبة هي عبارة عن دالة لامعلمية لمتغير واحد وهنا يمكن استخدام إحدى طرائق التمهيد في تشخيص النموذج، ولأجل تنفيذ ذلك يجب تحديد مجموعات المتغيرات المرشحة والتي يمكن أن تكون رتبة النموذج مجموعة جزئية منها.

أفرض أن المجموعة S_0 تحتوي على عدد من المتغيرات التوضيحية المرشحة وكالاتي

$$S_0 \subset \{1, 2, \dots, d\} \text{ فإذا كانت دالة الانحدار اللامعلمية بالشكل الآتي:}$$

$$\mu(x) = E(Y_t / X_t = x) \quad , \quad x = (x_{t_1}, x_{t_2}, \dots, x_{t_d}) \dots \dots \dots (11)$$

فإنها تكون دالة تجميعية في S_0 وهذا يعني أن $(x_i, i \in S_0)$.

وعندما تكون S_0 موجودة، فإن لأي مجموعة أخرى مثل S تحقق $S_0 \subset S \subset \{1, 2, \dots, d\}$ وأيضاً

تكون $\mu(x)$ دالة تجميعية في S ، $(x_i, i \in S)$ مع افتراض أن S_0 هي أصغر مجموعة مشتركة مرشحة مع خواص محددة.

تعريف 3: افرض أن S هي مجموعة مؤشرة من متغيرات محددة^[7]:

نقول أن S هي مجموعة "مطابقة صحيحة" إذا كان $S_0 = S$.

نقول أن S هي مجموعة "فوق المطابقة" إذا كان $S_0 \subset S$ but $S_0 \neq S$.

نقول أن S هي مجموعة "تحت المطابقة" إذا كان $S_0 \not\subset S$.

حيث أن فوق المطابقة (over fitting) تعني أن مجموعة المتغيرات المحددة تتضمن متغيرات أخرى فضلاً عن المتغيرات المعنوية، وأن تحت المطابقة (under fitting) تعني أن مجموعة المتغيرات المحددة لا تتضمن كل المتغيرات المعنوية. ولتعريف معيار تحديد المتغير لنماذج التجميع، نحتاج حسابه لكل مجموعة مثل $S \subset \{1, 2, \dots, d\}$ لذلك فإن التقدير لدالة الانحدار اللامعلمي يظهر مجموعة مؤشر للمتغيرات المعنوية وهي S .



5-2 الشرائح : Splines

أن طريقة الشريحة (spline) من الطرائق المهمة جداً في النمذجة اللامعلمية لأنها تعتمد على تقريبات عامة (Global Approximation) التي يمكن توسعتها الى أساليب الانحدار متعدد الحدود ، وأننا نسعى لإيجاد طرائق لامعلمية لتقريب دوال التمهيدي لنمذجة (NAARX) لذلك تعد هذه الطريقة من الأدوات المهمة في مجال بحثنا وفي نفس الوقت تعد هذه الطريقة من الأساليب الحديثة في الانحدار اللامعلمي . أن دالة متعدد الحدود (Polynomial function) تعد دالة غير مرنة بشكل جيد لدوال التقريب لأنها تمتلك جميع المشتقات في جميع المواقع مع اختلاف درجات التمهيدي عند مختلف المواقع^[9] .

الحل الوحيدة لزيادة مرونة التقريبات هو السماح لمشتقات دوال التقارب أن تكون غير مستمرة في بعض المواقع ، وهذه الصفة موجودة في متعدد الحدود المقطعي (Piecewise Polynomial) وهذا النوع من متعدد الحدود يسمى الشريحة (Spline) وفيها تكون المشتقات غير مستمرة في بعض المواقع التي تسمى العقد (Knots) وبهدف إجراء عملية التشخيص للنموذج بطريقة التوقع الشرطي والتي يمكن إيجادها عن طريق تقدير دالة الانحدار التجميعي اللامعلمي واستخدام شرائح متعدد الحدود (Polynomial Spline) كطريقة جيدة ومرنة في أسلوب تمهيدي دالة الانحدار اللامعلمي وهذا بدوره يتيح لنا استخدام طريقة الشرائح الانحدار (Regression Splines) بهدف الحصول على تمهيدي لمقدرات الشرائح .

6-2 تقدير الشرائح التجميعية (طريقة مقترحة) : Additive Splines Estimation

في هذا المبحث نتناول احد أهم الأساليب اللامعلمية لتقدير دالة الانحدار $E(Y_t / X_{it}, i \in s)$ لتحديد المتغيرات المعنوية المؤثرة في النموذج المدروس (NAARX) ولأي مجموعة مرشحة من تلك المتغيرات $\{X_{it}, i \in s\}$ لتقدير دالة انحدار الشريحة التجميعي .

أفرض أن دالة الانحدار متعدد الحدود يمكن كتابتها باستخدام الهيكلية التجميعية لدوال الانحدار اللامعلمية الأتي^[7] :

$$\mu(x) = \mu_0 + \sum_{i=1}^d \mu_i(x_i) \dots\dots\dots (12)$$

حيث أن $\mu(x)$: تمثل دالة الانحدار اللامعلمية . μ_0 : يمثل الحد الثابت .

$\mu_i(x_i)$: تمثل دوال التمهيدي ($i = 1, 2, \dots, d$) يمثل عدد الحدود ونفرض أن كل حد يمثل دالة

$$E(\mu_i(X_{ii})) = 0, \quad \forall i = 1, 2, \dots, d$$

أفرض أن $(\xi_1, \xi_2, \dots, \xi_{N_s})$ تمثل سلسلة من العقد (Knots) بحيث أن

$$-\infty < \xi_1 < \xi_2 < \dots < \xi_{N_s} < \infty$$

البيانات أو نقاط مجموعة البيانات نفسها ، أن الشريحة من الدرجة (P) مع نقاط العقد المذكورة انفا تقابل متعدد حدود من الرتبة (P) عند كل الفترات .

أن الشريحة من الدرجة (P) تمتلك (P-1) من الدوال القابلة للاشتقاق والمستمرة عند مواقع العقد

والمقيدة عند هذه الفترات. لأي دالة شريحة $\mu_i(x_i)$ من الدرجة (P) مع العقد $(\xi_1, \xi_2, \dots, \xi_{N_s})$ يمكن تمثيلها كالآتي^[9] :

$$\mu_i(x_i) \approx \sum_{j=1}^{J_i} \gamma_{ji} g_{ji}(x_i) \dots\dots\dots (13)$$



$$\text{where : } \begin{cases} g_j(x) = (x - \xi_j)_+^k, & j=1,2,\dots,N_s \\ g_{N_s+j}(x) = x^{j-1}, & j=1,2,\dots,P+1 \end{cases}$$

بحيث أن $J_i = P + N_s + 1$ ، حيث أن $P \geq 1$.

P : تمثل رتبة متعدد الحدود (order Polynomial) ، N_s : تمثل عدد العقد الداخلية (interior Knots)

$g_{ji}(x_i)$: تمثل أساس الفضاء لدوال الشريحة التي تعتمد على رتبة متعدد الحدود وسلسلة من العقد

وهذا النوع من الشرائح يسمى (B-Spline) لأنه يحتوي على (Basis function) .

γ_{ji} : تمثل المعاملات المرتبطة المطلوب تقديرها. ($i=1,2,\dots,d$ and $j=1,2,\dots,J_s$)

$$(x - \xi_j)_+ = \begin{cases} x - \xi_j, & x > \xi_j \\ 0, & \text{otherwise} \end{cases} \quad \text{حيث أن الجزء الموجب للـ } (x - \xi_j)_+ \text{ يعرف الأتي :}$$

عندما ($P=1,2,3$) فإنها توضح أساس قوة القطع والتي تقابل شرائح (3-) 1-linear , 2-quadratic cubic على التوالي . وعندما $P=3$ تسمى شريحة تكعيبية وهي دالة متعدد الحدود الأكثر أهمية والتي سوف نركز العمل عليها بالطريقة المقترحة .

7-2 اختيار عدد ومواقع العقد : Number Selection and Location of the Knots

أن دوال الشريحة تسمى شرائح متعددات الحدود والتي هي عبارة عن متعدد حدود قطعي باتصال تمهيدي في نقاط العقد الداخلية ، وعلى افتراض أن العقد هنا هي (fixed) ، لذلك هناك أسلوبين لتحديد مواقع هذه العقد ، الأول استعمال أسلوب المسافات المتساوية (المشاهدات المتساوية) بين كل عقدتين متجاورتين والثاني هو استعمال مقاييس تجزئية (Quantiles) أو الإحصاءات المرتبة مع عدد متساوي من المشاهدات بين كل عقدتين متجاورتين ، أما بالنسبة الى عدد العقد فإنها مسألة مهمة جداً ، إذ تعد بمثابة معلمة التمهيد في أساليب التمهيد الأخرى ، فقد اقترح الباحثين (J.Z.Huang and H.Shen(2004))^[6] استخدام معيار AIC لتحديد العدد الأمثل من العقد

وقد قدموا كل من (L.Xue and L.Yang 2005)^[8] صيغة محددة للعدد الأمثل للعقد والتي يرمز لها بالرمز

\hat{N}_s^{opt} من مجموعة من الأعداد الصحيحة والتي تكون ضمن الفترة المغلقة الآتية :

$$\hat{N}_s^{opt} \in [0.5N_r, \min(5N_r, M_s)]$$

$$\text{where : } N_r = n^{(2p+3)^{-1}} \ \& \ M_s = ((n/4) - 1) / d$$

n : تمثل عدد المشاهدات . p : تمثل رتبة متعدد الحدود . d : عدد المتغيرات التوضيحية .

ولأجل إيجاد المقدرات للمعالم $\{\mu_0, \gamma_{ji}, i=1,2,\dots,d ; j=1,2,\dots,J_i\}$ نستخدم طريقة المربعات الصغرى للصيغة الآتية :

$$\sum_{t=1}^n \left\{ Y_t - \mu_0 - \sum_{i=1}^d \sum_{j=1}^{J_i} \gamma_{ji} g_{ji}(x_i) \right\}^2 \dots\dots\dots(14)$$



نحصل على المقدرات الآتية : $\{\hat{\mu}_0, \hat{\gamma}_{ji}, i=1,2,\dots,d; j=1,2,\dots,J_i\}$ وعند التعويض نحصل على :

$$\hat{\mu}(x) = \hat{\mu}_0 + \sum_{i=1}^d \hat{\mu}_i(x_i) \dots \dots \dots (15)$$

$$\hat{\mu}_i(x_i) = \sum_{j=1}^{J_i} \hat{\gamma}_{ji} g_{ji}(x_i) \dots \dots \dots (16)$$

J_i : تمثل درجة الشريحة (Degree Spline) وهي رتبة متعدد الحدود (Polynomial) مضافة لها عدد العقد الداخلية (Interior Knots)

حيث أن الصيغة (15) تعد تقدير دالة الانحدار اللامعلمي التجميعي لنموذج (NAARX) وتطبق هذه الطريقة لكل مجموعة مرشحة من المتغيرات التوضيحية وبافتراض أن عدد المتغيرات التوضيحية الكلي (عدد متغيرات التأخر الزمني والمتغير الخارجي) محدد مسبقاً ($S_{\max} \leq d \Rightarrow S_{\max} = d$) وفي كل مرة نحسب (MSE) ومعايير (AIC و BIC و FPE) ولكل مجموعة من المتغيرات التوضيحية $\{x_i, i \in S\}$. أن تقدير الشريحة والذي يقابل المجموعة المؤشرة S يعطى وفق الصيغة الآتية :

$$\hat{\mu}_s(x) = \arg \min_{g \in G_s} \sum_{t=1}^n (y_t - \hat{\mu}(x_t))^2 \dots \dots \dots (17)$$

أن كل فضاء دوال الشريحة يعرف على مدى كل الـ X_{ii} مع رتبة متعدد الحدود وعدد العقد الداخلية ومن ثم فإن كل دالة في الفضاء G_s هي دالة لـ $(x_i, i=1,2,\dots,d)$ لذلك تمتلك الفضاءات

$$G_s \subset G_{s_0} \subset \dots \subset G_{\{1,2,\dots,d\}} \quad \text{for } S \subset S_0 \subset \{1,2,\dots,d\}$$

ولكل مجموعة جزئية للـ S يعرف متوسط مربعات الخطأ $(MSE_s)^{[7]}$.

المرحلة الأولى : حيث نبدأ من أنموذج العدم ($Y_t = \mu_0 + \varepsilon_t$) نضيف متغير توضيحي واحد الى النموذج في كل مرة ومن بين مختلف المتغيرات المرشحة التي لم تحدد نختار التي تقلل قيمة متوسط مربعات الخطأ ($MSE_s, s \in \{1,2,\dots,d\}$) ثم نضيف متغير توضيحي آخر وأيضاً نحسب (MSE_s) وهكذا نستمر بإضافة المتغيرات التوضيحية تباعاً ، مع ملاحظة أن المتغير الذي لا يخفض من قيمة (MSE_s) لا يتم إضافته للنموذج وتوقف عملية الإضافة عندما يكون عدد المتغيرات المضافة مساوي الى عدد المتغيرات المحدد مسبقاً ($S_{\max} = d$).

$$MSE_s = \frac{1}{n} \sum_{t=1}^n \{Y_t - \hat{\mu}_s(X_t)\}^2 \dots \dots \dots (18)$$

المرحلة الثانية : في هذه المرحلة نبدأ مع أكبر مجموعة من المتغيرات المختارة في المرحلة الأولى ثم نقوم بحذف متغير واحد في كل مرة عن طريق تقليل قيمة (MSE_s) ونتوقف عندما لا يكون هناك متغير بقية في النموذج .



المرحلة الثالثة : نحصل على مجموعة من النماذج الجيدة والمرشحة ، ثم نقوم بحساب المعايير
(AIC_s , BIC_s ; FPE_s) ولكل المجموعات المرشحة للمفاضلة بينها ثم نختار القيمة التي تقابل أقل
قيم للمعايير الاتية [6] :

$$AIC_s = \log(MSE_s) + 2 \frac{q_s}{n} \dots\dots\dots(19)$$

$$BIC_s = \log(MSE_s) + \frac{q_s}{n} \log(n) \dots\dots\dots(20)$$

$$\text{and } FPE_s = MSE_s \frac{n - q_s}{n + q_s} \dots\dots\dots(21)$$

وفي كل خطوة من المرحلتين الأولى والثانية نطابق أنموذج الشريحة التجميعي .

3- الجانب التجريبي :

1-3 المحاكاة : تعرف المحاكاة على أنها تقليد للواقع الحقيقي في عالم افتراضي من خلال تطبيق عدة تجارب
خاضعة لبعض الافتراضات والشروط التي يضعها الباحث ، والتي يجب ان تكون قريبة جداً من أصل الظاهرة
الحقيقية ثم يقوم الباحث بالعمل على تلك التجارب وتكرارها للعديد من المرات وبحسب حاجة البحث، والهدف
من ذلك متابعة التجربة في ظل ظروف افتراضية والتي سوف يطبقها في العالم الحقيقي . أن أغلب الباحثين
استخدموا المحاكاة في الإحصاء التطبيقي ضمن آلية محددة وهي افتراض أحجام عينات مختلفة وربما تباينات
مختلفة وهنا نذكر أننا سوف نستخدم تجارب المحاكاة لإثبات أفضلية عمل الطرائق المقترحة المقدمة لتحديد
الرتبة الأفضل في حالة الاعتماد على البيانات نفسها.

2-3 تنفيذ تجارب المحاكاة :

تم تنفيذ المحاكاة باستعمال برنامج (MATLAB R2015a) لغرض التحقق من طرق تحديد أفضل
رتب للنموذج بالاعتماد على خاصية التجميع واستخدام تقدير الشريحة التجميعي لتحديد الرتبة الأفضل
للأنموذج باستخدام ثلاثة معايير هما (AIC , BIC , FPE) . سيتم التنفيذ باستخدام ثلاثة أحجام مختلفة
وتكرار قدره (replicate=100) من العينات وهي كالآتي ($n_1 = 30$, $n_3 = 180$, $n_4 = 360$)
وللتخلص من تأثير القيم الأولية يتم توليد ($S + n_i$, $i = 1, 2, 3, 4$) حيث أن ($S = 100$) ولكل حجم
عينة ، ثم بعد ذلك سيتم إهمال أول ($S = 100$) من كل حجم .
يتم توليد متغيرين عشوائيين من التوزيع المنتظم القياسي العشوائي وكالآتي :

$$U_1, U_2 \sim unif(0,1)$$

ثم يتم تحويلهما الى التوزيع الطبيعي القياسي $N(0,1)$ من خلال استخدام طريقة (Box-Muller) ليمثل
الخطأ العشوائي .

يتم توليد المتغيرين المستقلين العشوائيين (X_1) و (X_2) من التوزيع المنتظم القياسي العشوائي وكذلك
متغير الخطأ العشوائي من التوزيع الطبيعي القياسي ويتم استعمالهما في النموذج الآتي :
($Y_t = 2 + 0.5X_1 - 0.3X_2 + e_t$) لتوليد بيانات السلسلة المستقرة $\{Y_t\}_1^n$. وبالطريقة نفسها
المذكورة انفا يتم توليد بيانات السلسلة المستقرة $\{U_t\}_1^n$ التي تمثل المتغير الخارجي .



يتم استخدام أنموذجين مختلفين في إجراء تجارب المحاكاة وهما :
1-النموذج الأول^[4] :

$$y_t = 0.8\log(1 + 3y_{t-1}^2) - 0.6\log(1 + 3y_{t-2}^2) + u_t y_{t-1} + \varepsilon_t$$

2-النموذج الثاني^[3] :

$$y_t = 2\text{Exp}(-0.1y_{t-1}^2) * y_{t-1} - \text{Exp}(-0.1y_{t-2}^2) * y_{t-2} + u_{t-2} + \varepsilon_t$$

4-3 تحديد رتبة الأنموذج بالطريقة المقترحة :

لأجل التحقق من أن الطريقة المقترحة تتأثر باختلاف أحجام العينات واستعمال نماذج مختلفة سوف نعمل التحليل الآتي :

1-4-3 الأنموذج الأول Model1 :

جدول رقم (1-1) يمثل تحديد الرتبة بالطريقة المقترحة (Additive Splines) لحجم العينة (30)

Lag's(y_t, u_t)	AIC	BIC	FPE
y_{t-1}	-0.58382	-0.30358	0.56082
y_{t-1}, y_{t-2}	-1.94587	-1.4321	0.102927
$y_{t-1}, y_{t-2}, y_{t-3}$	-1.7095	-0.9622	0.093415
$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}$	-2.46476	-1.53063	0.03362
$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}$	-3.38079	-2.25984	0.010303
$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}, y_{t-6}$	-3.64226	-2.33447	0.006076
y_{t-1}, y_{t-2}, u_t	-39.4639	-37.9693	1.29E-18
$y_{t-1}, y_{t-2}, u_t, u_{t-1}$	-39.5787	-37.8973	8.81E-19
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}$	-50.9851	-49.1169	7.51E-24
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}, u_{t-3}$	-53.9724	-51.9173	2.9E-25
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}, u_{t-3}, u_{t-4}$	-56.4766	-54.2347	1.81E-26
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}, u_{t-3}, u_{t-4}, u_{t-5}$	-56.6169	-54.1882	1.21E-26

جدول رقم (2-1) يمثل تحديد الرتبة بالطريقة المقترحة الأولى (Additive Splines) لحجم (180):

Lag's(y_t, u_t)	AIC	BIC	FPE
y_{t-1}	-1.23069	-1.10652	0.292101
y_{t-1}, y_{t-2}	-1.97448	-1.74388	0.129884
$y_{t-1}, y_{t-2}, y_{t-3}$	-1.94453	-1.53654	0.11976
$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}$	-1.90928	-1.4126	0.117353
$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}$	-1.87654	-1.29117	0.114705
$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}, y_{t-6}$	-1.88429	-1.21022	0.107669
y_{t-1}, y_{t-2}, u_t	-2.76366	-2.0009	0.042272



استعمال تقدير الشرائح التجميعية Additive Splines Estimation لتشخيص أنموذج الانحدار الذاتي التجميعي الاخطي بوجود متغير خارجي NAARX مع تطبيق عملي

$y_{t-1}, y_{t-2}, u_t, u_{t-1}$	-2.77391	-1.92245	0.03958
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}$	-2.73367	-1.79353	0.038978
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}, u_{t-3}$	-2.70405	-1.67521	0.037981
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}, u_{t-3}, u_{t-4}$	-2.65987	-1.54234	0.037551
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}, u_{t-3}, u_{t-4}, u_{t-5}$	-2.6256	-1.41937	0.03676

جدول رقم (3-1) يمثل تحديد الرتبة بالطريقة المقترحة (Additive Splines) لحجم (360):

Lag's (y_t, u_t)	AIC	BIC	FPE
y_{t-1}	-1.28348	-1.19712	0.277073
y_{t-1}, y_{t-2}	-1.94502	-1.7831	0.137531
$y_{t-1}, y_{t-2}, y_{t-3}$	-1.91905	-1.62759	0.132047
$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}$	-1.9086	-1.55237	0.12906
$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}$	-1.89037	-1.48017	0.127833
$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}, y_{t-6}$	-1.88196	-1.41778	0.125382
y_{t-1}, y_{t-2}, u_t	-2.58828	-2.07014	0.060175
$y_{t-1}, y_{t-2}, u_t, u_{t-1}$	-2.56402	-1.9919	0.059964
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}$	-2.54936	-1.92327	0.059182
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}, u_{t-3}$	-2.53326	-1.8532	0.058495
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}, u_{t-3}, u_{t-4}$	-2.52159	-1.78755	0.057561
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}, u_{t-3}, u_{t-4}, u_{t-5}$	-2.50234	-1.71433	0.057072

من خلال استعراض النتائج المعروضة في الجداول أعلاه للنموذج الأول للطريقة المقترحة الأولى فقد تبين الآتي:

نلاحظ أن قيم المعيار AIC لم يوفق في تحديد الرتبة الصحيحة للنموذج المستخدم في توليد البيانات عند حجم العينة (30) مشاهدة (حجم عينة صغير) وعند زيادة حجم العينة الى (180) مشاهدة نلاحظ أن قيم هذا المعيار كانت عند رتبة محددة وهي (2,1) ومتغيرات التأخر هي ($y_{t-1}, y_{t-2}, u_t, u_{t-1}$) وهي قريبة من الرتبة الصحيحة للنموذج المفترض، ولكن عند حجم العينة (360) مشاهدة فقد نجح المعيار بتحديد الرتبة الصحيحة للنموذج ومتغيرات التأخر كانت (y_{t-1}, y_{t-2}, u_t).

أما بالنسبة الى المعيار (BIC) فقد كان أكثر دقة بتحديد الرتبة، حيث أنه عند حجم العينة (30) مشاهدة لم ينجح بتحديد الرتبة الصحيحة للنموذج (حجم عينة صغير)، وعند زيادة حجم العينة الى (180) و (360) فقد نجح بتحديد الرتبة الصحيحة للنموذج المفترض ومتغيرات التأخر كانت (y_{t-1}, y_{t-2}, u_t).

إما بالنسبة للمعيار (FPE) فقد كان يميل الى أخذ أكبر عدد من متغيرات التأخر (lag's) من المجموعة، ومن ثم لا يمكن الاعتماد على هذا المعيار في تحديد الرتبة الصحيحة للنموذج ونلاحظ ثباته عند أكبر مجموعة من متغيرات التأخر (lag's) ولكل أحجام العينات المختلفة للنموذج المفترض.



استعمال تقدير الشرائح التجميعية Additive Splines Estimation لتشخيص أنموذج
الانحدار الذاتي التجميعي اللاخطي بوجود متغير خارجي NAARX مع تطبيق عملي

ومن خلال التحليل المذكور انفا فإن معيار (BIC) كان الأفضل والأكثر دقة في تحديد الرتبة الصحيحة للنموذج الأول المستخدم في توليد البيانات ويمكن ملاحظة ذلك في الجداول (2-1) و(3-1). وهذا يدل على أن هذا المعيار يكون دقيق بتحديد الرتبة عند أحجام العينات الكبيرة.

2-4-3 الأنموذج الثاني Model2 :

جدول رقم (1-2) يمثل تحديد الرتبة بالطريقة الأولى (Additive Splines) لحجم العينة (30)

Lag's (y_t, u_t)	AIC	BIC	FPE
y_{t-1}	-2.62272	-2.34248	0.073003
y_{t-1}, y_{t-2}	-2.58841	-2.07464	0.054135
$y_{t-1}, y_{t-2}, y_{t-3}$	-2.7923	-2.045	0.031635
$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}$	-3.32134	-2.38721	0.014276
$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}$	-5.18091	-4.05995	0.001703
$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}, y_{t-6}$	-5.87752	-4.56973	0.00065
y_{t-1}, y_{t-2}, u_t	-41.4156	-39.921	1.83E-19
$y_{t-1}, y_{t-2}, u_t, u_{t-1}$	-40.9938	-39.3124	2.14E-19
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}$	-53.6973	-51.829	4.98E-25
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}, u_{t-3}$	-56.2772	-54.2221	2.89E-26
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}, u_{t-3}, u_{t-4}$	-58.3279	-56.086	2.85E-27
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}, u_{t-3}, u_{t-4}, u_{t-5}$	-60.9727	-58.5439	1.55E-28

جدول رقم (2-2) يمثل تحديد الرتبة بالطريقة المقترحة الأولى (Additive Splines) لحجم (180):

Lag's (y_t, u_t)	AIC	BIC	FPE
y_{t-1}	-2.93475	-2.81057	0.053146
y_{t-1}, y_{t-2}	-3.27791	-3.04731	0.035276
$y_{t-1}, y_{t-2}, y_{t-3}$	-3.20202	-2.79403	0.034056
$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}$	-3.17304	-2.67636	0.033163
$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}$	-3.16089	-2.57552	0.031754
$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}, y_{t-6}$	-3.16689	-2.49283	0.029858
y_{t-1}, y_{t-2}, u_t	-3.49119	-2.72843	0.020422
$y_{t-1}, y_{t-2}, u_t, u_{t-1}$	-3.45771	-2.60625	0.019976
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}$	-3.41637	-2.47623	0.019694
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}, u_{t-3}$	-3.36602	-2.33718	0.019592
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}, u_{t-3}, u_{t-4}$	-3.33131	-2.21377	0.019188
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}, u_{t-3}, u_{t-4}, u_{t-5}$	-3.28209	-2.07587	0.019066



جدول رقم (3-2) يمثل تحديد الرتبة بالطريقة المقترحة الأولى (Additive Splines) لحجم (360):

Lag's (y_t, u_t)	AIC	BIC	FPE
y_{t-1}	-2.91895	-2.83259	0.053991
y_{t-1}, y_{t-2}	-3.11949	-2.95757	0.042495
$y_{t-1}, y_{t-2}, y_{t-3}$	-3.11406	-2.82261	0.039971
$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}$	-3.10796	-2.75174	0.038897
$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}$	-3.09788	-2.68768	0.038214
$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}, y_{t-6}$	-3.0941	-2.62992	0.037308
y_{t-1}, y_{t-2}, u_t	-3.50886	-2.99071	0.023967
$y_{t-1}, y_{t-2}, u_t, u_{t-1}$	-3.48458	-2.91246	0.023883
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}$	-3.49938	-2.87328	0.022888
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}, u_{t-3}$	-3.48046	-2.80039	0.022686
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}, u_{t-3}, u_{t-4}$	-3.4828	-2.74876	0.022013
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}, u_{t-3}, u_{t-4}, u_{t-5}$	-3.46673	-2.67871	0.021757

من خلال استعراض جداول نتائج النموذج الثاني للطريقة المقترحة تبين ما يلي :
نلاحظ أن قيم المعيار AIC كانت تميل الى أكبر مجموعة من متغيرات التأخر الزمني Lag's ، عند
حجم العينة (30) مشاهدة (حجم عينة صغير) وعند أحجام العينات (180 و 360) مشاهدة فإنها تتأثر بوجود
المتغير الخارجي لذلك تكون مجموعة متغيرات التأخر الزمني Lag's هي (y_{t-1}, y_{t-2}, u_t) وهي أقرب رتبة
محددة لرتبة النموذج المفترض.

اما بالنسبة الى المعيار BIC كان يميل الى أكبر عدد من مجموعة متغيرات التأخر الزمني Lag's عند
حجم العينة (30) مشاهدة ،، وعند حجم العينة (180) مشاهدة فإن أقل قيمة له تقابل مجموعة متغيرات التأخر
(y_{t-1}, y_{t-2}) أي الرتبة الثانية ، وعند حجم العينة (360) مشاهدة نلاحظ ان أقل قيمة له كانت تقابل نفس
المحددة بالمعيار AIC وهي (y_{t-1}, y_{t-2}, u_t) وهي أقرب رتبة للنموذج المفترض.

أما بالنسبة الى معيار FPE فنلاحظ أنه يميل الى أكبر مجموعة من متغيرات التأخر الزمني Lag's ومن
ثم لا يمكن الاعتماد على هذا المعيار في تحديد الرتبة الصحيحة للنموذج الثالث وهو الذي تمت الإشارة إليه في
النموذج الأول .

من خلال التحليل المذكور انفا للنماذج نلاحظ أن المعايير الثلاثة تكون غير فعالة في تحديد الرتبة
الصحيحة عند حجم العينة الصغيرة (30 مشاهدة) ، ونلاحظ أن المعيار BIC كان الأهم والأفضل المستخدم
في عملية تحديد الرتبة والذي حدد الرتبة بشكل عام وهو الأفضل من بقية المعايير الأخرى بالنسبة للنماذج
حيث كان الأقرب للرتبة الصحيحة للنماذج المفترضة في توليد البيانات . وأن المعيار FPE كان يميل الى أكبر
رتبة في كل النماذج المفترضة لذلك لا يمكن الاعتماد عليه .



4- الجانب التطبيقي :

1-4 تقدير الرتبة الصحيحة لبيانات الأحمال الكهربائية ودرجة الحرارة :

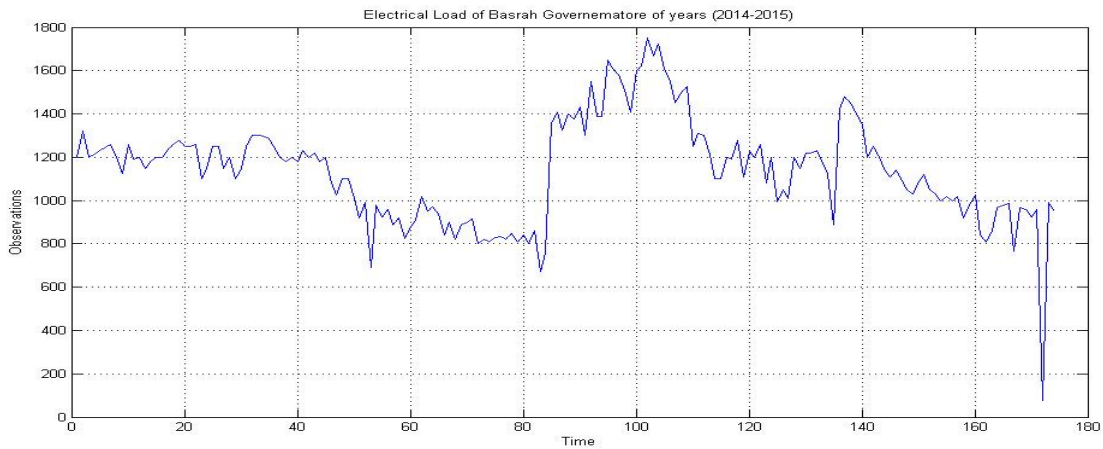
أن مجالات التطبيق للسلاسل الزمنية مختلفة ومتنوعة لكن الشرط الأساسي هو يجب أن تكون الظاهرة المدروسة هي عملية عشوائية مرتبطة بالزمن كأن تكون (حوادث يومية أو أسبوعية أو شهرية ...الخ)، وعند التحليل باستخدام سلسلة الانحدار الذاتي Autoregressive يجب ملاحظة أن الظاهرة سوف تدرس من خلال مشاهداتها داخل نفسها بأخذ الفروق (Difference) لبياناتها ، وعند إضافة متغير خارجي لها كسلسلة زمنية أخرى لكن مرتبة معها في ذات الوقت (الفترة الزمنية ذاتها) هنا ينصب جل اهتمامنا وهو دراسة بيانات ظاهرة محددة متعلقة معها بيانات متغير آخر في الفترات الزمنية ذاتها ، وفي العراق أصبح موضوع الطاقة الكهربائية من المواضيع المعقدة والمهمة في حياتنا اليومية وحاجتنا لها في تزايد مع مرور الزمن ، وفي العقدين الأخيرين أصبحت مشكلة نقص التجهيز للطاقة الكهربائية نتيجة زيادة الأحمال فعلاً مشكلة حقيقية تستدعي الباحثين للبحث فيها وإيجاد الحلول المناسبة لها أو إعطاء أفكار جديدة في كيفية ترشيد استهلاك الطاقة أو تقديم توصيات علمية في هذا الموضوع بالإضافة الى الخطوات العملية التي هي من شأن الجهات المعنية متمثلة بوزارة الكهرباء ومديرياتها المختلفة والتي ليس بصدها نحن الآن ولتناول موضوع الأحمال الكهربائية كجانب عملي لأنموذج الانحدار الذاتي اللاخطي مع متغير خارجي (NAARX) تم عرض حمل الذروة للاستهلاك اليومي للطاقة الكهربائية مع درجات الحرارة (درجات الحرارة مقاسه فقط لثلاثة محافظات هي بغداد والبصرة والموصل) بحسب التقرير الشهري دائرة توزيع الطاقة ووزارة الكهرباء وتبين أن أكثر محافظة مستقرة أمنياً وتعيش ظروف طبيعية (الظروف غير الطبيعية مثل الحروب والفيضانات) وفيها تأثير واضح للمتغير الخارجي (درجات الحرارة) وهناك قياس واستقرار لساعات التجهيز لبيانات الأحمال الكهربائية اليومية هي محافظة البصرة لذلك تم تناول موضع الأحمال الكهربائية ودرجات الحرارة لمحافظة البصرة للفصل الأول من الأعوام (2014-2015) والجدول الآتي يوضح بعض مقاييس الإحصاء الوصفي لتلك البيانات.

جدول رقم (1-4) يوضح بعض مقاييس الإحصاء الوصفي لبيانات الأحمال الكهربائية لمحافظة البصرة للفصل الأول للعامين (2014-2015)

Descriptive Statistics

	N	Range	Minimum	Maximum	Mean		Std. Deviation	Variance
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic
Yt	174	1080	670	1750	1142.60	16.920	223.187	4.981E4
Ut	174	22	12	34	23.17	.397	5.241	27.473
Valid N (listwise)	174							

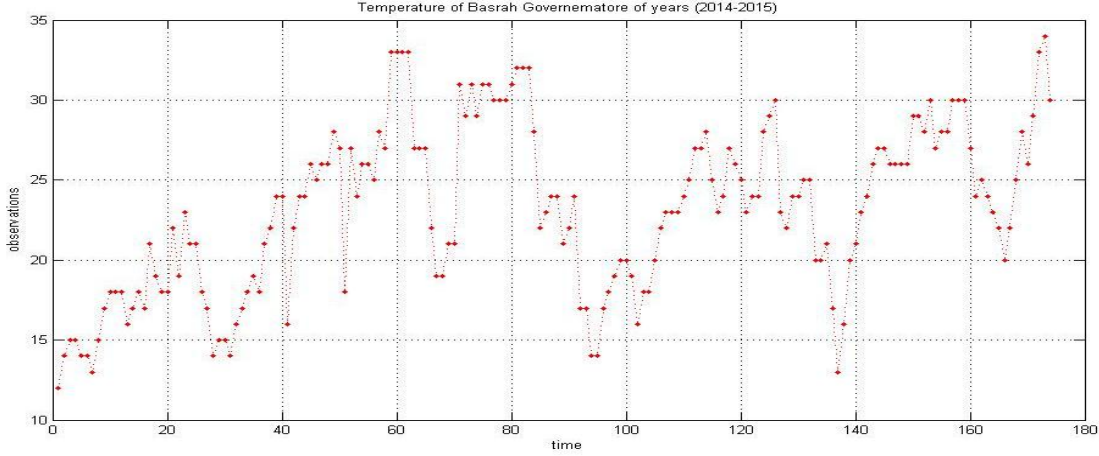
شكل رقم (1-2) يوضح الأحمال الكهربائية الحقيقية لمحافظة البصرة للفصل الأول من العامين (2014-2015)





استعمال تقدير الشرائح التجميعية Additive Splines Estimation لتشخيص أنموذج الانحدار الذاتي التجميعي الاخطي بوجود متغير خارجي NAARX مع تطبيق عملي

شكل رقم (2-2) يوضح درجات الحرارة الحقيقية لمحافظة البصرة للفصل الأول من العامين (2014-2015).



2-4 تحديد رتبة الأنموذج بالطريقة المقترحة الأولى :

بعد تحديد المجموعة المرشحة من متغيرات التأخر الزمني Lag's للسلسلة $\{y_t\}_{t=1}^n$ ولتكن $(p=6)$ وسلسلة زمنية أخرى $\{u_t\}_{t=1}^n$ كمتغير خارجي ومتغيرات التأخر الزمني Lag's التابعة لها $(q=6)$ كمتغيرات توضيحية ولو فرضنا أن $(d=p+q=12)$. وعند استعمال طريقة الشرائح التجميعية (Additive Splines) وباستعمال المعايير (AIC ,BIC ,FPE) لتحديد الرتبة الصحيحة للأنموذج للبيانات الحقيقية للأحمال الكهربائية ودرجة الحرارة كمتغير خارجي لمحافظة البصرة للفصل الأول من العامين (2014-2015) حيث ظهرت النتائج كما في الجدول الآتي :

جدول رقم (2-4) يمثل تحديد الرتبة الصحيحة لأنموذج NAARX(2,0) بالطريقة المقترحة (Additive Splines)

Regressor	AIC	BIC	FPE
y_{t-1}	-1.50951	-1.38242	0.221028
y_{t-1}, y_{t-2}	-1.64356	-1.40754	0.180417
$y_{t-1}, y_{t-2}, y_{t-3}$	-1.61928	-1.20170	0.164780
$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}$	-1.57399	-1.02932	0.159085
$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}$	-1.54422	-0.90878	0.154738
$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}, y_{t-6}$	-1.55072	-0.82450	0.145150
y_{t-1}, y_{t-2}, u_t	-1.37008	-0.55309	0.164175
$y_{t-1}, y_{t-2}, u_t, u_{t-1}$	-1.40729	-0.49952	0.149344
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}$	-1.39266	-0.41227	0.144735
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}, u_{t-3}$	-1.42221	-0.36919	0.134207
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}, u_{t-3}, u_{t-4}$	-1.44612	-0.32048	0.125147
$y_{t-1}, y_{t-2}, u_t, u_{t-1}, u_{t-2}, u_{t-3}, u_{t-4}, u_{t-5}$	-1.41567	-0.21741	0.123219



من الجدول أعلاه نلاحظ أن أقل قيمة لمعيار (AIC) هي (-1.64356) وهي عند الرتبة الثانية للسلسلة y_t كذلك نلاحظ أن أقل قيمة لمعيار (BIC) وهي (-1.40754) وهي أيضاً عند الرتبة الثانية ، أما بالنسبة الى معيار (FPE) فكانت أقل قيمة للمعيار هي (0.123219) والتي تقابل أكبر مجموعة من متغيرات التأخر Lag 's وهي نتيجة منطقية حيث أثبتت ذلك تجارب المحاكاة حيث أن معيار (FPE) كان دائماً يميل الى اختيار أكبر مجموعة من متغيرات التأخر لذلك لا يمكننا الاعتماد عليه بتحديد الرتبة الصحيحة للنموذج ، ومن خلال الجدول أعلاه فإن أفضل رتبة للنموذج بالاعتماد على البيانات الحقيقية هي الرتبة الثانية ومتغيرات التأخر (Lag's) المؤثرة في النموذج من السلسلة (y_t) وبالتالي يمكن الاعتماد على المعيارين (AIC , BIC) في تحديد الرتبة الصحيحة وهي أن رتبة الأنموذج هي (y_{t-1} , y_{t-2}) من السلسلة $\{Y_t\}_{t=1}^n$ وهذا يعني المتغيرات التوضيحية التي لها تأثير مباشر في دالة الانحدار للنموذج (NAARX) هما y_{t-1} و y_{t-2} وهذا يعني أن ($p=2$) وأن ($q=0$) وهذا يشير الى أن الأنموذج هو أنموذج انحدار ذاتي لاخطي من الرتبة الثانية ويرمز له ((NAR(2)). ومن هنا يمكن الإشارة الى ملاحظة جداً مهمة وهي أن التجارب عند حجم العينة (180) مشاهدة باختلاف النماذج والتي يكون فيها عدد المشاهدات قريب من العدد الحقيقي لبيانات الاحمال الكهربائية للفترة المدروسة فإن نتائج المعايير الثلاثة كانت قريبة من بعضها في حالة بيانات المحاكاة والبيانات الحقيقية وهذا مؤشر جيد لغرض المقارنة في النتائج . كذلك فإن رتبة النموذج التي حددت باستخدام البيانات الحقيقية هي قريبة جداً من التي تم تحديدها في حالة المحاكاة بالنسبة للنموذج الثالث للمعيار (BIC) ، والرتبة نفسها مع إضافة تأثير المتغير الخارجي في حالة النموذج الأول ، وبالنسبة الى معيار (AIC) أيضاً كانت قريبة من الرتبة الصحيحة للبيانات الحقيقية للسلسلة (y_t) .

5-الاستنتاجات والتوصيات :

1-5 الاستنتاجات :

- 1- أن التشخيص باستخدام الطريقة المقترحة يتأثر كثيراً بعدد العقد المستخدمة في التقدير وكذلك بحجم العينة ، حيث كلما كبر حجم العينة وأزداد عدد العقد أنتجت تقديرات دقيقة ومن ثم حصلنا على تشخيص دقيق للنموذج ومن ثم حددت الرتبة بشكل صحيح .
- 2- التشخيص الصحيح للنموذج يكون أكثر دقة عند زيادة حجم العينة .
- 3- تبين أن المعيار (BIC) هو أفضل معيار لتحديد الرتبة الصحيحة للنماذج المفترضة وهو معيار دقيق وأثبتت كفاءته باختلاف أحجام العينات .

2-5 التوصيات :

- 1- يمكن اعتماد الطريقة المقترحة الأولى للتشخيص وتحديد الرتبة لنماذج السلاسل الزمنية اللاخطية المختلفة وذلك لأنها حققت نجاح في التشخيص وتحديد الرتبة لأنموذج الانحدار الذاتي التجميعي اللاخطي بوجود متغير خارجي (NAARX) .
- 2- يمكن استخدام متغير الرطوبة كمتغير خارجي مع متغير الاحمال الكهربائية في النموذج نفسه الذي تم تناوله في هذه الدراسة ، وخاصة في الفصل السنوي الثالث وأخذ بيانات هذا المتغير من دائرة الأنواء الجوية لعدم وجود مثل هكذا بيانات في وزارة الكهرباء .
- 3- عند دراسة موضوع الاحمال الكهربائية لمحافظة أخرى (بغداد مثلاً) يمكن استخدام متغير معدل ساعات التجهيز وأخذ معدلات الاحمال الكهربائية المجهزة خلال اليوم الواحد ، لتكون الدراسة دقيقة جداً .
- 4- توصي الدراسة بضرورة قياس درجات الحرارة في كل محافظة وترفق مع التقرير اليومي للأحمال الكهربائية ولكل محافظة (حمل الذروة) ليتسنى للجهات ذات العلاقة متابعة زيادة الاحمال مع زيادة درجات الحرارة واتخاذ التدابير اللازمة من زيادة الإنتاج وحملات أعلانية للترشيد والعمل ببرنامج القطع المبرمج للسيطرة على النقص الذي يحصل الطاقة و ثم التوزيع للطاقة الكهربائية بشكل صحيح



استعمال تقدير الشرائح التجميعية Additive Splines Estimation لتشخيص أنموذج الانحدار الذاتي التجميعي اللاخطي بوجود متغير خارجي NAARX مع تطبيق عملي

المصادر :

- 1- العذاري، فارس مسلم و الوكيل، الوكيل، علي عبد الحسين (1990) "العمليات التصادفية"، كتاب منهجي، مطبعة جامعة الموصل.
- 2-Bradley, R. C. (2005). "Basic Properties of Strong Mixing Conditions", DTIC Document.
- 3-Chen, R., Liu, J. S. et al. (1995). "Additivity tests for nonlinear autoregression.", *Biometrika*, NO. 82(2):pp: 369-383.
- 4-Chen, R. and Tsay, R. S. (1993). "Nonlinear additive ARX models." *Journal of the American Statistical Association*, NO. 88(423): pp:955-967.
- 5-Damouras, S. (2008) "Nonparametric Time Series Analysis Using Gaussian Processes" A Dissertation submitted to Department of Statistics In Carnegie Mellon University.
- 6-Huang, J. Z. and Shen, H. (2004). "Functional Coefficient Regression Models for Non-linear Time Series: A Polynomial Spline Approach." *Scandinavian journal of statistics*, NO. 31(4):pp: 515-534.
- 7- Huang, J. Z. and Yang, L. (2004). "Identification of nonlinear additive autoregressive models.", *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, NO. 66(2):pp: 463-477.
- 8-Xue, L. and Yang, L. (2006). "Additive coefficient modeling via polynomial spline." *Statistica Sinica*:pp: 1423-1446.
- 9-Yao, Q. and Fan, J. (2003). "Nonlinear Time Series Nonparametric and Parametric Methods.", New York, Inc.

الملحق (1)

بيانات الأحمال الكهربائية (Y_t) ودرجات الحرارة (U_t) لمحافظة البصرة للفصل الأول للعامين (2014-2015) م. والتي استخلصت من التقرير الشهري دائرة توزيع الطاقة / وزارة الكهرباء

No.	Date	Yt	Ut	No.	Date	Yt	Ut
1-	2014/1/7	1200	12	61-	2014/3/8	910	33
2-	2014/1/8	1320	14	62-	2014/3/9	1020	33
3-	2014/1/9	1200	15	63-	2014/3/10	950	27
4-	2014/1/10	1210	15	64-	2014/3/11	970	27
5-	2014/1/11	1230	14	65-	2014/3/12	940	27
6-	2014/1/12	1245	14	66-	2014/3/13	840	22
7-	2014/1/13	1260	13	67-	2014/3/14	900	19
8-	2014/1/14	1200	15	68-	2014/3/15	820	19
9-	2014/1/15	1125	17	69-	2014/3/16	890	21
10-	2014/1/16	1260	18	70-	2014/3/17	900	21
11-	2014/1/17	1190	18	71-	2014/3/18	915	31
12-	2014/1/18	1200	18	72-	2014/3/19	800	29
13-	2014/1/19	1150	16	73-	2014/3/20	820	31
14-	2014/1/20	1180	17	74-	2014/3/21	810	29
15-	2014/1/21	1200	18	75-	2014/3/22	830	31
16-	2014/1/22	1200	17	76-	2014/3/23	835	31



17-	2014/1/23	1230	21	77-	2014/3/24	820	30
18-	2014/1/24	1260	19	78-	2014/3/25	850	30
19-	2014/1/25	1280	18	79-	2014/3/26	810	30
20-	2014/1/26	1250	18	80-	2014/3/27	840	31
21-	2014/1/27	1250	22	81-	2014/3/28	800	32
22-	2014/1/28	1260	19	82-	2014/3/29	860	32
23-	2014/1/29	1100	23	83-	2014/3/30	670	32
24-	2014/1/30	1150	21	84-	2014/3/31	760	28
25-	2014/1/31	1250	21	85-	2015/1/1	1355	22
26-	2014/2/1	1250	18	86-	2015/1/2	1410	23
27-	2014/2/2	1150	17	87-	2015/1/3	1325	24
28-	2014/2/3	1200	14	88-	2015/1/4	1400	24
29-	2014/2/4	1100	15	89-	2015/1/5	1375	21
30-	2014/2/5	1150	15	90-	2015/1/6	1430	22
31-	2014/2/6	1250	14	91-	2015/1/7	1300	24
32-	2014/2/7	1300	16	92-	2015/1/8	1550	17
33-	2014/2/8	1300	17	93-	2015/1/9	1390	17
34-	2014/2/9	1298	18	94-	2015/1/10	1390	14
35-	2014/2/10	1285	19	95-	2015/1/11	1650	14
36-	2014/2/11	1240	18	96-	2015/1/12	1600	17
37-	2014/2/12	1200	21	97-	2015/1/13	1580	18
38-	2014/2/13	1180	22	98-	2015/1/14	1510	19
39-	2014/2/14	1200	24	99-	2015/1/15	1410	20
40-	2014/2/15	1180	24	100-	2015/1/16	1600	20
41-	2014/2/16	1230	16	101-	2015/1/17	1620	19
42-	2014/2/17	1200	22	102-	2015/1/18	1750	16
43-	2014/2/18	1220	24	103-	2015/1/19	1670	18
44-	2014/2/19	1180	24	204-	2015/1/20	1725	18
45-	2014/2/20	1200	26	105-	2015/1/21	1610	20
46-	2014/2/21	1080	25	106-	2015/1/22	1550	22
47-	2014/2/22	1025	26	107-	2015/1/23	1450	23
48-	2014/2/23	1100	26	108-	2015/1/24	1500	23
49-	2014/2/24	1100	28	109-	2015/1/25	1525	23
50-	2014/2/25	1010	27	110-	2015/1/26	1250	24
51-	2014/2/26	920	18	111-	2015/1/27	1310	25
52-	2014/2/27	990	27	112-	2015/1/28	1300	27
53-	2014/2/28	690	24	113-	2015/1/29	1210	27
54-	2014/3/1	980	26	114-	2015/1/30	1100	28
55-	2014/3/2	925	26	115-	2015/1/31	1100	25
56-	2014/3/3	960	25	116-	2015/2/1	1200	23
57-	2014/3/4	890	28	117-	2015/2/2	1190	24
58-	2014/3/5	920	27	118-	2015/2/3	1280	27
59-	2014/3/6	825	33	119-	2015/2/4	1110	26
60-	2014/3/7	880	33	120-	2015/2/5	1225	25



استعمال تقدير الشرائح التجميعية Additive Splines Estimation لتشخيص أنموذج
الانحدار الذاتي التجميعي الاخطي بوجود متغير خارجي NAARX مع تطبيق عملي

No.	Date	Yt	Ut	No.	Date	Yt	Ut
121-	2015/2/6	1200	23	148-	2015/3/5	1050	26
122-	2015/2/7	1258	24	149-	2015/3/6	1030	26
123-	2015/2/8	1080	24	150-	2015/3/7	1088	29
124-	2015/2/9	1200	28	151-	2015/3/8	1120	29
125-	2015/2/10	995	29	152-	2015/3/9	1050	28
126-	2015/2/11	1050	30	153-	2015/3/10	1030	30
127-	2015/2/12	1010	23	154-	2015/3/11	1000	27
128-	2015/2/13	1200	22	155-	2015/3/12	1020	28
129-	2015/2/14	1150	24	156-	2015/3/13	1000	28
130-	2015/2/15	1220	24	157-	2015/3/14	1020	30
131-	2015/2/16	1220	25	158-	2015/3/15	919	30
132-	2015/2/17	1230	25	159-	2015/3/16	980	30
133-	2015/2/18	1170	20	160-	2015/3/17	1025	27
134-	2015/2/19	1125	20	161-	2015/3/18	840	24
135-	2015/2/20	890	21	162-	2015/3/19	810	25
136-	2015/2/21	1425	17	163-	2015/3/20	860	24
137-	2015/2/22	1480	13	164-	2015/3/21	965	23
138-	2015/2/23	1450	16	165-	2015/3/22	980	22
139-	2015/2/24	1400	20	166-	2015/3/23	985	20
140-	2015/2/25	1350	21	167-	2015/3/24	765	22
141-	2015/2/26	1200	23	168-	2015/3/25	965	25
142-	2015/2/27	1250	24	169-	2015/3/26	960	28
143-	2015/2/28	1200	26	170-	2015/3/27	925	26
144-	2015/3/1	1150	27	171-	2015/3/28	960	29
145-	2015/3/2	1110	27	172-	2015/3/29	75	33
146-	2015/3/3	1140	26	173-	2015/3/30	990	34
147-	2015/3/4	1100	26	174-	2015/3/31	954	30
148-	2015/3/5	1050	26	148-	2015/3/5	1050	26
121-	2015/2/6	1200	23	149-	2015/3/6	1030	26
122-	2015/2/7	1258	24	150-	2015/3/7	1088	29
123-	2015/2/8	1080	24	151-	2015/3/8	1120	29
124-	2015/2/9	1200	28	152-	2015/3/9	1050	28
125-	2015/2/10	995	29	153-	2015/3/10	1030	30
126-	2015/2/11	1050	30	154-	2015/3/11	1000	27
127-	2015/2/12	1010	23	155-	2015/3/12	1020	28
128-	2015/2/13	1200	22	156-	2015/3/13	1000	28
129-	2015/2/14	1150	24	157-	2015/3/14	1020	30
130-	2015/2/15	1220	24	158-	2015/3/15	919	30
131-	2015/2/16	1220	25	159-	2015/3/16	980	30
132-	2015/2/17	1230	25	160-	2015/3/17	1025	27
133-	2015/2/18	1170	20	161-	2015/3/18	840	24
134-	2015/2/19	1125	20	162-	2015/3/19	810	25
135-	2015/2/20	890	21	163-	2015/3/20	860	24



استعمال تقدير الشرائح التجميعية Additive Splines Estimation لتشخيص أنموذج الانحدار الذاتي التجميعي الاخطي بوجود متغير خارجي NAARX مع تطبيق عملي

136-	2015/2/21	1425	17	164-	2015/3/21	965	23
137-	2015/2/22	1480	13	165-	2015/3/22	980	22
138-	2015/2/23	1450	16	166-	2015/3/23	985	20
139-	2015/2/24	1400	20	167-	2015/3/24	765	22
140-	2015/2/25	1350	21	168-	2015/3/25	965	25
141-	2015/2/26	1200	23	169-	2015/3/26	960	28
142-	2015/2/27	1250	24	170-	2015/3/27	925	26
143-	2015/2/28	1200	26	171-	2015/3/28	960	29
144-	2015/3/1	1150	27	172-	2015/3/29	75	33
145-	2015/3/2	1110	27	173-	2015/3/30	990	34
146-	2015/3/3	1140	26	174-	2015/3/31	954	30
147-	2015/3/4	1100	26				

صور عن التقرير الشهري ليوم الأربعاء المصادف 2015/1/7 لدائرة توزيع الطاقة ببيان الاحمال الكهربائية ودرجات الحرارة لكل المحافظات

		2015/1/7		الاربعاء		التقرير اليومي ليوم						
		توليد العام الماضي د.ي		التردد (هيرتز)		الوقت		التوليد المتوقع وقت الذروة د.ي				
		11350		50		18:10		9965				
		معدلات العام الماضي 2014				الوطني والمستورد						
نسبة الزيادة % للحمل	نسبة الزيادة % للساعات	ساعات التجهيز	المعدل د.ي	حمل الذروة د.ي	عدد ساعات التجهيز	المعدل د.ي	المستورد د.ي	الوطني د.ي	الاحمال وقت الذروة د.ي	المحافظة	التسليم	
7.28	16.67	15:00	2460	2660	18:00	2639			2800	بغداد	1	
-100.00	-100.00	13:00	1031	1055	0:00	0			0	نينوى	2	
-6.59	9.09	22:00	1154	1200	24:00	1078	85	993	1300	البصرة	3	
-84.92	-50.00	14:00	431	460	7:00	65			85	الانبار	4	
5.81	41.18	17:00	671	705	24:00	710	248	462	690	كربلاء	5	
8.56	9.09	22:00	479	547	24:00	520			620	ذي قار	6	
-59.75	-57.14	14:00	569	580	6:00	229			235	صلاح الدين	7	
-19.31	-10.00	20:00	518	520	20:00	418			485	بابل	8	
-1.49	-5.00	20:00	404	440	20:00	398			430	التنجف	9	
43.93	31.25	16:00	321	440	21:00	462	165	297	430	ديالى	10	
18.75	20.00	20:00	336	360	24:00	399			560	واسط	11	
2.80	-5.00	20:00	429	450	20:00	441			475	كربلاء	12	
-2.82	-5.00	20:00	319	320	20:00	310			330	الديوانية	13	
-1.96	0.00	24:00	306	400	24:00	300	111	189	380	ميسان	14	
-8.85	4.55	22:00	226	250	23:00	206			225	العتيق	15	
			770	963		653			920	الضلع	35	
		####	11350			8828			9965	المجموع	36	
			2014								37	
			بغداد	الموصل	بغداد	الموصل	بغداد	الموصل	بغداد	الموصل	درجات الحرارة العظمى	39
			14	12	13				17	24	13	40
			3	3	1				3	10	6	41



**Use aggregate slide estimate additive splines estimation for the diagnosis of
non-linear composite model self-regression with practical application**

Abstract :

Nonlinear time series analysis is one of the most complex problems ; especially the nonlinear autoregressive with exogenous variable (NARX) .Then ; the problem of model identification and the correct orders determination considered the most important problem in the analysis of time series . In this paper , we proposed splines estimation method for model identification , then we used three criterions for the correct orders determination. Where ; proposed method used to estimate the additive splines for model identification , And the rank determination depends on the additive property to avoid the problem of curse dimensionally . The proposed method is one of the nonparametric methods , and the simulation results give a good identification of the NARX models and the BIC criteria is the best criterion used to determine the correct orders of the selected models.

Key Word : NARX , Additive , Splines Estimation , Interior Knots , Identification.