المجلة العراقية للعلوم الاحصائية (12) 2007
ص ص [77-97]

توظيف أسلوب المكونات الرئيسة في تقدير زمن التأخير في النظم الديناميكية

ظافر رمضان مطر* هيام عبد المجيد حياوي**

الملخص

يتناول هذا البحث توظيف أسلوب رياضي وهو تحليل المكونات الرئيسة Principle Components Analysis في تقدير زمن التأخير Delay Time للأنظمة الحركية التصادفية الخطية.ويعد لزمن التأخير أهمية خاصة في وصف التشخيص المناسب للنظام،لقد تم استخدام طرائق تقليدية ومقارنة نتائجها مع الأسلوب المقترح من خلال تجارب المحاكاة ، اذ تبين أن أسلوب المكونات الرئيسة يقدم إمكانية عالية للنجاح في تقدير زمن التأخير وبشكل خاص عند مقارنته مع دالة الارتباط التقاطعي ودالة الاستجابة النبضية.

Principle Components Approach Employment to Estimate the Delay Time in Dynamical Systems

Abstract

This research deals with a mathematical approach of principle components analysis to estimate the delay time in linear stochastic dynamical systems. The delay time has a statistical importance to describe the suitable identification. So classical methods are used and the results are compared with the proposed approach from the simulation experiments where the Principle Component approach gives a high degree of success to estimate the delay time especially with crosscorrelation and impulse response function.

* استاذ مساعد/ كلية علوم الحاسبات والرياضيات/قسم الإحصاء والمعلوماتية

** مدرس/ كلية علوم الحاسبات والرياضيات/قسم الإحصاء والمعلوماتية.

تاريخ التسلم : 3 /2006/9 ـــــــــــــــــــــــ تاريخ القبول : 22 /2006/11

1. مقدمة:- Introduction

تعد عملية التشخيص لخصائص النظم الديناميكية حجر الأساس الذي يبنى عليه تحديد النموذج المناسب الذي يمثل النظام أو العملية ولذلك فقد انصبت جهود الباحثين على كشف وتوظيف كل الأفكار العلمية في سبيل الوصول إلى التشخيص الدقيق للنظام، من هنا فقد جاء هذا البحث ليسلط الضوء على عملية تقدير زمن التأخير Delay Time لأهمية هذا الزمن وتأثيره الواضح في تشخيص بقية ملامح العملية الديناميكية.

2.هدف البحث

يهدف هذا البحث إلى توظيف أسلوب تحليل المكونات الرئيسة Principle Components Analysis في عملية تقدير زمن التأخير للأنظمة الحركية التصادفية الخطية من خلال التعامل مع دوال تقدير مختلفة لزمن التأخير ومقارنة النتائج المستحصلة مع نتائج الطرائق التقليدية وبالاعتماد على تجارب المحاكاة.

3.طرائق تقدير زمن التأخير:- Delay Time Estimation Methods

هناك العديد من الطرائق والأساليب الإحصائية والهندسية لتقدير زمن التأخير في النظم الحركية الخطية التصادفية والتي أشار إليها واستخدمها العديد من الباحثين ومن هذه الأساليب دالة الارتباط المتقاطع Crosscorrelation Function بين سلسلتي الإدخال Input والإخراج Output، ودالة الاستجابة النبضية Impulse Response Function، والأسلوب الذي استخدمه Thanoon & Ibrahim (1993) ليتم من خلاله تقدير زمن التأخير وهو أسلوب دليل خطّي مُزدوج لدالة الانحدار للسلاسل الزمنية المزدوجة ، في حين استخدم (Ljung 1995,) نماذج الانحدار الذاتي مع مدخلات إضافية ARX من الرتبة الثانية في

عملية تقدير زمن التأخير،واستخدم (البدراني ،2002) أسلوب نماذج فضاء الحالة Space State في تقدير زمن التأخير وكالآتي:-

1) دالة الارتباط المتقاطع *Crosscorrelation Function*

هي دالة تشير إلى الارتباط بين المتغير التابع ويمثل المخرجات(y_t)والمتغير المستقل ويمثل المدخلات(u_t) وتقيس هذه الدالة الارتباط بين سلسلتين زمنيتين عند فترات زمنية مختلفة ،وتُعرف معاملات الارتباط المضاعف كما يأتي:- (فاندل،1992) .

$$\rho_{uy}(k) = \frac{\gamma_{uy}(k)}{\sigma_u \sigma_y} \qquad , k = 0,\pm1,\pm2,... \qquad ... \quad (1)$$

إذ إن:-

$\rho_{uy}(k)$: يُعَبِرْ عن دالة الارتباط المضاعف بين المدخلات والمخرجات.

$\gamma_{uy}(k)$ تُعَبِرْ عن التغاير المتقاطع Cross covariance بين(y_t) و(u_t) ويرمز له اختصاراً ($\mathrm{Cov}(u_t y_t)$.

(σ_y) و(σ_u) : يشيران إلى الانحراف المعياري لسلسلة المخرجات والمدخلات على التوالي.

ويعرف التغاير المستعرض بين المدخلات والمخرجات بالشكل الآتي:-

$$\gamma_{uy}(k) = E[(u_t - \mu_u)(y_t - \mu_y)] \qquad , k = 0,\pm1,\pm2,... \qquad ... \quad (2)$$

إذ إن:-

(μ_y) و(μ_u) : يشيران إلى متوسط سلسلة المخرجات والمدخلات على التوالي.

وتعد دالة الارتباط المستعرض أحد الأساليب التي يمكن اعتمادها في تقدير زمن التأخير في الأنظمة الحركية (الديناميكية) اذ يتم الاعتماد على أول قيمة معنوية أو أول

قيمة مطلقة من قيم معامل الارتباط $\hat{\rho}_{uy}(k)$ لتمثل زمن التـأخير ويمكـن وصـفها كالآتي:-

$$\hat{d} = \max_{k=0,1,2,..} \left|\hat{\rho}_{uy}(k)\right| \quad ... (3)$$

إذ إن:-

$\hat{d}$: تُعَبرْ عن زمن التأخير المقدر.

2) دالة الاستجابة النبضية Impulse Response Function

تعد دالة الاستجابة النبضية مؤشراً للعلاقة بين المدخلات والمخرجـات ،حيـث تمثل الأثر الذي يحدث على المخرجات(y_t) نتيجة لتغير المدخلات(u_t)بمقدار وحدة واحدة ويتم التعبير عنها بوصف أن قيمة المخرجات(y_t)الحالية هي دالة في عدد من قيم المدخلات (u_t)السابقة ويعبر عنها كما ياتي :-

$$y_t = V(B)u_t + \varepsilon_t \quad ... (4)$$

إذ إن :-

y_t : تُعَبِرْ عن المخرجات ، u_t : تُعَبِرْ عن المدخلات ، ε_t : تُعَبِرْ عن الإزعاجـات Disturbances.

وتمثل $V(B) = v_0 + v_1 B + v_2 B^2 + ...$ دالة التحويل للمدخلات وتسمى $v_0, v_1, v_2, ...$ بأوزان الاستجابة النبضية.انظر(فاندل ،1992)،أما B فتمثل عامل الإزاحـة الخلفـي والذي يعرف بالشكل الآتي: $B^k u_t = u_{t-k}$

وتعد دالة الاستجابة النبضية أحد الأساليب المستخدمة في تقدير زمن التـأخير فـي الأنظمة الحركية(Makridakis et.al.,1983) اذ أن أول قيمة معنوية تقـع خـارج فترة الثقة ، تمثل زمن التأخير المقدر $\hat{d}$.

3) استخدام نماذج الانحدار الذاتي بمتغيرات إضافية
Autoregressive with Exogenous Variables (ARX)

اقترح الباحث (Ljung,1995) أسلوباً لتقدير زمن التأخير $\hat{d}$ اعتماداً على نماذج الانحدار الذاتي مع مدخلات إضافية ARX من الرتبة الثانية اذ يعرف نموذج ARXبالشكل الآتي:-

$$A(q)y_t = B(q)u_t + \varepsilon_t \quad \ldots (5)$$

إذ إن:-

$A(q)$: متعددة حدود من الرتبة na ومعرفة بالشكل الآتي:-

$$A(q) = 1 + a_1 q^{-1} + a_2 q^{-2} + \cdots + a_{na} q^{-na} \quad \ldots (6)$$

$B(q)$: متعددة حدود من الرتبة nb ومعرفة بالشكل الآتي:-

$$B(q) = b_1 q^{-1} + b_2 q^{-2} + \cdots + b_{nb} q^{-nb} \quad \ldots (7)$$

اذ تقسم مشاهدات المدخلات والمخرجات سوية إلى جزأين:الأول لتقدير النماذج Models Estimation والثاني لاختبار شرعية النماذج Models Validation، اذ يتم تقدير نماذج من نوع ARX برتب ثابتة لمتعددتي حدود المخرجات والمدخلات $A(q)$ و $B(q)$ على التوالي لتساوي(2) أي أن($na = nb = 2$) مع زمن تأخير يتغير من 1 إلى 10 ، وتستخدم بواقي النموذج الملائم بعد تطبيقه على بيانات الشرعية للحصول على قيمة دالة الخسارة Loss Function ، فيكون مقدر زمن التأخير $\hat{d}$ في هذه الحالة هو القيمة التي تقابل اقل مقدار لدالة الخسارة من بين النماذج المقدرة .(Ljung,2004)

4. المكونات الرئيسة:- Principle Components

يٌعَرَفْ تحليل المكونات الرئيسة Principle Components Analysis بأنـــه أسلوب استكشافي exploratory مفضل لتبسيط وصف مجموعـــة مـــن المتغيـــرات المرتبطة والتي تعالج بالتساوي ،ويستخدم هذا الأسلوب لفهم العلاقات المتداخلة بـــين المتغيرات ويعد طريقة لتحويل المتغيرات التفسيرية الأصلية المرتبطـــة خطيـــاً إلـــى متغيرات جديدة بشكل تراكيب خطية غير مرتبطة فيما بينها أي متعامدة والتي يكـــون عددها بعدد المتغيرات الأصلية اذ أن كل مكون رئيسي هو تجمع خطي من المتغيرات الأصلية، وتكون هذه المكونات مؤهلة لتفسير معظم التباين الكلي لذلك ترتب تنازليـــاً حسب تباينها ، أي أن كمية المعلومات التي يفسرها المكون الرئيس الأول أكبر ممـــا يفسره المكون الرئيس الثاني والمعلومات التي يفسرها المكون الرئيس الثاني أكبر مما يفسره المكون الرئيس الثالث وهكذا بالنسبة الى بقيـــة المكونـــات الرئيســـة الأخـــر. (Afifi,1984) و (الراوي ،1987) ويعد تحليـــل المكونـــات الرئيســـة أداة لتحويـــل الإحـــداثيات Coordinate Axis Transformation وتخفـــيض الأبعـــاد Dimensionality Reduction ، ومن خلاله يتم حســـاب مجموعـــة جديـــدة مـــن الإحداثيات من خلال تعظيم تباين عينة نقاط البيانات مع تلـــك الإحـــداثيات.(Nelles ,2001)

ومن اجل توضيح وتيسير فهم المكونات الرئيسة نأخذ متغيـــرين وليكونـــا(X_1) و(X_2) وبوجود N من المشاهدات، اذ يتم تحويل بيانات المتغيرين إلى بيانات بوسط حسابي صفر وتباين (S_2^2, S_1^2)على التوالي:- (Afifi,1984).

$$\left.\begin{aligned} x_1 &= X_1 - \bar{X}_1 \\ x_2 &= X_2 - \bar{X}_2 \end{aligned}\right\} \quad \ldots (8)$$

إن الفكرة الأساسية من تكوين المكونات الرئيسة هي الحصول على متغيــرين جديدين هما (C_1) و(C_2) وكل منهما هو دالة خطية بدلالة المتغيرين(x_1, x_2)على التوالي أي أن:-

$$\left.\begin{aligned} C_1 &= a_{11}x_1 + a_{12}x_2 \\ C_2 &= a_{21}x_1 + a_{22}x_2 \end{aligned}\right\} \quad \ldots (9)$$

إذ إن:-

($a_{11}, a_{12}, a_{21}, a_{22}$) تمثل معاملات المكونات الرئيسة أي المتجهات المميزة Egien Vectors وان متوسط المكونين الرئيسين(C_1)و(C_2) هو:-

$$mean\ C_1 = mean\ C_2 = 0 \quad \ldots (10)$$

في حين تباينهما :-

$$\left.\begin{aligned} Var\ C_1 &= a_{11}^2 S_1^2 + a_{12}^2 S_2^2 + 2a_{11}a_{12}rS_1S_2 \\ Var\ C_2 &= a_{21}^2 S_1^2 + a_{22}^2 S_2^2 + 2a_{21}a_{22}rS_1S_2 \end{aligned}\right\} \quad \ldots (11)$$

إذ إن:-

$$S_1^2 = Var\ X_1 \quad , S_2^2 = Var\ X_2$$

يتم اختيار معاملات المكونات الرئيسة لتحقق ثلاثة متطلبات:-

1- تباين المكون الأول C_1 يكون اكبر ما يمكن قدر الإمكان.

2- قيم المشاهدات في المكونات الرئيسة C_1 و C_2 تكون غير مرتبطة.

3- $a_{11}^2 + a_{12}^2 = a_{21}^2 + a_{22}^2 = 1$.

4. تجارب المحاكاة:- Simulation Experiments

أجريت تجارب المحاكاة على نموذجين احدهما خطي Linear والآخر غير خطي Nonlinear اذ تم توليد مشاهدات بحجم(n=1000) مشاهدة من كل من النموذجين وكانت مدخلات النموذج(u_t) إشارات عشوائية مولدة من التوزيع الطبيعي القياسي Random Gaussian signals "rgs" وباستخدام المولد randn ،أما التشويش فقد تم توليده باستخدام المولد randn لتوليد إشارات عشوائية تتبع التوزيع الطبيعي($\varepsilon_t \sim N(0,5)$) وتمت عمليات المحاكاة من خلال نظام MATLAB ،وقد استخدم النموذج الخطي الاتي وبزمن تأخير($d = 3$):-

$$y_t = 3 - 0.7 * u(t-3) + 0.4 * u(t-4) + 0.8 * y(t-4) + \varepsilon_t \qquad ...(12)$$

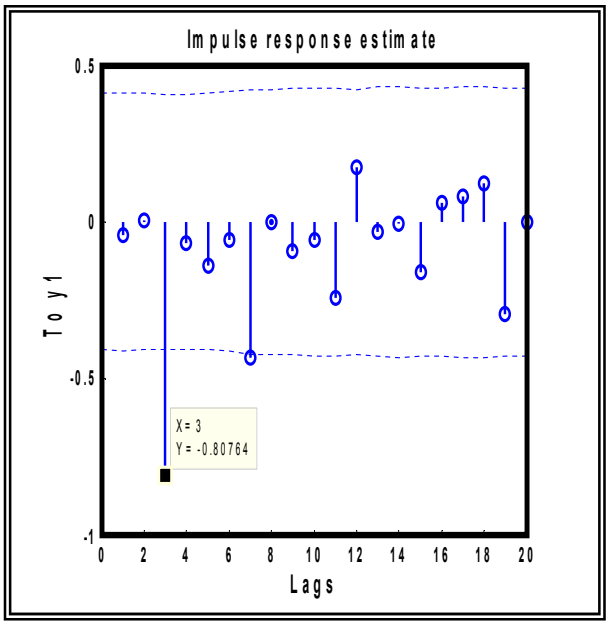
في حين استخدم النموذج غير الخطي التالي وبزمن تأخير($d = 2$):-

$$y_t = (1 + 0.5 * \exp(-0.6 * u(t-2)^\wedge 2)) * u(t-2) + 0.8 * y(t-1) + \varepsilon_t \qquad ...(13)$$

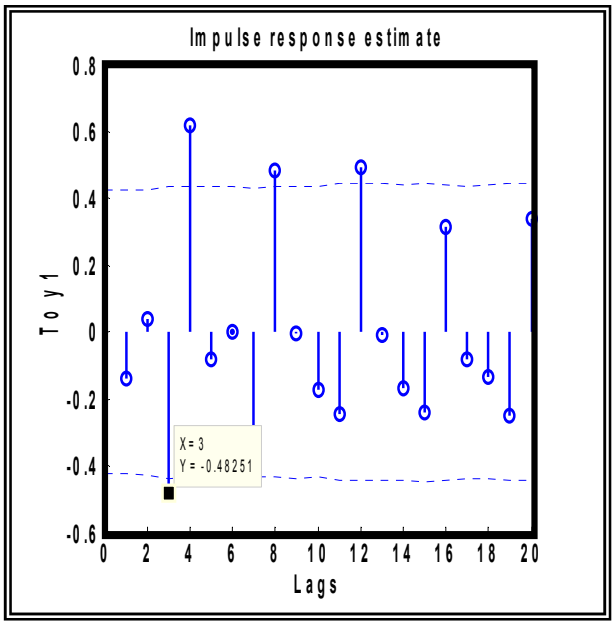
وبتطبيق طرائق تقدير زمن التأخير على البيانات الأصلية وعلى بيانات الإدخال-الإخراج بعد تحويلها إلى مكوناتها الرئيسة كانت نتائج عملية المحاكاة كالآتي:-

أ- النموذج الخطي مع سلاسل البيانات الأصلية وللمكونات الرئيسة .

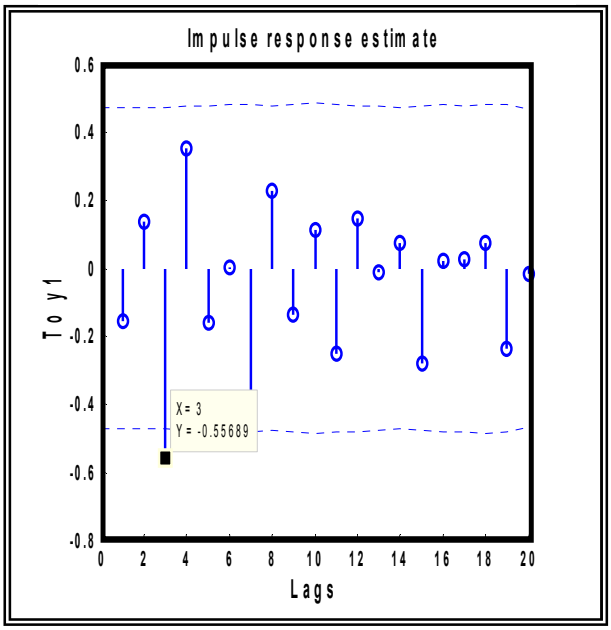
1-عند استخدام دالة الاستجابة النبضية مع سلاسل البيانات الأصلية وبتكرار التجربة (25) مرة، ظهر نجاحاً بنسبة 100% لهذه الدالة في تقدير زمن التأخير اذ أن أول قيمة معنوية لها تقع خارج فترة الثقة، كانت عند التخلف (lag=3) والتكرارات الأربعة الاتية في الشكل (1) توضح ذلك.



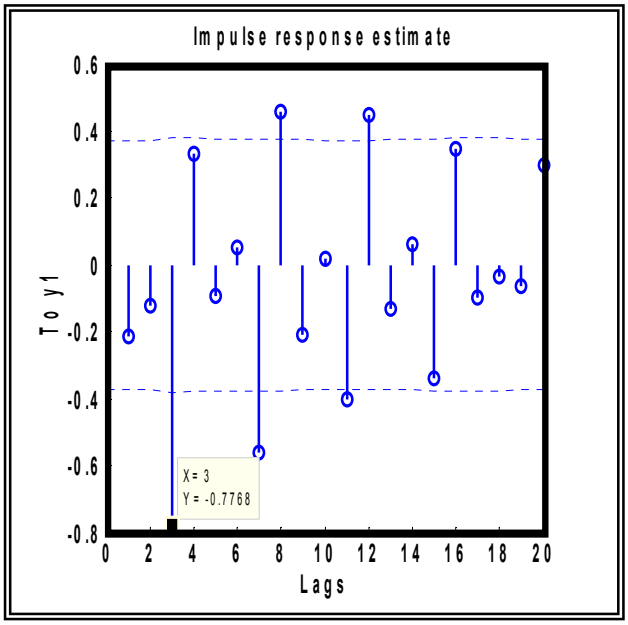
التكرار الثاني



التكرار الأول



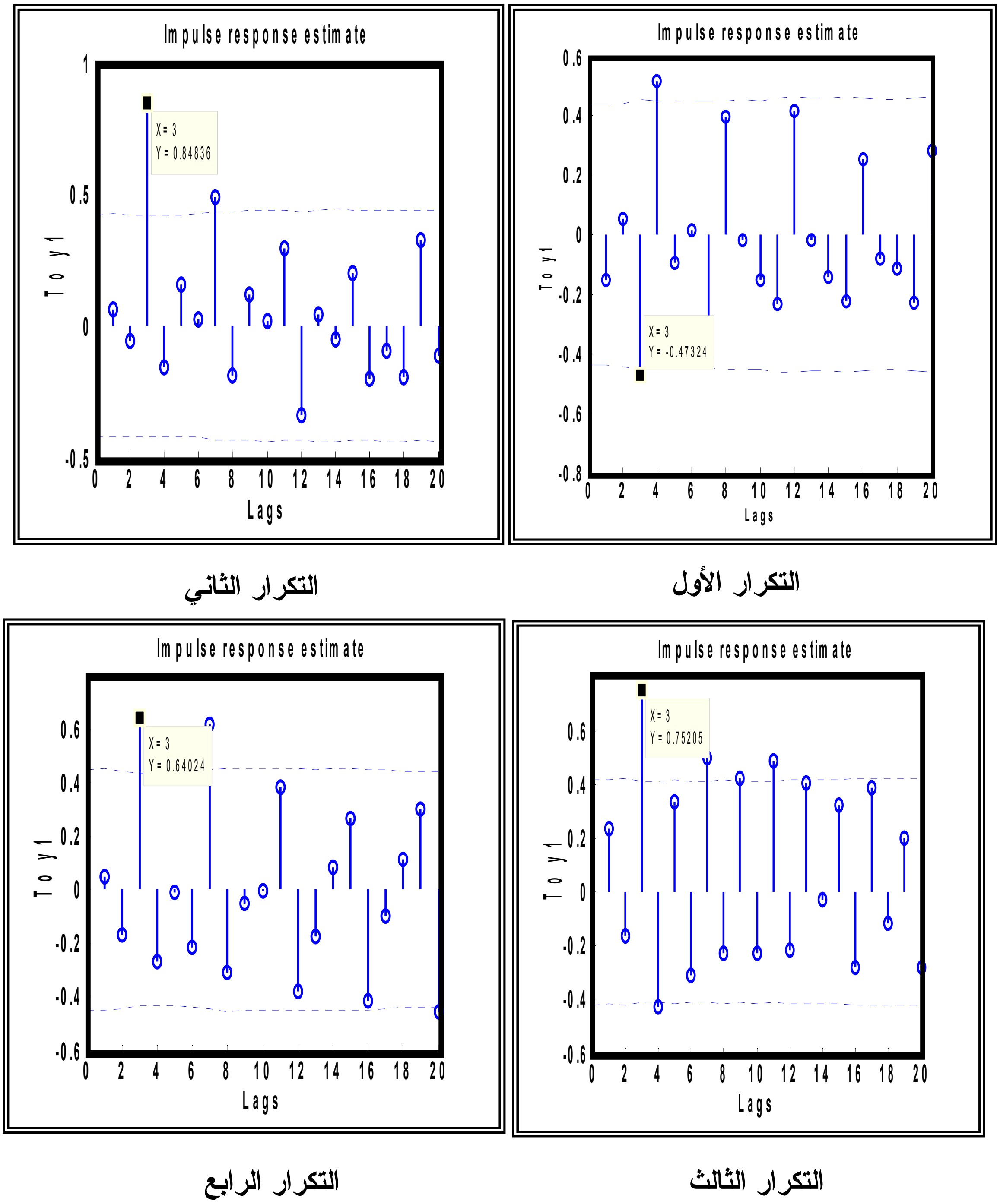
التكرار الرابع



التكرار الثالث

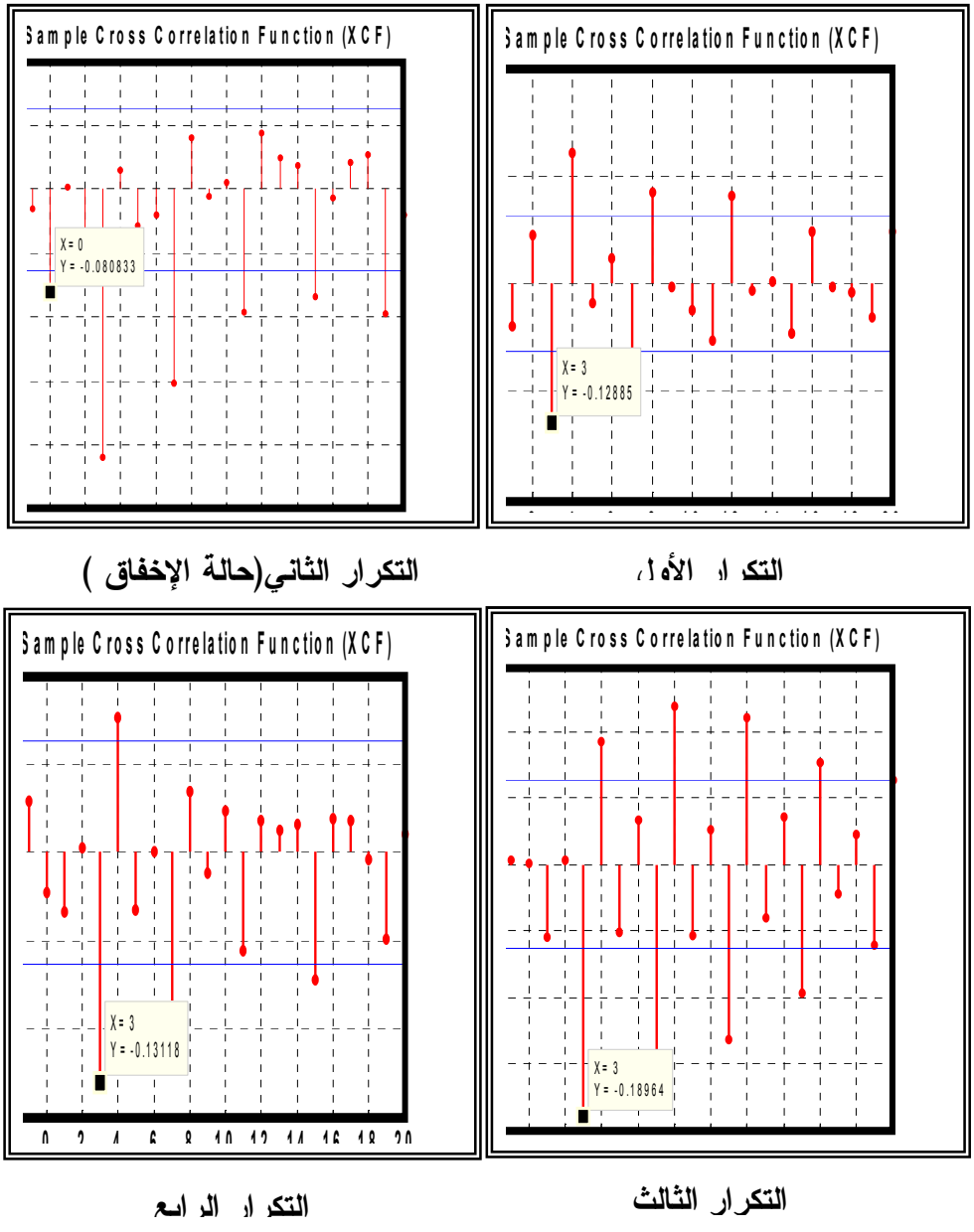
الشكل (1): مقدر دالة الاستجابة النبضية لسلاسل البيانات الأصلية الخاصة ببيانات النموذج الخطي (5).

أما عند استخدام المكونات الرئيسة فقد ظهر نجاحاً بنسبة 100% لهذه الدالة في تقدير زمن التأخير اذ أنها تبلغ ذروتها عند التخلف (lag=3) والشكل (2)يوضح ذلك.



الشكل (2): مقدر دالة الاستجابة النبضية لبيانات المكونات الرئيسة الخاصة ببيانات النموذج الخطي(5).

2-عند استخدام دالة الارتباط المتقاطع نلاحظ حصول بعض حالات الإخفاق لهذه الدالة في تقدير زمن التأخير عند استخدام البيانات الأصلية والشكل(3) يوضح ذلك.

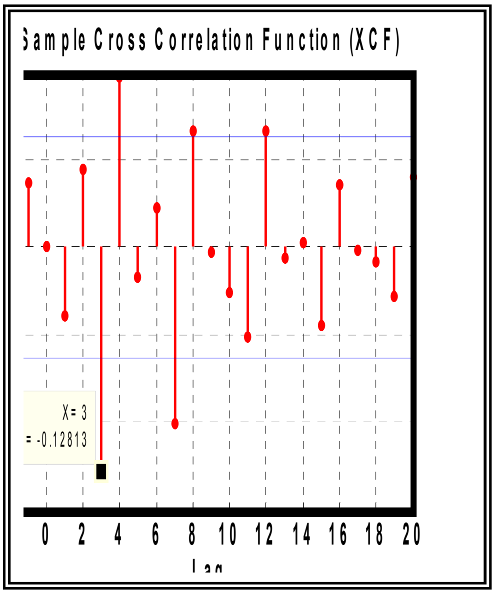


التكرار الأول التكرار الثاني(حالة الإخفاق)

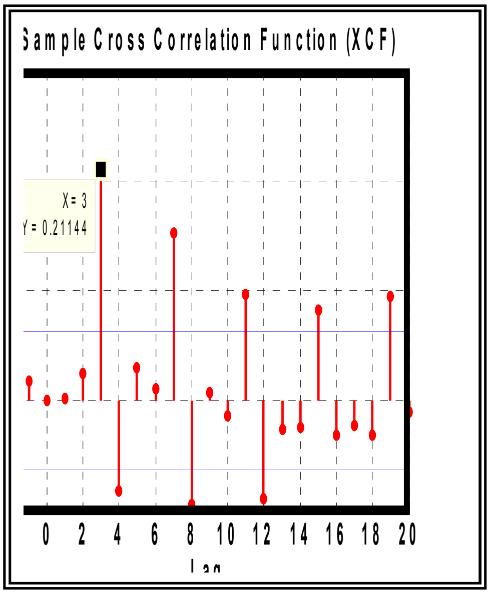
التكرار الثالث التكرار الرابع

الشكل(3):مقدر دالة الارتباط المتقاطع لسلاسل البيانات الأصلية الخاصة ببيانات النموذج

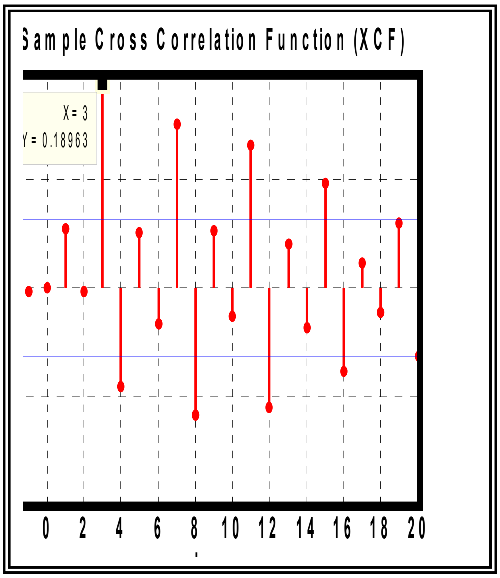
وعلى العكس من ذلك فقد نجحت دالة الارتباط المتقاطع نجاحا تاما في تقـدير زمـن التأخير للنموذج الخطي (5)عند تطبيق هذه الدالة على بيانات المكونـات الرئيسـة و الشكل (4) يبين ذلك.



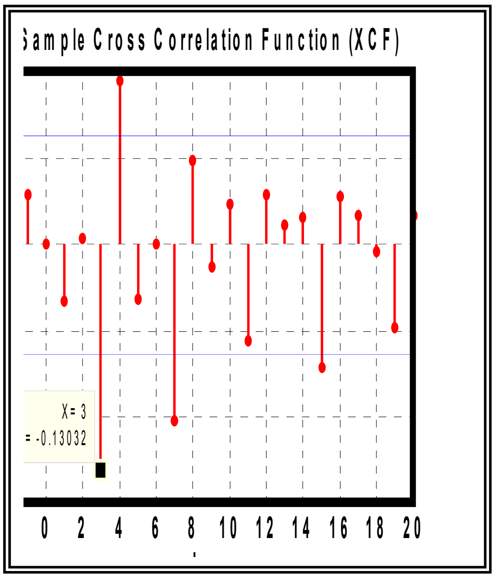
التكرار الأول



التكرار الثاني



التكرار الثالث



التكرار الرابع

الشكل(4): مقدر دالة الارتباط المتقاطع لبيانات المكونات الرئيسة الخاصة ببيانات النموذج

3- عند استخدام أسلوب Ljung لتوافيق نماذج ARXنلاحظ وجود تشابه في حالات الإخفاق بين التطبيق على البيانات الأصلية والمكونات الرئيسة وكما موضح في الجدولين (1) و(2).

الجدول (1) : قيم دالة الخسارة لنماذج ARX من الرتبة الثانية وبأزمنة تأخير مختلفة للبيانات الأصلية للنموذج الخطي (5)

زمن التأخير	قيم دالة الخسارة لنماذج ARX للنموذج الخطي ذي التشويش الطبيعي للبيانات الأصلية			
	التكرار الأول	التكرار الثاني	التكرار الثالث	التكرار الرابع
1	2.6281	2.8731	2.6231	2.3282
2	2.6116	2.7919*	2.6138	2.32*
3	2.5853**	2.7949	2.6217	2.3275
4	2.6079	2.8713	2.6179	2.3365
5	2.6412	2.8673	2.6005	2.3251
6	2.6288	2.8295	2.6038	2.3256
7	2.6166	2.8306	2.6295	2.3288
8	2.6351	2.8715	2.6349	2.3332
9	2.6373	2.8748	2.6089	2.3291
10	2.6342	2.8559	2.5863*	2.3322

(*) تعني الفشل في تقدير زمن التأخير. (**) تعني النجاح في تقدير زمن التأخير.

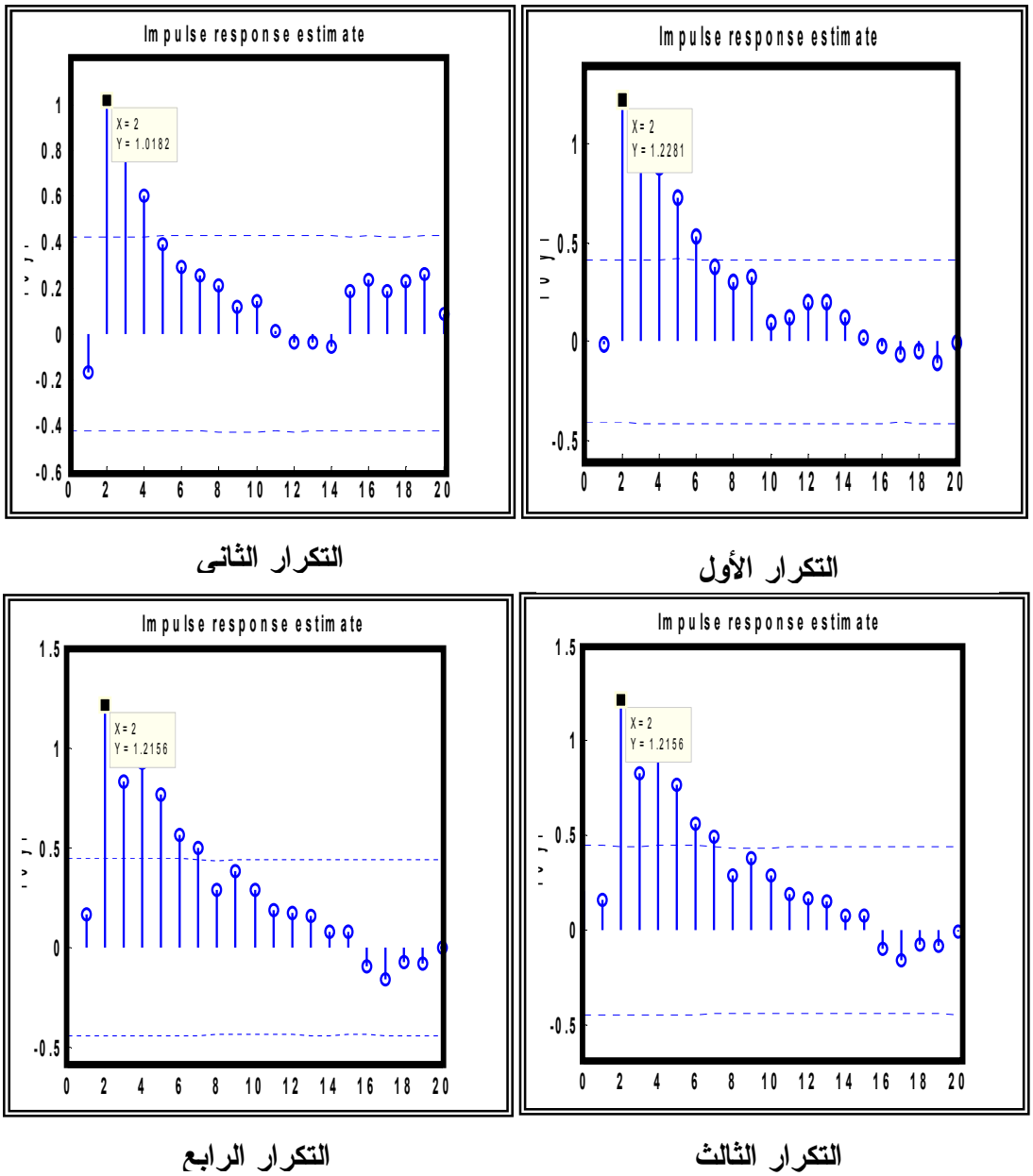
جدول (2): قيم دالة الخسارة لنماذج ARX من الرتبة الثانية وبأزمنة تأخير مختلفة للمكونات الرئيسة لبيانات للنموذج الخطي (5)

زمن التأخير	قيم دالة الخسارة لنماذج ARX للنموذج الخطي ذي التشويش الطبيعي للمكونات الرئيسة			
	التكرار الأول	التكرار الثاني	التكرار الثالث	التكرار الرابع
1	2.6285	2.8736	2.6231	2.3287
2	2.6132	2.7902	2.6138	2.3208*
3	2.5957**	2.7856**	2.6217	2.3294
4	2.6162	2.8631	2.6179	2.3386
5	2.6416	2.8679	2.6005	2.3256
6	2.6302	2.8301	2.6038	2.3264
7	2.6205	2.82	2.6295	2.3306
8	2.6374	2.8613	2.6349	2.3352
9	2.6374	2.8756	2.6089	2.3295
10	2.6352	2.8573	2.5863*	2.3326

(*) تعني الفشل في تقدير زمن التأخير . (**) تعني النجاح في تقدير زمن التأخير.

ب-النموذج غير الخطي مع سلاسل البيانات الأصلية والمكونات الرئيسة.

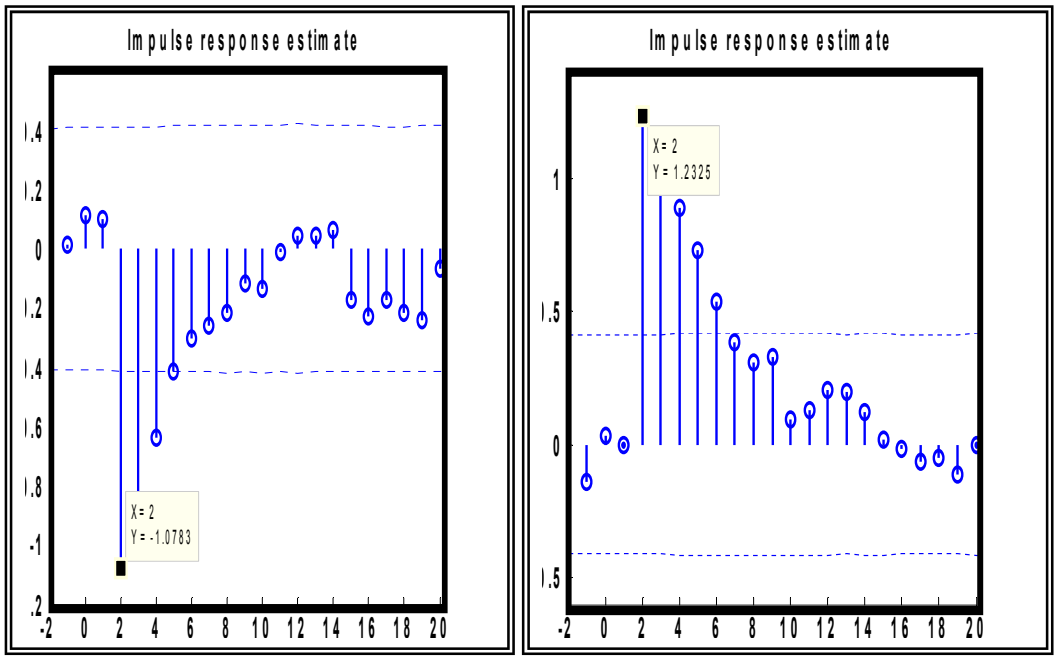
1- عند استخدام دالة الاستجابة النبضية نلاحظ النجاح التام لهذه الدالة في تقدير زمن التاخير للبيانات الأصلية للنموذج غير الخطي (6) وعند استخدام المكونات الرئيسة والشكلان (5) و(6) يوضحان ذلك.



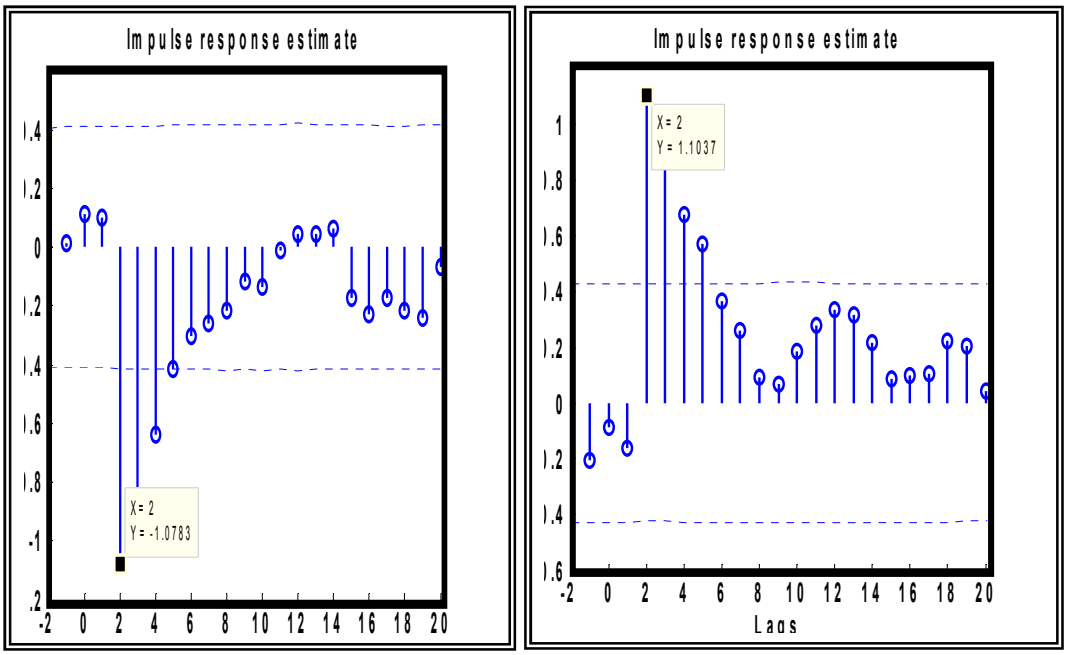
التكرار الثانى التكرار الأول

التكرار الرابع التكرار الثالث

الشكل (5): مقدر دالة الاستجابة النبضية لسلاسل البيانات الأصلية الخاصة ببيانات النموذج غير الخطي



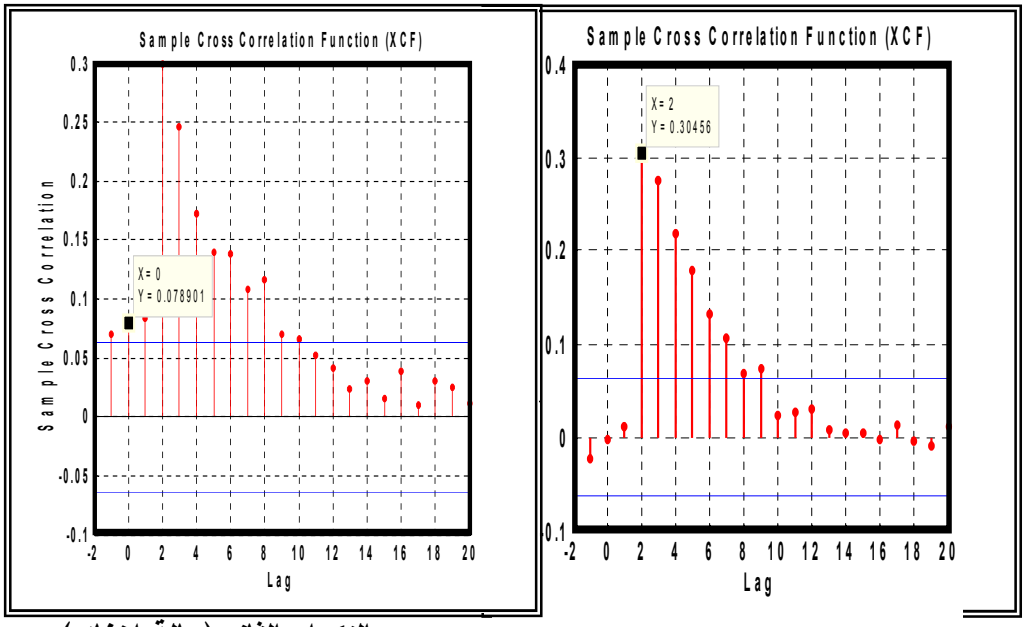
التكرار الاول التكرار الثاني



التكرار الثالث التكرار الرابع

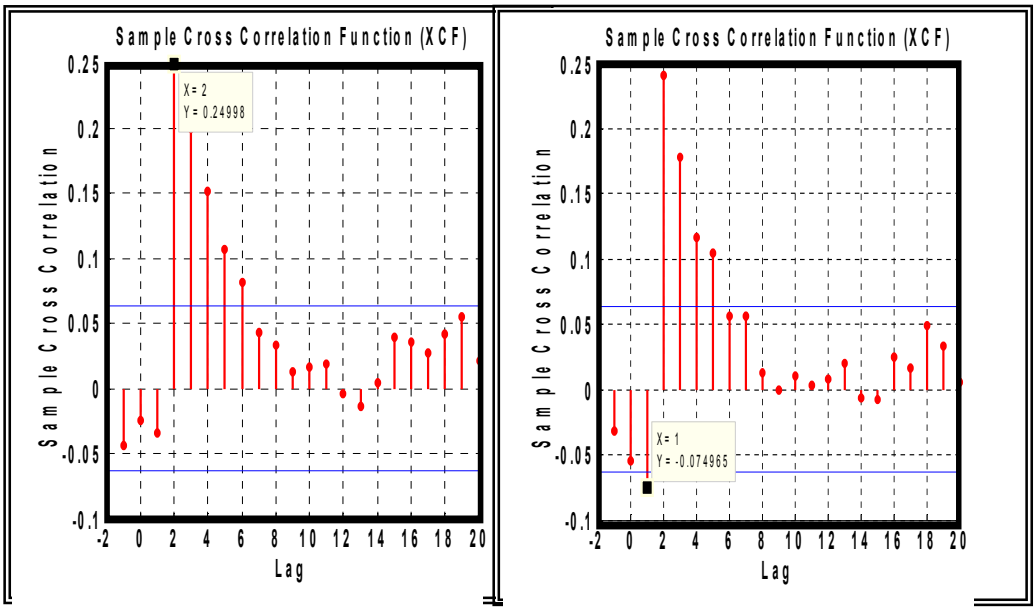
الشكل (6): مقدر دالة الاستجابة النبضية لبيانات المكونات الرئيسة الخاصة ببيانات النموذج غير الخطى

2- عند استخدام دالة الارتباط التقاطعي لوحظ وجود حالتي إخفاق لهذه الدالة في تقدير زمن التأخير عند استخدام البيانات الأصلية، ونجاحها بنسبة 100% مع المكونات الرئيسة والشكلان (7) و (8) يوضحان ذلك.



التكرار الأول

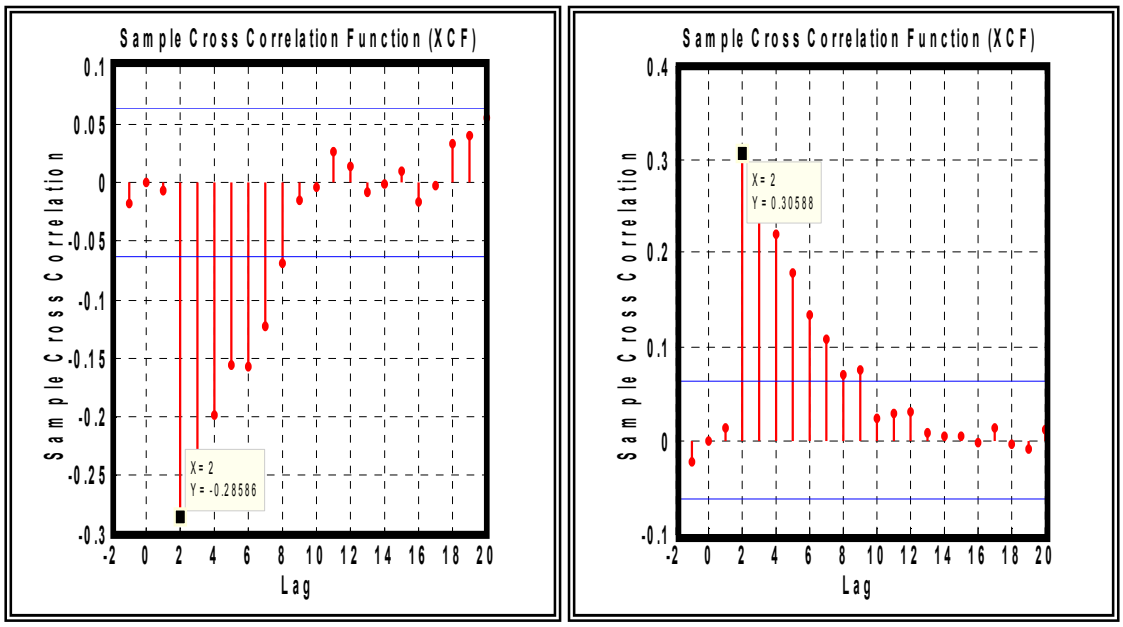
التكرار الثانى(حالة إخفاق)



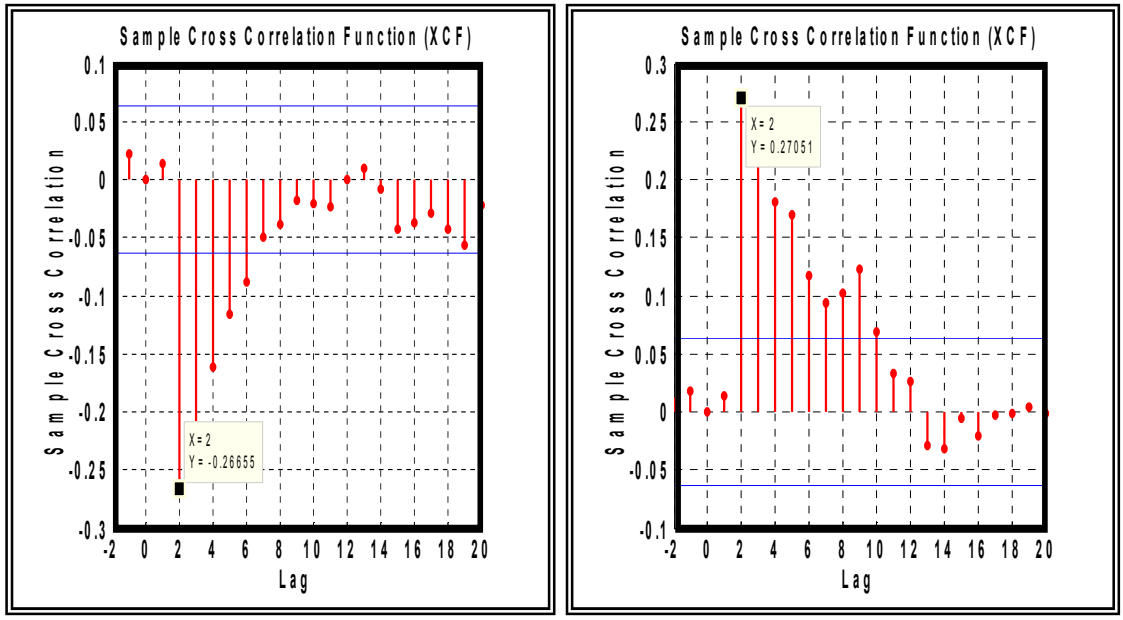
التكرار الثالث (حالة اخفاق)

التكرار الرابع

الشكل(7) : مقدر دالة الارتباط التقاطعي للبيانات الاصلية الخاصة ببيانات النموذج غير



التكرار الاول التكرار الثاني



التكرار الثالث التكرار الرابع

الشكل (8): مقدر دالة الارتباط المتقاطع **للمكونات الرئيسة الخاصة ببيانات النموذج غير الخطي (6).**

3- عند استخدام توافيق لنماذج ARX، لوحظ وجود حالات إخفاق عديدة لهذه النماذج في تقدير زمن التأخير للبيانات الأصلية وللمكونات الرئيسة والجدولان (3) و(4) يوضحان ذلك.

جدول (3) :قيم دالة الخسارة لنماذج ARX من الرتبة الثانية وبأزمنة تأخير مختلفة للبيانات الأصلية للنموذج غير الخطي (6).

التكرار الرابع	التكرار الثالث	التكرار الثاني	التكرار الأول	زمن التأخير
قيم دالة الخسارة لنماذج ARX للنموذج غيرا لخطي				
للبيانات الأصلية				
1.7165*	1.6777*	1.6777*	1.4864	1
1.736	1.6875	1.6875	1.4818**	2
2.073	1.9834	1.9834	1.7642	3
2.0658	1.9748	1.9748	1.7752	4
2.0681	1.9969	1.9969	1.7766	5
2.0684	1.9961	1.9961	1.7748	6
2.0783	1.9745	1.9745	1.7711	7
2.073	1.9763	1.9763	1.7807	8
2.0688	1.9853	1.9853	1.7719	9
2.0706	1.9878	1.9878	1.757	10

(*) تعني الفشل في تقدير زمن التأخير . (**) تعني النجاح في تقدير زمن التأخير.

جدول (4) :قيم دالة الخسارة لنماذج ARX من الرتبة الثانية وبأزمنة تأخير مختلفة للمكونات الرئيسة لبيانات النموذج غير الخطي (6).

التكرار الرابع	التكرار الثالث	التكرار الثاني	التكرار الأول	زمن التأخير
قيم دالة الخسارة لنماذج ARX للنموذج غيرا لخطي				
للمكونات الرئيسة				
1.7165*	1.6776*	1.5469*	1.4863	1
1.736	1.6874	1.5491	1.4848**	2
2.073	1.9833	1.7857	1.7642	3
2.0658	1.9747	1.7925	1.7753	4
2.0681	1.9968	1.7915	1.7767	5
2.0684	1.996	1.782	1.7748	6
2.0782	1.9746	1.7818	1.7711	7
2.073	1.9766	1.7835	1.7808	8
2.0688	1.9858	1.7834	1.7719	9
2.0706	1.9877	1.7915	1.757	10

(*) تعني الفشل في تقدير زمن التأخير. (**) تعني النجاح في تقدير زمن التأخير.