

اكتشاف وتقدير نقاط التغيير في النظم الحركية باستخدام أسلوب التقطيع

ظافر رمضان مطر البدراني ، نجلاء سعد ابراهيم الشرايبي

قسم الإحصاء والمعلوماتية ، كلية علوم الحاسوب والرياضيات ، جامعة الموصل ، الموصل ، العراق

الملخص

لأنه إن حدوث تغييرات مفاجئة في سلوك النظام الحركي يؤثر بشكل كبير على اختيار النموذج الملائم لتمثيل هذا السلوك ، من هنا تأتي أهمية اكتشاف نقاط التغيير في سلوك النظام الحركي مما يمكن الباحث من تقطيع هذا السلوك إلى مجاميع أو قطاعات بأسلوب التقطيع كل منها منسجم بحد ذاته وهذا يوفر للباحث إمكانية توفيق أنموذج ملائم لكل قطاع أو قسم يمثل البيانات أفضل تمثيل. أن اكتشاف نقاط التغيير وتقديرها تعد من المسائل بالغة الأهمية وفي هذا البحث تستخدم حزمة من المعايير الإحصائية والهندسية في اكتشاف نقاط التغيير وتقديرها، وتم التطبيق باستخدام أسلوب المحاكاة.

المقدمة

$$\begin{aligned} H_0 : \theta_1 = \theta_2 = \dots = \theta_m = \theta \\ H_1 : \theta_1 = \theta_2 = \dots = \theta_{k_1^*} \neq \theta_{k_1^*+1} = \theta_{k_1^*+2} \\ = \dots = \theta_{k_2^*} \neq \dots \neq \theta_{k_{m-1}^*+1} = \theta_{k_{m-1}^*+2} = \dots = \theta_{k_m^*} \dots \dots (1) \\ \neq \theta_{k_m^*} = \dots = \theta_r \end{aligned}$$

أن قيم $(\theta_i ; i = 1, 2, \dots, m)$ التي تمثل معلمات النموذج يمكن أن تكون معلومة أو غير معلومة وبالتالي فإن الهدف الأساسي هو اكتشاف وتحديد المواقع $(k_1^*, k_2^*, \dots, k_m^*)$ وأيضاً تقدير قيم (θ_i) عندما تكون غير معلومة.

(2) نماذج النظم الحركية الخطية: Linear Dynamical System Models

وتسمى أيضاً بنماذج الصندوق الأسود Black box أن هذا النوع من النماذج هو الأكثر شيوعاً وتطبيقاً ، بشكل عام إن نماذج النظم الحركية الخطية هي توسيع لنماذج السلسلة الزمنية وذلك من خلال إضافة متغير الإدخال (x) إليها لان نماذج السلسلة الزمنية عادة تكون غير دقيقة. وتقسّم نماذج النظم الحركية الخطية إلى قسمين: [3] و [4]

1- نماذج خطأ المعادلة Equation Error Model

تتميز هذه النماذج بوجود متعددة الحدود $A(z)$ كمقام حركي لدالة تحويل المدخلات ودالة تحويل التشويش وهذا ينسجم مع حقيقة إن التشويش لا يؤثر على مخرجات النموذج ولكن بدلا من ذلك يدخل النموذج من قبل المرشح $\frac{1}{A(z)}$. بعبارة أخرى أن التشويش يدخل إلى العملية متأخراً ومن ثم فإن ميزات التكرار تتشكل من خلال حركية العملية. ويشمل هذا النوع كلا من نماذج ARX و ARMAX وكما موضح ادناه :

أ) نموذج الانحدار الذاتي بمتغيرات خارجية إضافية Autoregressive with exogenous input model (ARX)

يعتبر نموذج ARX من النماذج الأكثر تطبيقاً لنماذج الأنظمة الحركية الخطية ويعبر عنه بالصيغة التالية: [5]

إن المهمة الرئيسية للتشخيص هي اشتقاق النماذج الرياضية من العمليات والإشارات، ومن النماذج الرياضية المهمة النماذج الخطية Linear Models والعمليات التي لا تتغير مع الزمن المتمثلة بإشارات المدخلات وإشارات المخرجات Input Signals and output signals وكذلك الإشارة التصادفية. يظهر النظام أحياناً تغييرات مفاجئة في سلوكه أثناء الزمن وللكشف عن هذه التغييرات وتحديد مواقعها يستخدم أسلوب التقطيع. تستخدم مسألة التقطيع عندما يكون النظام غير مستقر حيث يتم تقطيع سلسلتي المدخلات والمخرجات إلى فترات زمنية أطلق عليها تسميات عدة منها التقسيمات Partitions أو القطاعات Blocks أو القطع Segments بحيث يصبح النظام مستقر ضمن كل قطاع تقريباً. عموماً إن مسألة التقطيع تهدف للإجابة عن التساؤلات التالية: [1]

- أ- هل توجد تغييرات؟
 - ب- متى حدثت هذه التغييرات؟
 - ت- إذا حدث أكثر من تغيير واحد كيف يمكن تحديد أماكن هذه التغييرات؟
 - ث- عند تقسيم سلسلتي الإدخال والإخراج إلى عدة قطاعات هل سيتغير زمن التأخير ام سيبقى ثابت على طول السلسلتين؟
 - ج- عند تقسيم سلسلتي الإدخال والإخراج في النظم الحركية هل التغييرات تحدث بوقت واحد في كلتا السلسلتين (المدخلات والمخرجات) ام سيؤثر زمن التأخير على موقع نقاط التغيير.
- للإجابة عن الأسئلة أعلاه استخدمت معايير المعلومات الإحصائية لاكتشاف وتقدير وتحديد مواقع نقاط التغيير وذلك عن طريق مقارنة الإحصائيات المحسوبة تحت فرضية انه لا يوجد تغييرات في النظام الحركي. إن الفكرة الأساسية لعملية تقطيع النظم الحركية هي تقليل دالة الخسارة cost function [2].

لنكن (u_t, y_t) مدخلات ومخرجات النظام الحركي تحتوي على (m) من نقاط التغيير في المواقع $(k_1^*, k_2^*, \dots, k_m^*)$ حيث أن

$$(1 \leq k_1^* \leq k_2^* \leq \dots \leq k_m^* \leq T)$$

$$\left. \begin{aligned} e_t &= y_t - \hat{y}_{(t|t-1)} \\ &= y_t - B(z)u_t - (1 - A(z))y_t \\ &= y_t - B(z)u_t - y_t + A(z)y_t \\ e_t &= A(z)y_t - B(z)u_t \end{aligned} \right\} \dots (10)$$

ب (نموذج الانحدار الذاتي والمتوسطات المتحركة بمتغيرات خارجية إضافية:

Autoregressive Moving Average with Exogenous input model (ARMAX)

أن نموذج ARMAX ناتج من دمج نموذج الانحدار الذاتي AR الذي ينص على إن المخرجات الحالية لها علاقة بالقيم السابقة للمخرجات، ونموذج المتوسطات المتحركة MA الذي يستخدم التشويش الأبيض والمدخلات الإضافية X التي تعني أن النموذج لا يعتمد فقط على المدخلات الحالية. ويمكن توضيح هذا النموذج من خلال المعادلة الآتية:-

$$y_t = \frac{B(z)}{A(z)}u_t + \frac{C(z)}{A(z)}e_t$$

$$A(z)y_t = B(z)u_t + C(z)e_t \dots (11)$$

ويتم تقدير المعلمات a_i, b_i, c_i باستخدام طريقة المربعات الصغرى التي تم توضيحها في نموذج ARX. وإن التنبؤ بخطوة واحدة لنموذج ARMAX يكون حسب الصيغة التالية:

$$\hat{y}_{(t|t-1)} = \frac{B(z)}{C(z)}u_t + \left(1 - \frac{A(z)}{C(z)}\right)y_t$$

$$\hat{y}_{(t|t-1)} = B(z)u_t + (C(z) - A(z))y_t + (1 - C(z))\hat{y}_{(t|t-1)} \dots (12)$$

حيث أن الحد الأخير من المعادلة أعلاه يمثل المتنبأ السابق. إن متنبأ ARMAX يكون مستقراً على الرغم من عدم استقرار نموذج ARMAX لعدم استقرار متعددة الحدود $A(q)$ وان متعددة الحدود $C(q)$ يجب أن تكون مستقرة، عليه فان خطأ التنبؤ لهذا النموذج هو:-

$$e_t = y_t - \hat{y}_{(t|t-1)}$$

$$e_t = y_t - \frac{B(z)}{C(z)}u_t - \left(1 - \frac{A(z)}{C(z)}\right)y_t$$

$$e_t = y_t - \frac{B(z)}{C(z)}u_t - y_t + \frac{A(z)}{C(z)}y_t$$

$$e_t = \frac{A(z)}{C(z)}y_t - \frac{B(z)}{C(z)}u_t \dots (13)$$

2. نماذج خطأ المخرجات Output Error Models

وتشمل كل من نموذج خطأ المخرجات Output Error ويرمز له OE ونموذج بوكس-جنكنز Box-Jenkins ويرمز له BJ، ويمكن تمييز نماذج خطأ المخرجات بان نموذج التشويش فيها مستقل عن نموذج العملية المحددة:-

أ (نموذج خطأ المخرجات The Output Error model(OE)

بعد نموذج خطأ المخرجات نموذجاً ذا تركيب واسع الاستخدام ومن السهولة تمثله ضمن نماذج خطأ المخرجات التي غالباً ما تكون أكثر واقعية وأوسع استخداماً بشكل أكثر من نماذج خطأ المعادلة وذلك لان

$$y_t = \frac{B(z)}{A(z)}u_t + \frac{1}{A(z)}e_t$$

$$A(z)y_t = B(z)u_t + e_t \dots (2)$$

ويمكن كتابة المعادلة (2) بشكل أكثر توضيحاً:

$$y_t + a_1y_{t-1} + \dots + a_ny_{t-n} = b_1u_{t-m} + \dots + b_ny_{t-nb+1} + e_t \dots (3)$$

حيث إن (na و nb) هي رتب النموذج و (nk) هو زمن التأخير. ويمكن التنبؤ بقيمة المخرجات التالية بأخذ المشاهدات السابقة أي إن:

$$\hat{y}_t = -a_1y_{t-1} - \dots - a_ny_{t-n} + b_1u_{t-m} + \dots + b_ny_{t-nb+1} + e_t \dots (4)$$

ويمكن كتابة المعادلة أعلاه بشكل أكثر إيجازاً :

$$y_t = \phi'(t)\theta + e_t \dots (5)$$

حيث إن

θ : تمثل متجه معاملات النموذج ويعبر عنها بالشكل

$$\theta = [a_1 \dots a_n \quad b_1 \dots b_n]'$$

$\phi(t)$: تمثل متجه متغيرات الادخال والاخراج ويعبر عنه

$$\phi(t) = [-y_{t-1} \dots -y_{t-n} \quad u_{t-m} \dots u_{t-nb+1}]'$$

وأن تنبؤ المخرجات يمكن وصفه من خلال المعادلة التالية:

$$\hat{y}_{(t|\theta)} = \phi'(t)\theta \dots (6)$$

تدل المعادلة (6) على إن تقدير y_t يعتمد على البيانات السابقة وعلى متجه المعلمات θ علماً بأن قيم θ غير معروفة. ولتكن Z^N يمثل مجموعة من قياسات البيانات يمكن تمثيلها بالشكل الآتي:

$$Z^N = [u_{(-n)}, y_{(-n)}, \dots, u_{(N-1)}, y_{(N-1)}]$$

حيث إن: $n = \max[na, nb + nt + 1]$

وباستخدام طريقة المربعات الصغرى سوف نقدر المتجه $\hat{\theta}$ وذلك بجعل $\hat{y}_{(t|\theta)}$ مطابقاً لـ y_t حيث أن:

$$\hat{\theta} = \min [V(\hat{\theta}, Z^N)]$$

$V(\hat{\theta}, Z^N)$: تمثل دالة الخسارة وهي معادلة تربيعية من متجه المعلمات (θ) يتم تقديرها من خلال المعادلة التالية:

$$V(\hat{\theta}, Z^N) = \frac{1}{N} \sum_{t=0}^{N-1} (y_t - \hat{y}_{(t|\theta)})^2$$

$$V(\hat{\theta}, Z^N) = \frac{1}{N} \sum_{t=0}^{N-1} (y_t - \phi'(t)\theta)^2 \dots (7)$$

وبأخذ المشتقة الأولى للمعادلة أعلاه وتصغيرها نحصل على:

$$0 = \frac{d}{d\theta} V(\hat{\theta}, Z^N) = \frac{2}{N} \sum_{t=0}^{N-1} \phi(t)(y_t - \phi'(t)\theta)$$

$$\sum_{t=0}^{N-1} \phi(t)y_t = \sum_{t=0}^{N-1} \phi(t)\phi'(t)\theta$$

وان متجه المعلمات يمكن اختياره على انه:

$$\hat{\theta} = \left[\sum_{t=0}^{N-1} \phi(t)\phi'(t) \right]^{-1} \sum_{t=0}^{N-1} \phi(t)y_t \dots (8)$$

لنموذج ARX يمكن إيجاده من خلال المعادلة التالية:

$$\hat{y}_{(t|t-1)} = B(z)u_t + (1 - A(z))y_t \dots (9)$$

من خلال إيجاد قيمة المتنبأ في المعادلة (9) يمكن إيجاد خطأ التنبؤ لنموذج ARX كالاتي:

التوالي فإن دالة التغاير التقاطعي Cross-Covariance Function بين $U(t)$ و $Y(t+k)$ ، إذ إن k يمثل التأخر الزمني، تُعرّف على النحو الآتي:

$$C_{uv}(k) = E\{[U(t) - \mu_u]\{Y(t+k) - \mu_y\}\} \dots\dots(20)$$

إذ إن $\mu_u = E\{U(t)\}$ ويمثل معدل سلسلة المدخلات $U(t)$ و $\mu_y = E\{Y(t)\}$ ويمثل معدل سلسلة المخرجات $Y(t)$. أما دالة الارتباط التقاطعي بين المدخلات $U(t)$ والمخرجات $Y(t+k)$ فإن معامل الارتباط بينهما يُعرّف على النحو الآتي:

$$\rho_{u,y}(k) = \frac{C_{uv}(k)}{\sqrt{C_{uu}(0)C_{yy}(0)}} \dots\dots(21)$$

إذ إن $C_{uu}(0) = E\{[U(t) - \mu_u]^2\}$ ويمثل التباين لسلسلة المدخلات $U(t)$ و $C_{yy}(0) = E\{[Y(t) - \mu_y]^2\}$ ويمثل التباين لسلسلة المخرجات $Y(t)$. ويلاحظ أنه إذا كانت $U(t) = Y(t)$ فإن دالة الارتباط التقاطعي تكون مماثلة تماماً لدالة الارتباط الذاتي [6] و [7]. تعد دالة الارتباط التقاطعي من الأدوات المفيدة في تحديد زمن التأخير بين سلسلتين مترابطتين وإن قيمة زمن التأخير k هي التي تقابل أقوى علاقة خطية (موجبة أو سالبة) بين المتغيرين $U(t)$ و $Y(t+k)$ ، يتم إيجادها عن طريق إيجاد قيم دالة الارتباط التقاطعي لقيم مختلفة من k ، فيكون زمن التأخير هو قيمة k التي تقابل أكبر قيمة مطلقة من دالة الارتباط التقاطعي. فلو كان \hat{d} تمثل زمن التأخير، فيمكن تحديده كما يأتي: [8]

$$\hat{d} = \max_{\forall k} |\rho_{u,y}(k)| \dots\dots(22)$$

ب- استخدام نماذج الانحدار الذاتي بمتغيرات إضافية Autoregressive with Exogenous Variables (ARX) هناك أسلوب آخر لتقدير زمن التأخير \hat{d} يستند على نماذج الانحدار الذاتي مع مدخلات إضافية ARX، حيث تقسم سلسلتي المدخلات والمخرجات سوية إلى جزأين: يستخدم الجزء الأول للتقدير بينما يستخدم الجزء الثاني لاختبار شرعية التقدير، ويتم أخذ بيانات الجزء الأول وتقدير نماذج ARX الملائمة عن طريق إعطاء رتب ثابتة لمتعددي حدود المخرجات والمدخلات $A(q)$ و $B(q)$ على التوالي ولتكن $(na = nb = 2)$ مع زمن تأخير يتغير من (10-1) وبعدها يتم إيجاد قيمة دالة الخسارة Loss Function من خلال استخدام بواقي النموذج الملائم بعد تطبيقه على بيانات الشرعية، وبالتالي فإن مقدار زمن التأخير \hat{d} هو تلك القيمة التي تقابل أقل قيمة لدالة الخسارة، أن الأسلوب المذكور أعلاه تم اقتراحه من قبل الباحث Ljung (1999).

ج- استخدام نماذج فضاء الحالة: State Space Models

استخدم [9] هذا الأسلوب لتقدير زمن التأخير من خلال ملائمة العديد من نماذج فضاء الحالة برتب متصاعدة ابتداءً من الرتبة الأولى وحتى الرتبة العاشرة وملاحظة خصائص نموذج فضاء الحالة الملائمة للرتب العشرة كلاً على حدا وأخذ أول رتبة يصل فيها النموذج إلى

نماذج التشويش لا تتضمن المرشح $1/A(q)$ ، ويمكن وصف هذا النموذج بالشكل الآتي:-

$$y_t = \frac{B(z)}{F(z)} u_t + e_t \quad \dots (14)$$

والتنبؤ بخطوة واحدة لنموذج OE يتم تمثيله بالصيغة التالية:

$$\hat{y}_{(t|t-1)} = \frac{B(z)}{F(z)} u_t$$

$$F(z) \hat{y}_{(t|t-1)} = B(z) u_t$$

$$\hat{y}_{(t|t-1)} = B(z) u_t + (1 - F(z)) \hat{y}_{(t|t-1)} \quad \dots (15)$$

إن متنبأ OE يكون غير مستقر إذا كانت متعددة الحدود $F(z)$ غير مستقرة لذا لا يمكن استخدام نماذج خطأ المخرجات لنمذجة العمليات غير المستقرة. وأن خطأ التنبؤ لهذا النموذج هو:-

$$e_t = y_t - \hat{y}_{(t|t-1)}$$

$$e_t = y_t - \frac{B(z)}{F(z)} u_t \quad \dots (16)$$

ب) نموذج بوكس - جنكنز The Box-Jenkins Model

عرف نموذج بوكس -جنكنز من قبل الباحثين Box and Jenkins عام 1970، ويتميز عن نموذج OE بالمرونة من خلال ترشيح التشويش الأبيض بالمرشح ARMA، ويمكن تمثيله من خلال المعادلة الآتية:-

$$y_t = \frac{B(z)}{F(z)} u_t + \frac{C(z)}{D(z)} e_t \quad \dots (17)$$

BJ يتم تمثيله بالصيغة التالية:

$$\hat{y}_{(t|t-1)} = \frac{B(z)D(z)}{F(z)C(z)} u_t + (1 - \frac{D(z)}{C(z)}) y_t \quad \dots (18)$$

وإن خطأ التنبؤ لهذا النموذج يتم تمثيله بالصيغة الآتية:-

$$e_t = y_t - \hat{y}_{(t|t-1)}$$

$$e_t = y_t - \frac{B(z)D(z)}{F(z)C(z)} u_t + (1 - \frac{D(z)}{C(z)}) y_t$$

$$e_t = y_t - \frac{B(z)D(z)}{F(z)C(z)} u_t - y_t + \frac{D(z)}{C(z)} y_t$$

$$e_t = \frac{D(z)}{C(z)} y_t - \frac{B(z)D(z)}{F(z)C(z)} u_t \quad \dots (19)$$

(3) طرق تقدير زمن التأخير: Delay Time Estimation

Methods

قبل البدء بتحديد النموذج المناسب للنظم الحركية يجب أولاً تقدير زمن التأخير لأنه يؤثر بشكل مباشر على عملية تحديد النموذج الذي يصف النظام بصورة دقيقة ويوضح آلية عمل النظام، أن العامل الأكثر أهمية في أداء النظام هو التأخير بين المدخلات والمخرجات الذي يعود إلى طبيعة القصور الذاتي في النظام. حيث توجد عدة طرق وأساليب إحصائية لتقدير زمن التأخير منها:

أ- دالة الارتباط التقاطعي Crosscorrelation Function

وتسمى أيضاً بدالة الارتباط المضاعف وتستخدم لقياس قوة العلاقة الخطية واتجاهها بين عمليتين تصادفيتين. بافتراض إن كل من $U(t)$ و $Y(t)$ يمثلان مدخلات ومخرجات نظام حركي على

لمعظم قيم k . ولتقدير موقع نقطة التغيير k^* بـ \hat{k} حيث ان :

$$AIC(\hat{k}) = \min_{2-k-T} AIC_1(k) \dots (26)$$

لاكتشاف نقاط التغيير المتعددة بمعنى لإيجاد أكثر من نقطة تغيير في السلسلة يتم استخدام خوارزمية التقطيع الثنائي Binary Segmentation وهي طريقة موسعة لمسألة اكتشاف نقطة التغيير الوحيدة k^* . أن التقطيع الثنائي يستند على التقييم المتعاقب للإحصائيات عند أجزاء مختلفة من النظام، لاكتشاف نقاط التغيير وتحديد مواقعها بشكل آني تكون هنالك حاجة إلى اختبار وتحديد نقطة التغيير \hat{k} في كل خطوة ويكرر الاختبار في كل جزء من سلوك النظام الحركي حتى يتم قبول فرضية العدم، يمكن تلخيص الخوارزمية بالخطوتين التاليتين:

1- نحسب قيمة معيار اكاكي للمعلومات AIC للنموذج الذي تم ملائته للنظام الحركي بكامله ثم يتم البحث عن نقطة تغيير مهمة من خلال مشاهدة رسم سلسلتي المدخلات والمخرجات ونحسب إحصائية معيار المعلومات اكاكي للنموذج الذي تم ملائته للنظام عند نقاط التغيير المختارة ثم نختار النقطة التي يكون معيار اكاكي عندها اقل من بقية نقاط التغيير ونختبر المعادلة (24) فإذا رفضت فرضية العدم التي تنص بأنه لا يوجد نقطة تغيير في النظام سوف يتم تقسيم السلسلتين الأصليتين (المدخلات والمخرجات) إلى قسمين حيث النقطة \hat{k} تمثل الفاصل بين القسمين وننتقل إلى الخطوة التالية.

2- في كل قسم يتم إعادة الخطوة الأولى لاكتشاف نقطة التغيير التي قد تحدث بداخله وتستمر العملية حتى الوصول إلى نتيجة بعدم وجود نقطة تغيير إضافية في أي جزء من سلوك النظام الحركي.

(5) تجارب المحاكاة للبيانات المولدة:-

Simulation generator Data Experiments

لقد أجريت تجارب على بيانات مولدة للنماذج الحركية الخطية ، ومن ثم التعامل مع هذه البيانات لتشخيص الأنموذج الأكثر ملائمة وبعدها تم تطبيق عملية التقطيع عليها. تم توليد البيانات على أساس المعادلات أدناه وباستخدام البرنامج الجاهز MATLAB2013:

$$\left. \begin{aligned} y(t) &= a_1 y(t-1) + a_2 y(t-2) + b_1 u(t-1) + b_2 y(t-2) + e(t) \\ a_1(t) &= 0.32 \cos(1.5 - \cos(\frac{4\pi t}{n} + \pi)) & 1 \leq t \leq 256 \\ a_1(t) &= 0.32 \cos(3 - \cos(\frac{4\pi t}{n} + \frac{\pi}{2})) & 257 \leq t \leq 768 \\ a_1(t) &= 0.32 \cos(1.5 - \cos(\frac{4\pi t}{n} + \pi)) & 769 \leq t \leq 1024 \\ a_2(t) &= 0.4 \cos(\frac{4\pi t}{n}) & 1 \leq t \leq 1024 \\ b_1(t) &= 0.65 & 1 \leq t \leq 256 \\ b_1(t) &= 0.5 & 257 \leq t \leq 768 \\ b_1(t) &= 0.65 & 769 \leq t \leq 1024 \\ b_2(t) &= 0.6 & 1 \leq t \leq 1024 \end{aligned} \right\} \dots (27)$$

إن $n=1024$ وتمثل طول السلسلة و $\pi=180$ وقد تم توفير النماذج الاربعة الخطية لها وإجراء عملية التقطيع عليها حيث تم رسم سلسلتي المدخلات والمخرجات لمشاهدات النظام مع الزمن وهي خطوة أساسية

حالة الاستقرار، على سبيل المثال لو افترضنا بان أول رتبة تتحقق فيها الاستقرار هي عند الرتبة الرابعة وكذلك كانت ذات ملائمة جيدة عندئذ ننتقل إلى الخطوة التالية بان نوجد معيار اكاكي للمعلومات (AIC) للراتب من الرتبة الأولى إلى الرتبة الرابعة واخذ الرتبة التي تقابل اقل قيمة لمعيار (AIC) فإذا كانت اقل قيمة لمعيار (AIC) عند الرتبة الرابعة نوجد دالة الاستجابة النبضية (التي توصف بأنها مؤشراً جيداً لاستقرارية النموذج) عند هذه الرتبة، فان أول قيمة معنوية تقع خارج فترة الثقة ، تمثل زمن التأخير \hat{h} .

(4) اكتشاف وتقدير نقاط التغيير :-

Detecting and Estimation change-points

في هذا البحث اهتمامنا يتركز بنمذجة النظم الحركية غير المستقرة وذلك بتقطيع سلسلتي المدخلات والمخرجات إلى قطاعات (أو قطع) بحيث تكون السلسلة ضمن كل قطاع (قطعة) تقريباً مستقرة. إذ يمكن تعريف عملية التقطيع بأنها تقسيم السلسلة إلى قطع غير متداخلة وتحتوي كل قطعة على تمثيل بسيط للبيانات بحيث تكون البيانات داخل كل قطعة مستقرة تقريباً [10]. إن أماكن التغييرات في المعلمات التي تميز النموذج تعد مسألة بالغة الأهمية. تم استخدام نظرية المعلومات لاكتشاف وتقدير نقاط التغيير، فقد استخدمت معايير المعلومات ومن أبرزها معيار اكاكي للمعلومات (AIC) في هذا المجال.

معيار اكاكي للمعلومات:-

Akaike's Information Criterion

قدم معيار اكاكي في عام 1973 لاختيار الأنموذج الأفضل في تمثيل البيانات. إن هذا المعيار له تطبيقات عدة حيث طبق في السلاسل الزمنية وكشف التطورات وتحليل الانحدار. وتحسب قيمة AIC كما يأتي:

$$AIC = -2 \log L(\hat{\theta}) + 2P \dots (23)$$

حيث أن:

$$L(\hat{\theta}) : \text{تمثل نسبة الإمكان الأعظم للأنموذج.}$$

P: عدد معلمات النموذج المقدر.

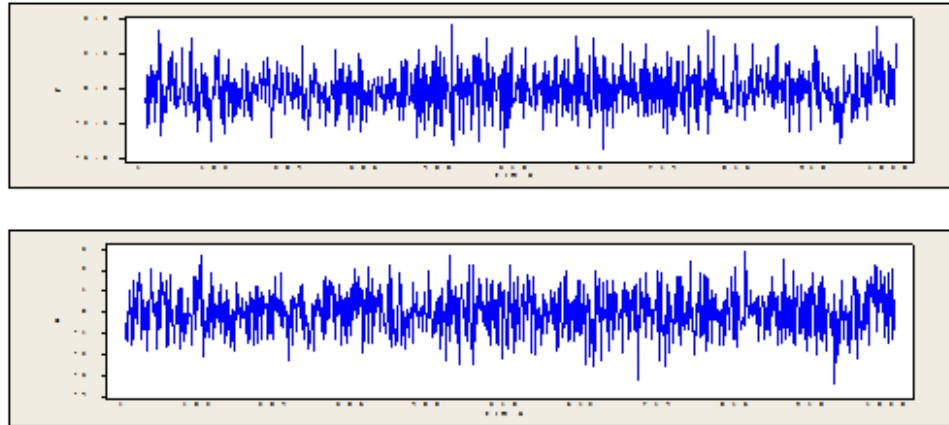
إن الأنموذج الملائم هو الذي يقابل اقل قيمة لـ AIC . لتكن $AIC_0(T)$ هي قيمة معيار اكاكي المحسوب تحت فرضية العدم H_0 الموضحة في المعادلة (1) التي تفترض بأنه لا توجد نقطة تغيير على طول النظام الحركي ، ولتكن $AIC_1(k)$ المحسوبة عند افتراض بأن هناك تغيير عند النقطة $(t=k)$ حيث ان $k=1,2,\dots,T$. حيث يتم رفض او قبول (H_0) بالاستناد على مبدأ تقليل معيار اكاكي للمعلومات. حيث يتم قبول فرضية العدم (H_0) إذا كانت:

$$AIC_0(T) < \min_k AIC_1(k) \dots (24)$$

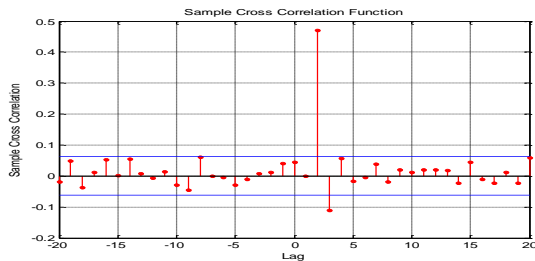
هذا يعني أن AIC التي حسبت بافتراض أنه لا يوجد تغييرات اصغر من AIC التي حسبت بافتراض وجود نقطة تغيير عند قيم k . وبالعكس ترفض فرضية العدم (H_0) عندما

$$AIC_0(T) > \min_k AIC_1(k) \dots (25)$$

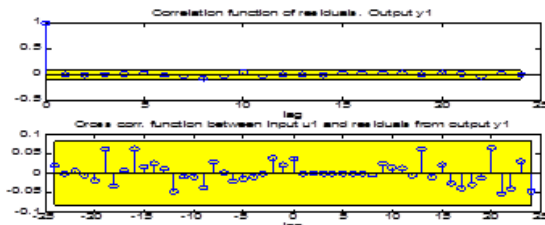
وهامة في التحليل، إذ يمكن من خلال الرسم الاطلاع على طبيعة التذبذب فيها وملاحظة فيما إذا كان النظام يحوي على نقاط تغيير.



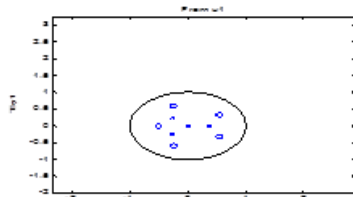
الشكل (1): الرسم الزمني لسلسلتي المدخلات والمخرجات المولدتين



الشكل (2): يوضح دالة الارتباط التقاطعي بين سلسلتي المدخلات والمخرجات.



دالة الارتباط الذاتي للأخطاء ودالة الارتباط التقاطعي لأفضل نموذج



دائرة الوحدة لأفضل نموذج

الشكل (3): يمثل بعض النتائج لأفضل نموذج وهو ARX(3,7,1)

نلاحظ من خلال الشكل أعلاه بأن هناك نقاط تغيير في سلسلتي المدخلات والمخرجات تقريباً عند المدى الزمني (285-400) وكذلك (700-800) لذا تم تطبيق معيار اكاكي للمعلومات لإيجاد نقاط التغيير في سلسلتي المدخلات والمخرجات وحسب الخطوات التالية: أولاً: إيجاد النموذج الملائم للبيانات لكي تتم المقارنة على أساسه. لقد تم إيجاد قيمة زمن التأخير باستخدام دالة الارتباط التقاطعي كما في الشكل (2) وكانت قيمة $\hat{d} = 2$ علماً إن البيانات ولدت بزمن تأخير $\hat{d} = 1$ من أنموذج خطي بمعلمات غير خطية وهذه النتيجة تدل على أن دالة الارتباط التقاطعي الخطية لا تعمل جيداً في هذه الحالة لذلك تم أيضاً استخدام نماذج الانحدار الذاتي بمتغيرات إضافية مع $\hat{d} = 1$ ، وعندما استخدمت نماذج فضاء الحالة لم تظهر قيمة زمن التأخير لان النظام المولد غير مستقر. ومن ملاحظة الجدول (1) الذي يمثل النتائج النهائية للنماذج التي تم اختيارها للبيانات المولدة ان افضل نموذج يحقق الشروط الاحصائية هو نموذج ARX بالرتب (3,7,1) حيث $AIC_0(T = 1024) = 0.4058$ والشكل (3) يوضح عدم معنوية كل من بواقي المخرجات ودالة الارتباط التقاطعي بين المدخلات وبواقي المخرجات ودائرة الوحدة الخاصة بنموذج المدخلات التي تدل على عدم استقرارية النموذج المختار وهي نتيجة متوقعة لنظام حركي يتضمن نقاط تغيير في سلوكه.

جدول(1): النماذج النهائية المختارة

Model	Aic	Loss	Fpe	Resid	Cross	Fitting
ARX(3,7,1)	0.4058	1.4962	1.5002	Random	Uncorrelated	11.84
ARMAX(1,2,3,1)	0.4131	1.4934	1.5114	Random	Uncorrelated	3.469
OE	لا يوجد نموذج مناسب إحصائياً					
BJ(10,1,2,9,1)	0.4138	1.4454	1.5111	Random	Uncorrelated	23.53

للمعلومات للنظام حتى نقطة التغيير المعينة. الجدول (2) يوضح النماذج النهائية المختارة لعدة نقاط تغيير ويبين قيم حزمة معايير المعلومات لكل رتبة و الشروط الإحصائية ونسبة الملائمة لكل نموذج.

ثانياً: بأخذ عدة نقاط ضمن الفترتين التي تم تحديدهم من الشكل (1) وإيجاد النماذج المناسبة لهذه النقاط واختيار النقطة التي تمتلك أقل قيمة لمعيار اكاكي ليتم عن طريقها اختبار فرضية العدم التي تنص على انه لا يوجد نقطة تغيير في النظام وذلك عن طريق مقارنة قيمة معيار اكاكي للمعلومات لكامل النظام الحركي مع قيمة معيار اكاكي

جدول (2): النماذج النهائية المختارة لقيم مختلفة من نقاط التغيير واختيار افضل نموذج بالاعتماد على حزمة المعايير.

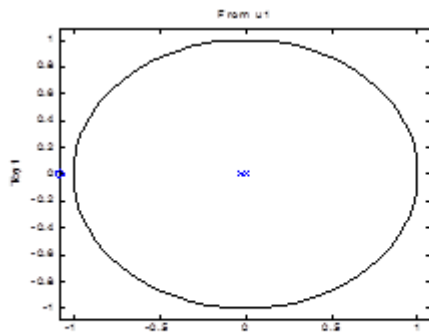
Value \hat{k}	na	nb	Nk	Aic	Loss	Fpe	Resid	Cross	Fitting	Unit Circle
1-291	1	2	1	0.0896	1.0901	1.0935	Random	Uncorrelated	17.34	Stable
1-300	1	2	1	0.0962	1.0974	1.1008	Random	Uncorrelated	15.59	Stable
1-322	1	2	1	0.1416	1.1485	1.1519	Random	Uncorrelated	-39.05	Stable
1-352	1	2	1	0.1841	1.1986	1.2019	Random	Uncorrelated	7.207	Stable
1-366	1	2	1	0.1653	1.1796	1.1765	Random	Uncorrelated	1.364	Stable
1-372	1	2	1	0.1674	1.1790	1.1820	Random	Uncorrelated	3.85	Stable
1-398	1	2	1	0.2374	1.2647	1.2677	Random	Uncorrelated	13.22	Stable
1-700	1	2	1	0.4022	1.4929	1.4951	Random	Uncorrelated	28.89	Stable
1-710	1	2	1	0.4020	1.4947	1.4925	Random	Uncorrelated	21.81	Stable
1-720	1	2	1	0.3968	1.4849	1.4870	Random	Uncorrelated	24.91	Stable
1-740	1	2	1	0.3963	1.4841	1.4862	Random	Uncorrelated	26.18	Stable
1-750	1	2	1	0.3859	1.4688	1.4708	Random	Uncorrelated	27.22	Stable
1-760	1	2	1	0.3828	1.4643	1.4663	Random	Uncorrelated	17.84	Stable
1-768	1	2	1	0.3738	1.4512	1.4531	Random	Uncorrelated	16.76	Stable
1-775	1	1	1	0.3732	1.4503	1.4523	Random	Uncorrelated	14.52	Stable
1-780	1	2	1	0.3687	1.4439	1.4458	Random	Uncorrelated	14.48	Stable
1-790	1	2	1	0.3767	1.4555	1.4574	Random	Uncorrelated	14.43	Stable
1-800	1	2	1	0.3841	1.4663	1.4682	Random	Uncorrelated	13.26	Stable

ثالثاً: بعد ان قسمت السلسلة إلى قسمين سيتم إجراء نفس العملية السابقة على كل قسم: بالنسبة للقسم الأول فإن $AIC_0(T=291) = 0.0896$ والشكل (4) يوضح دائرة الوحدة وعدم معنوية كل من بواقي المخرجات ودالة الارتباط التقاطعي بين المدخلات وبواقي المخرجات الخاصين بنموذج $ARX(1,2,1)$ لبيانات القسم الأول البالغ عددها (291). نلاحظ بان الانموذج الخاص بالبيانات القسم الأول مستقراً وعلى الرغم من ذلك تم أخذ نقاط عشوائية من السلسلة لمقارنتها، والجدول (3) بين خلاصة مقارنات للنماذج الاربعة عند قيم مختلفة لقيم \hat{k} لبيانات القسم الأول:

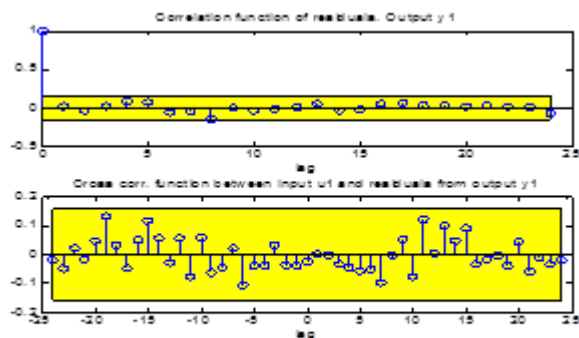
ملاحظة الجدول أعلاه نجد أن أقل قيمة لمعيار المعلومات اكاكي هي عند قيمة $\hat{k} = 291$ المقابلة لنموذج $ARX(1,2,1)$ وعند المقارنة نجد

$$AIC_0(T=1024) = 0.4058 > AIC_1(k=291) = 0.0896$$

لذا سوف يتم رفض فرضية العدم وقبول الفرضية البديلة التي تنص على أنه هناك نقطة تغيير عند $\hat{k} = 291$ وعند هذه النقطة سوف تنقسم السلسلتين (المدخلات والمخرجات) إلى قسمين وبذلك ننقل إلى الخطوة الثالثة.



ARX(1,2,1) دائرة الوحدة لنموذج



ARX(1,2,1) الاخطاء ودالة الارتباط التقاطعي لنموذج

الشكل (4): يمثل بعض المعايير لنموذج $ARX(1,2,1)$ لبيانات القسم الأول

جدول (3): النماذج النهائية المختارة لقيم مختلفة من نقاط التغيير واختيار افضل نموذج بالاعتماد على حزمة المعايير لبيانات

القسم الأول (1-291)

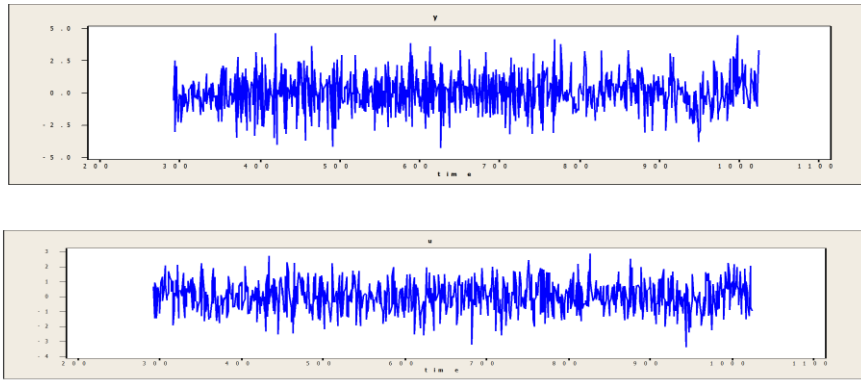
Value \hat{k}	Best model	Aic	Loss	Fpe	Resid	Cross	Fitting	Unit Circle
1-100	armax(1,1,8,1)	0.1368	0.9135	1.1211	Random	Uncorrelated	19.16	Stable
1-150	oe(1,1,1)	0.1908	1.1614	1.2092	Random	Uncorrelated	15.36	Stable
1-160	armax(1,1,8,1)	0.1344	0.9504	1.1265	Random	Uncorrelated	6.896	Stable
1-170	armax(1,2,1,1)	0.0941	1.0216	1.0959	Random	Uncorrelated	11.97	Stable

الارتباط التقاطعي بين المدخلات وبواقي المخرجات ولقد تم إيجاد قيمة زمن التأخير باستخدام دالة الارتباط التقاطعي كما في الشكل (6) وكانت قيمة $\hat{d} = 1$. الجدول (4) يوضح النتائج النهائية للنماذج التي تم اختيارها للبيانات القسم الثاني نلاحظ ان افضل نموذج يحقق الشروط الاحصائية هو نموذج ARX بالرتب (3,7,1). حيث $AIC_0(T = 733) = 0.3834$ والشكل (7) يوضح دائرة الوحدة وعدم معنوية كل من بواقي المخرجات ودالة الارتباط التقاطعي بين المدخلات وبواقي المخرجات الخاصين بنموذج ARX(3,7,1) لبيانات القسم الثاني.

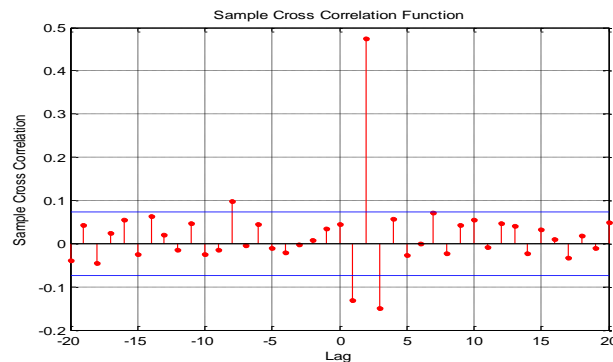
بملاحظة الجدول أعلاه نجد ان أقل قيمة لمعيار المعلومات اكاكي هي عند قيمة $\hat{k} = 170$ المقابلة لنموذج ARMAX(1,2,1,1)، وعند المقارنة نجد

$$AIC_0(T = 291) = 0.0896 < AIC_1(k = 170) = 0.0941$$

وهذا يقودنا الى قبول فرضية العدم التي تنص بأنه لا توجد نقطة تغيير في السلسلة وبالتالي التوقف عن تقسيم بيانات القسم الأول والانتقال إلى بيانات القسم الثاني. الشكل (5) يبين رسم سلسلة بيانات القسم الثاني (1024-292) بعدد يبلغ (733) قيمة لكل من سلسلتي الإدخال والإخراج ولقد تم إيجاد النموذج المناسب لهذه البيانات بالاعتماد على أقل قيم من معايير المعلومات والأخطاء العشوائية وعدم معنوية



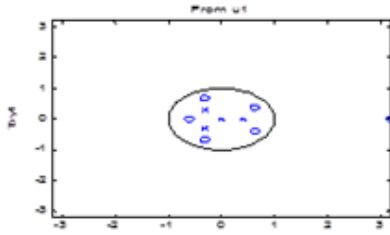
الشكل (5): يمثل الرسم الزمني لسلسلتي المدخلات والمخرجات لبيانات القسم الثاني



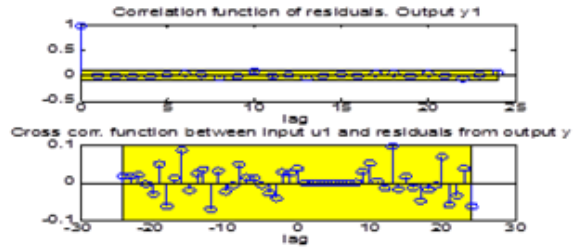
الشكل (6): يمثل دالة الارتباط التقاطعي لسلسلتي المدخلات والمخرجات لبيانات القسم الثاني البالغ عددهم (733).

جدول(4): النماذج النهائية المختارة لبيانات القسم الثاني البالغ عددها (733)

model	Aic	Loss	Fpe	Resid	Cross	Fitting
ARX(3,7,1)	0.3834	1.4260	1.4667	Random	Uncorrelated	19.07
ARMAX(1,2,3,1)	0.3933	1.4565	1.4817	Random	Uncorrelated	3.333
OE	لا يوجد نموذج مناسب إحصائياً					
BJ(2,1,2,8,1)	0.4100	1.4503	1.5057	Random	Uncorrelated	16.16



دائرة الوحدة لنموذج ARX(3,7,1)



ARX(3,7,1) الأخطاء ودالة الارتباط التقاطعي لنموذج

الشكل (7): يمثل بعض المعايير لأفضل نموذج وهو ARX(3,7,1) للقسم الثاني.

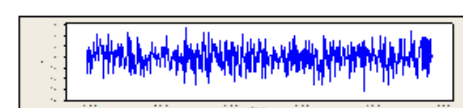
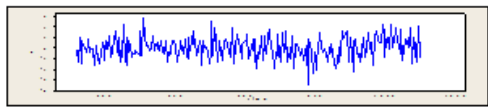
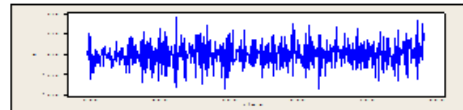
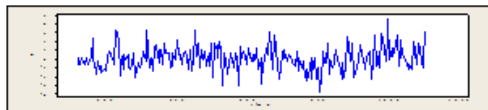
نلاحظ بيان النموذج الملائم لبيانات القسم الثاني لسلسلي المدخلات والمخرجات البالغ عددهم (733) كان غير مستقر مما يدل على أن هناك نقطة تغيير ضمن بيانات القسم الثاني، وكذلك من خلال مشاهدة الشكل (5) الذي يوضح الرسم الزمني لسلسلي المدخلات والمخرجات لبيانات القسم الثاني نلاحظ بأن هناك نقطة تغيير ضمن المدى الزمني

جدول (5): النماذج النهائية المختارة لقيم مختلفة من نقاط التغيير واختيار أفضل نموذج بالاعتماد على معيار اكاكي للمعلومات لبيانات القسم الثاني (1024-292).

قيم \hat{k} ضمن بيانات القسم الثاني	وتعادل من القيم الأصلية	Best Model	Aic	Loss	Fpe	Resid	Cross	Fitting	Unit Circle
1-361	292-652	armax(1,2,2,1)	0.2547	1.2520	1.2895	Random	Uncorrelated	33	Stable
1-373	292-664	arx(1, 2,1)	0.2430	1.2520	1.2748	Random	Uncorrelated	31.3	Stable
1-393	292-684	arx(1, 2,1)	0.2091	1.2127	1.2324	Random	Uncorrelated	42.31	Stable
1-404	292-695	arx(1, 2,1)	0.2101	1.2145	1.2337	Random	Uncorrelated	37.74	Stable
1-418	292-709	arx(1, 2,1)	0.1948	1.1959	1.2143	Random	Uncorrelated	31.34	Stable
1-424	292-715	arx(1, 2,1)	0.1878	1.1887	1.2065	Random	Uncorrelated	35.32	Stable
1-431	292-722	arx(1, 2,1)	0.1704	1.1685	1.1856	Random	Uncorrelated	29.81	Stable
1-450	292-741	arx(1, 2,1)	0.1718	1.1706	1.1873	Random	Uncorrelated	43.55	Stable
1-478	292-769	arx(1, 2,1)	0.1590	1.1566	1.1723	Random	Uncorrelated	32.15	Stable
1-487	292-778	arx(1, 2,1)	0.1537	1.1504	1.1660	Random	Uncorrelated	21.31	Stable

بملاحظة الجدول أعلاه نجد ان أقل قيمة لمعيار اكاكي للمعلومات هي عند قيمة $\hat{k} = 487$ المقابلة لنموذج ARX(1,2,1) (بمعنى $\hat{k} = 778$)، وعند المقارنة نجد $AIC_0 (T = 733) = 0.3834 > AIC_1 (k = 487) = 0.1537$ لذا سوف يتم رفض فرضية عدم وقبول الفرضية البديلة التي تنص على أنه هناك نقطة تغيير وهي $\hat{k} = 487$ عند هذه النقطة سوف

تتقسم السلسلة (أي سلسلة بيانات القسم الثاني) إلى قسمين، القسم الأول يتضمن البيانات من (1- 487) أي (292-778) والقسم الثاني يتضمن البيانات من (488-733) أي (779-1024) والشكل (8) يوضح رسم سلسلي المدخلات والمخرجات لبيانات القسمين بعد تجزئة القسم الثاني.

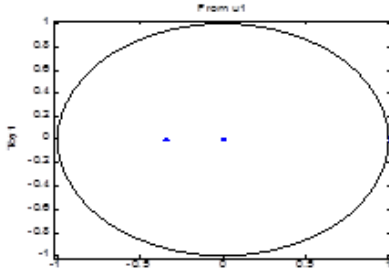


بيانات القسم الثاني (488-733) أي (779-1024)

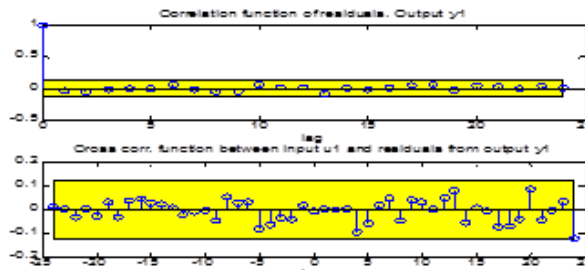
بيانات القسم الأول (1-487) أي (292-778)

الشكل (8): يمثل الرسم الزمني لسلسلي المدخلات والمخرجات بعد تجزئة القسم الثاني إلى قسمين

والمخرجات تعتبر $AIC_0(T = 487) = 0.1537$ والشكل (9) يوضح دائرة الوحدة وعدم معنوية كل من بواقي المخرجات ودالة الارتباط التقاطعي بين المدخلات وبواقي المخرجات الخاصة بنموذج $ARX(1,2,1)$ لبيانات هذا القسم. إن الجدول (6) بين قيم مختلفة لقيم \hat{k} لبيانات هذا القسم.



ARX(1,2,1) دائرة الوحدة لنموذج



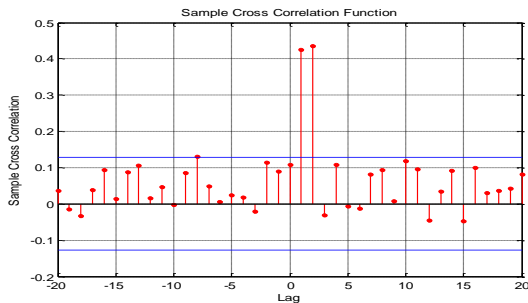
ARX(1,2,1) الأخطاء ودالة الارتباط التقاطعي لنموذج

الشكل (9): يمثل بعض النتائج لأفضل نموذج وهو $ARX(1,2,1)$ لبيانات القسم الأول البالغ عددهم (487) بعد تجزئة القسم الثاني.

جدول (6): النماذج النهائية المختارة لقيم مختلفة من نقاط التغيير واختيار أفضل نموذج بالاعتماد على حزمة من المعايير لبيانات القسم الأول (1-487) أي من (292-778) بعد تجزئة القسم الثاني لسلسلتي المدخلات والمخرجات

قيم \hat{k} ضمن بيانات القسم الأول بعد التجزئة	تعادل من القيم الأصلية	Best Model	Aic	Loss	Fpe	Resid	Cross	Fitting	Unit Circle
1-130	292-421	arx(2,2,1)	0.2133	1.1517	1.2347	Random	Uncorrelated	25.45	Stable
1-140	292-431	arx(1,2,1)	0.3074	1.2919	1.3582	Random	Uncorrelated	22.73	Stable
1-170	292-461	arx(2,2,1)	0.3451	1.3480	1.4107	Random	Uncorrelated	35.69	Stable
1-220	292-511	arx(2,2,1)	0.2694	1.2552	1.3080	Random	Uncorrelated	28.29	Stable
1-250	292-541	arx(2,2,1)	0.2785	1.2681	1.3201	Random	Uncorrelated	20.08	Stable
1-270	292-561	arx(1,2,1)	0.2529	1.2574	1.2874	Random	Uncorrelated	18.89	Stable
1-300	292-591	arx(1,2,1)	0.2646	1.2743	1.3026	Random	Uncorrelated	28.42	Stable
1-340	292-631	arx(2,2,1)	0.2517	1.2523	1.2857	Random	Uncorrelated	28.53	Stable

ARX بالرتب (2,2,1). حيث $AIC_0(T = 246) = 0.0400$. كذلك يوضح الشكل (11) استقرارية دائرة الوحدة وعدم معنوية كل من بواقي المخرجات ودالة الارتباط التقاطعي بين المدخلات وبواقي المخرجات الخاصة بنموذج $ARX(2,2,1)$ لبيانات القسم الثاني بعد تجزئة القسم الثاني. وإن الجدول (8) يبين قيم مختلفة لقيم \hat{k} لبيانات القسم الثاني بعد التجزئة.



الشكل (10): يوضح دالة الارتباط التقاطعي بين سلسلتي المدخلات والمخرجات لبيانات القسم الثاني بعد تجزئة القسم الثاني البالغ عددها (246)

من خلال مشاهدة الشكل (8) الذي يوضح رسم بيانات القسم الأول والثاني الناتجين من تجزئة القسم الثاني لسلسلتي المدخلات والمخرجات على التوالي نجد بأن السلسلتين تقريباً مستقرتين مما يدل بان ليس هنالك نقاط تغيير ضمن القسمين، وعلى الرغم من ذلك لقد تم اخذ نقاط عشوائية من كلتا السلسلتين لمقارنتها. بالنسبة لبيانات القسم الأول الناتجة عن تجزئة القسم الثاني لسلسلتي المدخلات

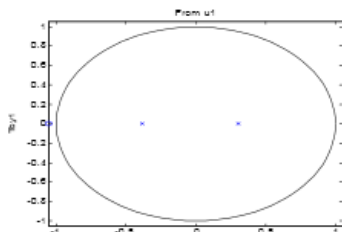
بملاحظة الجدول أعلاه نجد ان أقل قيمة لمعيار اكاكي للمعلومات هي عند قيمة $\hat{k} = 130$ المقابلة لنموذج $ARX(2,2,1)$ (أي بمعنى $\hat{k} = 421$ ، وعند المقارنة نجد أن

$$AIC_0(T = 487) = 0.1535 < AIC_1(k = 130) = 0.2133$$

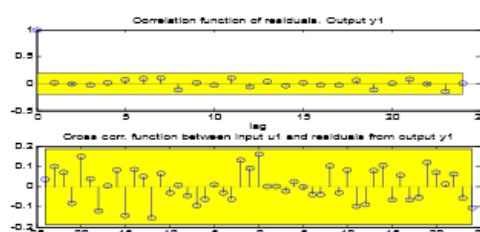
أي قبول فرضية عدم التنص بأنه لا توجد نقطة تغيير في السلسلة (1-487) أي (292-778) وبعد قبول فرضية عدم التنص يتم التوقف عن تقسيم بيانات القسم الأول والانتقال إلى بيانات القسم الثاني بعد تجزئة القسم الثاني لسلسلتي الإدخال والإخراج التي تبدأ من (488-733) أي (779-1024) بما يكافئ (246) قيمة لسلسلتي الإدخال والإخراج ولقد تم إيجاد النموذج المناسب لهذه البيانات بالاعتماد على أقل قيم من معايير المعلومات والأخطاء العشوائية وعدم معنوية الارتباط التقاطعي بين المدخلات وبواقي المخرجات ولقد تم إيجاد قيمة زمن التأخير باستخدام دالة الارتباط التقاطعي كما في الشكل (10) وكانت قيمة $d = 1$. ومن ملاحظة الجدول (7) الذي يوضح النتائج النهائية للنماذج التي تم اختيارها لبيانات القسم الثاني بعد تجزئة القسم الثاني ان أفضل نموذج يحقق الشروط الاحصائية والهندسية هو نموذج

جدول (7): النماذج النهائية المختارة لبيانات القسم الثاني بعد تجزئة القسم الثاني البالغ عددها (246).

Model	Aic	Loss	Fpe	Resid	Cross	Fitting
ARX(2,2,1)	0.0400	0.9921	1.0396	Random	Uncorrelated	31.85
ARMAX(1,2,2,1)	0.0684	1.0063	1.0688	Random	Uncorrelated	31.69
OE(4,5,1)	0.0938	0.9779	1.0914	Random	Uncorrelated	31.11
BJ(2,1,1,2,1)	0.0777	1.0028	1.0780	Random	Uncorrelated	32



دائرة الوحدة لتمودج ARX(2,2,1)



الأخطاء ودالة الارتباط التقاطعي لتمودج ARX(2,2,1)

الشكل (11): يمثل بعض المعايير لأفضل نموذج وهو ARX(2,2,1) لبيانات القسم الثاني بعد تجزئة القسم الثاني لسلسلتي الإدخال والخراج البالغ عددها (246)

جدول (8): النماذج النهائية المختارة لقيم مختلفة من نقاط التغيير واختيار أفضل نموذج بالاعتماد على حزمة من المعايير لبيانات القسم الثاني (488-733) أي بمعنى (779-1024) بعد تجزئة القسم الثاني لسلسلتي المدخلات والمخرجات.

قيم \hat{k} ضمن بيانات القسم الثاني بعد التجزئة	وتعادل من القيم الأصلية	Best Model	Aic	Loss	Fpe	Resid	Cross	Fitting	Unit Circle
1-140	779-918	arx(2,2,1)	0.0838	1.0205	1.0853	Random	Uncorrelated	35.97	Stable
1-150	779-928	arx(1,2,1)	0.0619	1.0159	1.0628	Random	Uncorrelated	36.28	Stable
1-170	779-948	arx(1,2,1)	0.0760	1.0337	1.0780	Random	Uncorrelated	10.28	Stable
1-180	779-958	arx(1,2,1)	0.0635	1.0224	1.0647	Random	Uncorrelated	15.11	Stable
1-200	779-978	arx(1,2,1)	0.0523	1.0124	1.0529	Random	Uncorrelated	20.59	Stable

لدينا فقط نقطتي تغيير وهما ($\hat{k}_1 = 291$, $\hat{k}_2 = 778$) ان اهم ما يمكن ملاحظته هو الحصول على نقطتي تغيير وان افضل النماذج التي تم اختيارها للأقسام الثلاثة التي تفصلها نقطتي التغيير هو نموذج ARX وهذا يعني انه تم التوصل الى نموذج نوع عمليته يوافق النموذج الذي تم توليد البيانات من خلاله على الرغم من ان نقاط التغيير لاتطابق النقاط التغيير ي عملية التوليد، الجدول (9) يبين النموذج المناسب للسلسلة ما قبل التقطيع والنماذج الملائمة للأقسام الثلاثة.

ملاحظة الجدول أعلاه نجد ان أقل قيمة لمعيار اكاكي للمعلومات هي عند قيمة $\hat{k} = 200$ المقابلة لنموذج ARX(1,2,1) أي عند النقطة ($\hat{k} = 978$)، وعند المقارنة نجد أن

$$AIC_0(T = 246) = 0.0400 < AIC_1(k = 200) = 0.0523$$

أي قبول فرضية عدم التنص بأنه لا توجد نقطة تغيير في السلسلة (779-1024) وبذلك تم التوقف عن تقسيم بيانات القسم الثاني بعد تجربة بيانات القسم الثاني. بعد ان تم فحص كافة الخيارات الممكنة ووجد بأنه لا توجد نقاط تغيير سنتوقف عملية التقطيع وبذلك اصبح

جدول (9): يوضح نماذج ARX المختارة للأقسام الثلاثة التي تفصلها نقطتي التغيير ($\hat{k}_1 = 291$, $\hat{k}_2 = 778$) وكذلك النموذج ما قبل التقطيع.

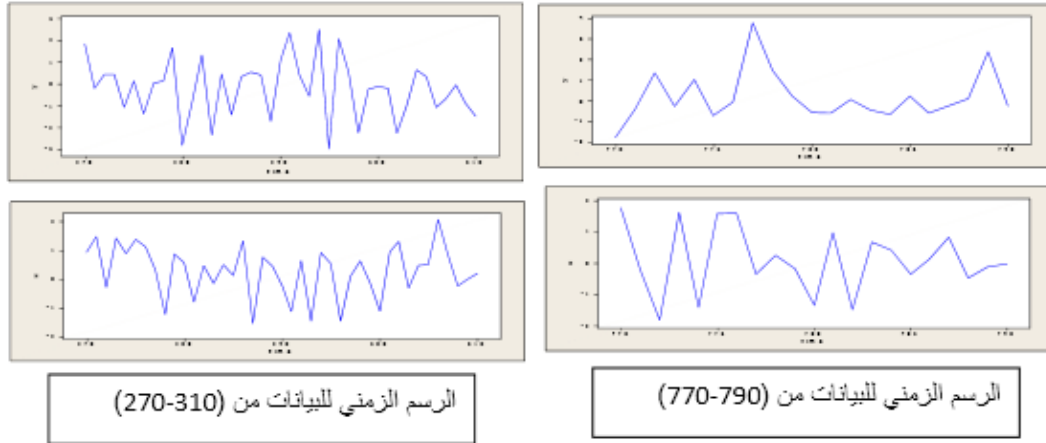
Value \hat{k}	na	nb	nk	Aic	Loss	Fpe	Resid	Cross	Fitting	Unit Circle
1-1024	3	7	1	0.4058	1.4962	1.5002	Random	Uncorrelated	11.84	Stable
1-291	1	2	1	0.0896	1.0901	1.0935	Random	Uncorrelated	17.34	Stable
291-778	1	2	1	0.1537	1.1504	1.1660	Random	Uncorrelated	21.31	Stable
779-1024	2	2	1	0.0400	0.9921	1.0396	Random	Uncorrelated	31.85	Stable

وكذلك كانت نسبة المطابقة جيدة بعد عملية التقطيع وهذه النتيجة تشكل نجاحاً جيداً لعملية التقطيع. وكذلك نلاحظ من خلال الجدول أعلاه بأن زمن التأخير ثابت للقطاعات الثلاثة بمعنى ان عملية تقطيع السلسلة إلى قطاعات لم تؤثر على قيمة زمن التأخير الشكل (2-15)

من خلال ملاحظة الجدول أعلاه الذي يوضح النماذج الملائمة للأقسام الثلاثة لسلسلتي المدخلات والمخرجات نجد ان دالة الخسارة وقيمة معيار اكاكي للمعلومات ومعيار خطأ التنبؤ النهائي لاكاكي للقطاعات الثلاثة (أي بعد عملية التقطيع) أقل منها قبل عملية التقطيع

التغيير. وكذلك نلاحظ من الرسم الزمني للبيانات (770-790) لسلسلتي المدخلات والمخرجات بان سلسلة المدخلات كانت مستقرة لغاية النقطة (777) وعند النقطة (778) فما فوق بدأت تأخذ شكلاً آخر للاستقرارية بينما سلسلة المخرجات نجد بان السلسلة كانت مستقرة لغاية النقطة (778) بدأت تأخذ شكلاً آخر للاستقرارية عند النقطة (779) فما فوق وهذا شيء متوقع لان قيمة زمن التأخير واحد.

يوضح الرسم الزمني لتقطتي التغيير بحيث نلاحظ من الرسم الزمني للبيانات (270-310) بان سلسلة المدخلات كانت مستقرة لغاية النقطة (290) وعند النقطة (291) فما فوق بدأت تأخذ شكلاً آخر للاستقرارية بينما لو نظرنا الى سلسلة المخرجات نلاحظ بان السلسلة كانت مستقرة لغاية النقطة (291) بدأت تأخذ شكلاً آخر للاستقرارية عند النقطة (292) فما فوق وهذا يعني عدم تأثير زمن التأخير بوجود نقاط



الشكل (2-12): يوضح الرسم الزمني للبيانات (770-790) وللبيانات (270-310).

التقاطع بين المدخلات وبيانات مخرجات النظام ونسبة مطابقة البيانات الأصلية مع النموذج المقدر، فضلاً عن معايير المعلومات، ومن الناحية الهندسية تحقق شرط استقرارية النموذج وهذه تعد المرحلة الثانية والأخيرة في وصولنا إلى حل مشكلة عدم الاستقرارية في النظم الحركية الخطية خارج الخط.

التوصيات

على ضوء الاستنتاجات التي تم الحصول عليها نوصي بما يأتي:

1- تطبيق عملية التقطيع على النماذج الحركية التصادفية الخطية متعددة المدخلات ومتعددة المخرجات والبحث في إمكانية التوصل إلى نتائج جيدة.

2- استخدام الشبكات العصبية في عملية تشخيص النظم الحركية الخطية مع عملية التقطيع للتشخيص خارج الخط .

الاستنتاجات

1- الاعتماد على حزمة من المعايير الإحصائية والهندسية مهد الطريق إلى النجاح في تشخيص نموذج تتفق عملياته مع عملية النموذج الذي تم توليد البيانات منه الا أن النموذج الذي حصلنا عليه كان غير مستقراً وهذه النتيجة تتفق مع وجود نقاط تغيير في سلوك النظام الحركي وهذه تعتبر المرحلة الأولى في التشخيص خارج الخط.

2- إن عملية التقطيع التي تم اقتراحها قسمت النظام غير المستقر الى ثلاث قطاعات بنماذج مستقرة نوع عمليات كل منها يوافق النماذج الذي تم توليد البيانات من خلاله.

3- الحصول على النماذج من خلال الاعتماد على حزمة من المعايير الإحصائية والهندسية مجتمعة، حيث تتحقق الشروط الإحصائية من حيث عشوائية الأخطاء وعدم معنوية دالة الارتباط

المصادر

- [1] Badagian, Ana Laura, "Time series segmentation procedures to detect, locate and estimate change-points". Universidad Carlos III de Madrid ,(2013). Thesis- Madrid – Spain.
- [2] Davis, R., T. Lee, and G. Rodriguez-Yam .Journal of the American Statistical Association 101 (473), 229–239. (2006). " Structural Break Estimation for Nonstationary Time Series Models".
- [3] Nelles, O. "Nonlinear System Identification from Classical Approach to Neural Network and Fuzzy Models", (2001), Springer Verlag Berlin Heidelberg Germany.
- [4] Rolf Isermann and Marco M'unchhof" Identification of Dynamic Systems : An Introduction

with Applications", (2011), Springer Heidelberg Dordrecht.Germany.

[5] Bemporad Alberto "System identification", (2011),University of Trento. Italia.

[6] فاندل ، والتر: " السلاسل الزمنية من الوجهة التطبيقية ونماذج بوكس- جنكز " (1983)، تعريب عبد المرصي حامد عزام (1992)، دار المريخ للنشر ، الرياض، المملكة العربية السعودية.

[7] Liu, L.M. "Time Series Analysis and Forecasting ",2nd ed., (2006).Scientific Computing Associates Crop., Illinois, USA.

[8] الخياط، باسل يونس ذنون " مدخل إلى نظرية الاحتمال وتطبيقاتها باستخدام MATLAB " (2013). مطبعة ابن الاثير - جامعة الموصل - العراق .

[9] البدراني ، ظافر رمضان " دراسة في تشخيص نظم السيطرة التصادفية مع إشارة خاصة إلى أسلوب فضاء الحالة والاستقرارية " جامعة الموصل ، (2002) ، أطروحة دكتوراه غير منشورة، موصل - العراق.

[10] Alkiviadis Tsimpiris and Dimitris Kugiumtzis, Journal of Statistical Software. Volume 33, Issue 5. (2010)."Measures of Analysis of Time Series (MATS): A MATLAB Toolkit for Computation of Multiple Measures on Time Series Data Bases".

Detecting and Estimation Change-point in Dynamic Systems using Segmentation Technique

Thafer Ramathan Muttar AL-Badrany , Najlaa Saad Abraham Al-Sharaby

Dep. of Statistics and Informatics, College of Computer Sciences and Mathematics, University of Mosul, Mosul, Iraq

thafer63@yahoo.com

najlaasa1986@yahoo.com

Abstract

There is no doubt that the sudden changes in dynamic behavior of the system greatly influences the choice of the appropriate model to represent this behavior, From here comes the importance of the discovery points that change in dynamic behavior of the system, Enabling the researcher of this behavior to segmentation it to groups or sectors in a manner cutting each line itself and this provides the Researcher the possibility of fitting appropriate model for each sector or department represents the best representation of the data. Discovery points of changing and estimation it is one of the critical issues, and in this research we use package of statistical and engineering criterion used in the discovery and estimation of the change points. The application by used simulation data.